HAL
open science

Évolution des gènes de la pigmentation chez les Vertébrés et développement pigmentaire chez un modèle émergent de poisson corallien, le poisson-clown Amphiprion ocellaris

Thibault Lorin

▶ **To cite this version:**

Thibault Lorin. Évolution des gènes de la pigmentation chez les Vertébrés et développement pigmentaire chez un modèle émergent de poisson corallien, le poisson-clown Amphiprion ocellaris. Embryologie et organogenèse. Université de Lyon, 2018. Français. NNT : 2018LYSEN027 . tel-02176028v2

HAL Id: tel-02176028

https://theses.hal.science/tel-02176028v2

Submitted on 7 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Numéro National de Thèse : 2018LYSEN027

THÈSE de DOCTORAT DE L'UNIVERSITÉ DE LYON

opérée par

l'École Normale Supérieure de Lyon

École Doctorale N° 340

Biologie Moléculaire, Intégrative et Cellulaire (BMIC)

Discipline de doctorat : Sciences de la Vie

Soutenue publiquement le 06/07/2018, par :

Thibault LORIN

Évolution des gènes de la pigmentation chez les Vertébrés et développement pigmentaire chez un modèle émergent de poisson corallien, le poisson-clown *Amphiprion ocellaris*

Devant le jury composé de :

Mme COURTIER-ORGOGOZO Virginie, Directrice de Recherche CNRS, Institut Jacques Monod	Rapporteure
M. LAUDET Vincent, Professeur à Sorbonne Université	Co-encadrant
Mme MANCEAU Marie, Chargée de Recherche CNRS, Collège de France	Examinatrice
M. NEGRUTIU Ioan, Professeur émérite à l'École Normale Supérieure de Lyon	Examinateur
Mme RÉTAUX Sylvie, Directrice de Recherche CNRS, Institut des Neurosciences Paris-Saclay	Examinatrice
M. ROULIN Alexandre, Professeur à l'Université de Lausanne	Rapporteur
M. VOLFF Jean-Nicolas, Professeur à l'École Normale Supérieure de Lyon	Directeur de thèse

RÉSUMÉ DE LA THÈSE

La couleur est un trait biologique essentiel et très variable qui permet à un organisme d'être adapté à son environnement biotique et abiotique. De multiples gènes influent sur la couleur ; leur évolution a été étudiée ici chez les Vertébrés en lien avec les duplications de génome survenues dans cette lignée. Alors que l'essentiel des gènes sont perdus lors de la rediploïdisation suivant une duplication totale de génome, ces travaux montrent que les gènes de la pigmentation appartiennent à des familles de gènes qui sont en moyenne plus retenues que le reste du génome. Ceci est particulièrement vrai pour les poissons Téléostéens, cette expansion du répertoire de gènes de pigmentation pouvant être en lien avec la grande diversité de cellules pigmentaires et de patrons de pigmentation observés chez ces animaux. Par ailleurs, l'étude de la pigmentation d'un organisme modèle émergent, le poisson-clown *Amphiprion ocellaris*, a été réalisée dans le cadre de cette thèse. La nature cellulaire de la bande blanche de ce poisson a été déterminée et a permis l'identification d'un nouveau gène présent presque exclusivement chez les poissons Actinoptérygiens, probablement acquis par transfert horizontal, et impliqué dans le développement d'un type de cellules pigmentaires, les iridophores. Enfin, l'acquisition de la pigmentation au cours du développement du poisson-clown *A. ocellaris*, un poisson corallien au cycle de développement marqué par une métamorphose, a été étudiée en rapport avec la signalisation thyroïdienne, qui joue un rôle majeur au cours du développement des Vertébrés.

PHD THESIS SUMMARY

Color is a key biological trait that allows individuals to adapt to their biotic and abiotic environment, and many genes have been shown to play an essential role in color acquisition in vertebrates. In this work, the consequences on this specific set of genes of the whole-genome duplications that occurred during vertebrate evolution were investigated. Whereas most genes are lost after a whole-genome duplication event, this work shows that pigmentation genes and other genes within their families have been more retained compared to the average observed genome retention. This was particularly the case for the teleost-specific whole-genome duplications. Hence, this high retention rate could be related to the high pigmentation diversity observed in the teleost fish lineage, both in terms of patterns and of pigment cell types. In addition, during this PhD thesis, pigmentation in an emerging model organism, the clownfish *Amphiprion ocellaris*, was investigated. The white color in clownfish was shown to be due to specific pigment cells, iridophores, and new iridophore genetic markers were detected. One such marker is almost exclusively restricted to ray-finned fish and could have been acquired through a horizontal gene transfer event that occurred at the onset of the evolution of this lineage. Last, in this work, the genetic basis of pigmentation acquisition was scrutinized during the development of clownfish *A. ocellaris*, and especially during metamorphosis – a major life history transition in coral reef fish. In particular, the deployment during development of thyroid hormone signaling, which is a key pathway in metamorphosis regulation in vertebrates, was examined.

Eusèbe

De Cape et de Crocs – tome 12

REMERCIEMENTS

Je commencerai cette reconnaissante anaphore par mon "coach", Jean-Nicolas. En premier lieu, un grand merci Jean-Ni pour m'avoir accueilli dans ton équipe. Merci également d'avoir su m'encadrer avec justesse, en me laissant une grande liberté scientifique tout en restant disponible et à l'écoute. Merci, enfin, pour ton soutien et tes conseils sur mes différents projets.

En d'antiques temps, tu as placé sous ton égide un marmot
Qui a grandi avec toi, un drôle de poisson et des bons mots
Ô poète catalan, daigne accepter ce murmure naissant
Qui enfle... enfle... enfle ! et te dit "merci, Vincent."

Merci à tous les membres du jury d'avoir accepté d'évaluer ce travail et particulièrement à Virginie Courtier-Orgogozo et Alexandre Roulin, rapporteurs de cette thèse.

Merci Magali de m'avoir accepté dans le bureau et Sophie pour ta bonne humeur contagieuse, et merci Corentin, la poutre ~~auvergnate~~ ardéchoise, d'avoir eu les épaules assez solides pour porter seul le poids de ma présence. Fréd', Laure, Delphine, merci pour votre gentillesse et votre disponibilité. Baptiste, Ema, Jérémy, Jossia et Sarah, bon courage pour la suite et merci pour ces bons moments passés pendant vos stages.

Carine, Marie, vous avez été là depuis le début de cette histoire. Merci infiniment pour votre pédagogie, votre patience, votre gentillesse et votre soutien ; je n'ai pas de mots. Merci également à toutes les cigognes pour avoir si souvent accepté de faire une place à un tupperware en plus.

À ma "deuxième" équipe banyulencque, un grand merci Natacha, Kévin et Pauline, pour ces agréables moments passés autour du poisson-clown et pour l'accueil pendant ces séjours toujours trop courts. Merci également Olivier et Patrick pour avoir su m'instiller votre passion du clown.

Merci Guillaume pour avoir été un encadrant de stage en or massif. Merci Sandrine et mon âne Benjamin : ce fut un plaisir d'essuyer les plâtres en votre compagnie, tout là-haut. Merci Anne pour tes précieux conseils, et merci aux "gens du premier" pour cette bonne ambiance de travail. Merci à la "A-team" pour votre gentillesse et vos compétences logistico-administratives. Enfin, merci à l'ENS de Lyon d'avoir financé ces travaux.

Merci Marc et Frédéric pour vos précieux conseils dispensés lors des comités de suivi de thèse ainsi que pour votre disponibilité. Merci aux "jodlers" Nicolas et Anna pour avoir rendu ces travaux possibles grâce au génome.

Merci Micky Love, Pierrot le Grand Faiseur de Awk One-liners, et toute cette communauté de passionnés ; vous êtes d'une aide précieuse pour les "noobs" en bio-informatique.

Parmi les "copains d'avant", et en restant dans un cadre professionnel, merci Hugo pour ton expertise réticulaire, merci Marco pour l'accueil et le "terrain" polynésien, et merci Germain pour tes clichés proprement stupéfiants.

Merci Déborah, Jean-Pierre, Cyril et Nathalie pour m'avoir permis de réaliser des enseignements dans d'aussi belles conditions.

D'un point de vue personnel, je ne peux pas ici remercier individuellement tous ceux qui me sont chers et qui ont contribué de plus ou moins loin à cette thèse. Je dis donc un grand merci aux géologues, aux profs, aux grimpeurs, aux bretons, aux naturalistes, aux randonneurs, aux cyclistes, aux basketteurs, aux fouilleurs, aux coureurs, aux brunchers, et aux écolo, qui se reconnaîtront sans mal.

François, un immense merci pour cette délicate "commande spéciale" au fusain qui représente si bien ma reconnaissance aux uns et aux autres.

Enfin, je remercie particulièrement ma famille, pour son indéfectible soutien. Et, bien sûr, Mathilde.

A
01
01
18

TABLE DES MATIÈRES

TABLE DES ILLUSTRATIONS 11
TABLE DES TABLEAUX 15
INTRODUCTION 17
POSITIONNEMENT SCIENTIFIQUE DE LA THÈSE 119
RÉSULTATS – Chapitre 1 125
RÉSULTATS – Chapitre 2 141
RÉSULTATS – Chapitre 3 167
RÉSULTATS – Chapitre 4 195
DISCUSSION GÉNÉRALE ET CONCLUSION 219
ANNEXES 239
RÉFÉRENCES BIBLIOGRAPHIQUES 269
LISTE DES ABRÉVIATIONS 287

TABLE DES ILLUSTRATIONS

Introduction

Chapitre 1

Figure 1 - Spectre de la lumière reçue à la surface de la Terre. 25
Figure 2 - Couleurs physiques et chimiques 27
Figure 3 - Classification fonctionnelle de la pigmentation illustrée par des poissons coralliens 30
Figure 4 - Radiation en trois étapes chez les poissons Cichlidés. 33
Figure 5 - Diversité des patrons de coloration des poissons coralliens. 35
Figure 6 - Le spectre du visible est peu à peu absorbé dans l'eau. 36
Figure 7 - Convergence des patrons de coloration chez les poissons coralliens. 36
Figure 8 - Changement du patron de coloration au cours du développement d'un poisson corallien. 38
Figure 9 - Voies de synthèse de la mélanine et leur contrôle 39
Figure 10 - Association entre kératinocytes et mélanocytes chez les Mammifères 39
Figure 11 – Organisation et ultrastructure des chromatophores chez un lézard Scincidae. 41
Figure 12 - Interprétation schématique de la base de la coloration verte des Lépidosauriens, Amphibiens et Téléostéens. 42
Figure 13 - Chromatophore mosaïque chez les Amphibiens. 43
Figure 14 - Les chromatophores décrits en 2018 chez les Téléostéens. 45
Figure 15 - Agencement des chromatophores dans la peau du Poisson-zèbre. 46
Figure 16 - Un même chromatophore peut être à l'origine de plusieurs couleurs. 46
Figure 17 - Couverture de *Nature* le 28 mai 1981. 48
Figure 18 - Développement des chromatophores chez le Poisson-zèbre. 49
Figure 19 - Motifs de Turing chez un poisson corallien. 51
Figure 20 - Les motifs des poissons du genre *Danio* répondent au modèle de Turing. 52
Figure 21 - Diversité des processus permettant l'acquisition d'une pigmentation fonctionnelle chez les mammifères. 53
Figure 22 - "Un seul être vous manque et tout est décoloré." 54
Figure 23 – La variation d'un trait pigmentaire n'est pas due qu'aux gènes. 55

Chapitre 2

Figure 24 - Indices synténiques permettant la détection d'un événement de duplication à grande échelle. 60
Figure 25 - Indices phylogénétiques d'une duplication de segment. 60
Figure 26 - Conservation de la synténie après une WGD. 61
Figure 27 - Mécanismes à l'origine d'une autopolyploïdisation. 62
Figure 28 - Mécanismes à l'origine d'une allopolyploïdisation. 62
Figure 29 - Conservation de la synténie entre blocs dupliqués chez la Levure et l'Arabette. 64
Figure 30 - Duplications de génomes chez les plantes à graines. 65
Figure 31 – Evénements de WGD chez les Métazoaires. 66
Figure 32 - Evolution schématique d'un chromosome de Vertébré. 68

Figure 33 - Les WGD chez les poissons Téléostéens 70
Figure 34 - Devenirs possibles d'un gène dupliqué. 73
Figure 35 - Conséquences de la duplication d'un facteur de transcription sur les réseaux de régulation de gènes. 76
Figure 36 - Modèles de perte de gènes après une duplication de génome. 79
Figure 37 - Importance de l'équilibre de dosage. 80
Figure 38 - Exemples d'organismes polyploïdes. 82
Figure 39- Isolement reproducteur suite à une duplication de génome. 83
Figure 40 - Diversité comparée des Holostéens et des Téléostéens au Permo-Trias. 85
Figure 41 - Modèle de la radiation post-WGD. 86
Figure 42 - Délai entre la rediploïdisation et la diversification des Salmonidés. 87

Chapitre 3

Figure 43 - Diversité morphologique chez les Labridae récifaux et non-récifaux. 91
Figure 44 - Cycle de vie biphasique des poissons coralliens. 92
Figure 45 – Deux espèces de poissons demoiselles. 94
Figure 46 - Phylogénie des Amphiprioninae. 96
Figure 47 - Distributions comparées d'*Amphiprion ocellaris* et *A. percula* 99
Figure 48 - Collection de morphes pigmentaires de Poissons-clowns. 99
Figure 49 - Table de développement du poisson-clown *A. ocellaris* 103
Figure 50 - Chromatophores chez *Amphiprion ocellaris*. 104
Figure 51 – Synthèse et niveaux d'expression de T4. 107
Figure 52 - La voie de signalisation des hormones thyroïdiennes. 108
Figure 53 - Développement pigmentaire et hormones thyroïdiennes chez le Poisson-zèbre. 109
Figure 54 - Effet des hormones thyroïdiennes sur la pigmentation de *A. ocellaris*. 110
Figure 55 - Principe général du RNA-Seq. 111
Figure 56 - Comparaison du développement au sein de différents phyla animaux. 114
Figure 57 - Analyse en composantes principales de données de RNA-Seq développemental. 115
Figure 58 - Principe des méthodes de clustering. 116

Discussion générale et conclusion

Figure 59 - Hypothèses pour l'évolution des types cellulaires pigmentaires chez les Vertébrés 231
Figure 60 - Niveaux d'expression relatifs des TR au cours du développement de différents poissons Téléostéens. 233

TABLE DES TABLEAUX

Introduction

Chapitre 1

Table 1 - Propriétés des gènes préférentiellement retenus après une WGD. 81

Chapitre 3

Table 2 - Avantages du RNA-Seq par rapport aux micro-arrays. 113

INTRODUCTION

PLAN DÉTAILLÉ DE L'INTRODUCTION

Chapitre 1 – La couleur chez les Vertébrés

I. La couleur est un trait biologique majeur des êtres vivants…

1. Lumière, couleur, pigmentation : quelques définitions
- a. La lumière est un phénomène physique double
- b. La couleur est une définition sensorielle, propre à chaque individu
- c. Les êtres vivants ne sont le plus souvent pas à l'origine de la lumière qu'ils émettent
- d. Un pigment est une molécule absorbant la lumière
- e. … à la différence des couleurs structurelles
- f. Conclusion : la pigmentation suppose une interaction entre pigment et structure du tissu

2. Les fonctions adaptatives de la couleur
- a. Une approche historique largement influencée par Wallace
- b. Les différentes fonctions adaptatives de la couleur selon Wallace, revisitées
  - *i. "Protective colors" : le camouflage*
  - *ii. "Warning colors" : les couleurs d'avertissement*
  - *iii. "Sexual colors" : les couleurs sexuelles*
  - *iv. "Typical colors" : une catégorie aujourd'hui largement remaniée*
- c. Les catégories de Wallace ne sont pas exhaustives

3. Evolution de la couleur
- a. La couleur est due à une interaction entre sélection naturelle, sexuelle et sociale
  - *i. La sélection naturelle*
  - *ii. La sélection sexuelle*
  - *iii. La sélection sociale*
  - *iv. Conflits entre modes de sélection multiples : application à la coloration*
- b. La couleur est un trait qui évolue souvent vite

II. … particulièrement diversifié chez les poissons Téléostéens

1. Les poissons Téléostéens constituent l'essentiel des Vertébrés…
- a. Les poissons Téléostéens sont très diversifiés…
- b. … et notamment les poissons coralliens

2. … et ont des patrons de coloration variés et variables…
- a. Diversité des patrons de pigmentation
  - *i. Il existe de nombreux patrons de pigmentation*
  - *ii. … souvent convergents*
- b. Changements associés au statut social ou à la reproduction
- c. Polyphénisme intraspécifique
  - *i. Chez l'adulte*
  - *ii. Au cours du développement*
- d. Bases physiologiques des changements pigmentaires

3. … notamment dus à un plus grand nombre de chromatophores et de pigments
- a. Mammifères
- b. Oiseaux
- c. Lépidosauriens
- d. Amphibiens
- e. Téléostéens

III. … et aux bases génétiques et développementales multiples

1. Importance de l'étude de la coloration dans l'histoire des sciences
2. Une meilleure compréhension de la différentiation des chromatophores grâce aux nouveaux modèles
- a. Origine développementale des chromatophores
- b. Les relations entre chromatophores
  - *i. Relations ontogénétiques entre les chromatophores*
  - *ii. Les relations cellulaires entre chromatophores sont à l'origine des patrons de pigmentation*

3. Vers une définition d'un "gène de la pigmentation" ?
- a. Origine du terme
- b. Les gènes de la pigmentation gouvernent des processus multiples
  - *i. Les gènes de la pigmentation sont parfois associés seuls à une fonction pigmentaire…*
  - *ii. … et ne sont pas uniquement associés à la pigmentation, ou pas seuls*
- c. De nouveaux gènes à découvrir…
- d. … même si les gènes ne sont pas les seuls acteurs de la pigmentation

Chapitre 2 – Les duplications de génomes et leurs conséquences évolutives

I. Les duplications complètes de génomes sont des événements fréquents chez les Eucaryotes

1. Pré-requis à la détection d'une duplication de génome
2. Comment détecter un événement de duplication complète de génome par l'étude des gènes ?
- a. Les témoins cytologiques
- b. Les témoins synténiques
- c. Les témoins phylogénétiques
- d. Conclusion : le génome, un palimpseste

3. Comment les WGD apparaissent-elles ?
- a. Autopolyploïdie
- b. Allopolyploïdie

4. Quelles sont les lignées concernées par les WGD ?
- a. Avant-propos
- b. Paramécie
- c. Angiospermes et plantes à graines (Spermatophytes)
- d. Levure
- e. Les WGD chez les Animaux non-Vertébrés

II. Les duplications complètes de génomes chez les Vertébrés

1. Bref historique de la découverte des WGD chez les Vertébrés
2. Les WGD chez les "poissons"
a. Importance de l'étude des WGD chez les Téléostéens dans la compréhension des WGD des Vertébrés
b. Mise en évidence de la Ts3R
c. D'autres WGD chez les Téléostéens
i. Salmonidés
ii. Cyprinidés
iii. Catostomidae et Cobitidae
d. Des WGD chez les "poissons non-Téléostéens"
3. Les WGD chez les autres Vertébrés
4. Conclusion sur les WGD chez les Vertébrés

III. Les conséquences évolutives des WGD diffèrent selon les lignées eucaryotes

A. Les WGD permettent l'apparition de nouveaux gènes qui fournissent un substrat propice à la diversification génétique

1. Les WGD sont une source de nouvelles séquences…
a. Les WGD constituent une source majeure de nouveaux gènes
i. Une WGD induit le doublement des gènes et des séquences régulatrices d'un individu…
ii. … qui ont des devenirs variables …
iii. … et pouvant être estimés au niveau génomique
b. L'état polyploïde est transitoire et la diploïdie est l'état stable
c. Les WGD peuvent modifier les réseaux de régulation de gènes
d. Les WGD ne sont pas les seuls processus à l'origine de la formation de nouvelles séquences
i. La formation de nouveaux gènes par duplication est majoritaire…
ii. … mais n'explique pas seule l'apparition des nouveaux gènes
2. L'essentiel des gènes issus d'une WGD est perdu
a. Les mécanismes de pertes après une WGD : perte rapide ou graduelle ?
b. Les facteurs à l'origine de la rétention préférentielle de certains gènes
i. La rétention peut être liée à un changement de fonction…
ii. … ou être corrélée à des propriétés intrinsèques des protéines…
iii. … et souvent inverses des SSD

B. La diversification génétique a parfois des conséquences sur la diversification phénotypique

1. Les WGD ont un impact possible sur la spéciation
a. La polyploïdisation peut entraîner des variations de fitness sur des temps évolutifs courts,..
b. … et être à l'origine de spéciation sympatrique par des pertes de gènes différentielles
2. Les WGD sont parfois corrélées à la diversification phénotypique d'un groupe…
3. … mais pas toujours, ou pas immédiatement
a. Les WGD ne sont pas nécessairement associées à une augmentation du taux de spéciation
b. La biodiversité actuelle n'est qu'un instantané de la paléo-biodiversité
c. La diversification peut être graduelle après une WGD
d. Une WGD peut induire une radiation sans diversification phénotypique
e. L'évolution phénotypique ne succède pas toujours immédiatement à un événement de WGD
f. Conclusion : les conséquences d'une WGD sur l'évolution d'un groupe sont hautement variables

Chapitre 3 – La pigmentation d'un modèle émergent de poisson corallien, le poisson-clown *Amphiprion ocellaris*

I. Le poisson-clown *Amphiprion ocellaris*, une espèce modèle émergente de poisson corallien

1. *A. ocellaris* est une espèce de poisson corallien
- a. Les récifs coralliens sont des écosystèmes…
- b. … où les poissons abondent pendant l'essentiel de leur cycle de vie…
- c. … mais pas pendant la phase larvaire

2. *A. ocellaris* est un poisson-demoiselle du genre *Amphiprion*
- a. Biologie des poissons-demoiselles, ou Pomacentridae
- b. Les poissons-demoiselles incluent les espèces du genre *Amphiprion*, ou "poissons-clowns"
  - *i. le genre* Amphiprion *contient une trentaine d'espèces…*
  - *ii. … à l'écologie différente du reste des Pomacentridae*

3. Spécificités de la biologie de *A. ocellaris*
- a. Le poisson-clown *A. ocellaris* : une icône en aquariophilie et dans l'imaginaire collectif
- b. La définition taxonomique d'*A. ocellaris* n'est pas claire
- c. La pigmentation adulte d'*A. ocellaris*

4. *A. ocellaris* est un intéressant modèle émergent
- a. Qu'est-ce qu'un organisme modèle ?
- b. Le poisson-clown *A. ocellaris*, un modèle émergent
- c. État de l'art scientifique sur *A. ocellaris*
  - *i. Structure sociale*
  - *ii. Physiologie et développement*

II. Le développement d'*Amphiprion ocellaris* est caractérisé par une métamorphose

1. Les changements morpho-anatomiques ayant lieu au cours du développement d'*A. ocellaris*…
- a. Les stades embryonnaires diffèrent des jours post-fécondation
- b. Acquisition de la pigmentation au cours du développement d'*A. ocellaris*

2. … indiquent qu'il est caractérisé par une métamorphose…
- a. Aspects tissulaires, morphologiques et écologiques de la métamorphose chez les Vertébrés
- b. Aspects physiologiques
  - *i. Un contrôle par les hormones thyroïdiennes*
  - *ii. Interaction de la voie des TH avec l'environnement*
  - *iii. Influence de la voie des hormones thyroïdiennes sur la pigmentation en particulier*

3. … qui est un phénomène biologique qui peut être étudié par RNA-Seq
- a. Le RNA-Seq est une méthode "omique" permettant d'étudier l'expression des gènes
  - *i. Principe général du RNA-Seq*
  - *ii. Avantages du RNA-Seq et application à l'étude du poisson-clown* A. ocellaris
  - *iii. Remarque conclusive*
- b. Le RNA-Seq est une méthode adaptée à l'étude du développement
  - *i. Apports du RNA-Seq à la comparaison du développement entre espèces*
  - *ii. Apports du RNA-Seq à l'étude du développement au sein d'une espèce*
- c. Le RNA-Seq est une méthode rarement utilisée pour étudier la métamorphose

INTRODUCTION

CHAPITRE 1

La couleur chez les Vertébrés

"The biology of color is a field that typifies modern research: curiosity-led, technology-driven, multilevel, interdisciplinary, and integrative." (Cuthill *et al.*, 2017).

I. La couleur est un trait biologique majeur des êtres vivants

1. Lumière, couleur, pigmentation : quelques définitions

a. La lumière est un phénomène physique double

Au cours de ma thèse, j'ai en partie travaillé sur les gènes de la pigmentation. J'aimerais donc, avant de parler plus avant de la pigmentation, revenir sur quelques processus physiques qui permettent de comprendre, sans entrer dans le détail des équations, différents termes tels que "lumière", "couleur" ou "pigmentation". Un ouvrage très complet, admirablement rédigé et non dénué d'humour détaille tout cela (Johnsen, 2012).

Avant de parler de pigmentation, il faut évoquer la lumière, ainsi que la couleur. La lumière est un phénomène physique double, ayant à la fois une nature ondulatoire (onde électromagnétique) et corpusculaire (photons), et j'utiliserai alternativement l'une ou l'autre pour comprendre un phénomène.

La **lumière**, en physique, est souvent considérée comme une onde électromagnétique pouvant prendre toutes les longueurs d'onde données, depuis les plus longues (ondes radio), jusqu'aux plus courtes (rayons gamma). Une partie de ce spectre, définie entre environ 400 et 800 nm de longueur d'onde, est la lumière dite "visible", car perceptible par notre œil, à nous, humains. Les bornes inférieure et supérieure de ce domaine sont variables en fonction des individus.

Les ondes lumineuses dont la longueur d'onde est de 400 nm ont des propriétés physiques différentes de celles dont la longueur d'onde est de 800 nanomètres : les photons correspondant à une lumière à 400 nanomètres ont une énergie plus élevée que ceux de la lumière à 800 nanomètres. Ces propriétés sont physiques et intrinsèques du rayon lumineux.

b. La couleur est une définition sensorielle, propre à chaque individu

La **couleur** est la façon dont chacun de nous perçoit un objet lumineux, quelle que soit sa longueur d'onde (ou ses longueurs d'ondes) dans le domaine du visible (400-800 nm). La couleur est donc une perception d'un signal environnemental et qui est propre à chaque individu : chaque être humain voit les couleurs différemment, avec des faibles différences de perception entre deux individus non atteints d'anomalies de vision, et, sinon, avec des plus fortes différences. Quoi qu'il arrive, nous, humains, voyons tous notre environnement différemment coloré.

La lumière visible est, on l'a vu, définie par la gamme perçue par l'œil humain, c'est-à-dire la gamme de longueurs d'ondes pour laquelle un signal capté au niveau de l'œil peut induire une réponse du cerveau. D'autres espèces, qui ont des capacités sensorielles différentes, peuvent "voir" (au sens de "être en mesure de répondre à des signaux électromagnétiques") à des longueurs d'ondes différentes. Par exemple, les abeilles sont bien connues pour pouvoir voir des signaux dans l'ultraviolet proche (à moins de 400 nm). Certains poissons, nous le verrons, le peuvent également (Losey *et al.*, 1999). De l'autre côté du spectre visible, certains serpents peuvent détecter les infrarouges, à plus de 800 nm (Gracheva *et al.*, 2010).

c. Les êtres vivants ne sont le plus souvent pas à l'origine de la lumière qu'ils émettent

La couleur est donc l'interaction entre un phénomène optique, la lumière, et la perception que chacun de nous en a. Pour qu'un objet, biologique ou non, nous apparaisse coloré, il faut donc qu'il "produise" de la lumière. Sauf dans le cas de la bioluminescence, où les organismes eux-mêmes produisent et émettent *vraiment* un photon par une réaction enzymatique (que nous ne détaillerons pas ici), tout objet biologique ne produit pas réellement de la lumière et a une couleur due à un phénomène "d'absorption" des rayons lumineux. D'où viennent-ils ?

Sur Terre, outre la bioluminescence, la seule autre source de lumière visible est le Soleil : celui-ci n'émet pas une lumière pure mais un spectre de nombreuses longueurs d'ondes avec un pic à 504 nm, dans le "vert" (pour nous), mais allant de l'ultraviolet à l'infrarouge. Pourquoi, en particulier, cette gamme de longueurs d'ondes ? Pourquoi le Soleil n'émettrait-il pas à des longueurs d'ondes inférieures ou supérieures ? C'est car tout corps chaud émet des radiations :

une loi physique, la loi de Wien, postule que la longueur d'onde de ces radiations dépend de la température de l'objet. Nous, humains, avec une température de 37°C, émettons dans l'infrarouge (à 9 350 nm). Le Soleil, ayant une température de surface de plus de 5 000 °C, émet à une longueur d'onde plus faible (i.e. une énergie plus élevée) ayant un pic autour de 500 nm mais émettant dans tout le visible.

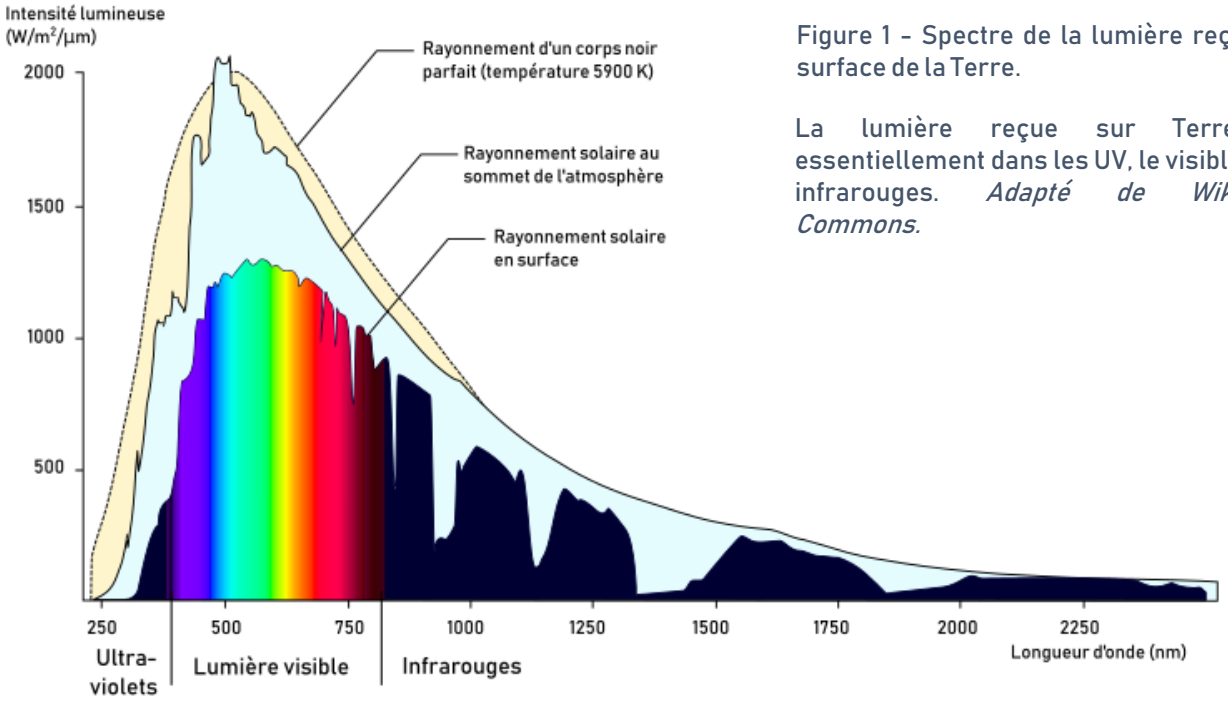


Figure 1 - Spectre de la lumière reçue à la surface de la Terre.

La lumière reçue sur Terre est essentiellement dans les UV, le visible et les infrarouges. *Adapté de Wikimedia Commons.*

Sur Terre, les êtres vivants sont donc adaptés à détecter des signaux électromagnétiques à des longueurs d'ondes "solaires" (Ensminger, 2001). Ainsi, même les Champignons, qui n'ont pas de système visuel, ont des photorécepteurs qui captent des longueurs d'ondes du visible ou du proche ultraviolet (Tisch and Schmoll, 2010). Dans un autre système stellaire, avec une étoile autre que le Soleil qui serait plus chaude ou plus froide, il est fort probable que, s'il y avait des êtres vivants et qu'ils étaient en mesure de voir, leurs systèmes de perception visuelle seraient décalés vers le maximum d'émission de leur étoile.

En conclusion, une feuille, un papillon ou un poisson, pourtant "colorés", ne sont pas eux-mêmes source de leur lumière et ne renvoient qu'une partie de la lumière (solaire) qu'ils reçoivent. Ce qui les rend visibles, à notre œil, est le résultat de l'interaction de leur structure physique avec la lumière incidente du Soleil.

d. Un pigment est une molécule absorbant la lumière…

Considérons un organisme qui reçoit de la lumière solaire constituée de toutes les longueurs d'onde visibles (lumière dite "blanche"). Si la lumière traverse cet objet de part en part, totalement, l'organisme nous paraîtra transparent : rien, dans cet organisme, ne le rend différent de l'environnement dans lequel il est et notre œil ne peut pas le distinguer. C'est le cas de certains Cnidaires, quasiment transparents.

Supposons maintenant qu'au sein de l'organisme soit présente une molécule, qui, par sa structure, peut absorber une partie du spectre de la lumière incidente. Cette molécule correspond à la définition d'un **pigment**. Une partie de la lumière qui traverse cet organisme sera alors absorbée par le pigment. L'autre partie du spectre, transmise à travers le

tissu pigmenté, permet alors de nous donner une impression de couleur. Par exemple, si le pigment absorbe toutes les longueurs d'ondes sauf celle correspondant au vert, l'organisme nous paraît vert (c'est l'exemple de la chlorophylle). On note donc qu'un pigment est défini par ce qu'il transmet, pas par ce qu'il absorbe.

Cela cache une subtilité. En effet, lors de l'absorption, une seule partie de la longueur d'onde est effectivement absorbée, mais *l'intensité* de la lumière diminue pendant la traversée du tissu : ainsi, si l'organisme est épais, celui-ci ne nous paraîtra pas vert, mais noir ! Un pigment ne permet, rappelons-nous, que d'absorber la lumière qui, peu à peu, perd en intensité. Si un mur noir est peint avec un pigment, le mur reste noir. Pour que le pigment donne un rendu coloré, il faut qu'il soit peint sur un mur… blanc, ou clair.

Cette propriété physique est également valable dans le vivant : pour qu'un organisme "épais" apparaisse coloré, et non pas noir, il faut que, sous la couche pigmentée, soit présente une surface qui réfléchit la lumière incidente. Chez les Animaux, l'ensemble des tissus conjonctifs sont, structurellement, réfléchissants. Cela explique, notamment, pourquoi les individus albinos ne sont pas transparents, mais blancs : dépourvus de pigment, seul leur tissu conjonctif, réfléchissant la lumière incidente, est visible et donc blanc. De même, chez les Oiseaux, la structure des plumes est constituée d'un mélange kératine/air qui, structurellement, réfléchit la lumière visible. Ainsi, bien qu'une plume isolée et dépigmentée soit fine, elle paraît blanche, et non transparente.

e. … à la différence des couleurs structurelles

Dans ce qui précède, la couleur est donc due à une molécule, appelée pigment, qui *absorbe* la lumière. Toutefois, il existe d'autres types de processus physiques permettant de "transformer" une lumière incidente blanche en couleur. Ces processus ne sont pas liés à *l'absorption* de la lumière par une molécule, mais à des propriétés de surface de la structure rencontrée. En rencontrant un point de la surface, la lumière est alors diffractée, ou réfractée, et l'onde interagit alors avec celle issue de la diffraction ou réfraction à un autre point de la surface. Les deux faisceaux lumineux peuvent alors physiquement interagir, *interférer*, et les interférences ainsi générées, selon l'angle d'observation et la structure initiale de la surface, peuvent être à l'origine de couleurs, voire d'iridescences. Ce phénomène est en jeu, par exemple, dans les ailes d'oiseaux ou de papillons (Prum *et al.*, 1998, 1999; Wasik *et al.*, 2014). Ce phénomène est également en jeu dans les cristaux de guanine de certains copépodes du genre *Sapphirina* qui paraissent, selon l'angle, totalement bleus ou totalement transparents. On note que les couleurs générées "structuralement" peuvent de plus s'ajouter à celles générées "chimiquement" par les pigments.

Ces couleurs structurales, sont, à l'instar des couleurs dues aux pigments, des couleurs qui ne sont pas dues à l'émission de lumière par les tissus. Les tissus sont "passifs" vis-à-vis de la lumière et n'en génèrent pas. Ils en exploitent toutefois ses propriétés.

f. Conclusion : la pigmentation diffère de la coloration et suppose une interaction entre pigment et structure du tissu

La **coloration** d'un organisme définit donc l'ensemble des facteurs physiques et chimiques qui conduisent à (i) la réception de certaines longueurs d'ondes par un tissu (ii) le renvoi par celui-ci d'une longueur d'onde vers l'œil, suite à des processus structuraux et/ou pigmentaires ; (iii) la transformation, par l'œil puis par le cerveau, d'un signal physique en une perception de couleur, propre à chacun.

La **pigmentation** est un terme plus spécifique qui caractérise le rôle joué par les pigments dans l'obtention d'une coloration. On note que la couleur n'est pas générée par les pigments seuls et que ces derniers agissent de concert avec la structure des tissus : le rendu d'une couleur vive requiert leur interaction physique proche avec un tissu sous-jacent qui est lui, réfléchissant. On parlera alors de couleur "chimique", même si le terme est partiellement impropre puisque, on l'a vu, il y a une intervention physique du tissu sous-jacent. De plus, une couleur peut être due à une interaction entre ces

couleurs "chimiques" et des couleurs purement physiques dues à des adaptations structurelles des tissus, produisant des interférences (Shawkey and Hill, 2005; Shamim *et al.*, 2014). Enfin, parfois, certaines cellules dites "pigmentaires" contiennent des molécules qui induisent une coloration "physique", c'est-à-dire sans absorption de la lumière par un pigment : c'est le cas des cristaux de purines au sein des iridophores, que nous verrons par la suite.

La limite entre couleur "chimique" et "structurale" peut donc être fine et la dichotomie souvent faite entre ces deux types de couleur est un raccourci sémantique pratique pouvant s'avérer fallacieux. Nous verrons par la suite que les couleurs chez les Mammifères sont quasiment uniquement basées sur des pigments : le terme de "pigmentation" est donc souvent employé comme un synonyme de coloration. Dans d'autres lignées, comme chez les Oiseaux ou chez les Insectes, s'ajoutent aux pigments des couleurs "physiques" très diverses : l'emploi du terme "coloration" est donc plus indiqué (Hill and McGraw, 2006; D'Alba *et al.*, 2012).

Figure 2 - Couleurs physiques et chimiques

Les couleurs dites "chimiques" sont en réalité dues à l'interaction entre des pigments et la structuration physique du tissu sous-jacent. Les couleurs "physiques" sont dues à des structurations particulières des tissus qui interfèrent avec la lumière sans l'absorber. *Cliché : T. Lorin.*

2. Les fonctions adaptatives de la couleur

a. Une approche historique largement influencée par Wallace

Quel que soit l'angle par lequel il s'intéresse au sujet, l'incroyable diversité de la couleur des êtres vivants a toujours intéressé l'Homme : le naturaliste, évidemment, mais également le chasseur cueilleur qui devait faire la différence entre une baie comestible et non comestible, ou l'artiste qui y voit une opportunité de retranscrire la beauté de la nature. La diversité des couleurs dans le monde animal a pendant longtemps été constatée, utilisée, voire étudiée, mais sa compréhension dans un cadre adaptatif n'a pas pu avoir lieu avant de comprendre les mécanismes en jeu dans l'évolution des populations naturelles au milieu du XIXème siècle. Si Darwin est souvent considéré comme le "père" de la théorie de la sélection naturelle, c'est un de ses contemporains, lui aussi convaincu par cette théorie (puisqu'il l'a élaborée indépendamment), qui, le premier, s'attache à étudier la "coloration" des animaux dans un cadre adaptatif : Alfred Russell Wallace (1823-1913) (Wallace, 1877). Darwin ne sera toutefois pas en reste sur l'étude de la pigmentation puisqu'il a de son côté travaillé sur les variétés de pigeons domestiques : à l'époque, "*Everyone is interested in pigeons*", et notamment pour la diversité des morphes pigmentaires (Darwin, 1868; Domyan and Shapiro, 2017). Membre de diverses sociétés

ornithologiques, sur les pigeons ou les canaris, Darwin fut en outre un ornithologue passionné. S'il est connu pour ses travaux sur les "pinsons" des Galapagos, Darwin s'est de longue date intéressé aux oiseaux, à leur forme et à leur couleur (Frith, 2016).

Wallace définit donc, à l'époque, des catégories de couleur ayant une utilité biologique. Cela va à l'encontre de la pensée victorienne de l'époque, statuant que les couleurs ont été créées pour l'agrément visuel qu'elles procurent, ou sont le résultat d'une exposition prolongée à la chaleur ou à la lumière (Caro, 2017). Wallace propose cinq grandes catégories de coloration pour les animaux et les plantes : couleurs de protection ; couleurs d'avertissement ; couleurs sexuelles ; "couleurs typiques" ; et couleurs d'attraction, chez les plantes et les insectes. Ces grandes catégories sont encore largement valables aujourd'hui, même si elles ont été quelque peu affinées.

b. Les différentes fonctions adaptatives de la couleur selon Wallace, revisitées

La couleur d'un organisme, à la fois au niveau de ses teintes comme de son patron de coloration, peut avoir plusieurs fonctions qui lui confèrent un avantage adaptatif, c'est-à-dire qu'il a un gain de fitness par rapport à s'il n'avait pas cette (ou ces) couleur(s). Une revue a récemment été publiée sur le sujet dans Science et décrit quelques grandes fonctions de la couleur, ainsi que les bases biologiques sous-jacentes et les "espaces à conquérir", c'est-à-dire les nouvelles pistes à explorer pour une meilleure compréhension de la couleur (Cuthill *et al.*, 2017). L'étude de la couleur d'un organisme, parce qu'elle est un trait biologique fondamental, a été réalisée à toutes les échelles : depuis les bases génétiques et épigénétiques jusqu'aux fonctions évolutives, en passant par le support moléculaire et cellulaire, l'encodage sensoriel et les aspects comportementaux. L'étude de la couleur des êtres vivants est donc un champ interdisciplinaire nécessitant l'interaction entre des spécialistes de différentes disciplines (Cuthill *et al.*, 2017).

Les fonctions adaptatives de la couleur ont déjà été décrites à maintes reprises chez divers groupes d'animaux (Caro, 2017; Cuthill *et al.*, 2017). Ici, puisqu'une grande partie de mon exposé leur sera consacrée, je prendrai des exemples chez les poissons Téléostéens pour illustrer chacune de ces catégories. Je les introduis dans le même ordre que Wallace, et donne leur nom anglais et une traduction française, lorsqu'elle existe. Les catégories de Wallace sont, quand il y a lieu, revisitées "façon XXIème siècle" (Caro, 2017).

i. "Protective colors" : le camouflage

La catégorie des couleurs de protection, les "protective colors" de Wallace, regroupe l'ensemble des stratégies de **camouflage**. La définition de ce dernier est variable selon les auteurs. Je prendrai ici celle de (Stevens and Merilaita, 2009), qui définit le camouflage comme l'ensemble des adaptations de couleur, de luminosité et de patron de couleur permettant à l'organisme de ne pas être détecté ou reconnu dans son milieu. Le cryptisme de coloration concerne les types de camouflage qui empêchent la détection. Cette catégorie inclut les organismes ayant une coloration permettant un camouflage spécialisé pour un arrière-plan donné (Emerson *et al.*, 1990), ou au contraire un camouflage généraliste leur permettant de s'adapter à divers arrière-plans, et reposant le plus souvent sur un changement de couleur. Les stratégies de contre-coloration (*countershading*) (Rowland, 2011; Manceau *et al.*, 2011), visibles chez les Requins ou chez le poisson *Chaetodon lunula* (Hamilton and Peterman, 1971), rentrent dans cette catégorie, de même que les stratégies de coloration perturbatrice (*disruptive coloration*) : l'individu, en arborant des signaux visuels très contrastés par rapport au reste de son corps, masque son contour et donc est paradoxalement camouflé. C'est le cas du Guépard, ou, chez les Poissons, de *Dascyllus aruanus* (Cuthill *et al.*, 2005; Phillips *et al.*, 2017). Les animaux transparents représentent un cas extrême de camouflage mais les exemples sont rares chez les Vertébrés adultes (voir toutefois le poisson amazonien *Microphilypnus amazonicus*). Ils concernent par contre les larves, notamment celles de poissons coralliens (Holzer *et al.*, 2017).

La mascarade (*masquerade*) est un autre type de camouflage où la détection a lieu mais où l'animal n'est pas reconnu car ressemblant à un objet de l'environnement (Stevens and Merilaita, 2009; Skelhorn *et al.*, 2010).

ii. "Warning colors" : les couleurs d'avertissement

Les *warning colors* de Wallace sont définies en deux grands groupes :

1. les couleurs portées par des "*creatures specially protected*" qui signalent ainsi leur toxicité, ou leur non-comestibilité, à l'égard des prédateurs. L'aposématisme en est un exemple : les couleurs colorées d'un individu permettent un apprentissage du prédateur qui se détourne de la proie (Speed, 2000). Etymologiquement, la proie "envoie le prédateur au loin par un signe distinctif" (le mot vient du grec *apo-*, "loin", et *sêma-*, "signe") (Poulton, 1890). Le mimétisme müllérien exploite l'aposématisme (Müller, 1879) : lorsque deux espèces toxiques se ressemblent, les prédateurs apprennent d'autant plus vite à les éviter, leur conférant un avantage sélectif.

2. les couleurs portées par les "*defenseless creatures*", qui miment les premières. C'est la définition du mimétisme batésien, premier mimétisme découvert par Henry W. Bates (Bates, 1862). Il existe des formes de mimétisme intermédiaires entre le mimétisme batésien et müllérien, que je ne détaille pas ici (Rowland *et al.*, 2010).

iii. "Sexual colors" : les couleurs sexuelles

Les *sexual colors* sont aujourd'hui considérées comme les couleurs dues au dimorphisme sexuel. Lorsque celui-ci est prononcé, un des sexes (le plus souvent, mâle), arbore des couleurs vives, alors que l'autre sexe est, lui, camouflé. Ce dimorphisme est un facteur clef de la sélection sexuelle, sur laquelle je reviendrai.

iv. "Typical colors" : une catégorie aujourd'hui largement remaniée

Les couleurs "typiques" de Wallace peuvent aujourd'hui être définies en deux sous-ensembles qui ne semblent pas être directement liés aux catégories sus-citées.

1. Les signaux colorés de communication intra-spécifique, en dehors du contexte de la sélection sexuelle : les signaux de compétition intra-spécifique (Tibbetts and Dale, 2004; Senar, 2006) ; des signaux attirant les conspécifiques, augmentant la probabilité collective de détecter et échapper à un prédateur (DE L. Brooke, 1998) ; des signaux servant à signaler sa jeunesse et donc sa compétitivité faible (Rohwer, 1978) ; enfin, des signaux permettant aux juvéniles de reconnaître leurs parents.

2. Les signaux de communication interspécifique, permettant soit d'éviter l'hybridation entre espèces proches (Wagner *et al.*, 2007; Santana *et al.*, 2012), soit de leurrer les prédateurs (Powell, 1982). Dans cette sous-catégorie seraient aujourd'hui "rangées" les *attractive colors*, couleurs permettant aux plantes d'attirer les pollinisateurs ou les espèces dispersant les fruits.

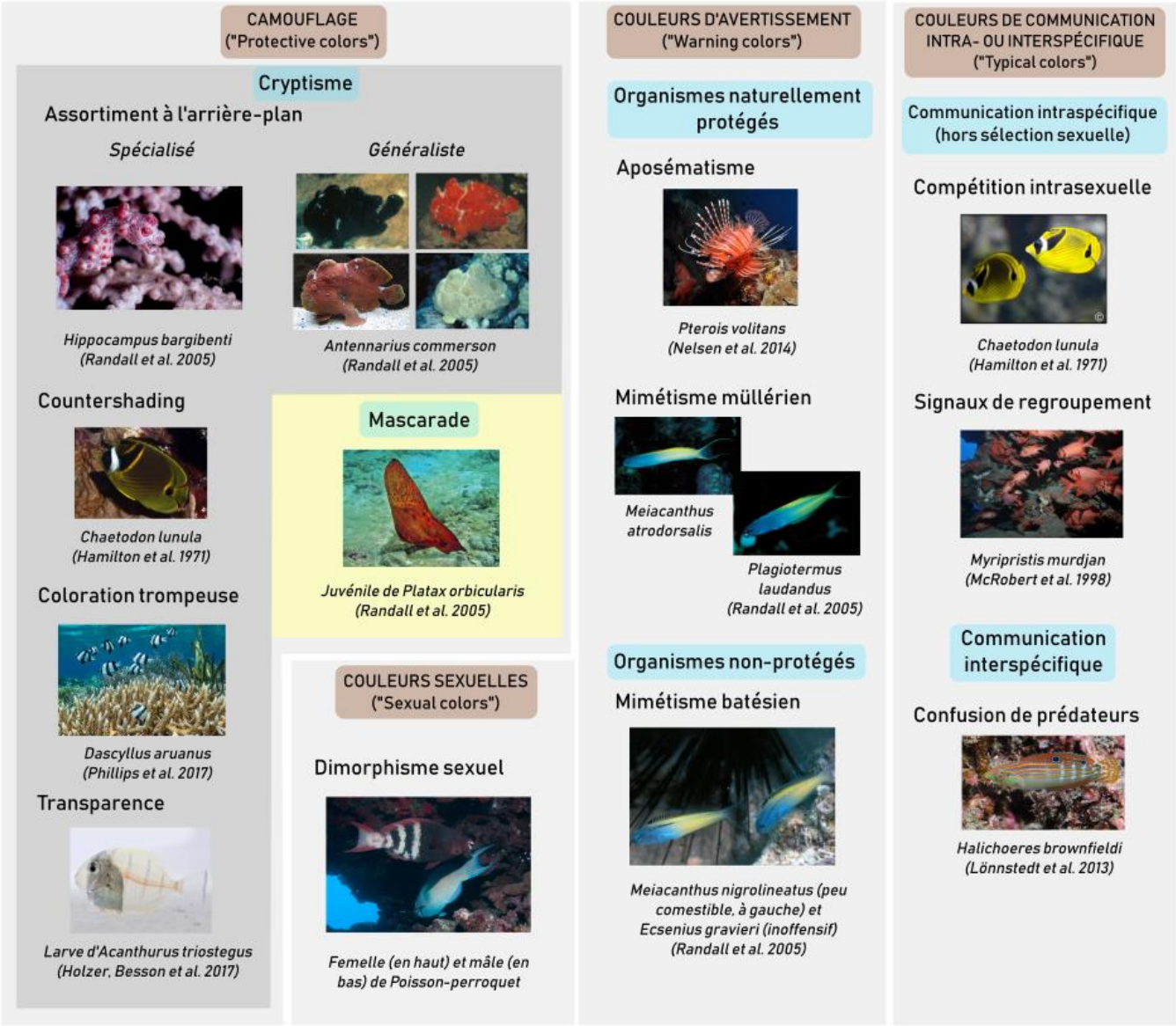


Figure 3 - Classification fonctionnelle de la pigmentation illustrée par des poissons coralliens

Les différentes catégories fonctionnelles définies par A. R. Wallace sont illustrées par des exemples de poissons coralliens. *Droits réservés : Alamy (D. aruanus), RecifAtHome (C. lunula), Marc Besson (A. triostegus), Randall et al. (2005) (P. orbicularis), A. commerson, espèces mimétiques), Autres clichés : Germain Boussarie.*

c. Les catégories de Wallace ne sont pas exhaustives

Une fonction adaptative de la pigmentation non envisagée par Wallace tient aux propriétés même des pigments : ceux-ci *absorbent* la lumière incidente et n'en renvoient qu'une partie. Cette absorption conduit rarement à l'émission d'un photon en retour (fluorescence) : dans la majorité des cas, l'énergie absorbée se dissipe sous forme de chaleur dans la molécule. Ainsi, chez les Animaux, la mélanine a un rôle de thermorégulation majeur. Les animaux sombres se réchauffent plus vite que les animaux clairs, qui absorbent moins de chaleur, et leur température corporelle, toutes conditions égales par ailleurs, peut être plus élevée (Hetem *et al.*, 2009). Toutefois, la thermorégulation est un processus complexe impliquant des facteurs physiologiques, comportementaux et écologiques, et la seule lumière incidente ne suffit pas à expliquer les variations de température (Dreiss *et al.*, 2016).

En outre, la coloration peut avoir une fonction de photoprotection : en absorbant les rayons ultraviolets de la lumière solaire, la mélanine neutralise l'effet mutagène de ses rayons (Brenner and Hearing, 2008).

3. Évolution de la couleur

Nous l'avons vu, il existe de multiples fonctions associées à la couleur. Celles-ci sont considérées comme "adaptatives", dans le sens où elles confèrent un gain de valeur sélective (fitness) à l'individu (Hubbard *et al.*, 2010). Quelles sont les forces évolutives sous-tendant ce trait ? Évoquées entre les lignes jusqu'ici, je les développe quelque peu maintenant. Je précise auparavant que les patrons de pigmentation ne sont pas qu'adaptatifs et des cas de polymorphisme de couleur soumis à la dérive ont été documentés chez la grenouille *Rana pipiens* (Hoffman *et al.*, 2006).

a. La couleur est due à une interaction entre sélection naturelle, sexuelle et sociale

i. La sélection naturelle

Deux forces évolutives principales sont souvent évoquées (et souvent opposés) pour expliquer l'évolution des patrons de coloration : la sélection naturelle, et la sélection sexuelle.

La sélection naturelle repose sur des facteurs écologiques : un individu mieux camouflé dans son environnement aura plus de chances de ne pas être vu par un prédateur, donc de survivre, donc de se reproduire, et de transmettre ses gènes associés à un camouflage efficace. La descendance sera donc plus à même d'être efficacement camouflée dans le même environnement, et, à terme, les individus seront adaptés à celui-ci. Ce mode de sélection fut développé de concert par Wallace et Darwin.

ii. La sélection sexuelle

L'explication des couleurs très voyantes portées par certains individus d'un sexe (souvent mâle) semble aller contre la logique énoncée ci-dessus. La "solution" à ce problème logique réside dans l'importance du choix de l'autre sexe (souvent femelle) vers les individus les plus colorés : c'est la sélection intersexuelle (qui agit sur d'autres traits que la couleur, comme certains attributs physiques, mais que je n'aborde pas ici). Il existe une autre forme de la sélection sexuelle dans laquelle les individus du même sexe sont en compétition pour l'accès à la reproduction : on parle de sélection intrasexuelle.

La première personne à suggérer l'importance de la sélection sexuelle et du choix des femelles est Darwin (Darwin, 1871), allant à l'encontre de la pensée majoritaire victorienne de l'époque, dans laquelle la femme n'est pas considérée comme à même de faire un choix. Darwin ayant postulé une continuité "psychologique, intellectuelle et morale" entre l'Homme et les autres êtres vivants, suggérer qu'une femelle d'un animal autre que l'Homme puisse choisir son partenaire, c'est accepter que les femmes soient également dotées d'un libre-arbitre (Milam, 2010) ! Le choix des femelles, selon Darwin, pousse les mâles à développer des caractères de plus en plus "esthétiques".

Pour Wallace, à l'inverse, le choix du partenaire est une marque de l'intelligence humaine, et les autres animaux ne sont donc pas en mesure de choisir leurs partenaires reproducteurs. En tant que "co-père" de la sélection naturelle, Wallace est biaisé vers cette dernière, et tente de tout expliquer à travers ce prisme. Il considère, par exemple, que les femelles sont moins colorées parce qu'elles ont été sélectionnées pour l'être moins, alors que les mâles le sont restés parce qu'ils sont plus "vigoureux", plus actifs, en s'appuyant pour cela sur le fait que certains de nos organes sont très colorés (Prum, 2012).

En cela, Wallace se trompait, et Darwin avait raison : la sélection sexuelle, à travers notamment le dimorphisme sexuel et le choix des femelles, est un facteur clef de l'évolution de la couleur des animaux. Pour autant, malgré cette

"erreur" de Wallace, l'essentiel des travaux menés sur la sélection sexuelle sont aujourd'hui considérés comme "néo-wallaciens". Comment expliquer ce renversement de situation ?

Darwin, en effet, en rapport avec la pensée de l'époque, considérait la sélection sexuelle comme une course à *l'esthétisme* : les mâles ne pouvaient être que plus beaux parce que les femelles le souhaitaient ainsi. Cette hypothèse était rejetée par Wallace, mais ce dernier pensait que le choix des femelles *pouvait* effectivement avoir lieu, à une condition : *"The only way in which we can account for the observed facts is by supposing that colour and ornament are strictly correlated with health, vigor, and general fitness to survive."* (Wallace, 1895, pp. 178-179). Ces propos permettaient à Wallace d'admettre la sélection sexuelle (qui faisait, à la fin du XIXème siècle, de plus en plus d'adeptes) à laquelle il s'était vigoureusement opposé, mais uniquement dans un cadre de sélection naturelle, the "admitted *vera causa*". Ironiquement, c'est pourtant lui qui a le premier proposé un mécanisme compatible avec une compréhension moderne des mécanismes de la sélection sexuelle (Prum, 2012).

Cette théorie a été reprise près d'un siècle plus tard, en 1975, par Zahavi, qui, apparemment, n'avait pas connaissance des propos de l'époque de Wallace . Il dit ainsi, dans le même article : "*Wallace dismissed altogether the theory of sexual selection by mate"* et *"I suggest that sexual selection is effective because it improves the ability of the selecting sex to detect quality in the selected sex*." (Zahavi, 1975, p. 207). Des propos bien comparables à ceux de Wallace ! Deux ans plus tard, en 1977, Zahavi publiait la théorie dite "des bons gènes" : les femelles tendent à se reproduire plus fréquemment avec les mâles les plus voyants, justement car ceux-ci sont repérables : s'ils survivent, c'est donc qu'ils sont plus à même d'échapper à des prédateurs (qui les repèrent pourtant mieux) et sont donc en meilleure santé, portent des "bons gènes", et il y a intérêt à se reproduire avec eux (Zahavi, 1977; Caro, 2005).

Cette théorie a depuis largement été étudiée. Après la confirmation que les femelles choisissaient effectivement les mâles (le plus souvent), les signaux qui les intéressent ont été de mieux en mieux compris et reposent sur l'intensité de la coloration ou son patron (Hill, 2006; Andersson and Simmons, 2006). Des cas inverses, où les mâles sont attirés par des signaux colorés chez les femelles, ont déjà été documentés chez les Oiseaux ainsi que chez les poissons Téléostéens (Amundsen, 2000; Amundsen and Forsgren, 2001). Lorsque les femelles sont colorées, ce n'est donc pas simplement une conséquence de la sélection pour certains caractères colorés chez les mâles (Amundsen and Forsgren, 2001). Enfin, la corrélation de ces signaux colorés avec certains traits liés à la condition phénotypique a également été effectuée, tels que la masse corporelle, la quantité de testostérone ou la masse de la progéniture (Roulin, 2009; Piault *et al.*, 2009; Béziers *et al.*, 2017).

En conclusion, alors que Darwin considérait que la sélection sexuelle était un facteur clef de l'évolution des espèces (à raison), le mécanisme sous-jacent a paradoxalement été proposé par Wallace, un opposant de longue date de cette hypothèse, dans un ultime sursaut de déni. Ainsi, l'histoire de l'acceptation de la sélection sexuelle est donc intéressante dans un cadre autre qu'épistémologique, et nous permet de reconsidérer les mécanismes biologiques sous-jacents.

Il est à noter que le déterminisme du choix du partenaire n'est pas uniquement génétique. Cela pourrait induire, à terme, une réduction de la diversité génétique par une sélection directionnelle (Roulin, 2015). Le choix peut également dépendre de facteurs environnementaux, comme cela a été montré chez des guppies *Poecilia reticulata* (Kodric-Brown, 1989).

iii. La sélection sociale

Un troisième mode de sélection est la sélection sociale (Tanaka, 1996; Wolf *et al.*, 1999). Ce type de sélection a été conceptuellement développé récemment, à la fin des années 1990 et est encore vivement débattu (Clutton-Brock, 2010; Milam *et al.*, 2011). La sélection sociale suppose que toutes les interactions au sein d'un groupe, et pas seulement pour l'accès à la reproduction, sont à prendre en compte et pourraient induire l'évolution de patrons de pigmentation. Ainsi, alors que les signaux issus de la sélection sexuelle induisent la mise à l'écart des rivaux, les signaux issus de la sélection sociale permettent au contraire à l'individu de ne pas se séparer de ses conspécifiques afin de maximiser ses chances d'échapper à la prédation (Senar, 2006).

iv. Conflits entre modes de sélection multiples : application à la coloration

Les différentes forces sélectives peuvent parfois rentrer en conflit. Fisher a ainsi montré que les préférences des femelles tendent à être corrélées aux attributs mâles. Cette corrélation peut conduire à un processus d'emballement (ou "runaway") dû à la sélection sexuelle, qui va alors à l'encontre de la sélection naturelle, puisque les mâles ne sont alors plus en mesure d'échapper à un prédateur : les caractères sexuels sont alors qualifiés "d'extravagants" (Fisher, 1958; Kirkpatrick, 1982). Chez les guppies *P. reticulata*, une étude a montré que les mâles les plus colorés, bien que préférés par les femelles, sont également plus susceptibles d'être prédatés car plus visibles (Godin and McDonough, 2003). Chez les lions, plus la crinière est foncée, plus les femelles sont attirées par les mâles ; cependant, une crinière foncée induit une hyperthermie (Caro, 2005). Il y a donc un équilibre entre un facteur de sélection "écologique" et "sexuel".

b. La couleur est un trait qui évolue souvent vite

Les différentes forces évolutives évoquées ci-dessus entrent en jeu en même temps lors de l'évolution sur le long-terme d'une espèce, ou d'une lignée. Cela a été démontré dans de nombreuses lignées, végétales comme animales, Protostomiens (comme les Lépidoptères) et Deutérostomiens (comme chez les Vertébrés), et a en particulier été démontré dans quelques radiations de Vertébrés actuels : les lézards du genre *Anolis* (Muñoz *et al.*, 2013), les poissons perroquets (Streelman *et al.*, 2002) et les poissons Cichlidés. Chez ces derniers, une évolution par sélection divergente de la coloration chez des paires d'espèces a été documentée (Allender *et al.*, 2003; Seehausen *et al.*, 2008; Wagner *et al.*, 2012). De plus, parfois, un caractère en particulier émerge et est vite propagé au sein d'une lignée, à l'instar des egg-spots au sein des Cichlidés, une innovation apparue il y a quelques millions d'années chez l'ancêtre de l'actuel groupe comptant le plus d'espèces, les Haplochromines (Santos *et al.*, 2014).

Selon Streelman *et al.*, une radiation évolutive se fait en trois temps : en premier lieu, les espèces s'adaptent à des habitats différents. Ensuite, ils se spécialisent par des morphologies différentes. Les signaux colorés, souvent liés à de la signalisation intra- et interspécifique, interviennent lors de la troisième étape, qui induit des barrières reproductives dues à des barrières de communication (Streelman *et al.*, 2002; Kocher, 2004).

À l'inverse, la couleur peut être un trait rapidement *perdu*. Au cours de l'évolution du poisson *Astyanax mexicanus*, plusieurs populations cavernicoles ont évolué de manière plus ou moins isolée avec les populations de surface, pendant des durées variables et difficiles à calculer, mais de l'ordre de quelques centaines de milliers d'années (Fumey *et al.*, 2018). Ces populations sont devenues aveugles et dépigmentées de manière convergente (Protas *et al.*, 2006; Casane and Rétaux, 2016). Ainsi, dans des conditions de réduction de la pression de sélection sexuelle basée sur la couleur, une perte de la pigmentation noire, mélanique, est rapidement observée. Cela souligne le coût du maintien d'un phénotype coloré pour les organismes de surface.

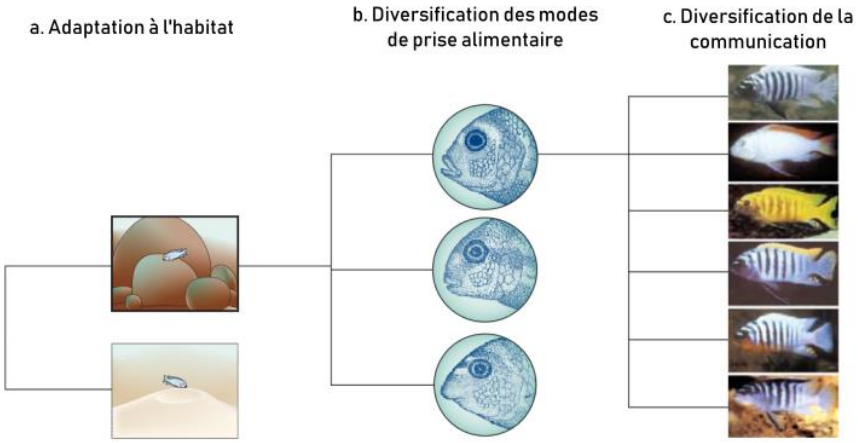


Figure 4 - Radiation en trois étapes chez les poissons Cichlidés.

(a) La première étape d'une radiation peut être l'adaptation à de nouveaux habitats. (b) Les morphologies trophiques peuvent alors se diversifier. (c) Au sein du nouvel habitat, les signaux de communication, tels que la couleur, peuvent alors évoluer rapidement. *Adapté de Kocher (2004).*

II. La couleur est particulièrement diversifiée chez les poissons Téléostéens

1. Les poissons Téléostéens constituent l'essentiel des Vertébrés…

"In their great numbers and degree of anatomical diversity, the modern ray-finned fishes may be considered the most successful of all vertebrates." (Robert L. Carroll)

a. Les poissons Téléostéens sont très diversifiés…

Les poissons Téléostéens constituent environ la moitié des espèces de Vertébrés actuelles, avec environ 30 000 espèces (Nelson *et al.*, 2016). La phylogénie de ce groupe s'est grandement améliorée ces dernières années, notamment grâce à l'apport de données moléculaires (Near *et al.*, 2012, 2013; Eytan *et al.*, 2015). Ce groupe est constitué d'environ 70 ordres et 500 familles qui partagent au moins 27 synapomorphies liées à différents aspects de leur morphologie (De Pinna, 1996) et dont la monophylie est confirmée par des données moléculaires (Near *et al.*, 2012). Ce taxon présente une incroyable diversité, à la fois en termes de taille, de morphologie, d'habitat, de couleur et de comportements. De plus, quelques lignées au sein de ce groupe ont connu une radiation adaptative. L'exemple des poissons Cichlidés des grands lacs africains est peut-être le plus emblématique : dans chacun des grands lacs (Malawi, Victoria, Tanganyika), entre 250 et 500 espèces sont apparues sur des périodes allant de 15 000 ans à 12 millions d'années, ce qui en fait le plus fort taux de spéciation connu chez les Vertébrés (calculé en nombre de nouvelle espèces par million d'année) (Kocher, 2004; Wagner *et al.*, 2012).

b. … et notamment les poissons coralliens

Alors qu'ils n'occupent que 0.1 à 0.2 % de la surface terrestre, les récifs coralliens présentent la plus forte diversité de poissons Téléostéens avec au moins 4 000 espèces (Price *et al.*, 2015). Je reviendrai sur ce groupe dans la troisième partie de cette introduction. Dans la suite j'essaierai cependant de prendre le plus d'exemples possibles parmi les poissons coralliens, qui est un des groupes (le groupe ?) les plus diversifiés au niveau de la couleur chez les Vertébrés.

2. … et ont des patrons de coloration variés et variables…

"Brilliantly-coloured fishes from warm seas are many of them well concealed when surrounded by the brilliant sea-weeds, corals, sea-anemones, and other marine animals, which make the sea-bottom sometimes resemble a fantastic flower-garden" (Wallace, 1879)

a. Diversité des patrons de pigmentation

i. Il existe de nombreux patrons de pigmentation

Les poissons Téléostéens, et les poissons coralliens en particulier, présentent une grande diversité de patrons de pigmentation, depuis des taches ponctuelles jusqu'à des alternances de bandes horizontales ou verticales, en passant par des patrons labyrinthiques ou uniformes.

Avant de proposer une hypothèse pour expliquer cette aussi grande diversité, il faut avant tout s'assurer d'un point : est-ce que, *du point de vue des poissons coralliens*, sous l'eau, la diversité de couleur est également perçue ? Les animaux que nous percevons comme colorés peuvent en effet ne pas l'être dans leur environnement, soit parce que ce dernier a des propriétés physiques différentes, soit car les systèmes de perception visuelle propres à l'espèce ne sont pas sensibles à ces signaux (Justin Marshall, 2000)

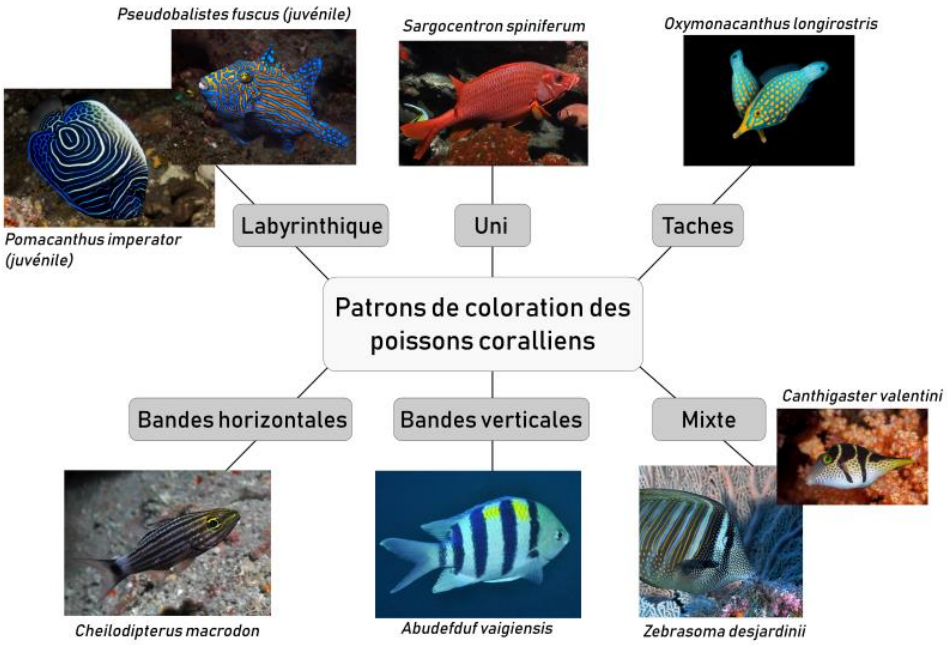


Figure 5 - Diversité des patrons de coloration des poissons coralliens.

Les patrons de coloration des poissons coralliens peuvent être plus ou moins complexes. Ce groupe écologique est probablement le plus divers en termes de coloration chez les Vertébrés. *Clichés : Germain Boussarie.*

La colonne d'eau absorbe les couleurs autre que le bleu, mais, à quelques mètres ou quelques dizaines de mètres, dans des eaux récifales souvent très claires, toutes les longueurs d'ondes sont encore présentes, et les couleurs sont donc encore *potentiellement* visibles. Certains poissons coralliens, comme le Baliste Picasso *Rhinecanthus aculeatus*, sont par ailleurs capables de percevoir une grande diversité de longueur d'ondes (Cheney *et al.*, 2013). Ainsi, il semblerait que les signaux que *nous* percevons comme très colorés le soient également pour les poissons coralliens, qui peuvent également, pour certains, voir les UV proches (Losey *et al.*, 2003; Marshall *et al.*, 2003).

Une autre caractéristique de l'habitat corallien est la non-homogénéité et la non-prédictibilité de l'arrière-plan : pour la majorité des espèces diurnes et mobiles, il est donc impossible d'avoir une coloration de "camouflage". Cela a également été décrit en 1940 par Cott, à propos des Oiseaux de la canopée : "*few birds - whatever their coloration - can be expected to harmonize cryptically with surroundings which vary constantly and widely, from moment to moment and from month to month*" (Cott, 1940). De plus, les eaux souvent translucides rendent la transmission du signal visuel efficace : c'est donc un monde "hyper-visible" (Alonso, 2016). Dans un tel environnement, selon Alonso (2016), quelle que soit la couleur, l'individu sera également repérable : la sélection pour les signaux de communication prend alors le dessus et permet d'expliquer l'apparition de patrons aussi bariolés.

Une autre hypothèse permettant peut-être une telle diversité est celle du "Carefree world" (Alonso, 2016). Les poissons coralliens mobiles et diurnes ayant souvent une bonne acuité visuelle, ils peuvent voir venir leur prédateur de loin, et, étant bons nageurs et à proximité d'un récif plein d'anfractuosités, ils peuvent facilement se mettre à l'abri. Dans un tel contexte, ils n'ont pas à "se préoccuper" (*care*) de leur patron de coloration : quel que soit ce dernier, ils peuvent relativement aisément échapper à un prédateur.

En conclusion l'explication d'une aussi grande diversité des couleurs des poissons coralliens ne peut se résumer à une ou deux hypothèses, et comme le souligne le spécialiste des poissons coralliens Justin Marshall, *"it is almost inconceivable for only one evolutionary force to be behind the colours of such a diverse assemblage, the final answer most probably being a compromise of many"* (Marshall, 2000).

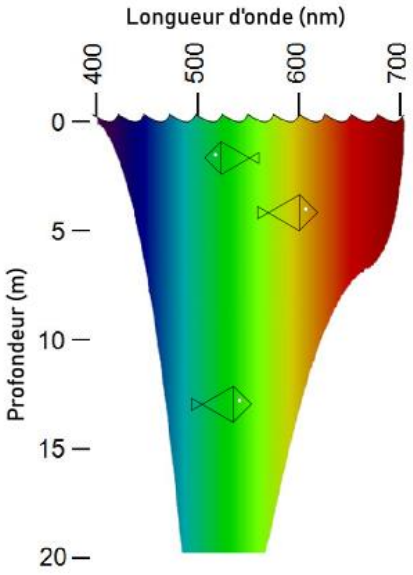


Figure 6 - Le spectre du visible est peu à peu absorbé dans l'eau.

Sous l'eau, les longueurs d'ondes sont différentiellement absorbées en fonction de la profondeur. Ainsi, à 10 mètres, un individu ne voit pas les mêmes couleurs qu'en surface. Les poissons coralliens sont pour la plupart à de faibles profondeurs (inférieures à 20 mètres, même si certaines espèces sont plus profondes) et leur environnement est donc encore coloré. *Adapté de Michiels et al. (2008).*

ii. Les patrons de pigmentation sont souvent convergents

À l'instar de nombreux animaux, les poissons coralliens ont des patrons de pigmentation souvent convergents. Par exemple, le patron de pigmentation avec des bandes noires verticales est apparu indépendamment chez les Pomacentridae (poissons-demoiselles) et les Acanthuridae (poissons-chirurgiens). Ce patron est même apparu trois fois indépendamment *au sein* des Pomacentridae (dans les genres *Abudefduf, Dascyllus* et *Chrysiptera*), suggérant que la convergence peut avoir lieu à des niveaux taxonomiques différents. Les bandes blanches horizontales sont également apparues à de multiples reprises, notamment chez les Apogonidae (poissons-cardinaux) et les Plotosidae (poissons-chats).

Figure 7 - Convergence des patrons de coloration chez les poissons coralliens.

Il existe des convergences entre des patrons de coloration chez des espèces phylogénétiquement distinctes. *Clichés d'A. meleagrides et d'O. meleagris : Wikimedia Commons. Autres clichés : Germain Boussarie.*

La convergence des patrons de coloration chez des animaux occupant le même environnement peut résulter de pressions sélectives mais peut aussi être due à des raisons autres que la sélection, comme la dérive génétique ou des contraintes développementales (Losos, 2011). La convergence n'est donc pas nécessairement un produit de contraintes sélectives communes.

b. Changements associés au statut social ou à la reproduction

La coloration des poissons coralliens est très diverse, à l'état adulte. Elle peut de plus varier, plus ou moins rapidement. Nous avons déjà vu l'exemple du dimorphisme sexuel chez les poissons perroquets du genre *Scarus*. En réalité, cette espèce étant hermaphrodite protogyne, il y a un changement de couleur entre les femelles (stade initial) et les mâles (stade terminal). De même, nous avons évoqué l'exemple du poisson-papillon *Chaetodon lunula*, un exemple de "couleur permettant une communication intraspécifique" : ce dernier change de couleur du côté dorsal (le countershading s'atténue) pour indiquer son statut social. Si deux individus de taille égale se rencontrent, le plus jaune sera le dominant, sans qu'il y ait de conflit intraspécifique (Hamilton and Peterman, 1971). Ce changement de couleur est donc éphémère mais socialement important. Chez le poisson-chirurgien *Acanthurus triostegus*, le changement de couleur et l'inversion du patron de pigmentation chez les juvéniles (bandes noires sur fond clair à bandes claires sur fond sombre) va de pair avec une augmentation de l'agressivité et de la territorialité (Besson *et al.*, 2018). Ce changement est réversible et a lieu en quelques *secondes*.

Le changement de couleur peut également intervenir sur une échelle de temps plus longue au moment de la reproduction ou en fonction de la saisonnalité (qui est faible en milieu tropical). Chez certaines espèces, avec ou sans dimorphisme sexuel pendant la majorité du temps, les mâles peuvent changer leur livrée au moment de la reproduction : la couleur donne alors des informations sur les capacités reproductives des mâles (Thresher, 1984; Kodric-Brown, 1998).

c. Polyphénisme intraspécifique

Outre les cas de changement de couleur dus aux interactions sociales au sein d'une même espèce, les poissons coralliens peuvent présenter différentes formes colorées stables : on parle de polyphénisme spécifique.

i. Chez l'adulte

Nous avons évoqué le camouflage généraliste de l'Antennaire *Antennarius commerson.* Ces individus, prédateurs, peuvent changer de couleur en quelques semaines pour imiter l'arrière-plan sur lequel ils se trouvent (Randall, 2005). Ainsi, au sein d'une même population, plusieurs formes colorées existent.

Des cas de mimétisme peuvent également expliquer un polyphénisme de coloration au sein d'une espèce corallienne. Chez le Pseudochromis sombre *Pseudochromis fuscus*, une espèce piscivore de juvéniles d'autres espèces, il existe deux formes : une forme sombre et une forme jaune, qui miment des poissons demoiselles. Par mimétisme, la forme jaune (respectivement sombre) se met en place au cours du développement si les individus des autres espèces sont jaunes (respectivement sombres), indépendamment de la couleur de fond de l'habitat (Cortesi *et al.*, 2015). Un individu peut passer d'une forme à l'autre en deux semaines environ.

ii. Au cours du développement

Au cours du développement des poissons coralliens (que nous décrirons plus dans le chapitre III de cette introduction), différents patrons colorés sont souvent présents. Les larves sont souvent transparentes alors que les adultes peuvent avoir des patrons très colorés (Leis and McCormick, 2002). Qui plus est, il peut y avoir, entre la larve et l'adulte, plusieurs morphes pigmentaires : chez le Pseudochromis sombre, au cours du développement, il y a une larve transparente, un juvénile cryptique et peu coloré, et des adultes on l'a vu, différentiellement pigmentés selon les espèces proies alentour (Cortesi *et al.*, 2016). On note que la présence de plusieurs morphes colorés au cours de différentes phases ontogénétiques n'est pas spécifique des poissons coralliens, et, chez les Vertébrés, les Oiseaux Laridés (Sternes, Mouettes, Goélands) en sont également un exemple… criant, puisqu'il peut y avoir plusieurs saisons avant l'acquisition du plumage adulte.

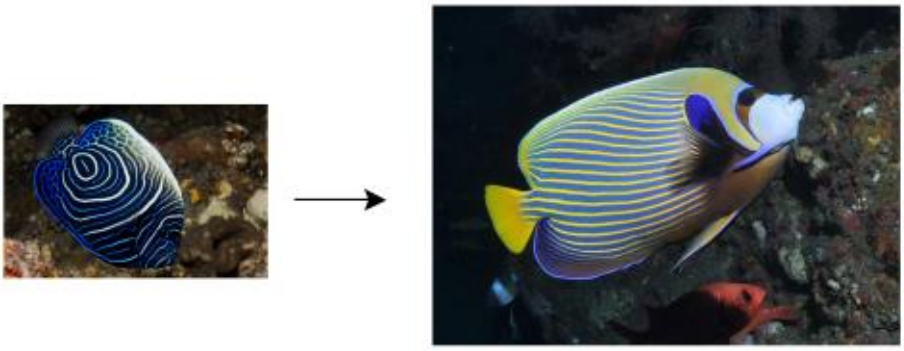


Figure 8 - Changement du patron de coloration au cours du développement d'un poisson corallien.

Exemple du poisson corallien Ange-empereur *Pomacanthus imperator.* Gauche : forme juvénile (~ 1 an). Droite : adulte. *Clichés : Germain Boussarie.*

d. Bases physiologiques des changements pigmentaires

Les changements ontogénétiques, ou saisonniers, ont lieu sur des échelles d'au moins quelques heures à quelques jours/semaines (Cortesi *et al.*, 2016). D'autres changements, tels que celui documenté chez le Poisson-chirurgien *A. triostegus* (Besson *et al.*, 2018) ou dans d'autres espèces pouvant rapidement se camoufler (Watson *et al.*, 2014), ont lieu en quelques secondes. Les bases physiologiques sous-jacentes diffèrent. Les changements lents sont souvent liés à des facteurs hormonaux induisant la prolifération d'un type cellulaire ou la synthèse d'un pigment particulier(Leclercq *et al.*, 2009). Les changements rapides sont, eux, liés à une innervation des cellules induisant une dispersion/agrégation des organites contenant la mélanine (les mélanosomes) ou une variation de l'espacement entre plaquettes réflectives des iridophores (Mäthger *et al.*, 2003; Nilsson Sköld *et al.*, 2013; Teyssier *et al.*, 2015). Ils peuvent également être liés à la transduction directe d'un signal lumineux (Ban *et al.*, 2005; Lin *et al.*, 2013).

3. Bases cellulaires et moléculaires de la diversité pigmentaire des Vertébrés

Dans cette partie, je ne traite que des Vertébrés, et n'aborde donc pas la diversité des bases structurelles et chimiques à la couleur dans d'autres taxons et notamment chez les Insectes – le groupe animal à la richesse spécifique la plus élevée (Shamim *et al.*, 2014).

Au sein des Vertébrés, nous l'avons vu, la diversité et la variabilité de la pigmentation sont stupéfiantes. Nous avons plus haut particulièrement mis en exergue un groupe de Vertébrés, les poissons Téléostéens, et notamment les poissons coralliens. Ce groupe présente peut-être la plus grande diversité de coloration chez les Vertébrés. Nous avons évoqué plus haut des raisons écologiques pouvant peut-être expliquer cela. Ici, je montre en quoi les bases moléculaires et cellulaires de la couleur des Téléostéens "potentialisent" une diversité phénotypique, en la comparant à ce qui est présent chez les autres grands groupes de Vertébrés.

a. Mammifères

Chez les Mammifères, la couleur est essentiellement due à une molécule, la mélanine, ou plutôt, à ses deux formes chimiques principales : l'eumélanine, sombre, et la phéomélanine, responsable des tons jaunes à rouges. Ce sont des polymères complexes dont la formule chimique précise n'est pas connue. Ces molécules sont des pigments, qui peuvent donc *absorber* une partie de la lumière incidente. Pour l'eumélanine, le pic d'absorption est à 335 nm, dans les UV, et toutes les longueurs d'ondes du visible sont absorbées (ce spectre diffère quelque peu pour la phéomélanine) (Kollias, 1994). La mélanine a donc un double rôle : absorption des UV, et donc photoprotection contre ces rayons mutagènes, et thermogenèse : en absorbant la lumière incidente, l'énergie est libérée sous forme de chaleur.

Ces pigments sont présents dans la peau ou dans les poils, ainsi que dans les yeux. Ils sont produits par le processus de mélanogenèse, dans des cellules dendritiques spécialisées de l'épiderme, les mélanocytes, au sein d'organites

particuliers, les mélanosomes. Chez les Mammifères, les mélanosomes sont ensuite exocytés puis endocytés par des kératinocytes receveurs de l'épiderme (Tarafder *et al.*, 2014; D'Mello *et al.*, 2016).

Chez certains Mammifères, comme la Souris, il n'y a pas de mélanocytes actifs dans l'épiderme du tronc (Slominski *et al.*, 2004). Pourtant, les Souris peuvent parfois être sombres, en raison de la présence de mélanocytes ailleurs que dans l'épiderme des Mammifères, et notamment dans les follicules pileux. La mélanine est alors directement exportée vers le poil en croissance et intégrée à celui-ci. Ainsi, les Chimpanzés, aux poils pourtant sombres, ont un épiderme très clair.

Chaque mélanocyte peut synthétiser eumélanine et phéomélanine, en proportion variable. La biochimie de ces polymères est complexe et variable mais ils proviennent de la transformation enzymatique de tyrosine (pour l'eumélanine) et de tyrosine et de cystéine (pour la phéomélanine) (Ito, 2003). La concentration de cystéine ainsi que l'activité relative des enzymes de la voie, notamment la tyrosinase, est à l'origine du ratio eumélanine/phéomélanine (Ito, 2003). Des signaux hormonaux, tels que ceux médiés par les MSH (Melanocyte-Stimulating Hormones) ou par ACTH, interviennent sur ce ratio (Sturm *et al.*, 2001; Wolf Horrell *et al.*, 2016).

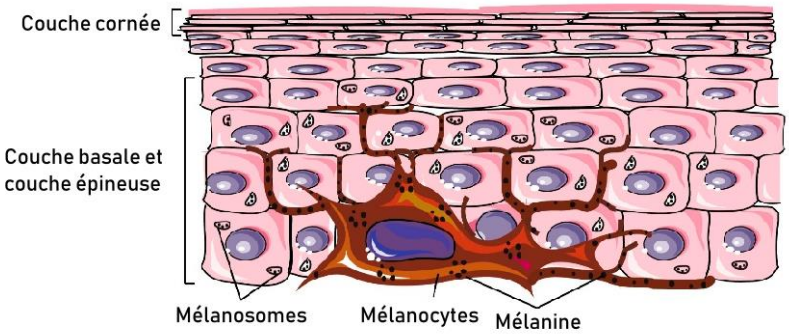


Figure 10 - Association entre kératinocytes et mélanocytes chez les Mammifères

Les mélanocytes, dendritiques, sont dans la couche basale de l'épiderme et produisent de la mélanine. Les mélanosomes sont transférés aux kératinocytes adjacents. *Adapté de D'Mello et al. (2016).*

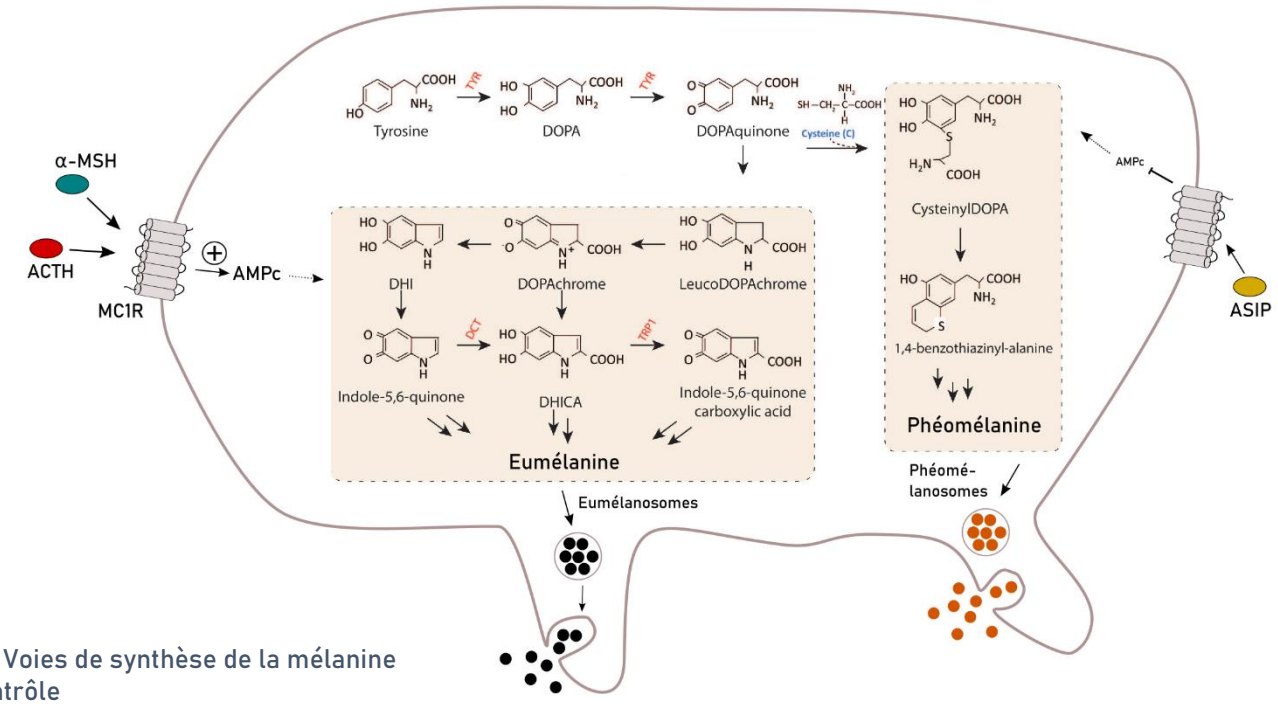


Figure 9 - Voies de synthèse de la mélanine et leur contrôle

La synthèse d'eumélanine ou de phéomélanine a lieu dans le mélanosome. Certaines enzymes de la voie, indiquées en rouge, sont à l'origine de phénotypes pigmentaires anormaux lorsqu'elles sont dysfonctionnelles. Le ratio de la synthèse entre eumélanine et phéomélanine dépend de signaux inducteurs activateurs (MSH ou ACTH) ou inhibiteurs (ASIP) de l'activation de MC1R. Activé, ce dernier induit une hausse de la teneur en AMPc qui active l'expression des gènes de synthèse de l'eumélanine et de formation des eumélanosomes. *Adapté de Wolf Horrell et al. (2016) et de Martin et Orgogozo (2017).*

b. Oiseaux

Chez les Oiseaux, les pigments sont bien plus divers et sont déposés tant dans les plumes que dans le tégument et les parties nues : le bec et les pattes. On y retrouve les mélanines (eu- et phéomélanine). Celles-ci sont transférées dans la plume selon un mécanisme mal connu : une étude récente a montré comment les mélanocytes étaient mis en place dans le follicule plumaire (Lin *et al.*, 2013), mais l'incorporation des mélanines dans la partie kératinisée de la plume reste inconnue (Galván and Solano, 2016). Toutefois, chez les Oiseaux, l'organisation des mélanosomes dans les barbules des plumes peut augmenter le spectre des possibilités de la seule coloration structurelle (qui donne des couleurs bleu-vert), et induire une coloration "structuralo-chimique" allant du violet au jaune (Stavenga *et al.*, 2011; D'Alba *et al.*, 2012).

Outre les mélanines, les pigments les plus répandus chez les Oiseaux sont les caroténoïdes (McGraw, 2006; Thomas *et al.*, 2014). Ces molécules, intégrées aux plumes, donnent une couleur allant du jaune au rouge. À l'exception d'une espèce de puceron (Moran and Jarvik, 2010), d'acariens d'araignées (Altincicek *et al.*, 2012) et de mouches de la famille des Cécidomyies (Cobbs *et al.*, 2013) ayant acquis la capacité de synthétiser les caroténoïdes par transfert horizontal de gène, ces molécules doivent être obtenues par l'alimentation chez les animaux ; ces composés sont donc souvent considérés comme un signal "honnête" de la qualité intrinsèque d'un individu (Simons and Verhulst, 2011; Simons *et al.*, 2012). Les caroténoïdes peuvent cependant être modifiées par les enzymes de l'individu.

De même que pour les mélanines, les couleurs parfois brillantes associées à ces pigments sont dues à une interaction de ces derniers avec la structure plumaire (Shawkey and Hill, 2005). Ceci est une règle générale chez les Oiseaux et dans les groupes de Vertébrés que nous verrons plus loin : couleurs "chimiques" et "physiques" interagissent bien souvent, et il est rare que la couleur ne soit due qu'à un pigment seul, ou à une structure seule. Ceci a conduit certains auteurs à publier un article au titre évocateur : "*You Can't Judge a Pigment by Its Color*" (McGraw and Nogare, 2004). À ces deux types de pigments s'ajoute une grande diversité de pigments "*uncommon*" (sic) (McGraw, 2006)) :

- les psittacofulvines, propres aux Perroquets, et de la même couleur que les caroténoïdes (McGraw and Nogare, 2004; Tinbergen *et al.*, 2013) ;
- les porphyrines, allant du rouge prononcé (turacine des Oiseaux Touraco) au bleu (biliverdine) en passant par le vert (turacoverdine) (McGraw, 2006)
- les ptérines (ou ptéridines), seulement présentes dans l'iris des Oiseaux (Oliphant, 1987)
- les sphéniscines, découvertes plus récemment chez les Sphenisciformes (Manchots), et reconnues comme un type pigmentaire différent. Ces molécules, jaunes et fluorescentes, diffèrent des autres familles de pigments mais leur formule chimique n'est pas connue (McGraw and Nogare, 2004; McGraw *et al.*, 2007; Thomas *et al.*, 2013)
- les oxydes métalliques, comme les oxydes de fer, donnant une couleur rouge (Toral *et al.*, 2008)

Si la diversité pigmentaire des Oiseaux est bien connue, il n'en est pas de même au niveau cellulaire et histologique. À part pour les mélanocytes (et encore), on ne sait en particulier pas, à ma connaissance, comment ces molécules sont déposées dans les tissus, et on ne connait d'ailleurs pas non plus d'autres types cellulaires que les mélanocytes qui seraient spécialisés dans la synthèse d'un type cellulaire. Ceux-ci *doivent* exister chez les Perroquets, les Manchots ou les Touracos, qui synthétisent des types pigmentaires propres et probablement pas dérivés de l'alimentation, mais ils ne sont pas documentés.

L'étude de la coloration chez les autres groupes de Vertébrés a été plus longue à émerger, et il a fallu attendre 1967 pour que les mélanocytes soient découverts ailleurs chez les Vertébrés (Bagnara *et al.*, 2007)

c. Lépidosauriens

Les bases génétiques, moléculaires, cellulaires et histologiques de la coloration chez les Lépidosauriens (lézards, serpents, iguanes…) sont très peu documentées. L'essentiel de la littérature s'appuie sur un ouvrage référence de 1973 de Joseph Bagnara sur l'histologie des cellules pigmentaires ainsi que quelques études éparses depuis (Kuriyama *et al.*, 2006; Teyssier *et al.*, 2015). Pourtant, outre la variabilité phénotypique de la coloration de ce taxon, les Lépidosauriens sont très

intéressants d'un point de vue histologique. Quatre types de cellules pigmentées (ou chromatophores) y ont été trouvés : des mélanophores, des xanthophores, des érythrophores et des iridophores (Olsson *et al.*, 2013).

Le terme de "mélanophores" est consacré par l'usage. Toutefois, une revue récente sur la pigmentation chez les Vertébrés a suggéré de ne plus l'utiliser, et d'employer "mélanocytes" pour toutes les cellules synthétisant la mélanine, qu'elle soit libérée par exocytose (mélanocytes s. s.) ou gardée au sein de la cellule (mélanophore) (Schartl et al., 2016). Par la suite, j'utiliserai donc autant que possible le terme de mélanocytes.

- Les mélanocytes des Lépidosauriens diffèrent des mélanocytes des Oiseaux et des Mammifères : ces cellules n'exocytent pas la mélanine et gardent leurs mélanosomes. De plus, les mélanocytes peuvent non seulement être dans l'épiderme mais également dans le derme.
- Les xanthophores (plutôt jaunes) et les érythrophores (plutôt rouges), distingués sur des critères visuels. Cette difficulté à distinguer ces deux types de cellules est également valable chez les poissons Téléostéens ; j'y reviendrai. La couleur de ces cellules provient d'organites particuliers, les ptérinosomes, contenant des ptéridines, et des vésicules de caroténoïdes (Bagnara and Matsumoto, 2006). Les deux types de vésicules peuvent co-exister au sein d'une même cellule.
- Un dernier type cellulaire, les iridophores, qui contiennent des plaquettes de purines et de ptéridines réfléchissant la lumière. Ces cellules ne sont donc pas des cellules "pigmentaires" au sens strict, car les plaquettes agissent comme des réflecteurs de lumière et ne sont donc pas en ce sens des "pigments". Toutefois, la quasi-totalité de la communauté les considère comme des cellules pigmentaires, ce que je ferai aussi.

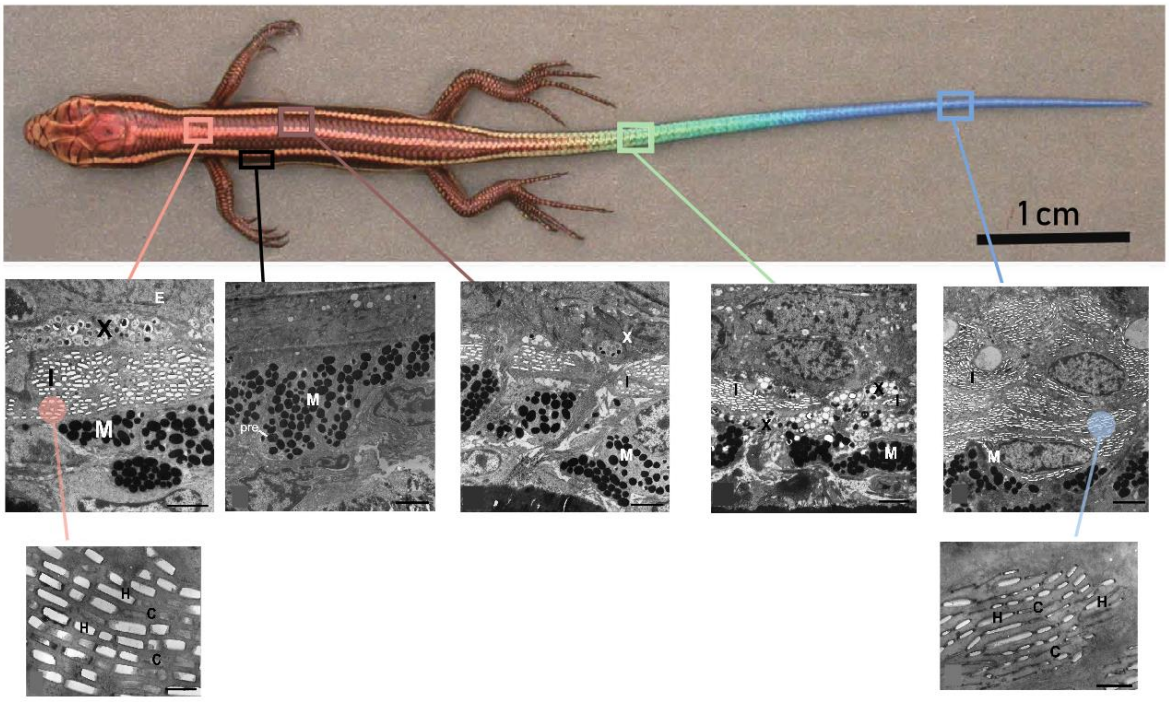


Figure 11 – Organisation et ultrastructure des chromatophores chez un lézard Scincidae.
Chez *Plestiodon latiscutatus,* plusieurs couleurs coexistent. Elles sont dues à une organisation structurale différente des chromatophores : mélanocytes (M), xanthophores (X) et iridophores (I). L'interaction entre chromatophores est donc à l'origine de la couleur du tissu. De plus, un même type cellulaire peut être variable entre différentes zones. Par exemple, les iridophores de la peau bleue (à droite) ont des plaquettes de purines plus fines que celles des iridophores de la peau claire (à gauche). H : plaquette non cristalline ; C : plaquette cristalline ; *pre* : pré-mélanosome (en cours de différentiation). Barre d'échelle : 2 micromètres. (milieu), 500 nm (bas). Orientation : haut = épiderme (E). *Adapté de Kuriyama et al. (2006).*

Chez le lézard scinque *Pestiodon latiscutatus*, une étude a montré que non seulement ces trois ou quatre types cellulaires intervenaient dans la coloration, indépendamment, mais qu'il fallait, à l'instar de ce que j'ai évoqué chez les Oiseaux, considérer *l'interaction* entre chromatophores dans le rendu coloré final (Kuriyama *et al.*, 2006). La disposition spatiale des chromatophores est en effet essentielle dans la transmission et la réflexion de la lumière. L'association intime entre différents types de chromatophores constitue une "unité dermique de chromatophores" ; ces cellules répondent en effet de concert à des signaux hormonaux et constituent une unité fonctionnelle (Bagnara and Matsumoto, 2006).

La forme, la taille, l'orientation, le nombre et la conformation des plaquettes, ainsi que la distance entre elles, sont des facteurs pouvant faire varier la longueur d'onde réfléchie et donc le rendu de couleur (Morrison *et al.*, 1995, 1996). La couleur des "reptiles" est donc bien souvent une interaction entre des couleurs "chimiques" (xanthophores / érythrophores / mélanocytes) et "physiques" (iridophores).

Outre les études de Morrison dans les années 1990, l'étude des cellules pigmentaires chez les Lépidosauriens s'est faite assez discrète. Quelques travaux ont été menés sur la mélanisation des squamates (Rosenblum *et al.*, 2010; Nunes *et al.*, 2011), mais, paradoxalement, très peu d'études se sont intéressées aux autres types cellulaires de cette lignée. Comme le regrettent (Saenko *et al.*, 2013) : *"Despite such a high potential for complexity and diversity, lizards and snakes remain relatively under-represented in evolutionary developmental studies in general, and in analyses of color-pattern evolution in particular."* Toutefois, quelques études récentes, notamment sur le changement de couleur du Caméléon panthère *Furcifes pardalis*, viennent raviver ce domaine d'étude (Teyssier *et al.*, 2015). Un nouveau pigment bleu, de nature inconnue, y a d'ailleurs été trouvé, preuve que ce taxon recèle de belles choses encore à découvrir.

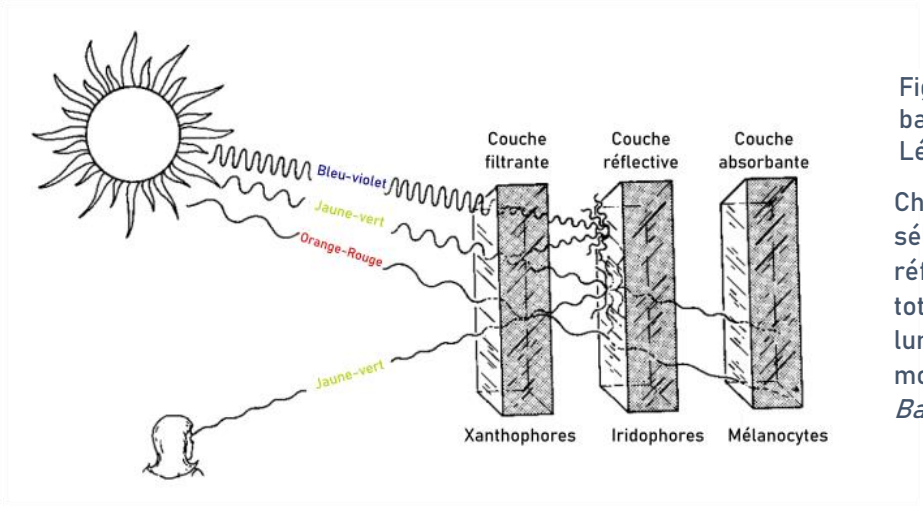


Figure 12 - Interprétation schématique de la base de la coloration verte des Lépidosauriens, Amphibiens et Téléostéens.

Chacune des couches cellulaires absorbe sélectivement (couche de xanthophores), réfléchit (iridophores) ou absorbe totalement (mélanocytes) la lumière. La lumière parvenant à l'observateur est donc modifiée par la traversée du tissu. *Adapté de Bagnara and Matsumoto (2007).*

d. Amphibiens

Le concept de "*dermal chromatophore unit*", bien que présenté plus haut chez les Lépidosauriens, a cependant d'abord été défini chez les Amphibiens, dès 1968, par M. Joseph Bagnara (Bagnara *et al.*, 1968). Les années 1960 et 1970 ont représenté les "années fastes" de l'étude de la pigmentation des Amphibiens, notamment sous l'impulsion de ce zoologiste. Dans l'ensemble, les types de cellules pigmentaires sont les mêmes que chez les "reptiles" : mélanocytes, xanthophores/érythrophores, et iridophores. C'est chez les Amphibiens que les changements de couleur dus à la dispersion et à l'agrégation des mélanosomes au sein des mélanocytes ont été pour la première fois mis en évidence (Bagnara and Hadley, 1973). Toutefois, à l'instar des Lépidosauriens, l'étude de la pigmentation chez les Amphibiens est rare. Peu d'études sont parues sur les cellules pigmentaires et certaines sont parfois contradictoires.

Ainsi, (Lunghi *et al.*, 2017) suggèrent qu'il y a des leucophores ainsi que des cyanophores chez les Amphibiens. Des cellules "leucophore-like" ont été décrites en 2004 chez les têtards de Xénope, et diffèrent – selon l'auteur – des iridophores : absence d'iridescence, fluorescence à la lumière verte, dispersion des pigments en réponse à la MSH

(Fukuzawa, 2004). Enfin, un autre type de cellule pigmentaire a été suggéré en 1985 chez les Amphibiens : les violeophores, qui seraient des xanthophores modifiés contenant des violeopterinosomes (Nishioka and Ueda, 1985).

Cependant, J. Bagnara, en 2006, pourtant spécialiste du domaine, n'évoque pas ce dernier type dans ses écrits (Bagnara and Matsumoto, 2006), et soutient une hypothèse différente : pour lui, les leucophores ne sont que des iridophores modifiés, et les cyanophores n'existeraient que chez les Téléostéens, contrairement à ce qu'affirme (Lunghi *et al.*, 2017). D'ailleurs, un ouvrage de référence du domaine, *Herpetology*, ne mentionne aucun de ces deux types de chromatophores (Vitt and Calwell, 2013).

Le débat n'est pas encore tranché, et deux hypothèses seraient à tester : (i) est-ce que les types cellulaires définis comme tels existent chez les Amphibiens et (ii) si oui, sont-ils homologues à ceux des poissons Téléostéens ? Au sein de ce taxon, des cas de chromatophores "mosaïques", arborant à la fois des mélanosomes, des plaquette de purines, des ptérinosomes et des vésicules de caroténoïdes, ont déjà été documentés (Bagnara, 1983). Alors que, comme nous le verrons plus loin, les relations ontogénétiques entre les différents chromatophores sont encore mal connues, la présence d'une telle cellule soulève des questions sur son origine développementale et sur les modes de différentiation du précurseur des chromatophores, ainsi que sur l'évolution de la différentiation des chromatophores au sein des différentes lignées de Vertébrés. Dans tous les cas, des efforts de recherche dans le domaine de la pigmentation amphibienne seraient bénéfiques à la communauté de ce groupe et d'autres groupes de Vertébrés.

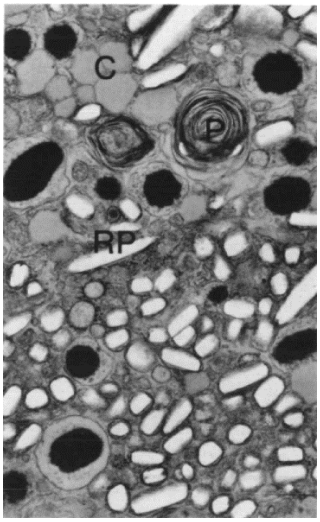


Figure 13 - Chromatophore mosaïque chez les Amphibiens.
Chromatophore chez la Grenouille *Phyllomedusa iheringii.* Au sein d'une même cellule on trouve des ptérinosomes (P) des mélanosomes, des vésicules de caroténoïdes (C) et des plaquettes réfléchissantes de purines (RP). Un tel chromatophore met en cause la distinction habituellement faite entre les différents types de chromatophores. Largeur de l'image : 3 micromètres. *Adapté de Bagnara (1983).*

e. Téléostéens

Au sein des Vertébrés, les poissons Téléostéens représentent le groupe où la diversité de pigmentation est probablement la plus élevée, tant en termes de patrons de pigmentation que de types cellulaires (Parichy, 2006; Braasch *et al.*, 2008). Ainsi, alors que chez les Mammifères et les Oiseaux il n'y a qu'un seul type de chromatophore, et qu'il y en a jusqu'à quatre ou cinq chez les "Reptiles" et Amphibiens, les poissons Téléostéens en comptent au moins huit, peut-être même 10. Les poissons étant souvent transparents pendant une partie de leur cycle de vie, l'étude de la pigmentation y est aisée (Wakamatsu *et al.*, 2001).

Parmi les chromatophores, on retrouve les mélanocytes, xanthophores, erythrophores et iridophores, déjà décrits plus haut et structurellement équivalents chez les poissons (Leclercq *et al.*, 2009). À ces quatre types cellulaires s'ajoutent les **leucophores** : ces cellules sont blanches chez l'adulte en raison de la présence de plaquettes réfléchissant la lumière au sein d'organistes spécialisés, les leucosomes (Menter *et al.*, 1979; Fujii, 1993). Les plaquettes, constituées d'acide urique (une purine), sont à l'origine du nom de "guanophore" initialement donné à ces cellules (Odiorne, 1933). Rapprochées pendant longtemps des iridophores, les leucophores sont dendritiques et diffèrent en cela des premiers, plutôt ronds (Menter *et al.*, 1979). Ces cellules sont également de taille comparable aux autres chromatophores

dendritiques, les mélanocytes, les érythrophores et les xanthophores. De plus, chez les embryons de Médaka *Oryzias latipes*, les leucophores prennent des teintes jaunes en raison de la présence d'un pigment, la drosoptérine, une ptéridine proche de la sépiaptérine des xanthophores (Oliphant and Hudon, 1993), avec des voies de biosynthèse proches (Braasch *et al.*, 2007). Ceci a conduit certains auteurs à suggérer que leucophores et xanthophores devaient être des types cellulaires proches (Nagao *et al.*, 2014; Kimura *et al.*, 2014).

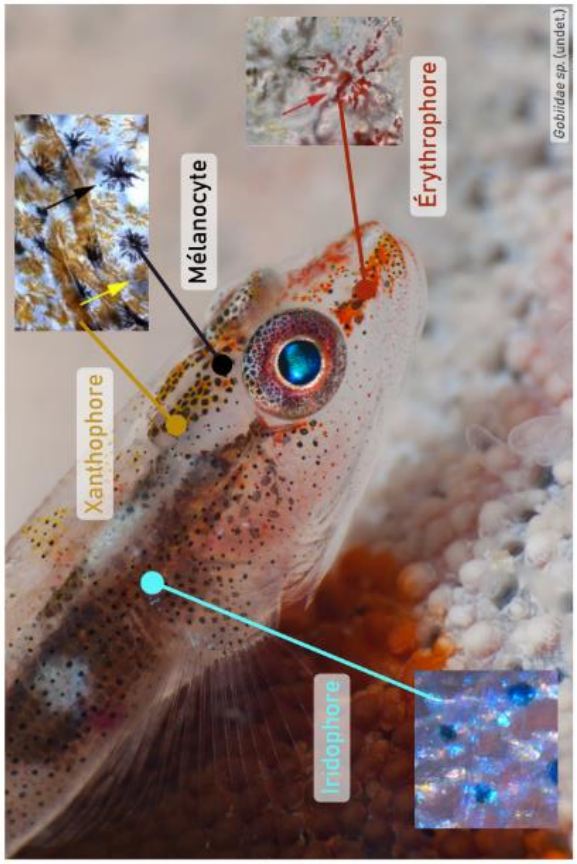
Outre les leucophores, les poissons Téléostéens comptent d'autres types cellulaires aux pigments souvent encore non caractérisés avec certitude mais dont les descriptions cellulaires sont robustes :

- Les cyanophores : ces cellules sont bleues et ont été décrites chez le Diamant mandarin *Synchiropus splendidus* (Goda and Fujii, 1995; Bagnara *et al.*, 2007)
- Les chromatophores fluorescents, peu colorés à la lumière visible mais qui fluorescent dans le rouge. Leur fluorescence est probablement due à des cristaux de guanine et ils diffèrent pour autant des leucophores et des iridophores (Michiels *et al.*, 2008; Wucherer and Michiels, 2012, 2014). Ils seraient impliqués dans la communication en eaux profondes (sous 10 mètres) (Michiels *et al.*, 2008; Meadows *et al.*, 2014)
- Les chromatophores "iridophore-like", ressemblant à des iridophores mais pour lesquels les plaquettes réflectives sont en périphérie de la cellule au lieu d'être dans l'ensemble du cytosol (Goda, 2017)
- Les érythro-iridophores, rouges/violets, et contenant à la fois des caroténoïdes et des plaquettes de purines (Goda *et al.*, 2011)
- Les cyano-érythrophores, bleus/rouges, contenant des ptéridines, des caroténoïdes et un pigment bleu inconnu (Goda *et al.*, 2013)

Certains de ces types cellulaires n'ont été décrits que chez les poissons coralliens suite aux travaux de M. Goda, "le Bagnara des poissons", qui documente régulièrement des nouveaux types de chromatophores. Ainsi, en 2004, dans une revue de référence sur le sujet, le spécialiste de la pigmentation Robert Kelsh annonçait que "*up to six different types of pigment cells*" avaient été documentés chez les Téléostéens (Kelsh, 2004). En 2015, Singh affirmait qu'il y avait "au moins sept différents types" (Singh and Nüsslein-Volhard, 2015), et, en 2016, R. Kelsh signait en dernier auteur une revue indiquant qu'au moins huit types de chromatophores existaient (Schartl *et al.*, 2016). Ce champ de recherche est donc plus actif que chez les Lépidosauriens et les Amphibiens. Encore récemment, d'ailleurs, une étude est parue chez les Salmonidés s'intéressant à la différence entre xanthophores et érythrophores et a montré que les érythrophores étaient uniquement colorés par des caroténoïdes ; les xanthophores, à l'inverse, contenaient également des ptérinosomes (Djurdjevič *et al.*, 2015). La distinction xanthophore/érythrophore, basée jusqu'à présent sur des purs critères de teinte, est donc peut-être "chimiquement réelle", et de nouvelles études sont nécessaires dans le domaine.

L'ensemble des types cellulaires décrits chez les poissons ne signifie pas qu'une seule espèce les possède tous. Par exemple, chez le Poisson-zèbre *Danio rerio*, il n'y en a "que trois" : xanthophores, mélanocytes et iridophores. C'est la lignée dans son ensemble qui est plus diversifiée que le reste des Vertébrés, et que les Tétrapodes en particulier.

Chez les Téléostéens, la coloration repose quasi-uniquement sur les cellules qui portent les pigments (ou les molécules réfléchissant la lumière). Ceci les distingue des Mammifères et des Oiseaux, groupes dans lesquels la coloration est non seulement due à des cellules, mais aussi à des pigments déposés dans des tissus, et pas *au sein* des cellules. Il en résulte, chez les Téléostéens, à l'instar de ce qui est observé chez les "reptiles" et Amphibiens, une couleur due à une interaction entre plusieurs types cellulaires. Par exemple, chez le Poisson-zèbre, les bandes claires sont dues à une superposition de xanthophores et d'iridophores très denses, alors que les bandes sombres sont constituées de xanthophores épars et peu colorés et d'iridophores peu denses au-dessus d'une couche de mélanocytes, elle-même au-dessus d'une autre couche d'iridophores (Singh and Nüsslein-Volhard, 2015; Irion *et al.*, 2016). Cet agencement relatif des chromatophores n'est pas toujours aussi structuré au sein des Téléostéens et il est plus variable dans ce groupe que chez les Lépidosauriens et Amphibiens (Bagnara and Matsumoto, 2006; Kottler *et al.*, 2013), si bien qu'on ne parle pas de "*dermal chromatophore unit*".



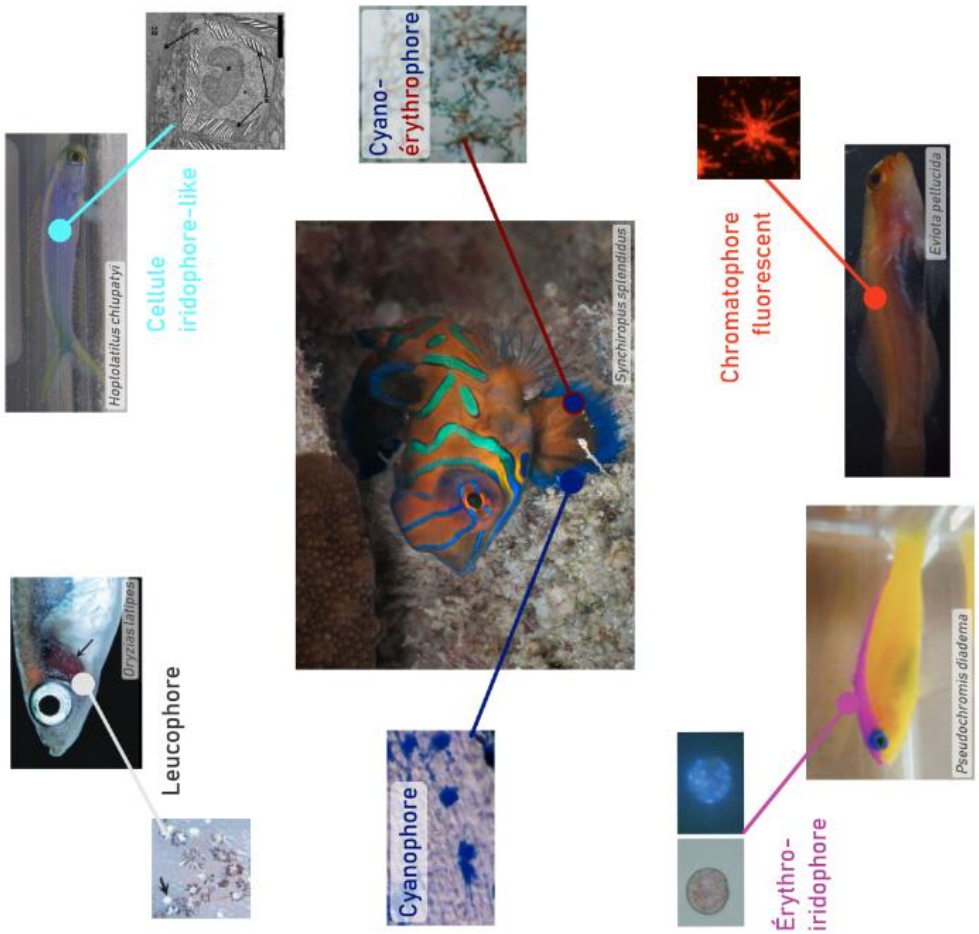


Figure 14 – Les chromatophores décrits en 2018 chez les Téléostéens.

Chez les poissons Teleosteens ont deja ete decrits au moins 10 types de chromatophores. Ils ne sont pas tous présents au sein d'une même espèce et il est rare qu'une espèce en présente plus de quatre types différents. Chez les poissons les chromatophores sont bien observables, notamment à l'état larvaire, en raison de la transparence des tissus. La larve de gobie (à gauche) mesure 2 cm de long.

Droits réservés pour les individus : Gobie et S. splendidus : Germain Boussarie ; O. latipes : Wakamatsu et al. (2001) ; H. chlupatyi : Goda (2017) ; P. diadema : Goda et al. (2011) ; E. pellucida : Michiels et al. (2008). Encarts : D. Parichy (iridophore), Cortesi et al. (2016) (xanthophores, érythrophores, mélanocytes), Wakamtsu et al. (2001) (leucophore), Goda (2017) (iridophore-like) ; Goda and Fujii (1995) (cyanophore) ; Goda et al. (2013) (cyano-érythrophore) ; Goda et al. (2011) (érythro-iridophore) ; Wucherer et al. (2012) (chromatophore fluorescent).

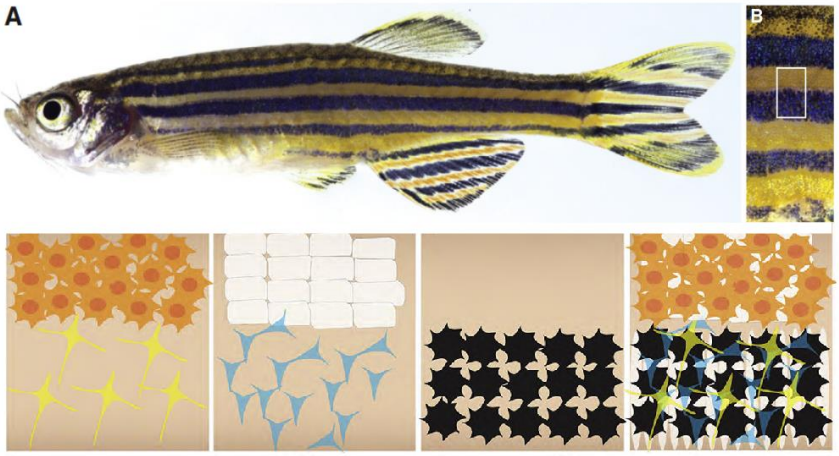


Figure 15 - Agencement des chromatophores dans la peau du Poisson-zèbre.
Haut : vue latérale de Poisson-zèbre (A) et détail sur les bandes claires et sombres (B).

Bas : De haut en bas (de gauche à droite), la bande claire est constituée de deux couches denses de xanthophores et d'iridophores. La bande sombre est elle constituée d'une couche de xanthophores épars surmontant une couche d'iridophores épars et une couche de mélanophores denses. Sous les mélanophores, il y a encore une couche d'iridophores.

Adapté de Singh and Nüsslein-Volhard (2015) et de Irion et al. (2016).

Chez les Mammifères, les tons plus ou moins clairs sont supportés par le ratio entre phéo- et eumélanine. On note que la phéomélanine n'est pas présente chez les poissons Téléostéens : certains auteurs l'avaient suggéré dans le papier sur le génome de la Carpe (Xu *et al.*, 2014a), mais ceci avait été remis en question dans la foulée (Kottler *et al.*, 2015). Cet exemple montre que "plus de types de cellules pigmentaires" n'est pas linéairement corrélé avec "une plus grande diversité de couleur". Outre les cellules, l'équipement enzymatique permettant la production d'un pigment ou d'un autre est également important, et, les Mammifères, avec un seul type de cellules pigmentaires et l'absence de couleurs "structurales", présentent tout de même une grande diversité de pigmentation (Caro, 2005) – même si on cherche encore un Mammifère vert…

Un même type de cellule pigmentaire peut parfois produire deux types de couleurs différentes : c'est le cas des iridophores, qui permettent de réfléchir les ultraviolets chez les poissons Téléostéens, et notamment chez les poissons coralliens Pomacentridae (demoiselles) (Siebeck *et al.*, 2010). Ces couleurs sont effectivement perçues par les conspécifiques et pas par les prédateurs, soulignant leur importance dans la communication intraspécifique au sein de ce groupe (Siebeck, 2014; Stieb *et al.*, 2017). La communication basée sur les UV n'est toutefois pas spécifique des Téléostéens et concerne aussi les Oiseaux (Barreira *et al.*, 2012).

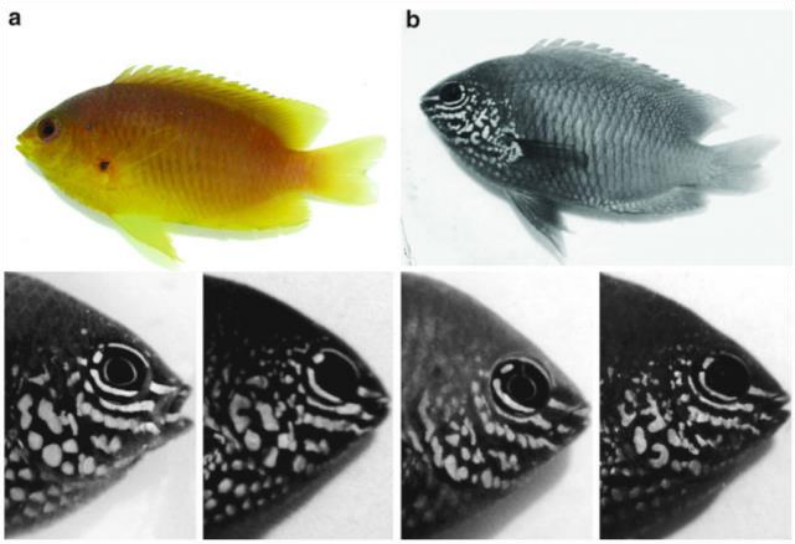


Figure 16 - Un même chromatophore peut être à l'origine de plusieurs couleurs.

Haut : *Pomacentrus amboinensis* vu en lumière visible (pour l'Homme) (a) et dans les UV (b).

Bas : De gauche à droite, quatre individus différents. Les iridophores, ressortant peu en lumière visible, sont clairement visibles. Les motifs sont utilisés pour de la reconnaissance inter- et intraspécifique. *Adapté de Siebeck (2014).*

Enfin, l'étude des chromatophores dans des lignées diverses chez les Téléostéens a également permis de caractériser certaines de leurs fonctions, jusqu'alors inconnues. Chez le tilapia *Oreochromis niloticus*, les érythrophores contiennent des photorécepteurs et agissent comme détecteurs de lumière, influençant donc probablement les cycles biologiques. Cela pourrait également jouer un rôle dans le camouflage (Ban *et al.*, 2005; Chen *et al.*, 2013, 2015).

III. Les bases génétiques et développementales de la coloration sont multiples

"*The only interesting thing about vertebrates is the neural crest*." – Peter Thorogood

1. Importance de l'étude de la coloration dans l'histoire des sciences

La redécouverte des lois de Mendel au début du XXème siècle s'est accompagnée d'une recherche de la base génétique des variations des attributs phénotypiques chez de nombreux organismes. Cette variation étant le préalable à l'adaptation, l'étude des gènes sous-tendant la variation phénotypique a permis une meilleure compréhension du processus adaptatif et de l'évolution des populations. Ainsi, comprendre comment la variation phénotypique émerge et est maintenue nécessite de relier les bases génétiques et développementales à des facteurs écologiques.

Pour de multiples raisons, l'étude de la coloration en général, et des animaux en particulier, a de longue date intéressé les naturalistes. La couleur, outre son intérêt esthétique, est également un trait visible aux variations relativement aisément mesurables, en faisant un bon modèle d'étude pour comprendre la variation phénotypique. Les "lois de Mendel" ont d'ailleurs été développées suite à l'étude de variations de formes et de couleurs des fleurs, graines et fruits des pois *Pisum sativum* (Bateson, 1902). Leur redécouverte a également hautement profité de la présence de deux morphes, aux yeux rouges ou blancs, chez la Drosophile (Morgan, 1910). Aujourd'hui, de nombreux mutants de la pigmentation sont disponibles en laboratoire, et la compréhension de la pigmentation au niveau génétique est avancée, permettant l'étude des bases génétiques de l'adaptation en milieu naturel. Cela n'a toutefois pas toujours été le cas et est le fruit d'un peu plus d'un siècle d'efforts.

L'étude de la pigmentation chez les Vertébrés a commencé au XVIIIème siècle en Asie, lorsque des passionnés ont commencé à sélectionner artificiellement des souches de couleur différentes chez les Souris. De la même manière, surtout en Europe, les races de pigeons, de chiens ou d'autres espèces domestiques ont été peu à peu sélectionnées pour des raisons autres que fonctionnelles, et les concours d'élevage ont très vite consacré les animaux qui remplissaient des critères morphologiques ou de couleur (Domyan and Shapiro, 2017). Le développement en laboratoire des espèces de Vertébrés était plus complexe que celui de la Drosophile mais T. H. Morgan lui-même s'essaya à l'étude des pigeons, et montra en 1911 que "*The white top of the head and the creast appear to be dominant; while the white wing bars are recessive"* (Morgan, 1911). Cependant, la praticité du modèle murin en laboratoire le consacra rapidement pour les études génétiques sur la pigmentation.

En 1903, avant même que le terme "génétique" soit inventé par Bateson (en 1905), des études de croisement de souris albinos ont montré que la transmission des gènes chez les mammifères était également mendélienne. La première carte génétique chez les Vertébrés (1915), les relations d'épistasie et la pléiotropie, puis les premières estimations des taux de mutations chez les Mammifères ont été déterminées par des études chez la Souris (Hoekstra, 2006) et références associées). L'étude de la pigmentation a donc considérablement influencé l'origine de concepts fondamentaux en génétique.

Le premier clonage d'un gène impliqué dans la pigmentation date de 1986 : le gène *Tyrp1*, originellement confondu avec le gène codant pour la Tyrosinase (Shibahara *et al.*, 1986; Jackson, 1988). Ce fut le début d'une longue série : vingt ans plus tard, environ 100 gènes étaient connus (Bennett and Lamoreux, 2003), et aujourd'hui, quasiment le double. *Mc1r*, un autre gène majeur de la pigmentation, a lui été cloné en 1992 (Mountjoy *et al.*, 1992).

L'importance de la pigmentation dans "l'histoire contemporaine" des sciences s'est également manifestée plus récemment. Les années 1970 ont ainsi vu l'émergence du modèle Poisson-zèbre. Ce nouvel organisme modèle, transparent pendant quasiment tout son développement, permettait d'étudier des processus développementaux jusqu'alors difficilement accessibles (Grunwald and Eisen, 2002). Il n'a donc pas été étonnant que la couverture de Nature du 28 mai 1981 y soit consacrée (Streisinger *et al.*, 1981). Or, parmi l'ensemble des photographies possibles pour choisir une couverture, c'est justement un mutant pigmentaire qui a été choisi : le mutant *golden*, aux embryons non-mélanisés. Ainsi, dès l'émergence de ce modèle, le déterminisme génétique de la pigmentation a été mis au premier plan, consciemment ou non (Haffter *et al.*, 1996; Kelsh, 2004).

Figure 17 - Couverture de *Nature* le 28 mai 1981.

Le 28 mai 1981, les embryons du mutant *golden* de Poisson-zèbre déficient en mélanocytes (le mutant adulte est à droite) font la une de la revue *Nature*. Parmi tous les mutants possibles pour illustrer l'émergence de ce nouveau modèle, c'est un mutant pigmentaire qui est choisi, soulignant le rôle prépondérant joué par la pigmentation dans les directions prises par la communauté scientifique, Echelle : 5 mm (encart : 0.5 mm). *Adapté de Grunwald and Eisen (2002) et de Lamason et al. (2005).*

En 2005, une étude a montré que le mutant *golden* était en fait dû à une mutation d'un gène codant pour une protéine d'échange de cations, *slc24a5* (Lamason *et al.*, 2005). De manière remarquable, ce gène découvert chez le Poisson-zèbre a pu être associé à la variation pigmentaire trouvée dans l'espèce humaine : à lui seul, ses variations expliquent entre 25 et 38% des différences dans l'indice mélanique entre Européens et Africains (qui est un trait hautement multigénique) (Barsh, 2003; Lamason *et al.*, 2005). L'étude du Poisson-zèbre permet donc de mieux comprendre les variations de pigmentation au sein de notre espèce et révèle que les voies permettant la production de mélanine sont étaient probablement déjà présentes chez l'ancêtre des Vertébrés. Cet organisme s'est sans nul doute avéré être un modèle précieux pour de très nombreux aspects de la biologie des Vertébrés (Grunwald and Eisen, 2002), mais peut-être que son essor n'aurait pas été le même sans le partage avec les Mammifères des bases génétiques de la pigmentation mélanique.

Toutefois, le Poisson-zèbre est également un modèle intéressant pour les autres espèces de Vertébrés qui partagent avec lui plusieurs autres types de cellules pigmentaires (Odenthal *et al.*, 1996). L'étude de cet organisme a donc certes permis de mieux comprendre comment se développent les mélanocytes, mais également les autres chromatophores.

2. L'utilisation de nouveaux modèles a permis une meilleure compréhension de la différentiation des chromatophores

a. Origine développementale des chromatophores

Les chromatophores sont des cellules qui dérivent des cellules des crêtes neurales (Le Douarin and Dupin, 2003). Ces cellules multipotentes sont à l'origine d'une grande diversité de types cellulaires qui se forment au cours du développement, et ont parfois été qualifiées de "quatrième feuillet embryonnaire Vertébré" (Hall, 2000; Knecht and Bronner-Fraser, 2002; Simoes-Costa and Bronner, 2015). Les cellules pigmentaires, en particulier, migrent jusque dans

le derme, voire l'épiderme (Mort *et al.*, 2015). Pour un même type de chromatophore, les voies de migration empruntées sont variables selon les différents taxons de Vertébrés (Kelsh *et al.*, 2009), et également, pour un même organisme, au cours du développement (Singh and Nüsslein-Volhard, 2015). Elles diffèrent, de plus, selon les types de chromatophores (Singh and Nüsslein-Volhard, 2015). Sans rentrer dans le détail des voies de migration permettant la mise en place des chromatophores au cours du développement, on peut retenir que chez le Poisson-zèbre, xanthophores, iridophores et mélanocytes se mettent en place selon des voies différentes :

1. Les mélanocytes larvaires migrent directement depuis la crête neurale dans le derme ; d'autres précurseurs migrent jusqu'aux ganglions de la racine dorsale de la moëlle épinière, et y restent indifférenciés (Dooley *et al.*, 2013). À la métamorphose, ces derniers se différencient en mélanocytes et migrent en passant par le myosepte horizontal (un espace entre les masses musculaires ventrale et dorsale), à partir duquel ils migrent vers le dos ou vers le ventre.

2. Les iridophores ne sont pas présents sur le côté du corps chez la larve et, à la métamorphose, migrent selon la même voie que les mélanocytes depuis les ganglions dorsaux (Singh *et al.*, 2014).

3. Les xanthophores, enfin, migrent à l'état larvaire dans le derme ; les xanthophores post-métamorphiques, adultes, sont issus de la prolifération des xanthophores larvaires directement sur place (Kelsh *et al.*, 2009; Mahalwar *et al.*, 2014).

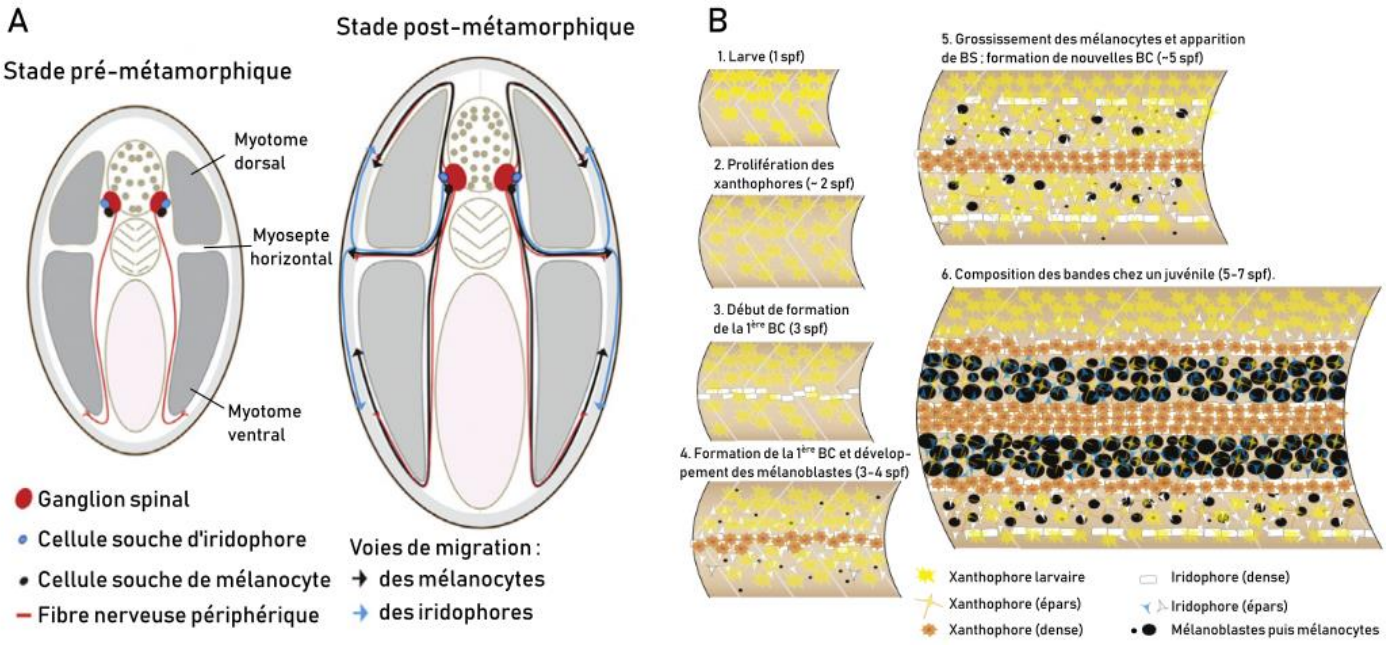


Figure 18 - Développement des chromatophores chez le Poisson-zèbre.
(A) Au cours de la métamorphose du Poisson-zèbre, il existe plusieurs voies de migration des précurseurs de chromatophores dont une voie à travers le myosepte horizontal (la seule pour les iridophores). Les xanthophores se différencient directement dans la peau alors que les autres chromatophores se différencient depuis des cellules souches des ganglions spinaux et migrent le long des fibres nerveuses. (B) La première bande claire (BC) se forme au bout de 3 semaines post-fécondation (spf), au milieu du corps. Peu à peu, les bandes sombres (BS) apparaissent et de nouvelles bandes claires se forment suite à l'arrivée par le milieu de nouveaux iridophores. Orientation : le dos est en haut, pour chaque figure. *Adapté de Singh and Nüsslein-Volhard (2015).*

b. Les relations entre chromatophores

i. Relations ontogénétiques entre les chromatophores

Les Mammifères ne présentant qu'un seul type de chromatophore, c'est vers le modèle "poisson" que la communauté s'est tournée pour l'étude des relations ontogénétiques entre chromatophores. Chez le Poisson-zèbre, trois

chromatophores sont présents, et mélanocytes et iridophores partagent un progéniteur commun, différent de celui des xanthophores (Kelsh *et al.*, 2009; Mahalwar *et al.*, 2014). Au cours du développement de celui-ci, l'expression relative des facteurs de transcription *mitfa* (induisant les mélanocytes) et *foxd3* (induisant les iridophores) permettait d'orienter le progéniteur vers l'un ou l'autre des deux destins (Curran *et al.*, 2010). Les xanthoblastes (précurseurs des xanthophores), à l'inverse, expriment le gène *csf1ra* (anciennement nommé *fms*), qui n'est pas exprimé dans la lignée des deux autres types de chromatophores (Parichy *et al.*, 2000; Curran *et al.*, 2010).

En 2016, cependant, une étude de l'équipe de C. Nüsslein-Volhard suggérait que les xanthophores adultes pouvaient provenir de la différentiation d'un précurseur de mélanocyte et xanthophores, mais pas d'iridophores (Singh *et al.*, 2016). Ces résultats récents suggèrent que les voies de différentiation ne semblent pas encore complètement caractérisées et qu'il est en réalité possible qu'il existe plusieurs types de progéniteurs ayant des potentialités sensiblement différentes (Walderich *et al.*, 2016; Singh *et al.*, 2016).

Quid, dès lors, des autres types de cellules pigmentaires ? L'étude du médaka *O. latipes*, un autre modèle de poisson, est intéressante, car elle permet d'étudier dans un même organisme non pas trois mais *quatre* chromatophores : en plus de ceux du Poisson-zèbre s'ajoutent les leucophores *(des leucophores ont déjà été décrits dans les nageoires caudale et dorsale du Poisson-zèbre, mais c'est anecdotique (Irion et al., 2016))*. Ces cellules sont, on l'a vu, blanches et constituées de plaquettes d'acide urique, ce qui les rapproche des iridophores. Toutefois, les leucophores embryonnaires et larvaires sont jaunâtres et contiennent des ptéridines rappelant les ptérinosomes des xanthophores (Menter *et al.*, 1979). De plus, leur morphologie rappelle celle de ces derniers, plus que celle des iridophores. Deux études parues en 2014 ont montré que les leucophores et les xanthophores partagent un même progéniteur, qui exprime le facteur de transcription *pax7a*, à l'inverse du progéniteur de commun aux mélanocytes et iridophores. Ensuite, les xanthophores se différencient du pool de progéniteurs exprimant *sox5* (Nagao *et al.*, 2014; Kimura *et al.*, 2014).

ii. Les relations cellulaires entre chromatophores sont à l'origine des patrons de pigmentation

Chez les poissons Téléostéens, les Lépidosauriens et les Amphibiens, la couleur "brute" est due à une interaction entre les chromatophores, qui partagent une origine développementale commune. Toutefois, leurs interactions interviennent à un autre niveau de la pigmentation : l'acquisition des motifs, ou patrons, de pigmentation. Nous avons vu à travers l'exemple des poissons coralliens en quoi ceux-ci pouvaient être très divers. Les poissons coralliens sont cependant peu utilisés (nous reviendrons là-dessus en troisième partie), mais de nombreux travaux menés chez les poissons Téléostéens permettent de mieux comprendre comment les interactions cellulaires permettent de générer des patrons de pigmentation variés.

L'étude des mécanismes permettant de générer des patrons développementaux, en bandes ou en taches, date des travaux d'Alan Turing et de son modèle de réaction-diffusion (Turing, 1952; Maini *et al.*, 2006). Un sous-type de ce modèle, le modèle de Gierer-Meinhardt, permet d'expliquer la formation de patrons lorsqu'un élément activateur (molécule ou cellule) s'auto-active localement, et qu'un autre élément, inhibiteur, l'inhibe depuis un endroit plus éloigné (Meinhardt and Gierer, 2000; Kondo and Miura, 2010). Si, dans l'absolu, un tel mécanisme est concevable pour des molécules diffusant librement, il est en revanche plus difficile d'imaginer comment il pourrait être mis en œuvre dans un système aussi complexe qu'un tissu biologique où cellules et molécules ne diffusent souvent pas librement.

Ces travaux ont donc constitué des avancées théoriques majeures au milieu du XXème siècle, mais il a fallu attendre ces deux dernières décennies pour qu'on puisse associer aux "éléments" une réalité moléculaire ou cellulaire (Maini *et al.*, 2006). C'est en particulier l'étude de la dynamique des patrons qui, sans entrer dans le détail des mécanismes moléculaires et cellulaires, a validé la théorie pour la première fois au cours de la croissance de poissons coralliens, les Poissons-anges *Pomacanthus* (Kondo and Asai, 1995; Painter *et al.*, 1999).

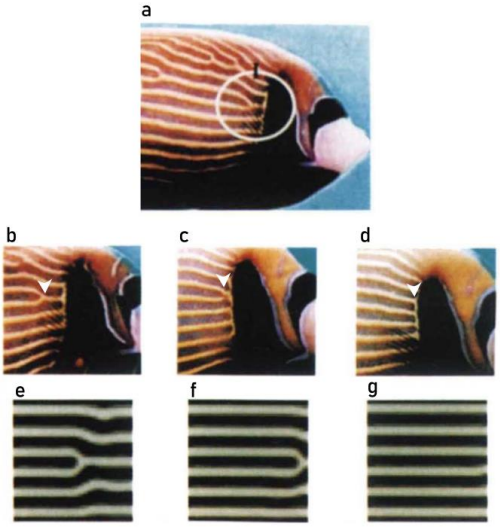


Figure 19 - Motifs de Turing chez un poisson corallien.

(a) Situation initiale du motif des bandes horizontales du Poisson ange-empereur. (b,c,d) Evolution du motif au cours de la croissance de l'individu (flèche blanche) : les deux bandes, se rejoignant initialement, se séparent. (e,f,g) Modèle numérique de la croissance d'un tissu gouverné par des équations de réaction-diffusion. Le modèle permet d'expliquer les observations. *Adapté de Kondo and Asai (1995).*

Chez les Mammifères, ce sont des études chez la Souris qui ont permis de comprendre comment ce modèle explique la disposition des follicules pileux (Sick *et al.*, 2006) : les éléments sont alors des molécules (Wnt et Dkk). Toutefois, loin de diffuser librement, leur position est au contraire activement contrôlée par les filopodes cellulaires, soulignant le rôle des cellules dans ce phénomène de "réaction-diffusion" (Stanganello *et al.*, 2015).

Chez les poissons, ce sont des études chez le Poisson-zèbre qui ont permis de mieux comprendre l'origine développementale de ces patrons (Hamada *et al.*, 2014). Le modèle n'est pas aussi "strict" que chez les Mammifères, puisque les "éléments" ne sont alors pas des molécules, mais des cellules (Maderspacher, 2012). Chez le Poisson-zèbre, mélanocytes et xanthophores se repoussent mutuellement à courte distance (Nakamasu *et al.*, 2009). Cette double inhibition résulte en une auto-activation locale des xanthophores. À l'inverse, les mélanocytes, dans la bande sombre, requièrent la présence de xanthophores dans la bande claire, à longue distance, pour survivre. Ils sont donc activés à longue distance, ce qui active leur action inhibitrice sur les xanthophores localement (Parichy and Turner, 2003). On retrouve donc bien le modèle de la réaction-diffusion, mais appliqué à des *cellules*, et non à des molécules : les xanthophores s'auto-activent localement et s'auto-inhibent à longue distance, par l'intermédiaire des mélanocytes (Watanabe and Kondo, 2015). Les bases cellulaires commencent à être connues et pourraient reposer sur des interactions Delta/Notch entre xanthophores et mélanocytes (Hamada *et al.*, 2014; Eom *et al.*, 2015). Les macrophages pourraient également agir en tant qu'intermédiaire entre ces deux types de chromatophores (Kornberg, 2017; Eom and Parichy, 2017). Le rôle des iridophores n'est lui toujours pas établi, mais sans eux, le patron ne se forme pas correctement (Watanabe and Kondo, 2015)

Un aspect intéressant de ce mécanisme est qu'une variation faible de l'intensité relative de l'inhibition ou de l'auto-activation peut être à l'origine de patrons de pigmentation divers. Par exemple, en induisant l'expression de variants différents d'un gène codant pour une connexine (une protéine de jonctions gap), *cx41.8*, Watanabe et Kondo ont réussi à reproduire une grande variation de mutants pigmentaires (Watanabe and Kondo, 2012).

En conclusion, ces mécanismes de réaction-diffusion peuvent expliquer en partie la variation des patrons pigmentaires observés dans la nature, et même certains patrons jusque récemment mal compris (Parichy, 2007; Mills and Patterson, 2009; Kondo and Miura, 2010). Ils permettent également d'expliquer les variations de patrons pigmentaires au cours du processus de croissance d'un individu (Kondo and Asai, 1995).

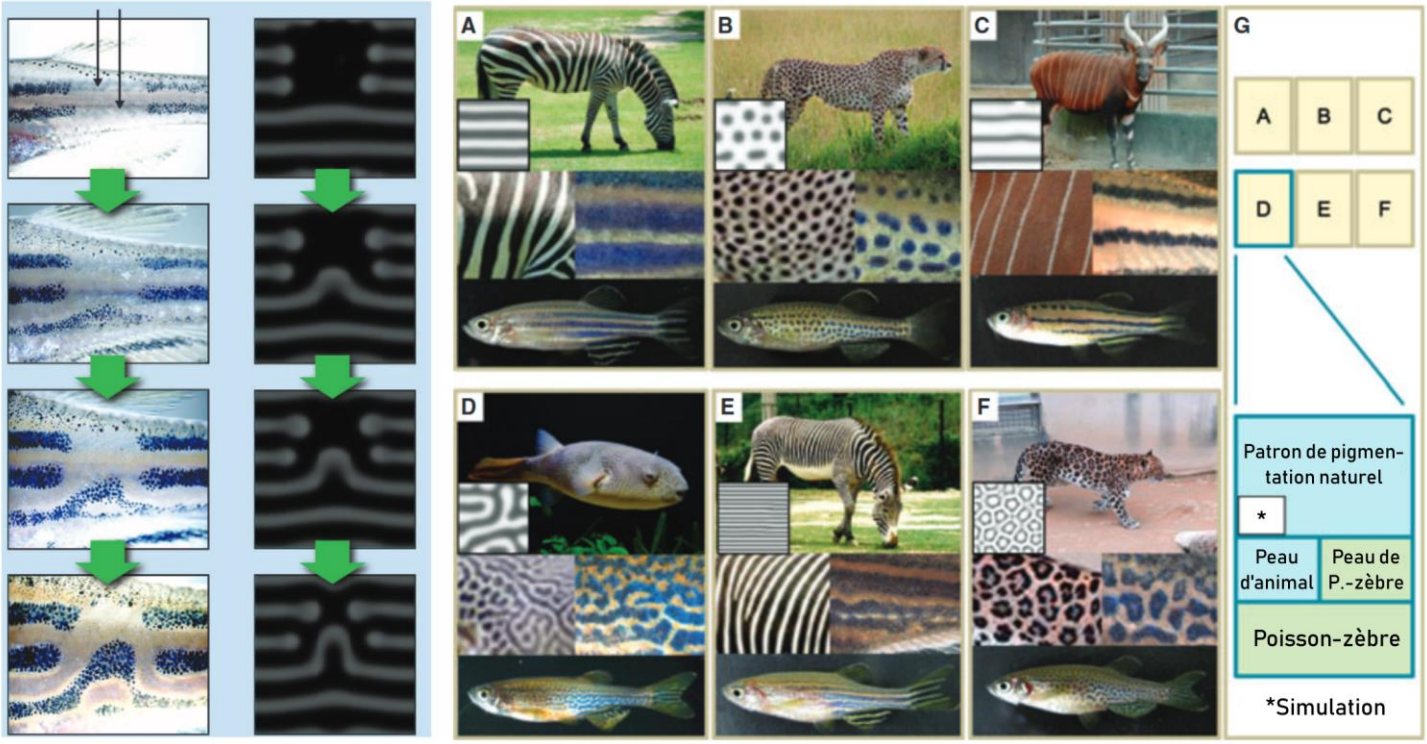


Figure 20 - Les motifs des poissons du genre *Danio* répondent au modèle de Turing.
Gauche : Le retour de la pigmentation après ablation laser des chromatophores du poisson-zèbre reproduit les prédictions des équations de réaction-diffusion. *D'après Kondo and Miura (2010).*

Droite : Comparaison entre le patron d'un animal, le patron modélisé par les équations de réaction-diffusion, et différents mutants de Poisson-zèbre pour le gène de la connexine *cx41.8*. Les variations de 'expression d'un seul gène peuvent induire des patrons très différents et peuvent expliquer les bases de la variation phénotypique. *Adapté de Watanabe and Kondo (2012).*

3. Vers une définition d'un "gène de la pigmentation" ?

J'ai jusqu'ici évoqué les cellules pigmentaires ainsi que les bases développementales permettant d'expliquer l'apparition de patrons colorés différents. J'ai brièvement évoqué quelques gènes associés, sur lesquels je reviens avec plus de détail.

a. Origine du terme

Originellement, les gènes impliqués dans l'acquisition du phénotype coloré ont été qualifiés de "color genes", ou "gènes de couleur", par Bennett et Lamoreux (2003), dans un article justement intitulé *The color loci of mice.* Bennett and Lamoreux désignaient par "*color loci*" "tous les loci génétiques affectant la *pigmentation* d'un vertébré, plus spécifiquement des cheveux, de la peau et/ou des yeux" (Bennett and Lamoreux, 2003). Ainsi, dans la définition d'un "gène de couleur" figure uniquement le terme "pigmentation". Or, comme je l'ai évoqué plus haut, la couleur ne dépend pas uniquement des pigments : chez les Mammifères, peut-être, mais chez les autres Vertébrés, non. D'autres gènes ont d'ailleurs été caractérisés chez les Oiseaux et sont connus pour affecter la couleur non pas en affectant la pigmentation de l'animal mais en affectant la structure de son plumage (Roulin and Ducrest, 2013). Dès lors, les gènes de coloration *incluent* les gènes de la pigmentation : ces derniers sont plus spécifiques. En 2006, trois ans plus tard, dans une autre revue de référence sur le sujet, H. Hoekstra utilisait d'ailleurs le terme de *"pigmentation gene*".

La définition initiale de Bennett et Lamoreux inclut les gènes impliqués dans la coloration des "yeux". Cependant, l'épithélium pigmenté de la rétine, un épithélium, n'a pas la même origine développementale que les chromatophores de la peau, qui dérivent des cellules des crêtes neurales. Qui plus est, ces cellules du fond de l'œil ne sont pas soumises aux

mêmes contraintes évolutives car elles n'influent que marginalement sur la valeur adaptative due à la couleur d'un individu.

Par la suite, je propose donc d'utiliser le terme de "gène de la pigmentation" strictement, pour qualifier "un gène affectant la pigmentation d'un animal à travers le développement de cellules pigmentaires dérivées de cellules des crêtes neurales". Les leucophores et iridophores, malgré leur absence de pigment *s. s.*, sont inclus dans cette définition car ils dérivent des cellules des crêtes neurales et sont fonctionnellement équivalents aux autres cellules pigmentaires.

b. Les gènes de la pigmentation gouvernent des processus multiples

i. Les gènes de la pigmentation sont parfois associés seuls à une fonction pigmentaire…

La pigmentation est, nous l'avons vu, un processus hautement multifactoriel. Les "gènes de la pigmentation" recouvrent donc une réalité variée et leurs produits peuvent être fonctionnellement très différents, bien que participant à la formation du même trait biologique. Ici, je ne les détaille pas tous, mais en illustre quelques-uns pour la majorité des différentes fonctions, afin d'illustrer leur diversité.

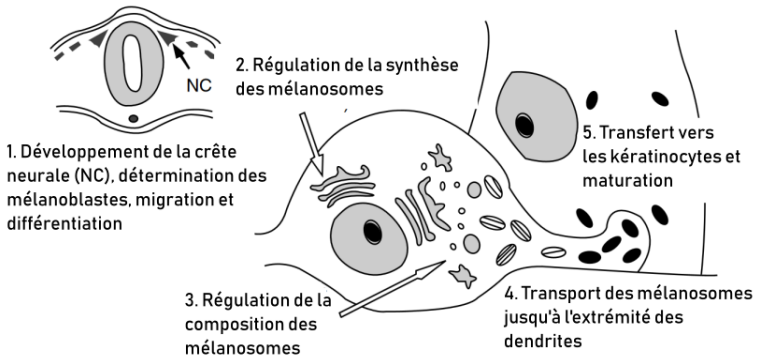


Figure 21 - Diversité des processus permettant l'acquisition d'une pigmentation fonctionnelle chez les mammifères.

Adapté de Bennett and Lamoreux (2003).

En premier lieu, les gènes de la pigmentation peuvent être des gènes impliqués dans la migration des précurseurs de chromatophores depuis la crête neurale. Par exemple, un défaut dans le gène *kit* induit une mauvaise migration de ces cellules et peut conduire à des individus leuciques car sans mélanocytes dans la peau (Geissler *et al.*, 1988).

Les gènes de la pigmentation peuvent également être des gènes permettant la différentiation en chromatophores depuis un précurseur de cellule de crête neurale. Par exemple, le gène *mitf* code pour un facteur de transcription permettant la différentiation en mélanocyte (Tassabehji *et al.*, 1994). Muté, l'organisme ne présente pas de mélanocytes (mutant *nacre* chez le Poisson-zèbre). De même, un mutant pour le gène *csf1ra* ne présente pas de xanthophores (mutant *panther*), et un mutant pour le gène *ltk* ne présente pas d'iridophores (mutant *shady*) (Parichy *et al.*, 2000; Lopes *et al.*, 2008). Les chromatophores interagissant, la perte d'un type cellulaire influe souvent sur les autres : par exemple, le mutant *panther* a également moins de mélanocytes car chez le sauvage les xanthophores envoient des signaux de survie aux mélanocytes (Parichy *et al.*, 2000).

Par ailleurs, les gènes de la pigmentation peuvent être impliqués dans le patterning, souvent à travers le développement d'une cellule pigmentaire. Par exemple, la connexine 41.8, appelée *cx41.8* ou *gja5b*, est à la fois impliquées dans le patterning et la survie des mélanocytes chez le Poisson-zèbre (Watanabe *et al.*, 2006; Watanabe and Kondo, 2012).

À l'extrême de gènes impliqués dans la prolifération ou la survie d'un type de chromatophore, les gènes de la pigmentation peuvent également être impliqués dans l'étape terminale de la voie de synthèse biochimique d'un pigment. Par exemple, le gène *tyrp1* est impliqué dans une des dernières étapes de la synthèse d'eumélanine. Muté, les souris sont appelées *brown* car le ratio eumélanine/phéomélanine est alors en faveur de la première (del Marmol and Beermann, 1996). Si, par ailleurs, le gène *tyr*, codant pour la Tyrosinase, une enzyme au début de la voie de synthèse de mélanine

(eu- et phéomélanine), est muté, alors l'individu sera albinos (Norman *et al.*, 2017). À l'inverse d'un individu leucique, un albinos ne produit pas de mélanine du tout, même dans l'épithélium pigmenté de la rétine. De nombreuses mutations peuvent provoquer l'albinisme car la voie de différentiation des mélanocytes ainsi que la voie de synthèse de la mélanine sont des processus multigéniques.

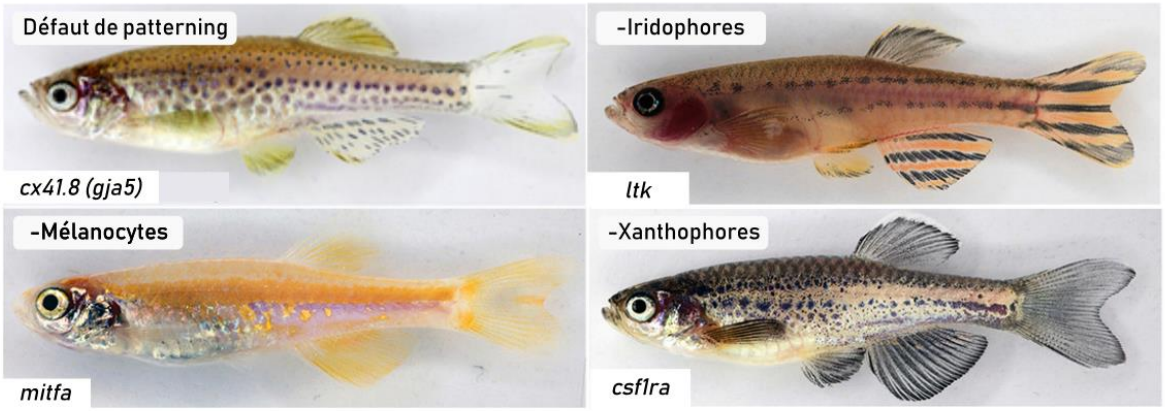


Figure 22 - "Un seul être vous manque et tout est décoloré."
Quelques gènes impliqués dans le développement des chromatophores de Poisson-zèbre : la déficience en l'un deux induit une perte ou une réduction du chromatophore correspondant qui se répercute sur le patron de pigmentation général de l'individu. *Adapté de http://www.eb.tuebingen.mpg.de/*

ii. Les gènes de la pigmentation ne sont pas uniquement associés à la pigmentation, ou pas seuls

1. Les gènes de la pigmentation ne sont pas seulement associés à une fonction pigmentaire

Il est à noter que certains gènes "de la pigmentation" sont impliqués dans d'autres processus physiologiques. Par exemple, la protéine ASIP, agoniste inverse de MC1R, est également un agoniste inverse de MC4R et, dans une moindre mesure, de MC3R et MC5R, tous trois impliqués dans une grande diversité de processus physiologiques (Ducrest *et al.*, 2008). Les gènes de la pigmentation peuvent donc être pléiotropes.

2. Parfois, un même trait pigmentaire est gouverné par plusieurs gènes

Chez les souris du genre *Peromyscus*, en Floride, plusieurs gènes déterminent le patron de coloration dorso-ventral (Steiner *et al.*, 2007), dont au moins *mc1r,* pour lequel une variation dans la séquence codante est à l'origine d'une variation du patron pigmentaire (Hoekstra, 2006), et *agouti*, dont le niveau d'expression et les isoformes varient (Steiner *et al.*, 2008; Manceau *et al.*, 2011; Mallarino *et al.*, 2017). Une base de donnée actualisée de la littérature corrélant les mutations aux variations phénotypiques, GePheBase, montre d'ailleurs que l'antagonisme *mc1r/agouti* est une "cible" régulière de l'évolution de la pigmentation chez les Tétrapodes (Martin and Orgogozo, 2013; Martin and Courtier-Orgogozo, 2017). *Mc1r* a d'ailleurs été qualifié de *"genetic hotspot of phenotypic variation"* (Martin and Orgogozo, 2013).

En conclusion, il n'y a pas de relation bijective entre "gène de la pigmentation" et "fonction pigmentaire" : un gène de la pigmentation peut avoir plusieurs fonctions, et un même trait peut être encodé par différents gènes (Manceau *et al.*, 2010).

c. De nouveaux gènes à découvrir…

En 2003, Bennett et Lamoreux estimaient qu'environ la moitié des gènes impliqués dans la couleur (une centaine) avaient été caractérisés (Bennett and Lamoreux, 2003). Aujourd'hui, on en connait presque deux fois plus, mais de nouveaux gènes sont sans cesse caractérisés. Par exemple, récemment, ont été découverts :

- Chez les Oiseaux, deux gènes impliqués dans l'oxydation endogène des caroténoïdes alimentaires jaunes en caroténoïdes rouges (Lopes *et al.*, 2016; Mundy *et al.*, 2016) et un gène impliqué dans le transport des caroténoïdes (Toomey *et al.*, 2017)
- Découvert chez l'Homme et validé chez le Poisson-zèbre, le gène *mfsd12*, au rôle encore peu clair mais probablement impliqué dans la différentiation des mélanosomes (Crawford *et al.*, 2017)
- Mis en évidence chez la Carpe, un gène du métabolisme de la tyrosine, un précurseur de la mélanine (Zhang, Liu, *et al.*, 2017)

De plus, certains gènes ayant des fonctions développementales mais non-pigmentaires sont parfois nouvellement co-optés dans des réseaux de régulation de gènes impliqués dans la pigmentation. C'est le cas de *wingless*, chez la Drosophile (Werner *et al.*, 2010). La liste des gènes de la pigmentation est donc variable évolutivement.

d. … même si les gènes ne sont pas les seuls acteurs de la pigmentation

On ne saurait toutefois réduire la pigmentation aux seuls gènes ou aux seules séquences codantes : la régulation de ces gènes est en effet primordiale. On peut penser en premier lieu à la mise en évidence historique des éléments transposables par B. McClintock : l'insertion aléatoire d'éléments au sein d'un gène impliqué dans la couleur des grains de Maïs conduit à un arrêt de l'expression du gène et à un patron de pigmentation mosaïque (McClintock, 1950).

De plus, le rôle des petits ARN non-codants, régulant l'expression d'autres gènes en *trans*, est également à prendre en compte, comme cela a pu récemment être montré chez le Muflier *Antirrhinum majus* (Bradley *et al.*, 2017).

Enfin, les changements de régulation en *cis* peuvent induire des variations de l'expression d'un gène de la pigmentation (Arnoult *et al.*, 2013). Cela peut être dû à l'insertion d'un élément transposable qui modifie la régulation du gène, comme cela a été montré pour expliquer l'apparition des egg-spots de la nageoire anale des poissons Cichlidés (Santos *et al.*, 2014). L'exemple emblématique de l'adaptation à l'obscurcissement des troncs par mélanisation de la Phalène du Bouleau pendant la Révolution Industrielle est due à l'insertion d'un élément transposable en amont du gène *cortex* (Hof *et al.*, 2016; Nadeau *et al.*, 2016). Il peut également y avoir des mutations induisant l'apparition ou la disparition de sites de fixation de facteurs de transcription, qui modifient ensuite les niveaux et domaines d'expression d'un gène (Gompel *et al.*, 2005). Chez les papillons du genre *Héliconius*, les morphes pigmentaires sont en partie dus à des variants alléliques de la séquence régulatrice du gène *optix* (Reed *et al.*, 2011).

En conclusion, nous avons précédemment vu comment des changements de séquence codante chez le Poisson-zèbre (dans le gène *cx41.8*) pouvaient induire des variations des patrons pigmentaires (Watanabe and Kondo, 2015). Les changements des réseaux de régulation sont également à l'origine de variations drastiques et rapides des patrons de pigmentation (Stern and Orgogozo, 2009).

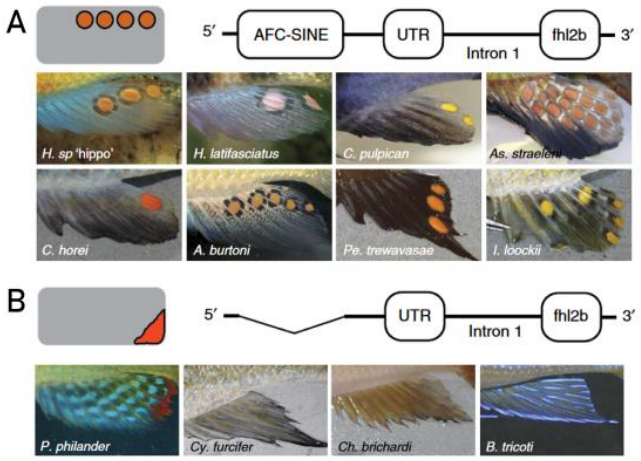


Figure 23 – La variation d'un trait pigmentaire n'est pas due qu'aux gènes.

Chez certains poissons cichlidés des grands lacs africains, l'insertion d'un transposon SINE a modifié la régulation du gène *fhl2b*, induisant l'apparition de « egg-spots » sur la nageoire anale des mâles (A). Sans cet élément, le patron de pigmentation de la nageoire anale est différent (B). Cette structure a un rôle sexuel. Après la ponte, la femelle commence une incubation buccale des œufs. Les taches mâles agissent comme des leurres d'œufs pour la femelle, qu'elle essaie de récupérer ; lorsqu'elle est proche, le mâle libère son sperme et la fécondation a alors lieu. *Adapté de Santos et al. (2014).*

Conclusion partielle et transition

Dans cette première partie, nous avons vu que la couleur est un trait biologique essentiel d'un individu. De nombreux forces évolutives sont à l'origine de la diversité des couleurs observées, et il est possible, dans une certaine mesure, de catégoriser les fonctions adaptatives. Les bases moléculaires, cellulaires et histologiques de la coloration diffèrent selon les taxons de Vertébrés. Les poissons Téléostéens représentent le groupe le plus diversifié, tant en termes de patrons de pigmentation que de types cellulaires. J'ai également évoqué les bases du déterminisme génétique de la pigmentation, et comment ce déterminisme génétique pouvait être à l'origine de l'évolution de la couleur, parfois rapide, au sein d'une lignée.

Le chapitre suivant est consacré à un aspect précis de l'évolution des gènes et des génomes : les duplications, et particulièrement les duplications de génomes. Celles-ci ont eu lieu à de multiples reprises chez les Eucaryotes, et notamment, au sein des Vertébrés, chez les poissons Téléostéens. Je présente les mécanismes à l'origine de ces duplications de génomes ainsi que leurs conséquences évolutives, qui diffèrent selon les lignées d'Eucaryotes.

INTRODUCTION

CHAPITRE 2

Les duplications de génomes et leurs conséquences évolutives

En 1998, quand les lignes suivantes ont été écrites, on pensait que le génome humain paraîtrait en 2005.

Take four, or maybe eight, decks of 52 playing cards. Shuffle them all together and then throw some cards away. Pick 20 cards at random and drop the rest on the floor. Give the 20 cards to some evolutionary biologists and ask them to figure out what you've done. For encouragement, tell them they can have the cards on the floor in 2005.
(Skrabanek and Wolfe, 1998)

I. Les duplications complètes de génomes sont des événements fréquents chez les Eucaryotes

1. Pré-requis à la détection d'une duplication de génome

L'étude des génomes s'est considérablement développée depuis l'avènement du séquençage et aujourd'hui des milliers d'espèces ont leur génome séquencé. L'essentiel des génomes d'êtres vivants disponibles aujourd'hui sont des génomes bactériens, plus petits et donc plus faciles à obtenir.

Ainsi, l'assemblage du premier génome de Vertébré, le génome humain en 2001 (Lander *et al.*, 2001; Venter *et al.*, 2001), a nécessité l'effort concerté de centaines de scientifiques dans le monde entier pendant plus de dix ans et a coûté plusieurs milliards de dollars. D'autres génomes ont peu à peu été disponibles, chez les Vertébrés en particulier mais également dans d'autres lignées évolutives (Végétaux, Bactéries, Eucaryotes unicellulaires…). Aujourd'hui, l'avènement des techniques de séquençage et d'assemblage permettent permet à des entreprises de d'obtenir votre génome pour moins de 1 000 dollars, en quelques semaines seulement.

Ce n'est qu'à partir de la fin des années 1990 et surtout du début des années 2000 qu'il a été (i) techniquement possible de séquencer des génomes avec l'amélioration des techniques de séquençage et (ii) computationnellement possibilité de traiter l'ensemble des données produites par séquençage en un ensemble cohérent de résultats exploitables scientifiquement. A quoi auraient servi des giga-octets de données de séquençage sans la possibilité matérielle de les assembler en génome dans un temps "relativement" court ? A l'inverse, des puissantes machines de calcul n'auraient pas été d'une grande utilité si les chimistes et biologistes moléculaires n'avaient pas, de leur côté, perfectionné et automatisé leurs techniques de séquençage. Cette double avancée, technique et informatique, a été un pré-requis nécessaire à l'étude systématique des génomes.

Les premiers génomes séquencés se sont très vite avérés être de formidables outils pour la recherche : chaque gène avait une séquence, et les séquences étaient elles-mêmes localisées précisément par rapport aux autres gènes du génome : gènes et séquences régulatrices alentour, position dans le chromosome… Les organismes modèles (Drosophile, *Caenorhabditis elegans*, Souris) furent parmi les premiers à être séquencés et leur génome devint un formidable outil.

Les biologistes moléculaires n'ont pas été les seuls à profiter de cette aubaine, et les biologistes computationnels se sont, eux aussi, amplement tournés vers ces nouvelles ressources. En particulier, pour les biologistes de l'évolution des Vertébrés, auparavant limités à l'étude de familles de gènes assez restreintes, telle que les gènes homéotiques (Meyer and Schartl, 1999), les études devenaient possibles à l'échelle du génome entier. Et de nombreux résultats, qui étaient déjà pressentis par des études au niveau "multi-génétique", se sont retrouvés confirmés par des études au niveau génomique. C'est le cas, en particulier, des duplications de génomes.

2. Comment détecter un événement de duplication complète de génome par l'étude des gènes ?

"Because of the paucity of available map and sequence data, it is premature to reach any conclusion about Ohno's original hypothesis for vertebrates" – (Skrabanek and Wolfe, 1998)

Comme son nom l'indique, une duplication complète de génome (que j'appellerai par la suite WGD, acronyme du terme anglais *Whole-Genome Duplication*), est un événement qui consiste à dupliquer totalement le contenu du génome d'un individu. Nous envisagerons les modalités permettant l'apparition d'une WGD par la suite : elles représentent des mécanismes qui n'ont été proposés ou démontrés que sur le tard, permettant d'expliquer des observations "de terrain" faites par les premiers chercheurs travaillant sur ces thématiques (Skrabanek and Wolfe, 1998). Quels sont donc les témoins, ou les indices, permettant de se rendre compte qu'un génome d'intérêt a connu un événement de WGD pendant l'évolution ? Comment les "pionniers" du domaine en sont-ils venus à imaginer que les WGD étaient un phénomène possible au cours de l'évolution d'une lignée ? Comment, de doutes initiaux sur la présence d'une WGD, en vint-on à confirmer son occurrence ?

a. Les témoins cytologiques

Historiquement, les premiers indices qui ont permis de se rendre compte qu'un génome avait été dupliqué sont cytologiques : un organisme au génome récemment dupliqué a un génome qui comporte deux fois plus de chromosomes, une quantité d'ADN deux fois plus importante et une taille des spermatozoïdes (qui est corrélée à la taille du génome) plus importante (Ohno, 1970; Gallardo *et al.*, 1999). Je reviendrai plus longuement sur ces aspects plus loin et me concentre ici sur les indices *génétiques*.

b. Les témoins synténiques

Lors d'une duplication d'un gène ancestral, le gène peut être soit dupliqué juste à côté de la copie initiale (on parle de duplication en "tandem"), soit ailleurs dans le génome, par exemple par rétrotransposition (Vinckenbosch *et al.*, 2006). Dans le second cas, on aura donc la copie initiale du gène X (appelons-la Xa), qui sera à un endroit dans le génome, et sa copie Xb ailleurs, peut-être même sur un autre chromosome. On conclura, à juste titre, que le gène X ancestral s'est dupliqué en Xa et Xb. Ces deux gènes seront qualifiés de gènes paralogues.

Imaginons maintenant que le gène voisin de Xa (appelons le Ya), ait lui aussi un paralogue Yb à côté de Xb, à des "encablures génomiques" de là. Cela signifierait donc que, par hasard, la duplication du gène ancestral Y (indépendante de celle de X) ait donné une copie juste à côté de celle de X. Il est très peu probable, étant donné la taille d'un génome, que cela ait lieu, mais pourquoi pas.

Ce mécanisme devient toutefois hautement improbable si on se rend compte qu'autour du tandem Xb-Yb, figure également le gène Zb, produit de la duplication d'un gène ancestral Z et dont la copie Za se situe juste à côté de Xa et Ya. Il faut alors imaginer un mécanisme alternatif, impliquant la duplication de l'ensemble du segment X-Y-Z en Xa-Ya-Za et Xb-Yb-Zb, plutôt que trois duplications indépendantes.

Si, de plus, ces duplications "segmentaires" sont observées à une échelle encore plus grande, et qu'on constate que des chromosomes entiers partagent des régions dupliquées avec d'autres chromosomes (régions appelées "paralogons"), cela donne un indice sur un événement passé de duplication de génome. Quelques études pionnières ont mis en évidence de tels paralogons, notamment chez la Levure en 1997 (Wolfe and Shields, 1997). Les auteurs avaient ainsi montré que les chromosomes 10 et 11 partageaient des segments avec de nombreux paralogues. Les auteurs observaient qui plus est, une conservation de l'ordre des gènes, c'est-à-dire de la "synténie" (du grec *syn-*, "avec", et *tén-*, "filament, chaîne"), indiquant là encore un probable événement de WGD.

c. Les témoins phylogénétiques

Considérons de nouveau les trois gènes X, Y et Z. Cette fois, nous ne connaissons pas leur position dans le génome. On trouve chez deux espèces A et B le gène X en deux copies : Xa et Xb. Il se trouve également qu'on connait une troisième espèce C, plus éloignée de A et B (groupe externe), chez laquelle le gène X n'est présent qu'en une seule copie. Si on aligne ces séquences et qu'on les compare, on peut obtenir une phylogénie simplifiée du gène X, qui topologiquement ressemble à celle indiquée **Fig. 25**. Une telle topologie indique que (1) le gène X ancestral a été dupliqué chez l'ancêtre de A et B après la divergence de la lignée de C (ii) ; il a ensuite été conservé en deux copies dans chaque lignée après la divergence entre A et B.

La conservation de cette topologie pour trois gènes peut s'expliquer par trois événements de duplication indépendants survenus entre la divergence ((A,B),C) et la divergence (A,B). C'est, de même, possible, mais d'autant plus rare que le nombre de gènes avec une topologie similaire est élevé. Une hypothèse alternative, sans connaissance préalable de la position des gènes, est alors d'imaginer un événement de duplication unique ayant dupliqué l'ensemble des gènes du génome.

Cette hypothèse parcimonieuse, appliquée à l'étude des gènes d'histones chez la Levure *Saccharomyces cerevisiae*, est suggérée par Smith dès 1987: "*the attractive simplicity of the duplication of both the H3-H4 and the H2A-H2B gene pairs having taken place in a single event, leads to the hypothesis that modern haploid* S.cerevisiae *may have evolved by*

diploidization or fusion of two ancient fungi." (Smith, 1987). Dix ans plus tard, en 1997, K. Wolfe et D. Shields identifient de nombreux gènes ayant une topologie de ce type.

Enfin, outre la topologie, les arbres phylogénétiques donnent une information sur les taux de divergence entre copies. Si on observe que les taux d'évolution sont comparables entre les paires de paralogues issues de X, Y et Z, cela peut également indiquer un événement passé de duplication du segment (Berthelot *et al.*, 2014).

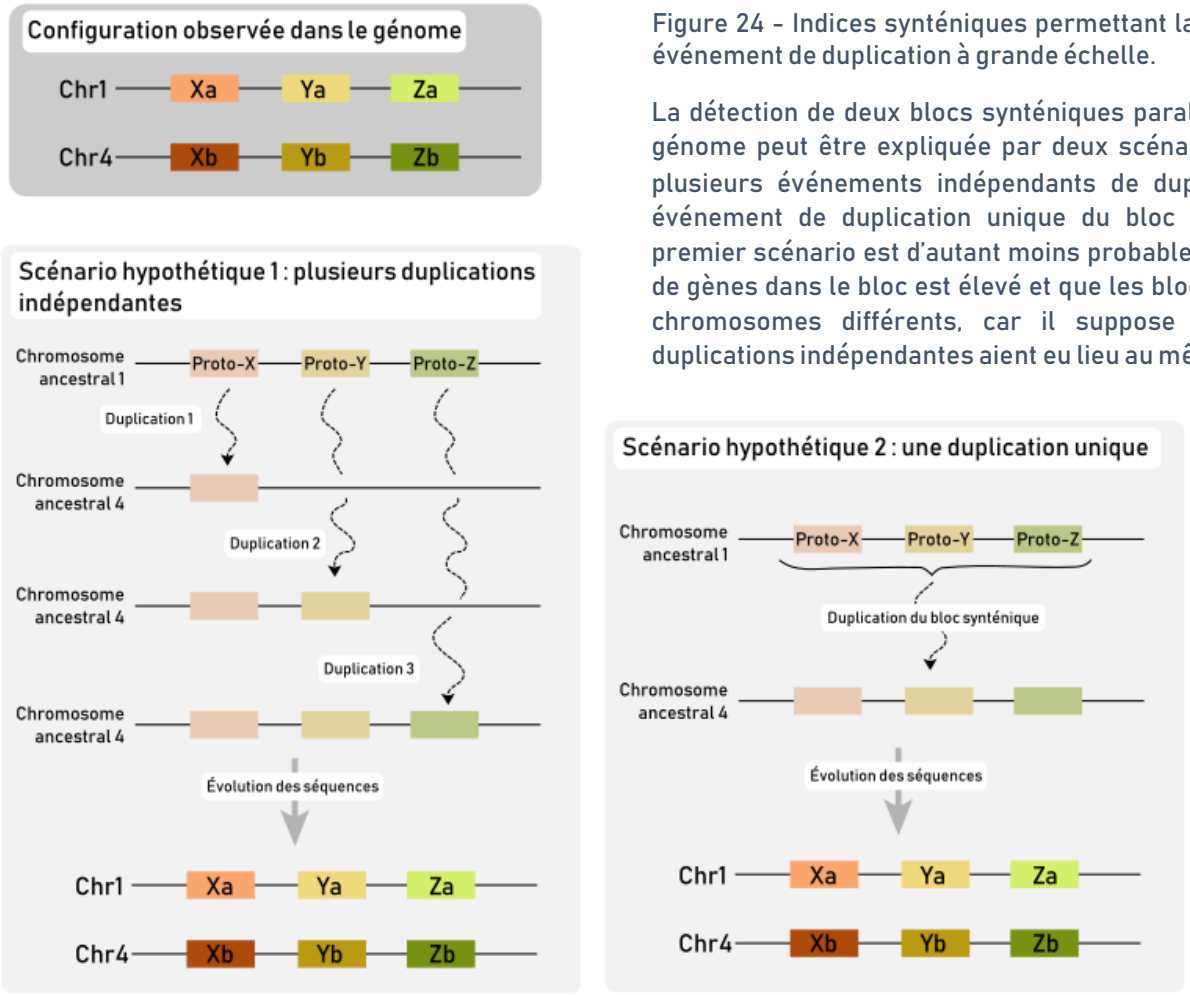


Figure 24 - Indices synténiques permettant la détection d'un événement de duplication à grande échelle.

La détection de deux blocs synténiques paralogues dans un génome peut être expliquée par deux scénarios extrêmes : plusieurs événements indépendants de duplication ou un événement de duplication unique du bloc synténique. Le premier scénario est d'autant moins probable que le nombre de gènes dans le bloc est élevé et que les blocs sont dur des chromosomes différents, car il suppose que plusieurs duplications indépendantes aient eu lieu au même locus.

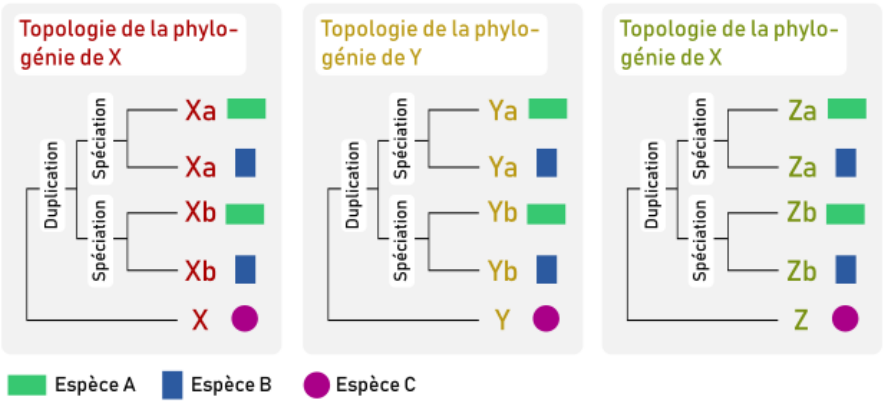


Figure 25 - Indices phylogénétiques d'une duplication de segment.

La reconstruction de la phylogénie des gènes X, Y et Z chez trois espèces produit des topologies similaires. Plus le nombre de gènes ayant cette topologie est élevé, plus la probabilité est forte que l'événement de duplication soit unique et qu'il concerne tous les gènes du génome.

d. Conclusion : le génome, un palimpseste

Nous avons jusqu'ici principalement raisonné sur trois gènes ancestraux théoriques. En croisant l'ensemble des indices (structuraux comme phylogénétiques) à l'échelle du génome entier, il a peu à peu été montré que des duplications de génomes avaient eu lieu. Au début, on ne parlait alors que de "large-scale duplications" ("duplications segmentaires à grande échelle"), car tous les gènes du génome n'avaient pas encore été analysés (Pebusque *et al.*, 1998).

Peu à peu, la communauté s'est rendue compte que toutes les régions du génome étaient affectées et, de "large-scale segmental duplication", on a commencé à parler de "Whole-Genome Duplication". Les paralogons ont même pu être observés pour certaines espèces, par exemple chez la Paramécie ou chez le Tétraodon (Jaillon *et al.*, 2004; Aury *et al.*, 2006). La correspondance théoriquement attendue n'était toutefois pas exacte. Ainsi, non seulement une duplication de génome avait pu avoir lieu, mais des réarrangements génomiques consécutifs avaient pu permettre de modifier la position relative des blocs synténiques, brouillant ainsi le palimpseste génomique.

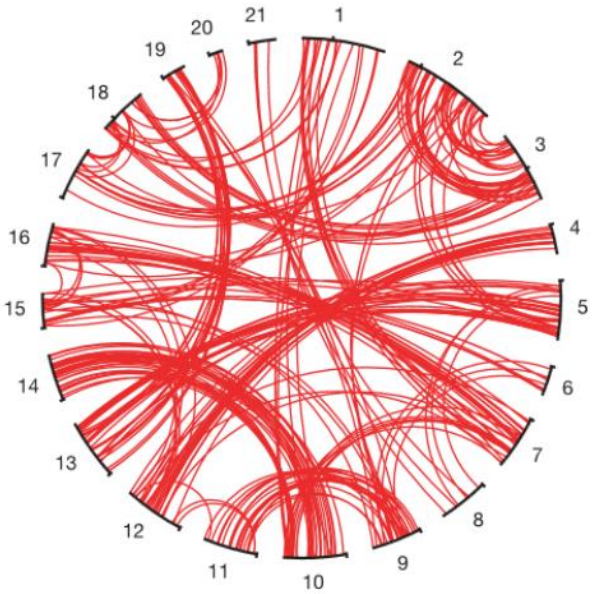


Figure 26 - Conservation de la synténie après une WGD.

Les 21 chromosomes du Tetraodon sont représentés en cercle et numérotés dans l'ordre. Chaque ligne rejoint deux gènes paralogues issus d'un événement de WGD, c'est-à-dire des gènes dont la topologie de la phylogénie et la synténie sont compatibles avec ce scénario. *D'après Jaillon et al. (2004).*

3. Comment les WGD apparaissent-elles ?

Une fois les premières WGD identifiées, de multiples mécanismes théoriques expliquant l'origine de ces observations génomiques ont été avancés. Les mécanismes permettant la multiplication du nombre de chromosomes (polyploïdisation, la ploïdie étant le nombre de chromosomes) étaient déjà connus depuis quelques années, notamment par des travaux chez les plantes. On peut en distinguer deux grands types : l'autopolyploïdisation et l'allopolyploïdisation, selon que le nouvel organisme possède deux génomes provenant d'individus de la même espèce (autopolyploïde) ou d'espèces différentes (allopolyploïde). Les chromosomes apparentés, au sein d'un organisme polyploïde, sont appelés des chromosomes homéologues. La polyploïdisation est un phénomène très fréquent chez les plantes, qui ont des barrières reproductives plus perméables en comparaison des animaux (Siroki *et al.*, 1999). Ce pan de recherche a donc grandement bénéficié des travaux en biologie végétale.

a. Autopolyploïdie

Les mécanismes principaux pouvant provoquer une autopolyploïdisation sont (i) une non-réduction de deux gamètes qui, lorsqu'ils se rencontrent, produisent un organisme autopolyploïde (ici, autotétraploïde) ou (ii) la formation intermédiaire d'un autotriploïde (Hegarty and Hiscock, 2008).

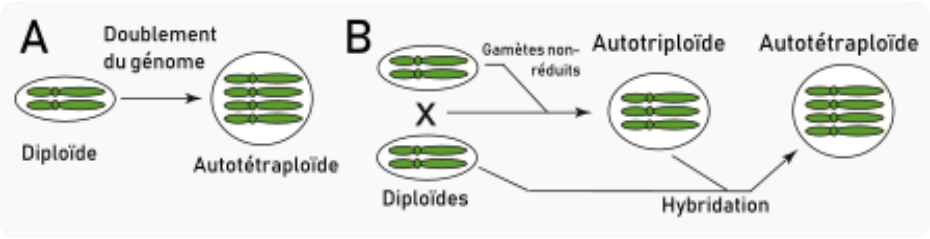


Figure 27 - Mécanismes à l'origine d'une autopolyploïdisation.

D'après Hegarty and Hiscock (2008).

b. Allopolyploïdie

L'allopolyploïdie implique deux espèces parentales différentes, mais suffisamment proches génétiquement pour que l'hybride qu'elles forment soit fertile. Les mécanismes permettant *in fine* un doublement (allotétraploïdie), ou plus rarement un triplement (allohexaploïdie) du génome, sont plus divers que pour l'autopolyploïdisation. Toutefois, chez les Végétaux, des cas d'allopolyploïdisation entre espèces de genres ou même de tribus différentes ont déjà été documentés, indiquant qu'une distance évolutive "moyenne" peut également permettre l'allopolyploïdisation (Joly *et al.*, 2009). On parle alors de "transfert horizontal de *génome*".

On note que les frontières entre auto- et allopolyploïdie sont labiles et les classifications peuvent parfois être problématiques : l'autopolyploïdie peut en effet parfois impliquer des croisements entre deux populations génétiquement divergentes de la même espèce (polyploïdie "interraciale") mais dont la structuration génomique est similaire. Elle peut également concerner des allopolyploïdes récents qui "s'autotétraploïdisent" ensuite (cas A et B de la **Fig. 28**). Au-delà de la nomenclature, l'important est que de multiples mécanismes peuvent induire l'apparition d'un organisme polyploïde avec un génome doublé, voire triplé.

Enfin, des travaux récents chez les plantes suggèrent que l'allopolyploïdisation peut également avoir lieu de manière asexuée (à la différence des processus ci-dessus). L'hybride peut ainsi être formé par fusion directe de noyaux après mise en contact de tissus au cours d'une greffe (Fuentes *et al.*, 2014).

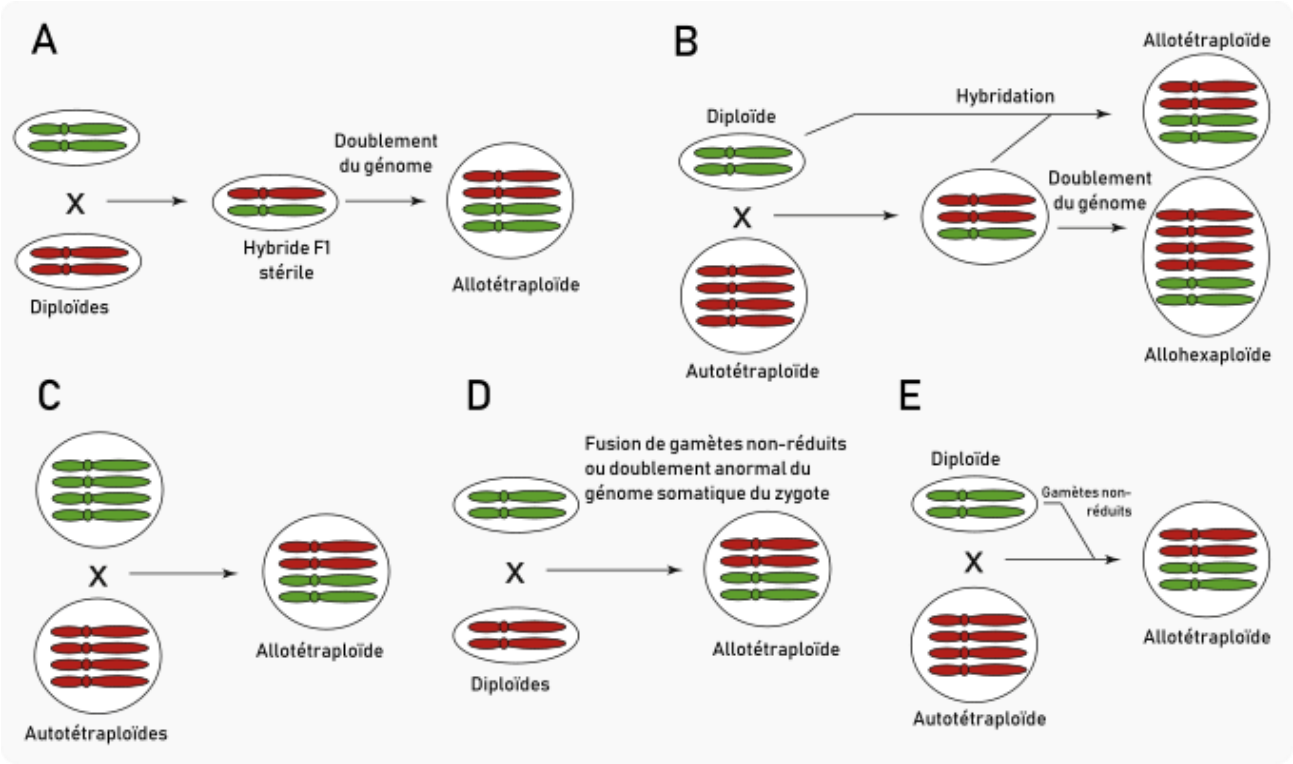


Figure 28 - Mécanismes à l'origine d'une allopolyploïdisation.

Les mécanismes à l'origine d'une allopolyploïdisation sont plus variés que pour l'autopolyploïdisation. Les couleurs représentent les deux génomes parentaux. *D'après Hegarty and Hiscock (2008).*

4. Quelles sont les lignées concernées par les WGD ?

a. Avant-propos

Jusqu'ici, j'ai essentiellement évoqué les duplications de génomes ayant eu lieu chez les plantes, la Levure, la Paramécie et un poisson Téléostéen. Ces taxons ne sont pas les seuls concernés par des WGD parfois multiples. Je laisse pour l'instant de côté les poissons Téléostéens et les autres Vertébrés, qui ont fait l'objet de l'essentiel de mes travaux de thèse et auxquels je consacrerai une partie indépendante de ce deuxième chapitre introductif **(cf. II. 2)**. Je présente auparavant les WGD dans les autres taxons et prends un recul par rapport à mon objet d'étude qui me semble nécessaire. L'étude des WGD chez les Vertébrés s'inscrit en effet dans un cadre plus large, et les apports des études des autres WGD Eucaryotes a été majeur à la compréhension des WGD et de leurs conséquences évolutives, phénotypiques et génomiques chez les Vertébrés (Otto and Whitton, 2000).

Les duplications de génomes n'ont pas été documentées, à ma connaissance, chez les Bactéries. Ceci est probablement dû à leur évolution trop rapide qui ne permet pas de conserver des traces de paléoploïdisations ; de plus, la forte fréquence de transferts horizontaux entre espèces proches est probablement un facteur qui réhomogénéise rapidement les génomes après duplication. Nous n'envisagerons ici que des organismes eucaryotes : paramécies, spermatophytes, levures et autres animaux. L'idée n'est pas ici d'être exhaustif sur chacune des lignées mais de présenter quelques-unes de leurs spécificités. Nous verrons au cours de la partie III en quoi ces études ont permis de mieux éclairer les processus en jeu au cours de l'évolution des lignées suite à une (ou plusieurs) WGD.

b. Paramécie

Les Paramécies sont des Eucaryotes unicellulaires du taxon des Ciliés, à l'évolution du génome très particulière : en 2006 un consortium dirigé par une équipe française a montré que le génome de *Paramecium tetraurelia* (une des espèces du complexe d'espèces *P. aurelia*) était le résultat de trois événements de duplication de génome ayant eu lieu au cours de l'évolution de cette lignée (Aury *et al.*, 2006) : une récente, une intermédiaire et une ancienne. Les datations de ces duplications ne sont pas encore certaines, mais la plus récente pourrait dater de 320 millions d'années (McGrath, Gout, Johri, *et al.*, 2014). La duplication intermédiaire pourrait, elle, dater de 1.5 *milliards* d'années, mais les résultats sont encore discutés (McGrath, Gout, Doak, *et al.*, 2014). Dans tous les cas, ces deux événements sont temporellement très éloignés. La duplication la plus ancienne n'a, à ma connaissance, pas été datée précisément. L'événement de duplication originel serait une autopolyploïdisation (Gout and Lynch, 2015).

c. Angiospermes et plantes à graines (Spermatophytes)

Les génomes des plantes à graines (Spermatophytes) varient beaucoup en taille, essentiellement par la grande variabilité de leurs contenus en séquences répétées (Leitch and Leitch, 2013). Toutefois, ce n'est pas le seul facteur expliquant la grande variabilité de la taille des génomes de plantes : les WGD y sont également pour beaucoup, et sont également à l'origine de variations de la structure et du contenu génique des génomes (Soltis *et al.*, 2015). L'étude des génomes des plantes est donc intimement liée à celle des WGD.

Les Spermatophytes sont constitués de deux subdivisions : les Angiospermes (plantes à fleurs) et les Gymnospermes (plantes à ovule nu). Les Angiospermes en particulier constituent un groupe d'espèces très diversifié (presque 400 000 espèces) et représentent l'essentiel de la diversité des végétaux sur les continents. Indispensables à l'Homme pour de multiples raisons, notamment alimentaires, et étudié depuis des millénaires, l'avènement de la génomique a très vite concerné ce groupe. Ce n'est d'ailleurs pas surprenant si le Riz et le Maïs (des Poacées) ont été parmi les premières plantes à être séquencées dans les années 1990 (Ahn and Tanksley, 1993; Gaut and Doebley, 1997). Les résultats de cartographie génétique chez le Maïs montrèrent déjà que de nombreuses régions semblaient être dupliquées

ailleurs, sur d'autres chromosomes, avec une conservation de la synténie – le terme de "bloc synténique" n'était pas encore utilisé : on parlait de "*non-overlapping duplicated regions*" (Skrabanek and Wolfe, 1998).

Quelques années plus tard, en 2000, le génome d'*Arabidopsis thaliana*, qui appartient à une autre lignée d'Angiospermes (les Brassicacées), est publié (Arabidopsis Genome Initiative, 2000). Deux ans avant sa publication, en 1998, le récent auteur du papier sur les WGD du génome de la Levure, K. Wolfe, s'étonnait : *"Recent results from the Arabidopsis genome sequencing project indicate unexpectedly that it, too, has substantial regional duplications."* (Skrabanek and Wolfe, 1998). Ainsi, une autre lignée d'Angiospermes semblait concernée par les WGD. Cette remarque est intéressante à double titre puisqu'elle indique également que, avant même la parution de la séquence complète d'un génome, une des premières questions auxquelles on cherche à répondre est : y détecte-t-on des signes de grands événements de duplication ? C'est la fin des années 1990 et la duplication de génome est entrée dans les mœurs.

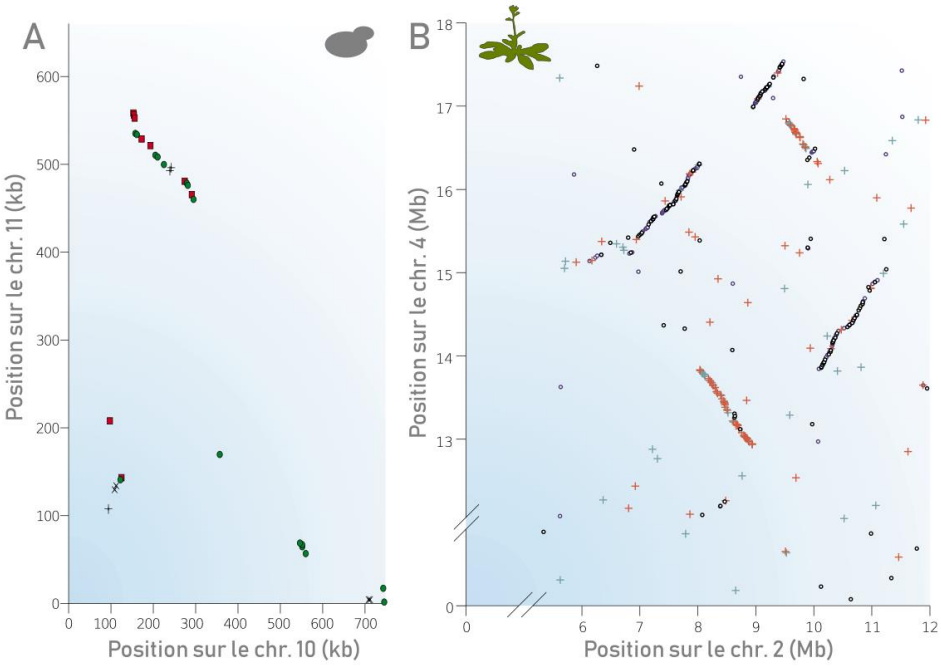


Figure 29 - Conservation de la synténie entre blocs dupliqués chez la Levure et l'Arabette.

Comparaison de la conservation des séquences entre paires de chromosomes chez la Levure (A) et l'Arabette (B). Chaque point représente les paires de gènes identifiées par meilleur hit blast réciproque. Les génomes de l'Arabette et de la Levure portent donc la trace de la conservation de blocs synténiques d'échelle sub-chromosomique (paralogons). *D'après Wolfe (2002).*

Le reste des études génomiques menées depuis une vingtaine d'années chez les Angiospermes ont mis en évidence des événements de polyploïdisations dans de très nombreuses lignées, ce qui a même fait dire à Soltis et al. : "la question n'est pas tant quelle est la proportion d'Angiospermes qui sont polyploïdes, mais plutôt combien d'événements de polyploïdisation ont eu lieu dans chaque lignée d'Angiospermes ?" (Soltis *et al.*, 2009). De plus, des taxons qu'on croyait exempts de WGD sont à leur tour "démasqués". Ainsi, en 2009, dans ce même article, les auteurs affirmaient : "*However, there is currently no evidence for WGD in Amborella*" – une lignée ayant divergé très tôt du reste des Angiospermes (Soltis *et al.*, 2009). Deux ans plus tard, en 2011, la parution du génome d'*Amborella* montrait qu'en réalité ce taxon a lui aussi été affecté par la WGD ancestrale des Angiospermes (Adams, 2013; Chamala *et al.*, 2013).

La propension aux WGD chez les Angiospermes était déjà présente chez leur ancêtre puisqu'un événement de WGD a eu lieu chez l'ancêtre des Spermatophytes (Jiao *et al.*, 2011). De même, au cours de l'évolution des Gymnospermes – le groupe frère des Angiospermes – d'autres WGD ont eu lieu (Li, Baniaga, *et al.*, 2015).

Toutefois, même si les Spermaphytes dans leur ensemble sont une lignée caractérisée par de nombreuses WGD, c'est au sein des Angiospermes que les événements de polyploïdisation ont été les plus nombreux (Badouin *et al.*, 2017). Est-ce que cela témoigne de propriétés génomiques ayant facilité la conservation de ces événements, ou est-ce qu'on en a plus trouvé parce qu'on en a plus cherché, en raison de l'intérêt économique de ces espèces ? Il est difficile de répondre à cette question.

Les Angiospermes sont également les seules espèces où, à ma connaissance, ont été documentées des Whole-Genome *Triplications*. Ainsi, à la base des Eudicotylédones, une hexapolyploïdisation a eu lieu (voir (Salse, 2016) pour une revue), et d'autres WGT ont également été plus récemment mises en évidence, notamment chez les Brassicacées (Wang *et al.*, 2011). Le génome du Coton a quant à lui même peut-être été affecté par une dodécapolyploïdisation récente (Salse, 2016), soit une multiplication du génome par 6, montrant là encore que les génomes sont hautement plastiques chez les Angiospermes.

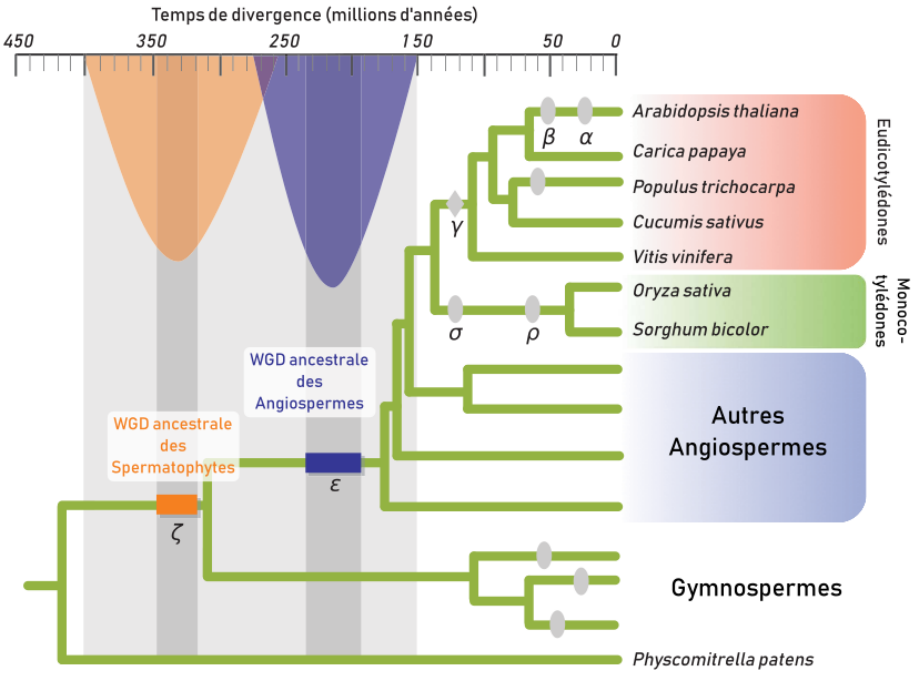


Figure 30 - Duplications de génomes chez les plantes à graines.

Au sein des plantes à graines (Spermatophytes), les WGD sont très fréquentes, en particulier chez les Angiospermes. Les lettres grecques symbolisent chacune des WGD. *Adapté de Jiao et al. (2011) et de Li et al. (2015).*

d. Levure

Quittons le monde des végétaux et faisons un bond évolutif dans la direction des Métazoaires (ou animaux). Au sein des Opisthokontes, les organismes dont les cellules mobiles sont propulsées par un flagelle postérieur (*opisthō* = derrière), le groupe frère des Métazoaires est le groupe des Mycètes, ou champignons. À l'instar des plantes, certains champignons ont un intérêt économique majeur et ont très tôt suscité l'intérêt, des chercheurs certes, mais également des industriels. C'est le cas en particulier de la Levure *Saccharomyces cerevisiae,* dont la respiration ou la fermentation sont exploitées en cuisine (fabrication du pain) ou dans la production d'alcool (vin, bière – d'où elle tient d'ailleurs son nom).

Qui plus est, le génome de *S. cerevisiae*, d'une longueur estimée à 12 Mpb, est relativement court par rapport aux génomes d'autres Eucaryotes (le génome humain est estimé à plus de 3000 Mpb). Ainsi, il n'est pas surprenant que le génome de la Levure soit le premier génome Eucaryote à avoir été séquencé, après plusieurs années de travaux impliquant plus de 600 chercheurs dans quelque 100 laboratoires du monde entier (Dujon, 1996; Wolfe, 2015). Ce projet fut également un des premiers à utiliser internet dans un cadre collaboratif, et il fut même indiqué dans un article paru dans *Science*, de manière qui paraîtrait aujourd'hui désuette, que *"[The daily notebook of highly specialized scientists and technicians] was put on the World Wide Web, was fully accessible to the scientific community and was (...) completed when new information became available.*" (Goffeau *et al.*, 1996).

Ces travaux s'inscrivaient dans la continuité des efforts précédents : après quelques décennies d'analyses génétiques sur cet organisme et la publication en 1992 du premier chromosome Eucaryote (le chromosome III de – déjà – la Levure) (Oliver *et al.*, 1992), en 1996 le génome de la Levure est enfin publié dans un article au nom évocateur de *Life with 6 000 genes* (Goffeau *et al.*, 1996). Dans l'article, les auteurs soulignent qu'alors qu'on ne connaissait que 1 000 gènes environ jusqu'à présent, le nombre total a été étendu à 5 885. Parmi tous ces nouveaux gènes jusque-là inconnus, nombreux sont ceux qui sont présents en deux copies "identiques ou très similaires, ce qui fournit le matériel pour l'évolution de nouvelles fonctions" (sic) (Goffeau *et al.*, 1996). On sait aujourd'hui que seulement 8-14% des gènes

dupliqués après WGD sont aujourd'hui toujours présents dans le génome, et que le reste a été perdu – je reviendrai plus loin sur les taux de rétention (Wolfe and Shields, 1997; Scannell *et al.*, 2007).

Dans l'article "génome", la possibilité de grandes duplications segmentaires était déjà évoquée, mais il a fallu attendre l'année suivante (1997) pour qu'une duplication totale de génome soit réellement envisagée chez la Levure *S. cerevisiae* (Wolfe and Shields, 1997), et 2004 pour que l'hypothèse soit réellement validée par la publication des génomes d'espèces sœurs pré-WGD (Dujon *et al.*, 2004). Aujourd'hui, on pense que c'est un événement d'hybridation interspécifique (allopolyploïdisation) ayant eu lieu il y a environ 100 millions d'années qui a donné naissance à la lignée de *S. cerevisae* (Marcet-Houben and Gabaldón, 2015).

Enfin, on peut noter que les champignons dans leur ensemble sont caractérisés par de multiples événements de polyploïdisation et la Levure *S. cerevisiae* n'est pas un cas unique dans cette lignée (Albertin and Marullo, 2012). Le Rosé des prés *Agaricus campestris* en est un autre exemple d'intérêt culinaire.

e. Les WGD chez les Animaux non-Vertébrés

Chez l'ensemble des Animaux étudiés au niveau génomique jusqu'à présent, et hormis le cas des Vertébrés, sur lesquels nous reviendrons plus loin, quelques rares événements de WGD (par rapport au nombre d'espèces) ont été documentés : un chez les Rotifères bdelloides (Flot *et al.*, 2013), un chez les Xiphosures (taxon peu diversifié incluant la Limule) il y a 135-300 Ma (Nossa *et al.*, 2014; Kenny *et al.*, 2016), et un chez les Arachnopulmonés (Araignées + Scorpions) (Schwager *et al.*, 2017). On note que des cas de polyploïdie, certes rares, ont déjà été signalés chez les insectes (Lokki and Saura, 1979) et qu'une étude récente vient de montrer que des multiples WGD auraient eu lieu au cours de l'évolution des hexapodes (Li *et al.*, 2018).

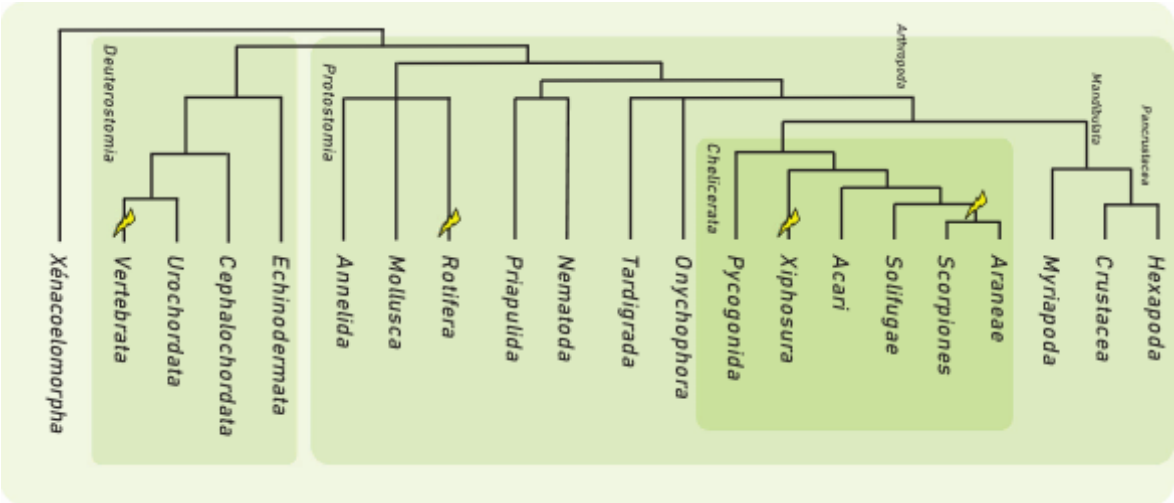


Figure 31 – Evénements de WGD chez les Métazoaires.

Les événements de WGD sont relativement rares au sein des Métazoaires autres que les Vertébrés. *Adapté de Kenny et al. (2015).*

II. Les duplications complètes de génomes chez les Vertébrés

Il a fallu attendre 2005 pour que soit publié un article assumant pleinement l'occurrence de deux événements de WGD à la base des Vertébrés (1R et 2R). L'article était intitulé : *Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate* (Dehal and Boore, 2005). Ces travaux, toutefois, sont le résultat de plusieurs années de discussions sur le sujet – qui ne sont pas d'ailleurs pas encore totalement éteintes à ce jour.

1. Bref historique de la découverte des WGD chez les Vertébrés

Dès 1970, dans son livre *"Evolution by Gene Duplication*", "souvent cité mais rarement lu" (sic) (Hahn, 2009), et plus particulièrement dans le chapitre 19, Susumu Ohno suggère qu'une augmentation de la taille du génome a accompagné l'émergence des Vertébrés : "*I venture a guess that paedomorphosis from tunicate-like creatures to amphioxus-like creatures which took place either during the Ordovician period or the Silurian period was accompanied by a two to threefold increase in the genome size"* (Ohno, 1970). Des mesures de la taille du génome, du nombre d'isozymes et de la morphologie des chromosomes lui permettent de dire que les Vertébrés ont émergé à partir du groupe qui, morphologiquement, était paradoxalement considéré par Ohno comme le plus simple de l'époque : les Urochordés. Il dit ainsi : "*Of the various invertebrates which existed nearly 500 million years ago in Cambrian times, it was the simplest sessile form which gave rise to the first vertebrate through tunicate-like creatures*". Par des études cytogénétiques, Ohno avance déjà, en 1970, que c'est peut-être un événement de duplication de génome qui est l'origine de cette augmentation de taille : "*it is our contention that the ancestors of reptiles, birds, and mammals have experienced at least one tetraploid evolution either at the stage of fish or at the stage of amphibians*."

Jusqu'à la publication du génome humain, on pensait également que l'Homme avait un génome avec bien plus de gènes que ceux des "inVertébrés" les plus proches (Lander *et al.*, 2001; Venter *et al.*, 2001). Le chiffre de 100 000 gènes fut parfois évoqué (Chaudhari and Hahn, 1983), et cela est même monté jusqu'à 142 634, pas un de plus, pas un de moins (Dickson, 1999) !

Pendant une quinzaine d'années, du début des années 1990 au milieu des années 2000, de nombreuses études ont été publiées sur l'origine du génome des Vertébrés, indiquant (i) que l'hypothèse de la 1R/2R était vraie ("one-to-four rule") (Meyer and Schartl, 1999) (ii), que seule la 1R avait eu lieu (McLysaght *et al.*, 2002) ou que (iii) aucune des deux n'avait eu lieu et que les duplications étaient le résultat de grandes duplications indépendantes ayant eu lieu avant ou après la séparation des Vertébrés (Friedman and Hughes, 2003). McLysaght et al., au milieu de cette période, résument le sentiment de l'époque : "*Opinions on the hypothesis that ancient genome duplications contributed to the vertebrate genome range from strong skepticism to strong credence*" (McLysaght *et al.*, 2002).

Les principaux points de divergence concernaient les deux types d'indices qui permettent de retracer une duplication de génome, que j'ai présentés plus haut : des informations structurales (conservations des paralogons) et phylogénétiques. Dans le cas d'une double duplication de génome, en supposant les 4 paralogues X_a, X_b, X_c et X_d d'un gène X ancestral, on s'attendrait à avoir des topologies du type $(X_a,X_b)(X_c,X_d)$ plus fréquentes que $(X_a)(X_b,X_c,X_d)$ ou que $(X_a,X_b,X_c)(X_d)$. Or, cela n'est pas toujours le cas. De plus, seulement 5% des gènes homologues sont dans des blocs synténiques de type 1:4, c'est-à-dire avec un paralogon chez une espèce anté-1R/2R qui correspond à quatre paralogons chez un Vertébré.

La double WGD a été confirmée en regardant non pas chacun de ces facteurs indépendamment, mais en combinant les deux types d'informations (Dehal and Boore, 2005). En effet, pendant les 450-500 millions d'années qui ont suivi ces deux événements, de nombreux événements de réarrangement de génome (duplications, inversions, translocations) ont pu avoir lieu, masquant les liens entre les paires de paralogues (Dehal and Boore, 2005; Putnam *et al.*, 2008; Smith and Keinath, 2015). La présence de quatre complexes Hox chez les Mammifères par rapport à un unique chez la Drosophile a également fait partie des argument forts en faveur de la 1R/2R (Wolfe, 2001).

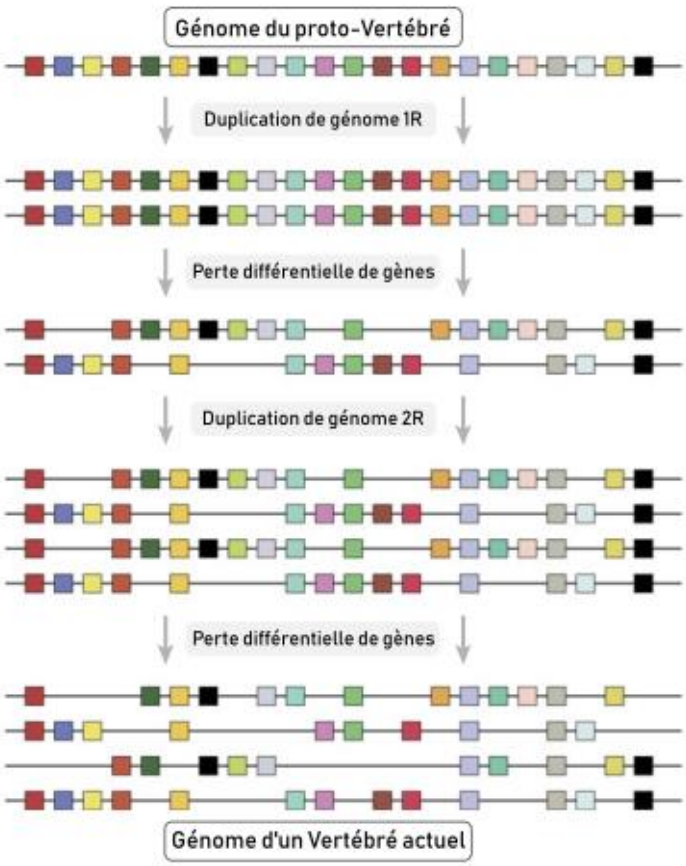


Figure 32 - Evolution schématique d'un chromosome de Vertébré.

Au cours de l'évolution des Vertébrés, deux duplications de génomes suivies d'événements de pertes de gènes ont conduit à une conservation partielle des paralogons qui rend difficile de retracer le signal des WGD par la seule étude de la synténie. D'autres événements de remaniements chromosomiques peuvent de plus intervenir (non illustrés). *Adapté de Dehal and Boore (2005).*

Si la double WGD ayant eu lieu à la base des Vertébrés (appelée 1R/2R) est aujourd'hui communément admise, quelques points sont encore discutés suite, notamment, à l'analyse du génome de lamproies, des Cyclostomes. Même si leur position systématique a pendant longtemps été discutée, les Cyclostomes ("Vertébrés sans mâchoires") sont aujourd'hui considérés comme des Vertébrés (Janvier, 2011), groupe frère des Gnathostomes ("Vertébrés avec mâchoire") (**cf. Annexe 1**).

L'étude du génome de lamproies permet d'éclairer l'évolution du génome des premiers Vertébrés. Les premières études sur les génomes de lamproies sont parues en 2013 pour *Petromyzon marinus* (Smith *et al.*, 2013) et pour *Lethenteron japonicum* (Mehta *et al.*, 2013). Elles ont mis en évidence l'existence de six complexes Hox (au lieu de quatre chez l'Homme) et ont été jusqu'à suggérer qu'un troisième événement de WGD avait eu lieu chez les Cyclostomes, tout en constatant que de nombreux gènes avaient été perdus par rapport au reste des Vertébrés. Cependant, ces deux publications ont émergé dans un contexte où les deux WGD 1R/2R étaient communément admises : les lamproies, en tant que Vertébrés, devaient donc en porter les stigmates. En 2015, une étude menée par un des auteurs de la publication du génome de *P. marinus* a montré qu'une hypothèse alternative expliquant la structure du génome des lamproies était d'envisager que des duplications de chromosomes entiers, à grande échelle, aient eu lieu avant un événement unique de WGD à la base des Vertébrés (Smith and Keinath, 2015). Ensuite, une autre WGD aurait eu lieu dans la lignée des Lamproies. Cela remettrait donc en cause l'hypothèse d'une 1R/2R ancestrale pour les Vertébrés.

L'histoire ne s'arrête pas là : le génome des lamproies, par sa forte teneur en GC et éléments répétés, est difficile à assembler (Benjamini and Speed, 2012; Smith *et al.*, 2013), et donc n'est toujours pas complet. Qui plus est, il a été découvert en 2009 que les lamproies ont un "programme de réarrangement génomique" : au cours du développement embryonnaire, 0.5 Gb des 2.3 Gb du génome sont éliminés dans la lignée somatique et ne sont conservés que dans la lignée germinale (Smith *et al.*, 2009). Or, le génome avait été obtenu suite à l'assemblage d'ADN somatique, et en 2018 l'analyse du séquençage du génome "germinal" a été réalisée (Smith *et al.*, 2018). Les résultats de cette équipe semblent

confirmer l'hypothèse de trois événements de duplications chromosomiques puis d'une WGD (unique !) à la base des Vertébrés, suivies de deux duplications chromosomiques des chromosomes portant le complexe Hox chez les lamproies.

Concernant la lamproie, ce scénario semble plus compatible avec certaines observations qu'une troisième WGD. Par exemple, il n'y a que deux complexes *ParaHox* dans cette lignée, au lieu de 4 chez l'Homme (Zhang, Ravi, *et al.*, 2017). On en attendrait 8 dans le cas d'une WGD surnuméraire.

Concernant les Vertébrés, ces résultats récents viennent réouvrir un débat qu'on pensait clos : "1R/2R or not 1R/2R" ? La majorité de la communauté est toutefois en faveur de la 1R/2R, et ces travaux sur la Lamproie ne sont pour l'instant le fait d'un seul et même groupe de recherche. En attendant que ces travaux soient confirmés, par exemple par le séquençage de génomes de myxines (d'autres Cyclostomes), la 1R/2R ancestrale est le scénario retenu (Spring, 1997).

2. Les WGD chez les "Poissons"

a. Importance de l'étude des WGD chez les Téléostéens dans la compréhension des WGD des Vertébrés

Environ la moitié des espèces de Vertébrés sont des espèces de poissons Téléostéens. Outre leur importance taxonomique, l'étude des génomes de ce taxon a pour beaucoup contribué à une meilleure compréhension des duplications de génome chez les Vertébrés.

En premier lieu, l'essentiel des analyses cytogénétiques menées par Ohno dans les années 1960-70, et qui ont contribué de manière majeure à l'étude des duplications de gène en général et notamment des WGD, ont été menées chez les poissons. Ils ont donc un rôle "historique", épistémologique, dans l'analyse des WGD. Le livre de S. Ohno, un livre traitant largement de cytogénétique des poissons mais en réalité discutant de concepts valables pour tous les Vertébrés, a été peu lu à l'époque de sa parution mais a vu ses citations multipliées par 3 entre 1990 et 2000 (Wolfe, 2001).

Par ailleurs, les premières études sur les génomes de poissons, avant même que des génomes entiers soient disponibles, sont venues alimenter la réflexion sur la 1R/2R. Chez le Poisson zèbre, la mise en évidence de sept complexes Hox ainsi que leur cartographie a fait suggérer aux auteurs de l'étude : "*These results suggest two rounds of HOX chromosome duplication (probably whole genome duplication) before the divergence of ray-finned and lobe-finned fishes, and one more in ray-finned fish."* (Amores *et al.*, 1998). Ainsi, dès 1998, soit bien avant la parution du génome complet en 2013 (Howe *et al.*, 2013), les premières cartes génétiques chez le Poisson zèbre sont en faveur d'une 1R/2R puis d'une "3R", un troisième round, chez les Téléostéens (Postlethwait *et al.*, 1998).

b. Mise en évidence de la Ts3R

Cependant, ces résultats ne concernaient qu'un nombre restreint de gènes chez quelques espèces seulement, ce qui conduit certains auteurs à mettre en doute la possibilité d'une WGD supplémentaire (Robinson-Rechavi *et al.*, 2001). La publication de génomes "compacts" de Téléostéens permettra de trancher. En 2002 paraît le génome du Fugu *Takifugu rubripes* (Aparicio *et al.*, 2002), le deuxième génome de Vertébré après le génome humain, mais les auteurs ne mettent pas en évidence un événement de WGD supplémentaire par rapport aux autres Vertébrés ; cela est probablement dû au faible nombre de clusters Hox chez cette espèce (4, de même que chez l'Homme), qui ne permet pas de conclure à la 3R. C'est en 2004, deux ans plus tard, qu'une "*definitive answer*" (sic) à la question est apportée, avec la publication du génome du Tétraodon *Tetraodon nigroviridis* (Jaillon *et al.*, 2004). Les auteurs montrent en particulier pour la première fois la conservation à l'échelle génomique de blocs synténiques dupliqués (paralogons), souvent sur des chromosomes différents.

Cette duplication a été alternativement appelée FSGD (Fish-Specific Genome Duplication), TSGD (Teleost-Specific Genome Duplication) ou encore Ts3R, par allusion aux deux premiers événements de duplication de génome ayant eu lieu à la base des Vertébrés : 1R et 2R. Par la suite, je n'utiliserai que le terme Ts3R, qui est selon moi plus "parlant" car il indique le nombre de duplications de génomes. Si d'aventure la "1R" venait à être remise en question suite aux travaux

chez la Lamproie, peut-être parlerons-nous dans quelques années de… Ts2R. La temporalité de cet événement est discutée et on pense aujourd'hui qu'elle a eu lieu il y a environ 300 millions d'années (les estimations vont de 225 à 333 Ma) (Hurley *et al.*, 2007; Santini *et al.*, 2009; Near *et al.*, 2012).

Le taxon frère des Téléostéens, les Holostéens, est composé d'une plus faible diversité d'espèces : huit seulement. Ce taxon actuellement peu divers n'a pas connu d'événement de duplication de génome (nous reviendrons plus loin sur les liens entre WGD et richesse spécifique). Une espèce en particulier, le Lépisosté tacheté *Lepisosteus oculatus*, est une espèce modèle de ce groupe considérée comme un bon groupe externe aux Téléostéens (Braasch *et al.*, 2015). Le génome de cet organisme, qualifié de "fossile vivant" par Darwin, a été publié en 2016 et montre que son évolution a été lente en comparaison des génomes des Téléostéens (Braasch *et al.*, 2016). Par ailleurs, c'est le seul assemblage au niveau chromosomique (assemblage appelé "chromonome") aujourd'hui disponible pour un "non-tétrapode non-Téléostéen", ce qui en fait un modèle de choix pour des études de génomique comparative chez les Vertébrés. En particulier, contrairement à des hypothèses précédentes (Nakatani *et al.*, 2011), l'étude du génome du Lépisosté tacheté a permis de montrer que des événements de fusion de chromosomes avaient eu lieu chez les Téléostéens avant la Ts3R mais après la divergence Holostéens/Téléostéens. Ces résultats suggèrent que la Ts3R n'a pas été le seul facteur à l'origine des forts taux de réarrangement génomique observés chez les poissons Téléostéens (Braasch *et al.*, 2016).

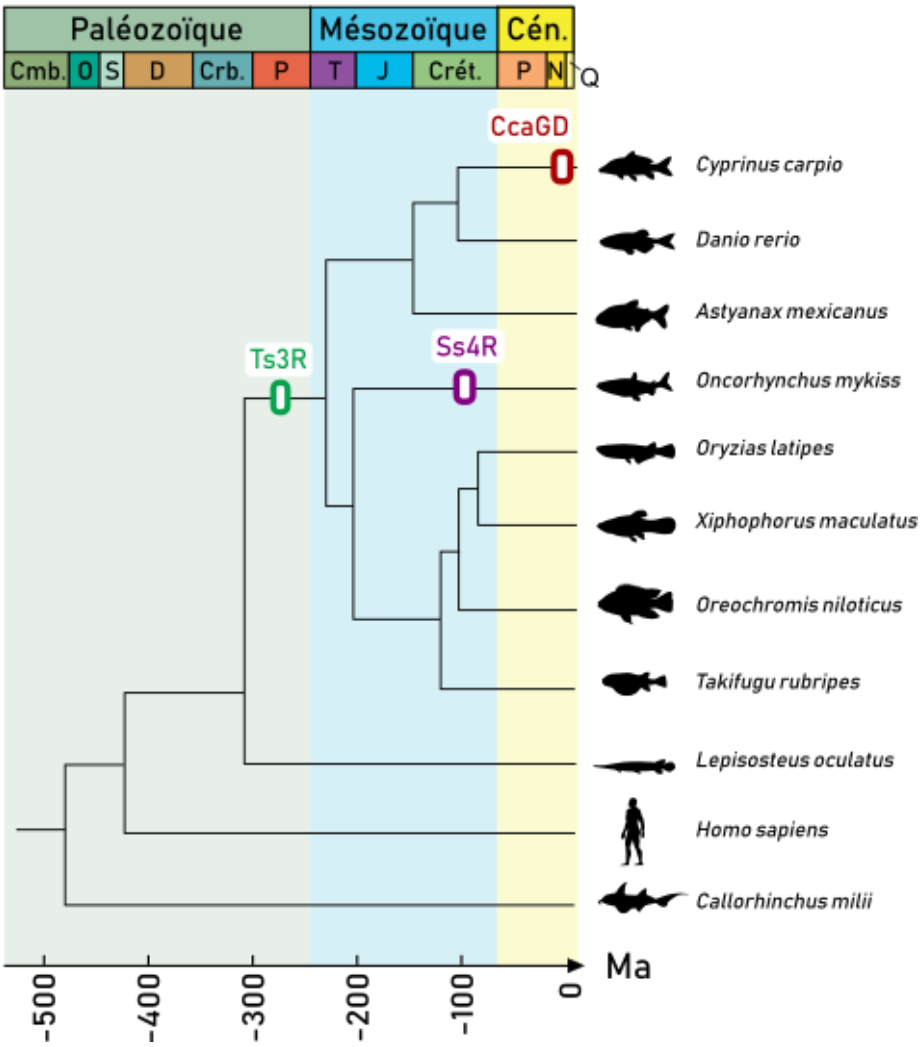


Figure 33 - Les WGD chez les poissons Téléostéens

Au cours de l'évolution des Téléostéens, plusieurs événements de WGD ont eu lieu : un à la base de cette lignée (Ts3R), un chez les Salmonidés (Ss4R), un chez les Cyprininae (CcaGD).

Les temps de divergence entre les différentes lignées ont été estimés par la plate-forme TimeTree,org. Silhouettes : PhyloPic.org

c. D'autres WGD chez les Téléostéens

Les Téléostéens dans leur ensemble sont concernés par la Ts3R, ayant eu lieu il y a environ 300 Ma. Toutefois, au sein même de ce taxon, d'autres événements de WGD ont été décrits.

i. Salmonidés

Peut-être parce que les espèces concernées sont des poissons d'élevage, donc particulièrement étudiées (Thorgaard *et al.*, 2002), un autre cas documenté de WGD chez les Téléostéens est celle spécifique des Salmonidés. Appelée alternativement SSGD (*Salmonid-Specific Genome Duplication*) ou Ss4R ("4ème" round de WGD), je l'appellerai par la suite Ss4R. Cet événement de duplication, plus récent, a lui été daté à environ 100 Ma (entre 80 et 100 Ma selon les auteurs) (Macqueen and Johnston, 2014; Berthelot *et al.*, 2014; Lien *et al.*, 2016). Le processus ayant conduit à une WGD est probablement une autotétraploïdie (Lien *et al.*, 2016). Deux génomes de Salmonidés sont aujourd'hui disponibles : le génome de la Truite arc-en-ciel (*Oncorhynchus mykiss*) et celui du Saumon (*Salmo salar*). Ces deux génomes portent la trace de nombreux réarrangements, mais le partage de grands paralogons entre ces génomes suggère que ces réarrangements ont eu lieu avant la séparation de ces deux lignées (Lien *et al.*, 2016).

ii. Cyprinidés

Les Cyprinidés, famille au sein des Cypriniformes, incluent le Poisson zèbre et les Carpes, dont la Carpe Commune (*Cyprinus carpio*), espèce d'importance majeure en aquaculture (en 2011, son marché était estimé à 5.3 milliards de dollars – source FAO). La Carpe commune (2n = 100) compte le double des chromosomes des autres Carpes (2n ≈ 50), notamment la Carpe de roseau *Ctenopharyngodon idella* (Wang *et al.*, 2015), ce qui a fait suggérer dès les années 1960 qu'une WGD avait récemment eu lieu dans cette lignée. L'auteur qui avançait cette hypothèse était un certain… S. Ohno (Ohno *et al.*, 1967).

La date exacte de cet événement de duplication dans la lignée de la Carpe commune, appelée "CcaGD", était encore discutée au début des années 2000 : 58 Ma (Zhang *et al.*, 1995) ? 16 Ma (Larhammar and Risinger, 1994) ? En 2003, une étude a daté l'événement de WGD à 12 Ma (David *et al.*, 2003) avant qu'en 2012 une étude menée sur le transcriptome de cette espèce ne date cet événement à 5.6-11.3 Ma (Wang *et al.*, 2012). Le génome de la Carpe commune *Cyprinus carpio*, publié en 2014, a estimé que la WGD était due à un événement d'allotétraploïdisation survenu il y a 8.2 Ma (Xu *et al.*, 2014b). La "récente" occurrence de cet événement de WGD par rapport aux autres WGD chez les poissons Téléostéens permet en particulier d'étudier l'évolution post-duplication d'un génome sur une échelle de temps différente.

iii. Catostomidae et Cobitidae

D'autres événements de duplication de génome, moins étudiés, ont été documentés chez deux autres familles de l'ordre des Cypriniformes : les Catostomidae (Uyeno and Smith, 1972) et les Cobitidae (Ferris and Whitt, 1977).

d. Des WGD chez les "poissons non-Téléostéens"

Les WGD ont également eu lieu chez les "poissons non-Téléostéens", et notamment les Acipenseriformes (la famille des Esturgeons et des Poissons-spatules). Ce clade est marqué par une WGD ancestrale ayant eu lieu il y a 190 Ma et des événements multiples de polyploïdisation indépendantes dans chacune des lignées (Trifonov *et al.*, 2016), ayant conduit à une explosion du nombre de chromosomes chez les Esturgeons (plus de 200)(Havelka *et al.*, 2016).

3. Les WGD chez les autres Vertébrés

Hormis la 1R/2R, peu de WGD ont été documentées chez les autres Vertébrés, à l'exception de la lignée du Xénope qui contient plusieurs espèces avec des degrés de ploïdie différents (Kobel and Du Pasquier, 1986). L'espèce *Xenopus tropicalis* est diploïde et les autres espèces ont un nombre de chromosomes qui varie (en ordre de grandeur) d'un facteur

2 (tétraploïde, comme *X. laevis*), 4 (octoploïde) ou 6 (dodécaploïde). L'étude récente du génome de *Xenopus laevis* (2n = 36) a montré qu'il était issu d'un événement d'allotétraploïdisation à partir de deux espèces ancestrales diploïdes ayant chacune 2n = 18 chromosomes (Session *et al.*, 2016).

Les duplications de génomes ainsi que les WGD sont plus rares chez les autres Tétrapodes, et notamment chez les Mammifères et les Oiseaux. Les mécanismes "bloquant" ces phénomènes ne sont pas connus avec précision : "*No satisfactory explanation of the rarity of polyploidy among animals exists*", affirme Orr en 1990 (Orr, 1990), mais il a depuis été suggéré que des contraintes sur la compensation de dosage des chromosomes sexuels empêchaient probablement la polyploïdisation (Wertheim *et al.*, 2013). Un cas de duplication de génome chez un rongeur, le Rat-viscache roux d'Argentine, en 1999 (Gallardo *et al.*, 1999), a ouvert la porte à de nouvelles théories, mais il a vite été remis en question (Svartman *et al.*, 2005). Retour au point de départ !

4. Conclusion sur les WGD chez les Vertébrés

En conclusion, les événements de WGD sont relativement fréquents chez les Vertébrés. Toutefois, on peut constater que les WGD ont souvent été détectées là où on les a "cherchées" : le plus souvent, dans des lignées de poissons d'élevage (Carpe, Truite, Saumon, Esturgeon) ou chez d'importants organismes modèles (Poisson-zèbre, Xénope). S'il existe chez les Mammifères et les Oiseaux des vraies indispositions biologiques à la polyploïdisation, probablement dues au déterminisme du sexe, il y a fort à parier que les WGD soient plus fréquentes chez le reste des Vertébrés que ce qui a été observé jusqu'à présent. Nul doute que l'émergence de nouveaux modèles, notamment chez les poissons, permettra dans les années à venir d'explorer de nouveaux génomes (Braasch *et al.*, 2015).

III. Les conséquences évolutives des WGD diffèrent selon les lignées eucaryotes

A. Les WGD permettent l'apparition de nouveaux gènes qui fournissent un substrat propice à la diversification génétique

1. Les WGD sont une source de nouvelles séquences…

a. Les WGD constituent une source majeure de nouveaux gènes à partir desquels les changements de fonction sont possibles

i. Une WGD induit le doublement des gènes et des séquences régulatrices d'un individu…

La duplication d'un génome conduit à un doublement immédiat de l'ensemble du génome d'un individu, et en particulier des gènes et des séquences régulatrices. Certains auteurs ont suggéré que cette augmentation soudaine du nombre de gènes fournissait le "*raw material*", le matériau brut, à partir duquel l'acquisition de nouveaux traits devenait possible. Cela a en particulier conduit Susumu Ohno à déclarer, dès les années 1970, que "*Natural selection merely modified while redundancy created*" (Ohno 1970). On note que Ohno utilise le terme de "*redundancy*", et non pas "*duplication*" : le phrasé suggère qu'au-delà de la duplication en elle-même (aspect structurel) c'est effectivement la redondance (aspect fonctionnel) des fonctions des différents produits qui est la base de l'innovation phénotypique.

ii. … qui ont des devenirs variables

Considérons un gène ancestral X, dans un génome G. Immédiatement après la duplication de G, deux gènes ayant exactement les mêmes propriétés co-existent dans le génome. Plusieurs devenirs sont alors possibles pour les deux copies, et si on observe le génome descendant de G, bien des générations plus tard, on pourrait observer différents scénarios :

1. L'une des deux copies a conservé la fonction ancestrale, pendant que l'autre copie a accumulé des mutations et dégénéré (pseudogénisation), ou avoir été complètement éliminée : on parle de non-fonctionnalisation.

2. les deux copies se sont partagés les fonctions ancestrales du gène : on parle de sub-fonctionnalisation

3. plus rarement, une des deux copies a conservé la fonction ancestrale du gène et l'autre copie a acquis une nouvelle fonction : on parle de néo-fonctionnalisation (Thompson *et al.*, 2014) (terme inventé par Force *et al.*, 1999). Ohno, à qui on attribue souvent ce concept, avait lui parlé de "*The Creation of a New Gene from a Redundant Duplicate of an Old Gene*"… périphrase condensée en "*néo-fonctionnalisation*" quelque trente années plus tard.

4. Les deux copies sont présentes, quasi-identiques, et ont gardé leur fonction ancestrale (on parle de gène à effet dose) ; c'est le cas de certaines protéines ribosomiques chez la Levure (Wolfe, 2001)

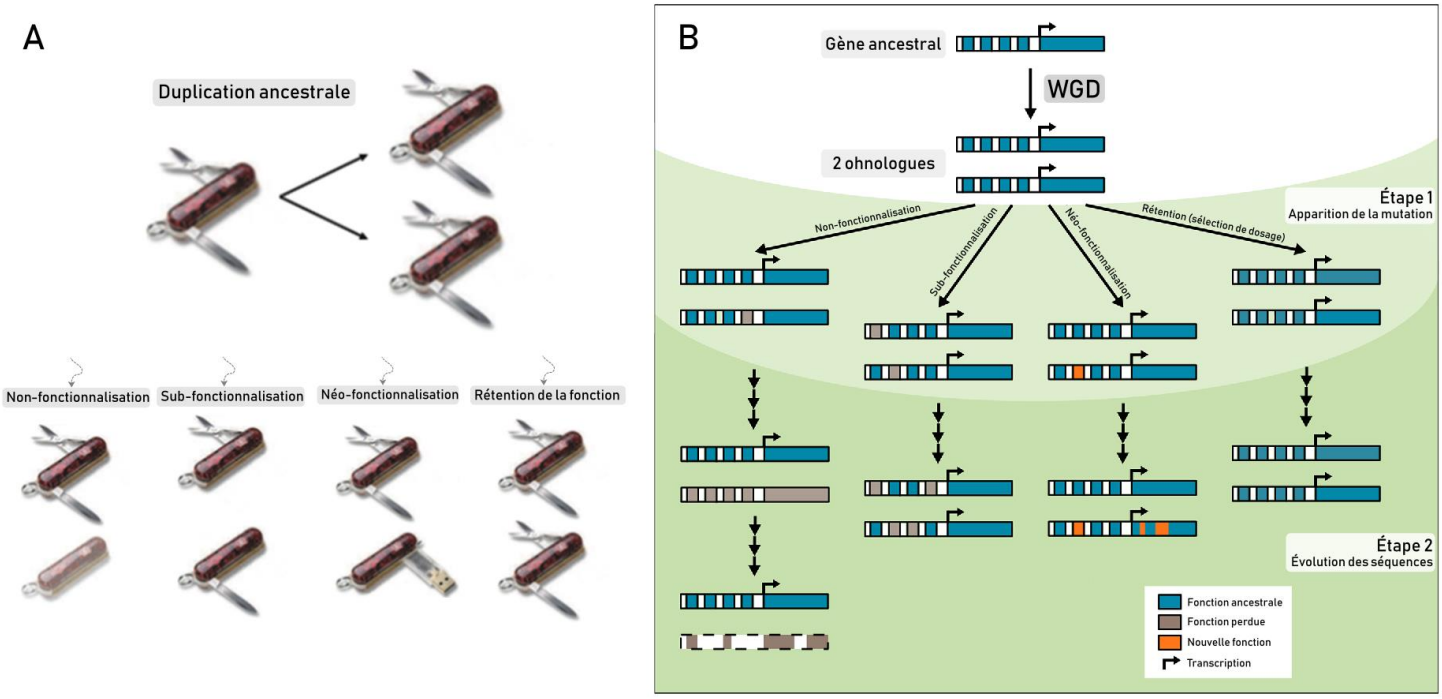


Figure 34 - Devenirs possibles d'un gène dupliqué.

(A) Représentation schématique du devenir des deux paralogues après une duplication (WGD par exemple). Plusieurs issues sont possibles : perte de l'un des deux paralogues (non-fonctionnalisation), division des fonctions ancestrales (sub-fonctionnalisation), gain d'une nouvelle fonction par l'un des deux (néo-fonctionnalisation) ou conservation des deux et rétention de la fonction ancestrale.

(B) Représentation des événements ayant lieu au niveau de la séquence. Les changements potentiels peuvent concerner soit la séquence codante, soit les séquences régulatrices. *Adapté de Glasauer and Neuhaus (2014).*

Les trois derniers processus conduisent à une préservation à long terme des deux paralogues. Si ces derniers sont issus d'une WGD, ces paralogues sont appelés *ohnologues*, en hommage à S. Ohno, précurseur du concept. Ces processus ne sont d'ailleurs pas mutuellement incompatibles : si les deux copies se partagent la fonction ancestrale et une des

nouvelles copies a acquis une nouvelle fonction, on parlera de néo-sub-fonctionnalisation (He and Zhang, 2005). On note que les changements de fonction peuvent soit concerner les séquences codantes, soit les séquences régulatrices.

De nombreux modèles ont été développés au cours des vingt dernières années expliquant les facteurs à l'origine de la rétention ou non d'un gène après une duplication et/ou prédisant les devenirs possibles des gènes dupliqués. Une synthèse exhaustive de ces modèles est disponible dans (Innan and Kondrashov, 2010). Dans le cas de la sub-fonctionnalisation, deux modèles principaux existent :

1. Le modèle DDC, pour "*Duplication-Degeneration-Complementation*" (Force *et al.*, 1999) : dans ce modèle, l'accumulation de mutations restreint la fonction de chacun des deux gènes et, une fois ces mutations fixées, aucune copie ne peut à elle seule assurer la fonction ancestrale. Un exemple de sub-fonctionnalisation par DDC est le cas de la protéine CRALBP (cellular retinaldehyde-binding protein) (Fleisch *et al.*, 2008). Chez les Tétrapodes, cette protéine importante dans la vision est exprimée à la fois dans les cellules de l'épithélium pigmenté de la rétine et dans les cellules gliales de Müller. A l'inverse, chez le Poisson-zèbre, il y a deux paralogues, chacun exclusivement exprimé dans un des deux types cellulaires. Dans ce cas, c'est un changement de séquence régulatrice qui est à l'œuvre, mais des cas de changements de séquence codante ont déjà été documentés, par exemple dans le cas du gène *pomc* codant pour la proopiomélanocortine (de Souza *et al.*, 2005).

2. Le modèle EAC, ou "*Escape from adaptive conflict*" (Hughes, 1994; Des Marais and Rausher, 2008). Dans ce modèle, les deux paralogues n'évoluent pas seulement par accumulation de mutations délétères ou neutres, mais par sélection positive. Cette situation est possible lorsque le gène ancestral portait deux fonctions qui ne pouvaient être améliorées simultanément au sein de la même séquence, mais qui peuvent être améliorées dans les deux copies filles, de manière indépendante. Les exemples illustrant ce modèle sont rares car ils requièrent de montrer que (i) les fonctions sont améliorées par rapport au gène ancestral et que (ii) cette amélioration était vraiment contrainte avant la duplication. (Des Marais and Rausher, 2008; Huang *et al.*, 2012).

iii. Tendances à la sub- ou néo-fonctionnalisation au niveau génomique

Les mécanismes ci-dessus s'appliquent à tout événement de duplication. Dans le cas des WGD, ces processus s'appliquent à tous les gènes du génome, et il est possible, en réalisant des études génomiques, de mettre en évidence que les paralogues issus d'un événement de WGD n'évoluent pas à la même vitesse (en moyenne). Cela a été réalisé en comparant les gènes humains à ceux de poissons Téléostéens. L'effet de la Ts3R a été en particulier quantifié, et des taux d'évolution asymétriques ont été identifiés entre paralogues issus de la Ts3R (Brunet *et al.*, 2006) et interprétés comme majoritairement dus à des événements de sub-fonctionnalisation. Les plus forts taux d'évolution constatés chez les Téléostéens comparé à l'Homme ont même fait suggérer aux auteurs qu'il pourrait en général s'agir de néo-fonctionnalisation, ou d'un mélange des deux (sub-néo-fonctionnalisation), associé à de la sélection positive.

Chez les Paramécies, de même, des taux d'évolution différents ont été trouvés entre les paralogues (McGrath, Gout, Doak, *et al.*, 2014). Cependant, à l'inverse des Téléostéens, les auteurs concluent que dans cette lignée, c'est essentiellement la pseudogénisation (associée à une relaxation de la sélection purifiante), plutôt que la sub-fonctionnalisation, qui explique ces variations des taux d'évolution entre paralogues. Chez le Saumon, la perte de gènes est elle aussi liée principalement à la pseudogénisation post-Ss4R puisque des fragments homéologues pseudogénisés sont fréquemment retrouvés en position attendue par la synténie (Lien *et al.*, 2016).

b. L'état polyploïde est transitoire et la diploïdie est l'état stable

Considérons de nouveau le génome G. Immédiatement après duplication, par exemple une allotétraploïdisation, le génome néo-formé G' est tétraploïde. Or, l'étude des génomes Eucaryotes montre qu'ils sont aujourd'hui tous diploïdes, à de rares exceptions près (certaines Levures, certaines plantes). Ainsi, la conservation de la polyploïdie semble être difficile sur des temps longs, et l'état stable pour un génome semble être la diploïdie. Toutefois, comme le soulignait K.

Wolfe en 2001, "*Diploidization, the evolutionary process whereby a tetraploid species 'decays' to become a diploid (...) is one of the most interesting but unclear aspects of genome evolution.*" (Wolfe, 2001).

La diploïdisation est un processus graduel de réarrangements chromosomiques et, au sein d'un même organisme, il est possible d'observer certaines régions génomiques diploïdes alors que d'autres sont encore tétraploïdes. Quand un organisme est parfaitement retourné à l'état diploïde (exemple de l'Homme), on parle de diploïdie paléopolyploïde. Les réarrangements chromosomiques font que le nombre de chromosomes n'est pas uniquement croissant au cours de l'évolution. D'ailleurs, il a été estimé que l'ancêtre des Téléostéens avait vraisemblablement entre 24 et 26 chromosomes (Glasauer and Neuhauss, 2014). Chez les Salmonidés, la rediploïdisation consécutive à la duplication de génome ayant eu lieu il y a environ 100 millions d'années n'est pas encore totalement effectuée et certaines espèces de truite, donc la truite arc-en-ciel *Oncorhynchus mykiss*, présentent une ségrégation de tétravalents à la méiose (Allendorf and Danzmann, 1997; Lien *et al.*, 2011) avec potentiellement 10% résiduels du génome encore à l'état tétrasomique (Lien *et al.*, 2016).

Les mécanismes de la rediploïdisation sont donc, à l'heure actuelle, assez mal compris. Certains auteurs ont suggéré un processus graduel (Berthelot *et al.*, 2014) alors que d'autres supposent de grands réarrangements chromosomiques (Lien *et al.*, 2016)et une réduction de la pression de sélection (Douglas *et al.*, 2015).

c. Les WGD peuvent modifier les réseaux de régulation de gènes

Toute duplication de gène induit la multiplication du nombre de copies, et les facteurs de transcription sont parmi les gènes dont la duplication a le plus de conséquences (Teichmann and Babu, 2004). Après une duplication locale (Small-Scale Duplication, ou SSD) du gène *tf* en *tf* et *tf'*, si le gène *tf* code pour un facteur de transcription TF, il y aura immédiatement deux fois plus de TF pour un même gène cible. Le nombre de séquences régulatrices auxquelles TF se fixe étant en quantités constantes, toutes les protéines ne peuvent se fixer. La duplication implique donc, le plus souvent, une diminution globale de l'expression des deux paralogues de *tf* pour revenir à au niveau d'expression ancestral : (i) soit les deux paralogues sont moins exprimés (effet dose), (ii) soit, au hasard, l'un des deux est non-fonctionnalisé, (iii) soit les deux se spécialisent dans des types cellulaires différents (sub-fonctionnalisation).

À l'inverse d'une SSD, une WGD duplique non seulement *tf*, mais l'ensemble de ses gènes cibles *tg* ("target gene") ou de ses cofacteurs. Cela peut permettre une évolution plus lente des deux paralogues de *tf* : toutes les quantités étant doublées, les deux protéines TF et TF' sont nécessaires, et, à terme, les deux protéines peuvent diverger tout en restant fonctionnelles. La divergence des séquences de régulation peut également s'effectuer de concert, comme cela a été montré chez la Levure (Ihmels *et al.*, 2005; Hittinger and Carroll, 2007). En agissant à la fois au niveau des séquences mêmes des facteurs de transcription (modification en *trans*) et au niveau des séquences de régulation (modifications en *cis*), les WGD peuvent induire une néo-fonctionnalisation et l'apparition de nouveaux modules de régulations de gènes impliqués dans une nouvelle fonction. Les duplications de facteurs de transcription jouent donc un rôle prépondérant dans les modifications des réseaux de régulation de gènes (Voordeckers *et al.*, 2015). Sur le long terme, les conséquences d'une WGD sur le "*rewiring*" (rebranchement) des réseaux de régulation de gènes sont bien plus fortes que celles d'une SSD (Thompson *et al.*, 2013).

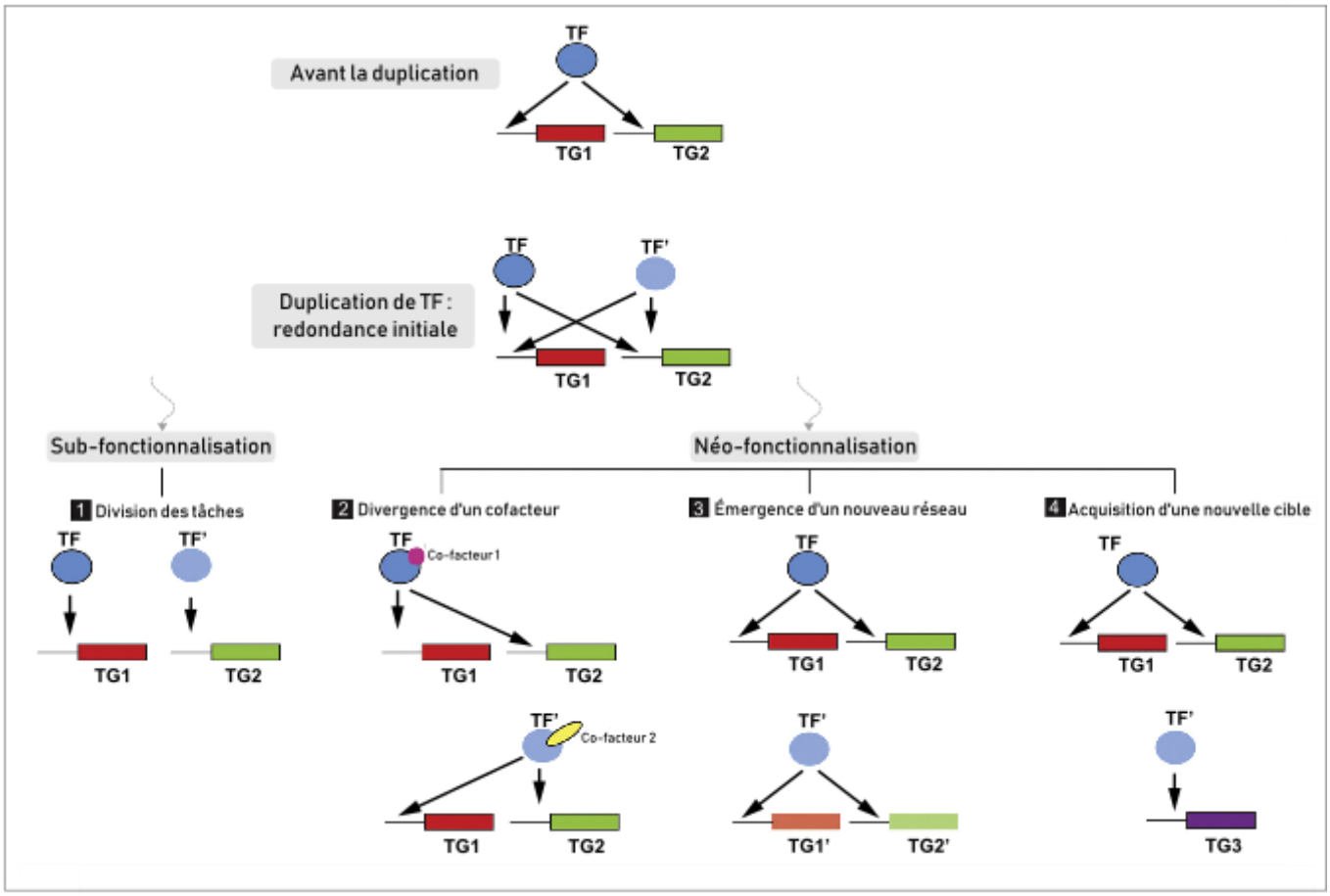


Figure 35 - Conséquences de la duplication d'un facteur de transcription sur les réseaux de régulation de gènes.

Considérons un facteur de transcription TF qui régule deux gènes cibles, TG1 et TG2. Aussitôt après la duplication de TF en TF et TF', les paralogues ont une fonction redondante, mais la duplication peut induire un changement de régulation des gènes cibles. (1) S'il y a une division des tâches entre paralogues, il y a sub-fonctionnalisation. (2) Chacun des deux paralogues peut également se spécialiser dans la reconnaissance d'un cofacteur spécifique, induisant une néo-régulation des gènes cibles en fonction des conditions cellulaires. (3) Dans le cas d'une WGD, les gènes cibles sont également dupliqués. Si un des deux paralogues de TF se spécialise dans la régulation de ces gènes, un nouveau réseau parallèle émerge et est susceptible d'acquérir une nouvelle fonction. (4) Enfin, si TF'change de fonction et régule un nouveau gène cible, il y a également néo-fonctionnalisation. *Adapté de Voordeckers et al. (2015).*

d. Les WGD ne sont pas les seuls processus à l'origine de la formation de nouvelles séquences

i. La formation de nouveaux gènes par duplication est majoritaire…

Nous avons jusque-là beaucoup évoqué les WGD. Toutefois, celles-ci ne sont qu'un cas, bien que majeur, d'apparition de nouveaux gènes. Les SSD, évoquées brièvement dans la partie précédente, en sont un autre. Dans l'ensemble, les gènes issus de processus de duplication constituent une grande majorité des gènes : 50% des gènes procaryotes en sont issus. Pour les Eucaryotes, ce taux s'élève à 90% des gènes (Voordeckers *et al.*, 2015).

ii. … mais n'explique pas seule l'apparition des nouveaux gènes

1. Des nouveaux gènes à partir de séquences codantes déjà existantes

Outre les duplications, d'autres mécanismes sont à l'origine de l'apparition de nouveaux gènes à partir desquels l'adaptation est possible (Kaessmann, 2010). La domestication d'éléments transposables en est un, tels que les gènes *gin*, des intégrases de la famille des Gypsy (Warren *et al.*, 2015). Les événements de fusion de gènes pré-existants constituent

un autre mécanisme ; le gène *jingwei* chez la Drosophile, probablement impliqué dans la biosynthèse des phéromones, est issu de la fusion du gène de l'alcool déshydrogénase *adh* et du gène *yande* (Kaessmann *et al.*, 2009).

Les transferts horizontaux (ou latéraux) de gènes (HGT) sont un autre cas de formation de nouveaux gènes qui contribuent à l'acquisition de nouveaux traits par les lignées concernées (Soucy *et al.*, 2015). Ces mécanismes correspondent à l'acquisition d'un gène depuis un autre organisme qui n'est pas l'ascendant direct. Alors que les HGT sont fréquents chez les procaryotes, il est aujourd'hui reconnu que les HGT ont grandement contribué à l'évolution des Eucaryotes, unicellulaires certes, mais également multicellulaires, et notamment des Plantes, chez lesquelles les HGT sont très documentées, et les Métazoaires (Boto, 2014). Chez ces derniers, de nombreux taxons sont concernés (annélides, mollusques, éponges, cnidaires, arthropodes, rotifères), mais pour ce qui est des chordés, les cas documentés sont plus rares.

Chez les Urochordés, un possible HGT impliquant l'acquisition de la cellulose synthase a été documenté (Nakashima *et al.*, 2004), mais chez les Vertébrés, tous les cas documentés ont été sujet à controverse. C'est le cas en particulier des protéines anti-gel chez les poissons, pour lesquelles un mécanisme de HGT a été proposé en 2008 (Graham *et al.*, 2008). Les auteurs proposent que le mécanisme le plus probable ayant généré le HGT soit l'absorption d'ADN environnemental (par exemple issu de la lyse des spermatozoïdes d'une espèce donneuse) par les spermatozoïdes d'une espèce receveuse, qui se serait peu à peu intégré au génome après une première phase extrachomosomale. Cet événement est très peu probable, étant donnée la demi-vie de quelques heures de l'ADN dans l'eau et l'incompatibilité génétique entre l'espèce donneuse et receveuse, mais les auteurs suggèrent que l'avantage adaptatif procuré par les protéines anti-gel est tel qu'un événement de ce type aurait pu se produire à de multiples reprises (Graham *et al.*, 2008). Chez l'Eperlan , il est possible que les protéines anti-gel soient issues d'un gène de Hareng transmis par HGT (Graham *et al.*, 2012).

Ce scénario est discuté et un scénario alternatif qui n'est pas à exclure serait que ces gènes aient été présents de manière ancestrale chez tous les poissons et qu'ils aient été perdus chez les espèces pour lesquels ils ne conféraient pas d'avantage adaptatif (Boto, 2014). Quel que soit le scénario sous-jacent, cet exemple illustre une propriété du milieu aquatique (au sein duquel évoluent les Téléostéens) : aqueux, il permet la conservation transitoire et la dispersion de molécules d'ADN, susceptibles d'être transférées à d'autres espèces. Ce milieu pourrait également permettre à des parasites peu spécifiques comme des rétrovirus (à l'instar du médiatisé virus de la grippe aviaire) de co-infecter des espèces dans des taxons différents et de participer à la transmission horizontale de séquences (Naville *et al.*, 2014; Naville and Volff, 2016).

Des transferts horizontaux entre parasites ou symbiontes et l'hôte ont également été fréquents au cours de l'évolution des Eucaryotes, et sont également possibles chez l'Homme. Ils ont été récemment identifiés dans le cas de certains cancers dont les cellules présentent des gènes issus de bactéries du microbiote intestinal (Riley *et al.*, 2013). Ils sont, eux aussi, sources de nouveaux gènes.

2. Des nouveaux gènes à partir de séquences non-codantes

Enfin, quelle est la probabilité qu'un nouveau gène apparaisse *de novo*, à partir de rien ("from scratch"), ou plutôt de séquences nucléotidiques ne codant pour aucune protéine ? Selon les mots mêmes de François Jacob, pour qui l'évolution travaille comme un "bricoleur" (sic) à partir des fonctions déjà disponibles, cet événement est trop improbable : *"The probability that a functional protein would appear de novo by a random association of amino acids is practically zero….creation of entirely new nucleotide sequences could not be of any importance in the production of new information"* (Jacob, 1977). Pourtant, des travaux récents suggèrent que la formation d'ARN fonctionnel, voire même de peptides, est possible à partir de séquences aléatoires insérées dans des plasmides dans la bactérie *Escherichia coli* (Neme *et al.*, 2017). Chez l'Homme, plusieurs gènes encore non-caractérisés ont également été formés *de novo* après la divergence avec le chimpanzé (Knowles and McLysaght, 2009).

On note que les différents processus à l'origine de la formation de nouveaux gènes interagissent parfois dans la génération de l'adaptation : par exemple, les WGD sont parfois associées à des événements de prolifération des éléments transposables (Oliver *et al.*, 2013), comme cela a pu être montré dans le cas de la WGD récente ayant eu lieu chez les

poissons Téléostéens du groupe des Corydoradinae (Marburger *et al.*, 2018). Il existe des exceptions à cette tendance : à l'échelle des Téléostéens, aucun lien entre la Ts3R et l'expansion des éléments transposables n'a pu être mis en évidence puisque le Lépisosté tacheté, qui a divergé des Téléostéens avant la Ts3R, possède une diversité en éléments transposables équivalente (Chalopin and Volff, 2017).

2. L'essentiel des gènes issus d'une WGD est perdu

L'étude du génome de la Levure montre qu'entre 8 et 14% des gènes dupliqués il y a 100 Ma ont été retenus (Wolfe and Shields, 1997; Scannell *et al.*, 2007) : le reste a donc été perdu. De même, entre 4 et 18% des gènes dupliqués à la Ts3R, ayant eu lieu à la base des Téléostéens il y a environ 300 Ma, ont été conservés, même si ce chiffre varie selon les auteurs et selon les organismes (Jaillon *et al.*, 2004; Inoue *et al.*, 2015; Pasquier *et al.*, 2017). Ainsi, alors que j'ai principalement présenté jusqu'ici les devenirs possibles des gènes retenus, il semblerait que le destin prépondérant d'un gène après un événement de WGD ne soit pas la rétention, mais la perte. D'ailleurs, le titre d'une publication citée plus haut à propos des taux d'évolution asymétriques entre paralogues issus de la Ts3R est "***Gene loss** and evolutionary rates following whole-genome duplication in teleost fishes"* (Brunet *et al.*, 2006). Cette sémantique illustre que la perte de gènes est l'événement prépondérant, mis en avant. La question du devenir des paralogues conservés, bien qu'importante (nous l'avons vu plus haut), ne concerne qu'un nombre plus restreint de gènes.

De plus, au sein des Téléostéens, une autre duplication de génome a eu lieu plus récemment, il y a 100 Ma : la Ss4R, à la base des Salmonidés (Berthelot *et al.*, 2014; Lien *et al.*, 2016). À l'échelle génomique, la rétention est estimée à 48-55% suite à cet événement. Cela semble donc indiquer qu'une duplication plus récente est associée à une rétention plus élevée, et pose des questions sur les mécanismes de perte après une WGD : les gènes sont-ils perdus graduellement ? Sont-ils tous perdus d'un coup, ou sont-ils perdus "par vagues", en masse ? Et enfin, y a-t-il des gènes plus facilement retenus, ou plutôt, moins facilement perdus, après une WGD ?

a. Les mécanismes de pertes après une WGD : perte rapide ou graduelle ?

L'étude des processus de perte a été menée chez les Téléostéens. Dans une étude récente, Pasquier *et al.* montrent que chez le Poisson-zèbre, 13% des gènes ont été gardés en duplicats après la Ts3R, et 12% chez le médaka (Pasquier *et al.*, 2017). Le reste des gènes correspond à des singletons issus de la Ts3R, ce qui suggère des taux de rétention génomiques post-Ts3R inférieurs à 15%. Les auteurs ont ensuite regardé quelle proportion de singletons sont partagés par ces deux espèces. (i) Une proportion élevée suggèrerait que les gènes étaient déjà à l'état de singletons avant la spéciation Poisson-zèbre/Médaka ; (ii) une proportion faible montrerait que les gènes ont été perdus après la divergence de ces lignées, il y a plus de 250 Ma (Near *et al.*, 2012). Les résultats sont en faveur du scénario (i) et indiquent donc que ces gènes ont été perdus en moins de quelques dizaines de millions d'années. Une autre étude de 2015 suggérait déjà qu'entre 70 et 80% des gènes avaient été perdus en moins de 60 Ma après la Ts3R (Inoue *et al.*, 2015).

Ces résultats semblent en concordance avec une étude de 2000 qui suggérait qu'à l'échelle des Eucaryotes dans leur ensemble, la perte de gènes était rapide (quelques millions à quelques dizaines de millions d'années) après un événement de duplication (Lynch and Conery, 2000).

Toutefois, l'étude de l'évolution des génomes chez les Paramécies viennent nuancer ces propos. En effet, il a été montré dans cette lignée que certains gènes perdent probablement rapidement leur état dupliqué alors que de nombreux autres sont conservés à l'état dupliqué et ne sont perdus que graduellement, linéairement, au cours du temps (Aury *et al.*, 2006; McGrath, Gout, Doak, *et al.*, 2014; McGrath, Gout, Johri, *et al.*, 2014).

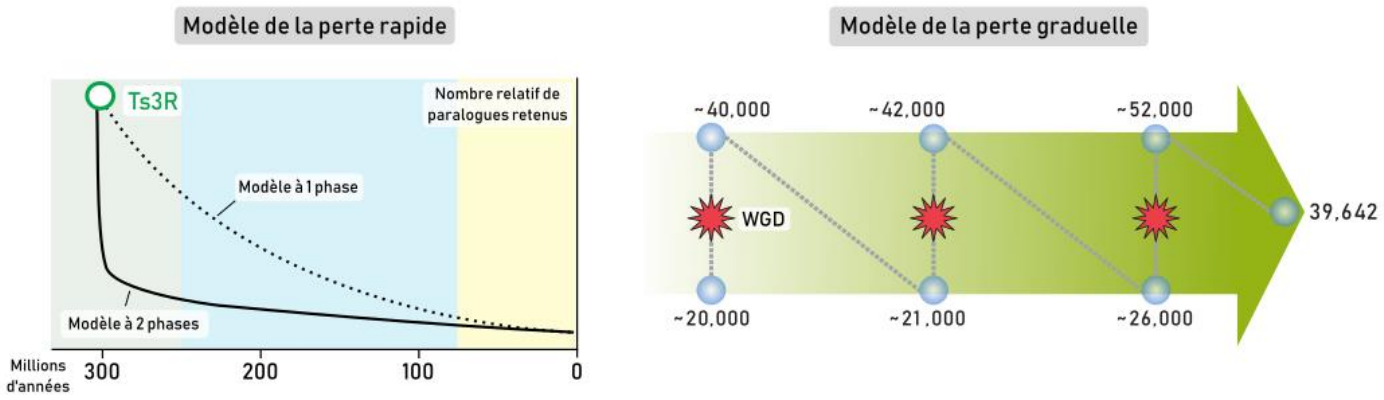


Figure 36 - Modèles de perte de gènes après une duplication de génome.
Le modèle de la perte rapide ou exponentielle de gènes est reconnu après la Ts3R (gauche). Le modèle en 1 phase suppose que la perte de gènes soit une exponentielle décroissante. Le modèle tendant à être accepté aujourd'hui est un modèle en deux phases : une première phase de perte massive, et une seconde phase régulière. Chez les Paramécies, le modèle retenu est une perte graduelle entre les WGD successives. *D'après Inoue et al. (2015) et Aury et al. (2006).*

b. Les facteurs à l'origine de la rétention préférentielle de certains gènes

Si de nombreux gènes sont perdus après un événement de WGD, les gènes retenus ne le sont pas au hasard, et correspondent à certaines caractéristiques.

i. La rétention peut être liée à un changement de fonction (néo- et sub-fonctionnalisation)…

En premier lieu, comme nous l'avons déjà évoqué, les gènes sub- ou néo-fonctionnalisés, concernés par un changement de fonction, sont retenus.

ii. … ou être corrélée à des propriétés intrinsèques des protéines

Les gènes dupliqués peuvent parfois être retenus sans changement de fonction, et il existe de nombreuses propriétés des gènes qui sont dans l'ensemble corrélées à une rétention élevée. J'en détaille ici quelques-unes et évoque uniquement les autres.

Je souligne toutefois le terme "dans l'ensemble corrélées", puisque les liens de cause à effet entre une rétention élevée et l'un ou l'autre de ces paramètres ne sont pas non plus une règle immuable : de nombreuses exceptions existent, que ce soit au niveau des gènes (certains gènes ayant les propriétés requises ne sont pourtant pas retenus) ou au niveau des lignées (certains paramètres s'avèrent valables après un événement de WGD dans une lignée mais pas dans une autre).

Un paramètre majeur contrôlant la rétention d'un gène après un événement de WGD est "l'équilibre de dosage". Après un événement de WGD, tout le contenu du génome est multiplié par deux, et donc en particulier tous les gènes codant pour des facteurs de transcription ainsi que les séquences régulatrices. La stœchiométrie des protéines est conservée. Imaginons que X', un des deux ohnologues issus du gène ancestral X codant pour un facteur de transcription, soit perdu rapidement, alors que les séquences auxquelles il se fixe sont toutes conservés. Si l'expression de l'autre ohnologue X'' n'est pas dans le même multipliée par deux pour assurer la fonction commune aux deux gènes, la perte de X' est contre-sélectionnée. Ainsi, la pression de sélection s'opposant à la perte d'un gène est forte et les deux paralogues sont conservés (Papp *et al.*, 2003; Birchler *et al.*, 2005; Conant *et al.*, 2014).

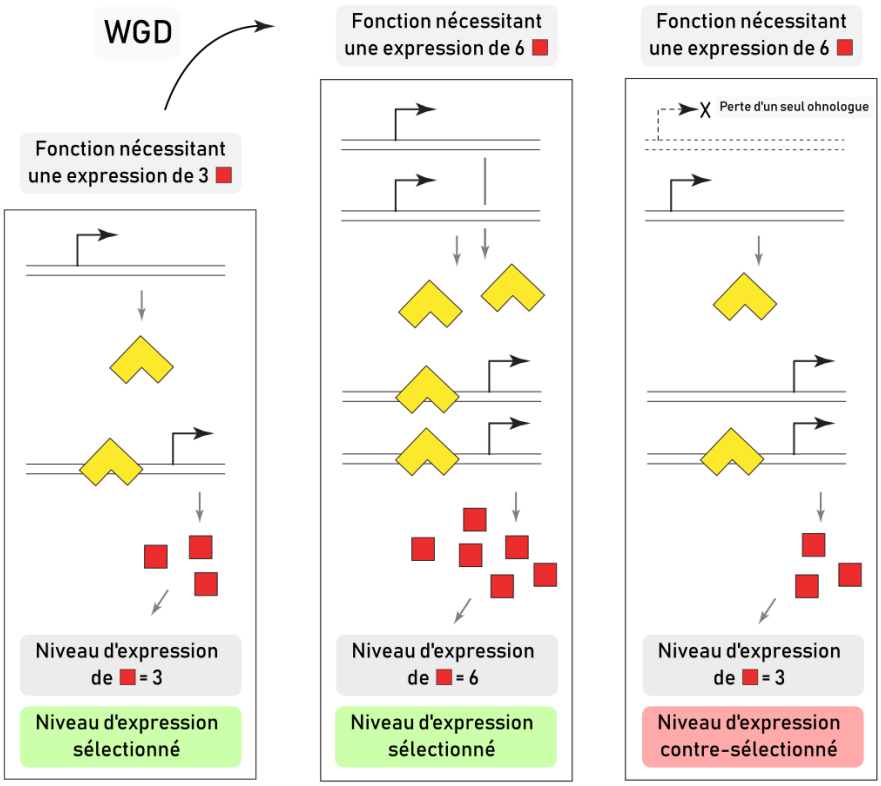


Figure 37 - Importance de l'équilibre de dosage.

Considérons une fonction biologique devant quantitativement être assurée par une certaine « dose » d'une protéine rouge (par exemple, une enzyme). Les niveaux d'expression de la protéine rouge sont dus à l'induction de son expression par un facteur de transcription jaune (gauche). Après une WGD (milieu), toutes les quantités sont doublées (jaune et rouge) et il faut 2 doses de la protéine rouge. Toute perte d'un gène induisant l'expression de celle-ci (par exemple, le facteur de transcription jaune) est contre-sélectionnée car la fonction ne peut alors plus être assurée (droite).

On s'attend donc à ce que les gènes codant pour des facteurs de transcription, ou au cœur de réseaux d'interactions protéiques, soient plus retenus, car leur perte a des conséquences sur l'équilibre du dosage. Cette "prédiction" est effectivement vérifiée pour ces deux paramètres, mais d'autres paramètres sont aussi liés à une forte rétention (voir le **tableau 1**).

De la même manière, certaines catégories sont plus facilement perdues. Chez le Xénope, les gènes de réparation de l'ADN, les gènes du Complexe Majeur d'Histocompatibilité ainsi que les gènes des récepteurs olfactifs ont été perdus rapidement après l'événement de WGD (Session *et al.*, 2016).

Ces propriétés sont parfois liées : ainsi, il a été montré que les gènes qui sont plus exprimés ont aussi tendance à avoir des taux d'évolution plus faibles (deux paramètres corrélés à une forte rétention) (Semon and Wolfe, 2008). Autre exemple : un contenu en GC élevé corrèle avec un taux de conversion génique élevée (McGrath, Gout, Johri, *et al.*, 2014). De même, les catégories fonctionnelles ne sont pas exclusives, puisqu'une kinase peut également intervenir dans la régulation du cycle cellulaire. Ainsi, il n'est pas toujours évident de démêler les facteurs expliquant une rétention plus élevée d'un gène après un événement de WGD.

En raison de leurs propriétés biologiques, il est fréquent qu'au cours de multiples événements de WGD, certains gènes aient une propension à être conservés après chaque événement. Cela a en particulier été montré pour la succession 1R/2R-Ts3R-Ss4R chez la Truite (Berthelot *et al.*, 2014) ou chez les trois WGD dans la lignée de la Paramécie (McGrath, Gout, Johri, *et al.*, 2014). Lorsque ce n'est pas le cas, par exemple chez le Saumon, les auteurs emploient le terme "*surprisingly*" pour présenter leurs conclusions, signe que cette tendance est régulièrement observée (Lien *et al.*, 2016). Ainsi, l'histoire évolutive du gène, et les événements de duplication précédents, semblent être corrélés au taux de rétention à un événement de WGD donné. Toutefois, il ne faudrait pas conclure que la rétention à une WGD précédente est un paramètre qui prédit nécessairement la rétention à la WGD suivante.

Table 1 - Propriétés des gènes préférentiellement retenus après une WGD.

Propriété de la séquence, du gène ou de la protéine corrélée à une rétention plus élevée que la moyenne après une WGD		Références
Expression élevée		(Aury *et al.*, 2006; Gout *et al.*, 2010; Session *et al.*, 2016)
Catégories fonctionnelles	Facteurs de transcription	(Session *et al.*, 2016; Voldoire *et al.*, 2017)
	Protéines ribosomales	(McGrath, Gout, Johri, *et al.*, 2014)
	Kinases	(Brunet *et al.*, 2016; Session *et al.*, 2016)
	Composants de voies de signalisation du développement (TGFβ, Wnt, Hedgehog)	(Session *et al.*, 2016)
	Protéines de régulation du cycle cellulaire	(Session *et al.*, 2016)
	Rab GTPases	(Bright *et al.*, 2017)
Eléments non-codants conservés (CNE) *cis*-régulateurs		(He and Zhang, 2005; Session *et al.*, 2016)
Nombre de sites de phosphorylation élevé		(Amoutzias *et al.*, 2010)
Longueur du gène et nombre de domaines fonctionnels élevés		(He and Zhang, 2005; Session *et al.*, 2016; Guo, 2017)
Nombre d'interactions protéiques élevé		(Wu and Qi, 2010)
Contenu en GC élevé		(McGrath, Gout, Johri, *et al.*, 2014)
Taux de conversion génique élevé		(McGrath, Gout, Johri, *et al.*, 2014)
Faible taux d'évolution		(Chapman *et al.*, 2006; Semon and Wolfe, 2008)
"Essentialité" forte (la délétion du gène a un effet létal ou provoque une infertilité) – valable pour les Mammifères, pas pour la Levure		(Makino *et al.*, 2009)

iii. Les taux de rétention après une WGD sont souvent inverses de ceux des SSD

Des études menées chez *Arabidopsis* ou chez la Levure ont montré que les catégories fonctionnelles qui étaient plus retenues après une WGD étaient à l'inverse *moins* retenues après un événement de SSD – facteurs de transcription notamment (Maere *et al.*, 2005; Davis and Petrov, 2005; Hakes *et al.*, 2007; Wapinski *et al.*, 2007). De plus, les protéines des paralogues issus d'une SSD ont en moyenne moins d'interactions que celles issues d'une WGD (Hakes *et al.*, 2007).

Ces observations sont compatibles avec l'hypothèse de l'équilibre de dosage. Une WGD conserve la stœchiométrie ; à l'inverse, après une SSD, les quantités respectives ne sont pas conservées et la rétention de la mutation, en particulier si elle affecte un gène de régulation, est délétère. Les SSD affectent en particulier les gènes qui ont des effets délétères plus faibles que la moyenne (Conant and Wolfe, 2008).

B. La diversification génétique a parfois des conséquences sur la diversification phénotypique

Certains auteurs ont suggéré que la rétention préférentielle de gènes de régulation (facteurs de transcription, gènes de contrôle du cycle cellulaire…) après un événement de WGD est un des facteurs qui permet l'apparition de nouvelles fonctions (Freeling and Thomas, 2006; Glasauer and Neuhauss, 2014). Les "modules fonctionnels" conservés après une WGD seraient les précurseurs de réseaux de gènes co-adaptés et formeraient une unité nouvelle soumise à la sélection. Ils pourraient également être à l'origine d'une complexification morphologique (Freeling and Thomas, 2006), ayant par exemple permis la formation de types cellulaires nouveaux, comme les cellules des crêtes neurales à la base des Vertébrés, après la 1R/2R (Green, Simoes-Costa, *et al.*, 2015). Dans cette partie nous allons nous intéresser aux conséquences phénotypiques des événements génétiques que nous avons abordés plus haut.

1. Les WGD ont un impact possible sur la spéciation

La présence de duplications de génomes à la base de multiples radiations chez les Eucaryotes (plus de 25 000 espèces de Téléostéens, plus de 350 000 Angiospermes) a conduit certains auteurs à voir un lien causal entre les WGD, la diversification morphologique et le succès évolutif (Van de Peer *et al.*, 2009).

a. La polyploïdisation peut entraîner des variations de fitness sur des temps évolutifs courts,..

Les conséquences phénotypiques de la polyploïdisation ont été envisagées dès 1911, presque 60 ans avant les travaux de Susumu Ohno (Taylor and Raes, 2004). Kuwada concluait déjà que le Maïs était un ancien tétraploïde et que la production de "races innombrables" de Maïs était reliée à ces événements de "duplications chromosomiques". En 1915, Tischler observa, chez des plantes apparentées, que des variations du nombre de chromosomes induisaient des différences morphologiques. Des études chez la Levure ont montré que par une augmentation du dosage de tous les gènes d'une voie métabolique, une WGD est susceptible d'induire un gain de fitness dans des conditions environnementales qui requièrent des flux métaboliques élevés, par exemple dans un milieu riche en glucose (van Hoek and Hogeweg, 2009).

L'allopolyploïdisation peut également parfois conduire à de nouvelles espèces invasives prenant le pas sur les espèces indigènes : c'est le cas de la Spartine *Spartina anglica*, sur les côtes bretonnes (Baumel *et al.*, 2001). De même, le Xénope *X. laevis*, allotétraploïde récent, est lui fertile et invasif dans les habitats anthropisés. Il tolère en particulier mieux la sécheresse, le froid et la privation alimentaire que l'espèce apparentée diploïde, *Silurana tropicalis* (Van de Peer *et al.*, 2009).

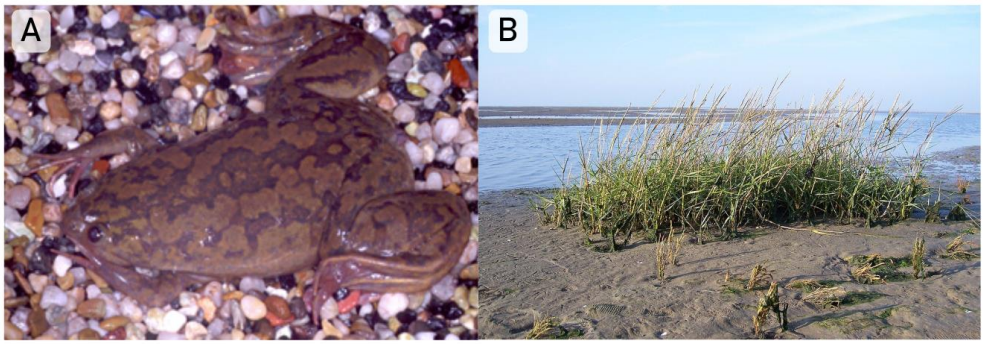


Figure 38 – Exemples d'organismes polyploïdes.

Le Xénope *X. laevis* et la Spartine *Spartina anglica* sont deux espèces issues d'une allopolyploïdisation récente. *Clichés : Wikimedia Commons.*

Les bases génétiques de cette adaptation à des nouveaux milieux sont mal connues et de nombreux modèles existent (Chen, 2010). Certains auteurs suggèrent que des copies redondantes des gènes peuvent atténuer les effets délétères des mutations et conférer ainsi à l'espèce une "résistance mutationnelle" (sic) (Otto and Whitton, 2000). D'autres auteurs suggèrent que ce sont les variations d'expression entre paralogues, combinées aux variations génétiques qui suivent l'événement de polyploïdisation (de nombreux fragments d'ADN étant gagnés puis perdus), ainsi qu'aux variations

épigénétiques (changements des profils de méthylation de l'ADN notamment), qui permettent une adaptation plus rapide aux nouveaux environnements (Osborn *et al.*, 2003).

b. ... et être à l'origine de spéciation sympatrique par des pertes de gènes différentielles

"Loss is nothing else but change, and change is Nature's delight" — Marcus Aurelius, AD 121–180

La polyploïdisation peut induire un isolement reproducteur, et à terme, une spéciation, par des mécanismes liés aux pertes de gènes différentielles entre deux populations.

Considérons deux gènes X' et X'' issus de la duplication d'un gène ancestral X. Si deux espèces A et B ont perdu la même copie (par exemple, X'), et ont conservé X'', les deux paralogues aujourd'hui conservés sont orthologues 1 à 1. Si X' a été perdu dans l'ancêtre commun à A et B avant la spéciation, on parle de "résolution ancestrale" ; s'il a été perdu indépendamment dans chacune des lignées, on parle de "résolution parallèle" de la duplication. À l'inverse, si l'espèce A a conservé uniquement le gène X', et l'espèce B le gène X'', les deux gènes X' et X'' sont des paralogues 1 à 1 et on parle de "résolution divergente", de "résolution différentielle" ou encore de "perte de gène réciproque" (*reciprocal gene loss*, ou RGL) (McGrath, Gout, Johri, *et al.*, 2014).

La RGL est un mécanisme permettant un isolement reproducteur rapide entre des populations différentes. Quand deux individus de deux populations différentes se reproduisent, le gène (par exemple, X') est manquant dans 1/4 des gamètes de F1 et 1/16 des gamètes de F2. Si chaque population perd des gènes à un taux relativement élevé, le processus de RGL mène rapidement à un isolement reproducteur total entre les populations puisqu'une partie significative des individus de F2 ont une valeur sélective faible (Lynch and Force, 2000). Sans en arriver à la perte complète du gène, la sub-fonctionnalisation seule peut elle aussi induire un isolement reproducteur entre deux populations (Volff, 2005). On note que ces processus pouvant mener à un isolement reproducteur entrent en jeu dans le cas d'une "simple" duplication locale (cela a été montré chez la Drosophile (Masly *et al.*, 2006) et le Riz (Yamagata *et al.*, 2010). Dans le cas d'une WGD, tous les loci étant concernés, l'isolement reproducteur est plus probable.

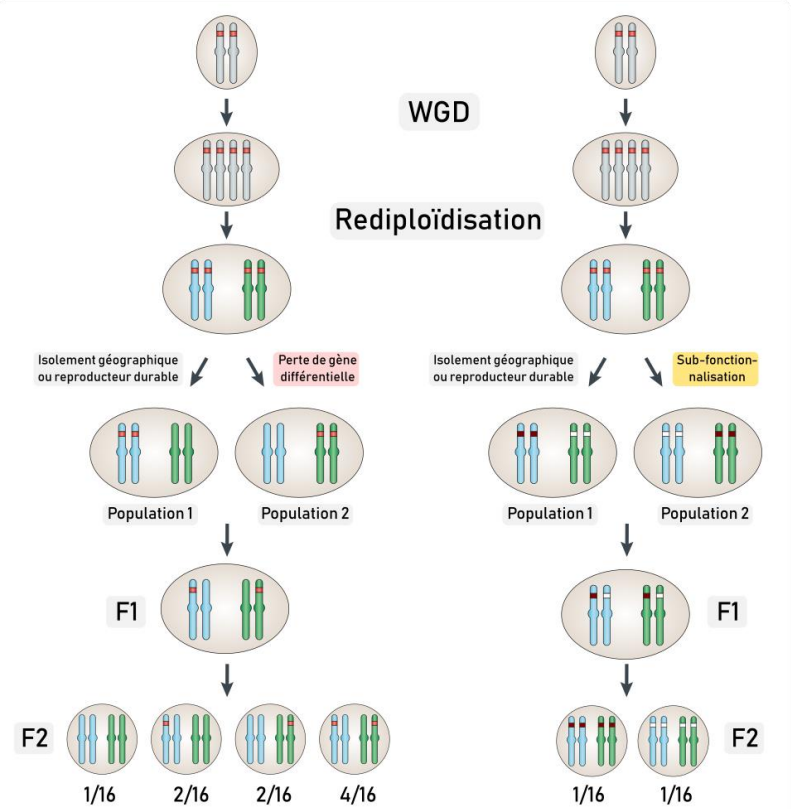


Figure 39- Isolement reproducteur suite à une duplication de génome.

L'isolement reproducteur peut avoir lieu suite à une perte différentielle de gène (à gauche) ou suite à une sub-fonctionnalisation différentielle (à droite) dans les deux populations isolées. Lorsque le contact entre les deux populations est de nouveau établi, une partie non négligeable des individus de F2 a une fitness plus faible. Par exemple, dans le cas d'une perte différentielle, 9/16 individus de F2 ont 0, 1 ou 2 copies du gène ancestral au lieu de 4. Dans le cas d'une sub-fonctionnalisation, 2/16 des individus de F2 ont une seule copie spécialisée et pas l'autre. À terme, ces mécanismes peuvent conduire à un isolement reproducteur et à une spéciation puisque la descendance entre les deux populations est moins viable qu'au sein de chacune des populations. *Adapté de van de Peer et al. (2009).*

Entre le Poisson-zèbre *D. rerio* et le fugu *T. rubripes*, il a été montré que la RGL a eu lieu à hauteur de 8% des gènes issus de la WGD, contribuant probablement, ancestralement, à un isolement reproducteur entre les ancêtres de ces deux lignées (Volff, 2005; Semon and Wolfe, 2008). Ainsi, ce n'est pas seulement la perte qui est importante dans le processus de spéciation, mais la perte *différentielle*, rendant paradoxalement le processus de perte d'un gène au cœur des facteurs à l'origine de la divergence entre lignées (Albalat and Cañestro, 2016).

2. Les WGD sont parfois corrélées à la diversification phénotypique d'un groupe

Imaginons que la séparation des sub-populations présentées plus haut et la RGL se poursuivent au cours du temps : le processus de spéciation peut continuer, produisant ainsi une augmentation du nombre d'espèces à partir d'une espèce ancestrale unique. C'est la définition d'une radiation évolutive. La radiation peut de plus être adaptative si les nouvelles espèces sont adaptées à différentes niches écologiques.

Les WGD, en modifiant les réseaux de régulation de gènes, sont à l'origine de néo-fonctionnalisations qui peuvent induire l'apparition de nouveaux traits phénotypiques et d'une complexification morphologique (Ohno, 1970; Freeling and Thomas, 2006). En particulier, cela a été suggéré pour les 15 espèces du complexe *aurelia* chez la Paramécie (Aury *et al.*, 2006), pour la radiation adaptative des 350 000 Angiospermes (Van de Peer *et al.*, 2009; Tank *et al.*, 2015) ainsi que pour celle des Téléostéens, qui comptent presque la moitié des quelque 60 000 espèces de Vertébrés (Amores *et al.*, 1998; Meyer and Van de Peer, 2005). De même, chez les Salmonidés, il a été suggéré que la Ss4R a précédé la transition à l'anadromie (retour en eau douce pour la reproduction) (Alexandrou *et al.*, 2013).

3. La corrélation WGD/diversification phénotypique n'est pas toujours vérifiée, ou pas immédiatement

On l'a vu, duplication de génome et radiation évolutive (voire radiation adaptative) sont parfois liées, à tel point que certains auteurs ont vu un rôle causal dans cette association, la WGD *induisant* une radiation évolutive. La réalité est un peu plus complexe et cette association, et *a fortiori* le lien causal, bien qu'intellectuellement séduisants, doivent être relativisés. La littérature abonde aujourd'hui de contre-exemples ou de nuances sur ces notions, pour diverses raisons que nous allons maintenant aborder.

a. Les WGD ne sont pas nécessairement associées à une augmentation du taux de spéciation

Chez les Métazoaires, quelques événements de duplication de génome ont été décrits ailleurs que chez les Vertébrés. Aucun de ces événements n'a été associé à une radiation, ni même à une augmentation du taux de spéciation (Schwager *et al.*, 2017).

b. La biodiversité actuelle n'est qu'un instantané de la paléo-biodiversité

Lorsqu'on dit, par exemple, que "les poissons Téléostéens, avec presque 30 000 espèces, sont plus divers que leurs plus proches parents, les Holostéens, qui n'en comptent que 8 et ont été qualifiés de "fossiles vivants" par Darwin en raison de leur ressemblance phénotypique avec des fossiles", c'est partiellement vrai : il manque la précision "*actuellement* plus divers". En effet, pour comparer la diversité des deux groupes, il faut prendre en compte les variations de l'abondance des différents taxons au cours des 300 derniers millions d'années, depuis l'événement de WGD à la base des Téléostéens. Et l'étude de cette variation indique que l'emploi du terme "actuel" n'est pas superflu.

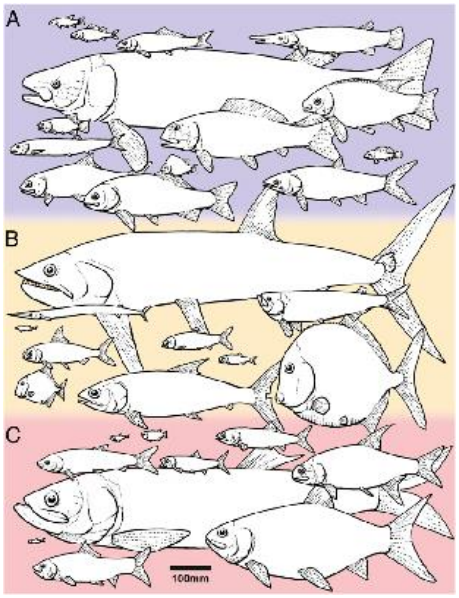


Figure 40 - Diversité comparée des Holostéens et des Téléostéens au Permo-Trias.

Une étude récente a montré qu'il n'y a pas, au Permo-Trias (~300-200 Ma), de différence de diversité morphologique entre les Holostéens, groupe frère des Téléostéens (A), le groupe-tronc des Téléostéens (B) et le groupe-couronne des Téléostéens (C). Ainsi, alors qu'actuellement les Holostéens représentent seulement 8 espèces et sont considérés comme des « fossiles vivants », il n'en a pas toujours été de même. *D'après Clarke et al. (2016).*

Une étude de 2016 montre en effet que les premiers Téléostéens (après la WGD) n'ont pas montré d'évolution phénotypique plus importante que les paléo-Holostéens (n'ayant pas eu de WGD), l'inverse étant parfois même vrai (Clarke *et al.*, 2016). Les auteurs montrent également une hétérogénéité, au cours des temps géologiques, des taux de spéciation au sein même du taxon des Téléostéens, soulignant que parler de "la" radiation des Téléostéens doit être envisagé avec précaution. Ainsi, une WGD corrélant aujourd'hui avec un nombre élevé d'espèces ne signifie pas que le groupe qu'elle a marqué a toujours été plus divers par rapport aux taxons apparentés.

c. La diversification peut être graduelle après une WGD

À un niveau taxonomique plus élevé, à l'échelle des Vertébrés, les duplications de génomes ont été elles aussi associées à des innovations anatomiques, telles que les cellules des crêtes neurales et leurs dérivés (Green and Bronner, 2013; Green, Simoes-Costa, *et al.*, 2015). Cependant, les données paléontologiques suggèrent que la diversification du groupe n'a pas été soudaine, mais plutôt graduelle (Donoghue and Purnell, 2005). De même que pour les Téléostéens, les auteurs soulignent : "*analyses based on* ***extant diversity alone*** *reveal an apparent pattern of bursts in character evolution and morphological complexity that is misleading [...]. The pattern is a consequence of the incomplete sampling of taxonomic and morphological diversity*".

d. Une WGD peut induire une radiation sans diversification phénotypique

De même, l'association parfois faite entre WGD et diversification phénotypique peut être fallacieuse. Ainsi, à l'inverse de ce qui a pu être montré chez les Angiospermes, la Levure ou même les Vertébrés, les duplications de génome chez les Paramécies, même si elles ont provoqué une radiation (pour la troisième), ne sont pas allés de pair avec une diversification phénotypique puisque les espèces sont en effet indiscernables morphologiquement alors qu'il peut y avoir jusqu'à 4 000 gènes différents entre deux espèces (Sonneborn, 1975). Il est certes possible que des innovations phénotypiques soient encore à découvrir, mais il est tout de même intéressant de constater qu'autant de différences au niveau génétique n'aient pas induit de différences phénotypiques visibles (McGrath, Gout, Doak, *et al.*, 2014).

e. L'évolution phénotypique ne succède pas toujours immédiatement à un événement de WGD ("time-lag")

Enfin, les diversifications au sein d'un groupe peuvent parfois avoir lieu longtemps après l'événement de WGD. C'est le cas par exemple chez les Salmonidés, pour qui la diversification principale en différentes familles a eu lieu il y a 40-50 millions d'années, soit quelques 40 à 50 millions d'années après la Ss4R (). Dans ce groupe, les plus forts taux de diversification corrèlent avec les périodes de climat froid à l'Ecoène/Oligocène et les auteurs suggèrent que l'anadromie aurait évolué en réponse à ces variations environnementales (Macqueen and Johnston, 2014). Dans le cas des Salmonidés, le principal paramètre à l'origine de la diversification semble donc être environnemental. Si la Ss4R a joué un rôle, il n'a probablement été qu'indirect, en potentialisant des réseaux de régulation de gènes ayant permis, par exemple, l'évolution de l'anadromie – (Alexandrou *et al.*, 2013).

Le "time-lag" entre l'événement de WGD et la diversification spécifique est également un trait caractérisant l'évolution des Angiospermes et qui a conduit certains auteurs à proposer le modèle du "time-lag entre une WGD et la radiation" (*WGD radiation lag-time*) (Schranz *et al.*, 2012). Dans ce modèle, (1) une WGD a lieu (2) et contribue à l'évolution de quelques traits essentiels du taxon T. (3) L'évolution de T continue pendant un temps *t* sans augmentation du taux de spéciation. (4) Au bout du temps *t* (quelques milliers d'années à quelques dizaines de millions d'années), un événement de dispersion majeur a lieu, déplaçant quelques sous-taxons de T vers de nouvelles niches (5) où la radiation peut avoir lieu. Les sous-taxons non adaptés à la nouvelle niche, eux, ne se diversifient pas.

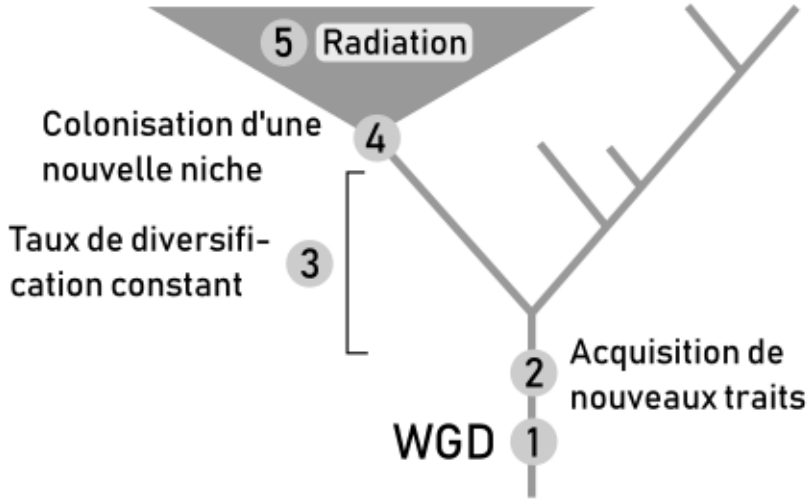


Figure 41 - Modèle de la radiation post-WGD. Dans ce modèle, la radiation est permise par les nouveaux traits acquis lors de la WGD ancestrale. Toutefois, la WGD seule ne suffit pas ; il faut un événement de colonisation d'une nouvelle niche, qui permet à la lignée de se diversifier. Ce modèle est plutôt adapté aux taxons végétaux, moins mobiles. *D'après Schranz et al. (2012).*

Ce modèle, théorique, est nécessairement imparfait dans les détails, et peut-être même inadapté à certains taxons plus mobiles, notamment chez les Téléostéens (Laurent *et al.*, 2017). Il intègre toutefois deux événements importants dans l'évolution d'une lignée : (i) la WGD initiale, qui "potentialise" certains réseaux de gènes qui sont à l'origine d'innovations évolutives, pas nécessairement très avantageuses au début, mais pas désavantageuses non plus et qui se fixent peu à peu ; (ii) des changements environnementaux brusques (biotiques ou abiotiques) qui permettent à quelques taxons seulement de coloniser des nouvelles niches, où l'innovation se révèle être hautement avantageuse et induit une radiation.

Enfin, des travaux récents montrent que ce n'est pas la WGD en tant que telle qui permet de générer la diversification au sein d'un groupe, mais plutôt la rediploïdisation post-WGD, qui permettrait également d'expliquer le "time-lag". Suite à un événement de polyploïdisation, le retour à l'état diploïde s'accompagne de réarrangements génomiques majeurs (fusion et/ou fission de chromosomes, pertes de séquences répétées). La RGL peut être considérée comme la fin du processus de diploïdisation, même si elle en est parfois dissociée (Van de Peer *et al.*, 2009). Ce ne serait donc pas la WGD en elle-même qui serait responsable du succès évolutif d'une lignée, mais la rediploïdisation consécutive. Ce time-lag a été récemment montré chez les Angiospermes (Dodsworth *et al.*, 2016) et chez les Salmonidés (Lien *et al.*, 2016; Robertson *et al.*, 2017).

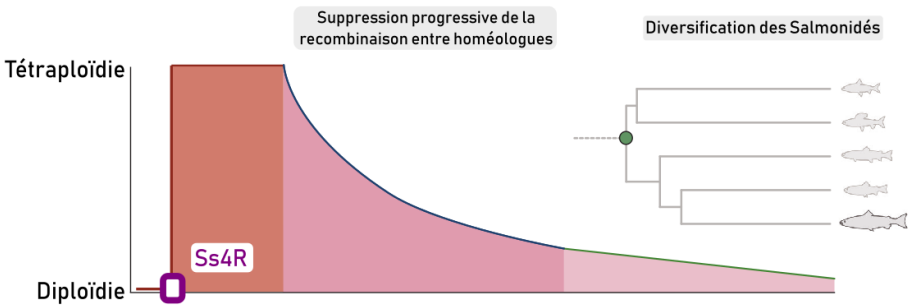


Figure 42 - Délai entre la rediploïdisation et la diversification des Salmonidés.
Après une première phase pendant laquelle les chromosomes homéologues peuvent encore recombiner, le génome se rediploïdise partiellement. La diversification de la lignée des Salmonidés n'a eu lieu qu'après (la temporalité des phénomènes est peu connue). *Adapté de Lien et al. (2016).*

f. Conclusion : les conséquences d'une WGD sur l'évolution d'un groupe sont hautement variables et d'autres facteurs sont à envisager pour expliquer la radiation

Nous l'avons vu, il n'y a pas de règle générale entre une WGD et l'évolution d'un groupe, même si certains auteurs proposent des modèles intéressants. Comme le concluent Green et al. (2015), on ne peut parfois que constater la chose suivante : "*Thus, the timing of the acquisition of these traits* ***correlates*** *nicely with inferred instances of genome duplication, although one cannot distinguish cause from effect."* (Green, Simoes-Costa, *et al.*, 2015)

Dans le cas particulier des Téléostéens, d'autres facteurs génomiques permettent peut-être d'expliquer l'origine d'une si grande radiation. En effet, récemment, des auteurs ont tiré profit des génomes maintenant bien assemblés du Lépisosté tacheté ainsi que de la Chimère, un Chondrichthyen. Les études ont montré que les génomes des Téléostéens, par rapport à ceux des autres Vertébrés, présentent un taux d'évolution accéléré des séquences codantes (Jaillon *et al.*, 2004), un plus fort turnover des introns (Venkatesh *et al.*, 2014), la perte de petites séquences *cis*- régulatrices conservées (Lee *et al.*, 2011), une importante diversité en éléments transposables (Volff, 2005) ainsi que des blocs synténiques plus courts dus à des réarrangements chromosomiques plus nombreux (Braasch *et al.*, 2016). C'est peut-être une combinaison de tous ces facteurs qui a fait de ce groupe le taxon le plus diversifié parmi les Vertébrés actuels (Ravi and Venkatesh, 2018).

Conclusion partielle et transition

Ce deuxième chapitre introductif s'est attelé à montrer comment les duplications de génome apparaissaient, dans quelles lignées et selon quelles modalités, et également quelles sont les conséquences évolutives de ces événements. En particulier, les liens entre duplications de génome, innovation génotypique et phénotypique et diversification taxonomique ont été abordés. La première partie des résultats s'attellera à étudier les conséquences des duplications de génome ayant eu lieu chez les Vertébrés sur le répertoire des gènes de la pigmentation dans ce groupe.

Auparavant, le chapitre introductif suivant vise à présenter un modèle émergent d'organisme : le poisson-clown *Amphiprion ocellaris*. Après une présentation du cycle de vie et de la biologie de cette espèce de poisson corallien, au développement caractérisé par une métamorphose typique d'un organisme Vertébré, je présente quelques aspects de "l'état de l'art" sur ce modèle émergent, sur lequel beaucoup reste à faire. Je souligne en particulier les connaissances relatives à la pigmentation et à ses contrôles hormonaux, et comment je les ai abordées au cours de mes travaux de thèse, à travers les méthodes de RNA-Seq.

INTRODUCTION

CHAPITRE 3

La pigmentation d'un modèle émergent de poisson corallien, le poisson-clown *Amphiprion ocellaris*

"the fish is a fish… is a frog… is a chicken… is a mouse"

(Kimmel, 1989)

I. Le poisson-clown *Amphiprion ocellaris*, une espèce modèle émergente de poisson corallien

1. *A. ocellaris* est une espèce de poisson corallien

a. Les récifs coralliens sont des écosystèmes…

Les récifs coralliens bordent les côtes de 109 pays du monde où quelque 450 millions de personnes habitent (Pandolfi *et al.*, 2011). Ils attirent chaque année des dizaines de millions de touristes souhaitant profiter de leurs atouts : des eaux chaudes, claires, et à la biodiversité marine sans commune mesure. L'attractivité des récifs est à l'origine du développement massif du tourisme de certaines régions du monde ayant permis leur essor économique.

Cet écosystème représente 0.1 à 0.2 % de la surface des océans et concentre à lui seul plus d'un tiers de leur biodiversité (Spalding and Grenfell, 1997). Parmi toutes les espèces récifales, les poissons coralliens possèdent des rôles écologiques majeurs. Certaines espèces permettent en particulier de réguler la prolifération des algues qui empêcheraient les coraux de croître (Bellwood *et al.*, 2004; Mumby, 2006). Outre leur importance écologique, les poissons coralliens sont également des espèces majeures pour l'Homme. Ils sont en effet la principale source de protéines animales pour des dizaines de millions de personnes et génèrent une économie considérable (Salvat, 1992). Au travers de leur rôle de barrière, ils offrent une protection des littoraux contre les tempêtes (Costanza *et al.*, 2014). Selon certaines estimations, 36 milliards de dollars seraient rapportés chaque année par les récifs et les services écosystémiques qu'ils offrent (Brumbaugh, 2017).

Enfin, les poissons, requins, raies et tortues, ont une influence culturelle considérable, non seulement pour les populations locales, mais également dans le monde entier, et chacun de nous, même sans être nécessairement déjà allé sur une île corallienne, a des images de films, documentaires, romans, musiques, en rapport avec ces contrées exotiques.

b. … où les poissons abondent pendant l'essentiel de leur cycle de vie…

Les poissons coralliens, comme nous l'avons vu dans le premier chapitre introductif, sont très divers en termes de pigmentation. Ce n'est toutefois qu'une des multiples facettes d'une diversité à plusieurs niveaux. On note avant tout que la définition de "poisson corallien" est vague et quelquefois discutée (Bellwood, 1998). Le seul consensus qui émerge est une définition tautologique : "est considéré comme corallien un poisson qui, à au moins une étape de son cycle de vie, interagit avec un récif corallien." (Bellwood, 1998; Robertson, 1998).

Les poissons coralliens sont très divers taxonomiquement et contiennent la plus grande diversité spécifique de poissons sur Terre (Harmelin-Vivien, 2002). Jusqu'à 1 000 espèces co-existent même parfois à un seul endroit (Bellwood *et al.*, 2005). Les Téléostéens représentent l'essentiel des "poissons" trouvés dans les récifs (Price *et al.*, 2015). L'évolution de ce groupe, apparu il y a environ 300 millions d'années, a été marquée par l'émergence des coraux actuels (Scléractiniaires) au Trias (250-200 Ma). Des travaux récents montrent que les récifs ont eu un rôle dans la diversification de certaines lignées de Téléostéens, et qu'ils sont un moteur de la cladogenèse (Alfaro *et al.*, 2007; Cowman and Bellwood, 2011). De plus, il a été suggéré qu'au cours du Paléozoïque (du Cambrien à la fin du Permien, soit 542-252 Ma) le nombre de genres marins apparus a été plus rapide dans les environnements coralliens qu'ailleurs (Kiessling *et al.*, 2010). Les récifs auraient de plus contribué à exporter leur diversité aux environnements non-coralliens (Kiessling *et al.*, 2010).

Quelques taxons, tous de l'ordre des Perciformes, sont en particulier représentés parmi les poissons coralliens (Choat and Bellwood, 1991; Wainwright and Bellwood, 2002) :

- La famille des Gobiidae (> 1 500 espèces), dans le sous-ordre des Gobioidei
- Les Labridae (> 450 espèces) et Pomacentridae (> 300 espèces), au sein des Labroidei
- Les Chaetodontidae et Pomacanthidae, au sein des Percoidei (> 200 espèces)
- Les Acanthuridae, Siganidae, Zanclidae, au sein des Acanthuroidei > 70 espèces)

L'essentiel des poissons de chacun de ces groupes sont coralliens, soit de manière facultative (le Chaetodon *Chaetodon marleyi*), soit exclusive (le Pomacanthidé *Pomacanthus imperator*).

Par ailleurs, les poissons coralliens sont très divers en termes d'écologie et de morphologie. Les stratégies de prise alimentaire et la place dans le réseau trophique diffèrent selon les espèces, avec des espèces détritivores (certains Poissons-chirurgiens), planctivores (Fusiliers), carnivores (Murènes), omnivores (poissons-clowns), herbivores et/ou corallivores (Poissons-perroquets). Pour une famille donnée, la diversité morphologique et trophique ("morphospace") est également plus importante que dans les environnements non-récifaux. Une étude systématique chez les Labridae a ainsi étudié l'ensemble des valeurs prises par de multiples traits liés à la forme du crâne : hauteur, longueur et angle de la mâchoire... Dans une analyse en composante principale, qui représente l'ensemble des variations de tous ces paramètres dans un espace en deux dimensions, l'aire ainsi obtenue des espèces récifales est plus vaste que celle des espèces non-récifales (Price *et al.*, 2011). Le "morphospace" récifal, c'est-à-dire l'ensemble des variations possibles des phénotypes en milieu récifal, est donc plus étendu qu'en milieu non-récifal.

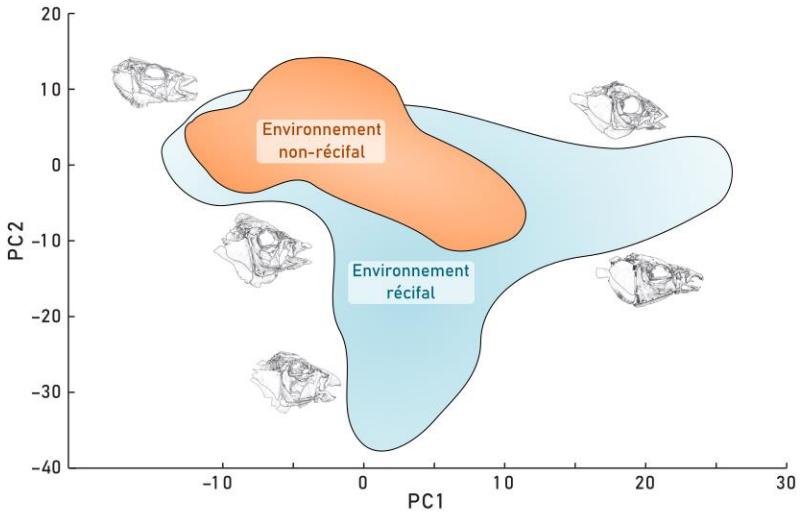


Figure 43 - Diversité morphologique chez les Labridae récifaux et non-récifaux.

Plusieurs paramètres morpho-anatomiques ont été mesurés sur des crânes de diverses espèces de labres coralliens et non-coralliens. Dans une analyse en composante principale, qui représente l'ensemble des variations de tous ces paramètres dans un espace en deux dimensions, l'aire ainsi obtenue pour les espèces récifales est plus vaste que celle des espèces non-récifal. Le "morphospace" récifal, c'est-à-dire l'ensemble des variations possibles des phénotypes en milieu récifal, est donc plus étendu qu'en milieu non-récifal. Les dessins représentent la forme approximative des crânes à l'endroit du morphospace correspondant. *Adapté de Price et al. (2011).*

Nous avons vu en première partie que la diversité *taxonomique* de ce groupe est probablement due à des caractères liés à la sélection sexuelle ou à la communication intra- ou interspécifique. Cependant, la diversité *morphologique* qui l'accompagne est vraisemblablement permise par la forte productivité de l'écosystème corallien, à même de supporter une forte densité d'individus de nombreuses espèces (Fraser and Currie, 1996; Genin *et al.*, 2002). De plus, le récif est une structure topologiquement complexe, riche en micro-habitats colocalisés, qui permet la co-existence de niches différentes propices à accueillir de manière stable de nombreuses espèces (MacArthur and Levins, 1964; Schoener, 1974).

c. ... mais pas pendant la phase larvaire

Nous avons jusqu'ici discuté des poissons coralliens à l'âge adulte. Toutefois, une autre spécificité de ces organismes est leur cycle de développement majoritairement biphasique, avec une phase récifale (l'âge adulte) et une phase pélagique (la phase larvaire) (Leis and McCormick, 2002; Cowen and Sponaugle, 2009; Clobert *et al.*, 2012).

Les modalités de la reproduction varient selon les espèces de poissons coralliens. Certaines espèces se reproduisent lors de regroupements massifs dans la colonne d'eau qui sont dépendants du cycle lunaire : c'est le cas par exemple du poisson chirurgien *Acanthurus triostegus*, un pondeur pélagique (*pelagic spawner*) (Y. Sadovy de Mitcheson and P.L. Colin, 2012). À l'inverse, chez certaines espèces qualifiées de "*substrate spawners*", comme les poissons demoiselles, la femelle pond ses œufs sur le substrat avant fécondation par un seul mâle (Lieske and Myers, 2002).

Si les œufs sont pélagiques, l'embryon est directement dans la colonne d'eau et la larve éclot au bout de quelques dizaines d'heures. S'ils sont fixés au substrat, l'alevin éclot après quelques jours. Dans les deux cas, les larves ne restent pas au niveau du récif : elles sont brassées par les courants (processus physiques) ou nagent activement vers le large (processus biologiques) et commence alors pour eux une phase dite "pélagique" (Cowen and Sponaugle, 2009; Green, Maypa, *et al.*, 2015; Brooker and Dixson, 2016). Cette phase a une durée variable selon les espèces : entre environ deux semaines chez les poissons demoiselles (Thresher *et al.*, 1989) et jusqu'à plusieurs mois – le record a été documenté à plus de 64 semaines chez un Poisson Porc-épic (Ogden and Quinn, 1984).

Au terme de cette période, les larves reviennent vers le récif et intègrent le stock de poisson récifal : cette étape est appelée le "recrutement larvaire" (Richards and Lindeman, 1987; Doherty, 2002). La terminologie de "recrutement" est ambiguë et initialement associée à un terme de pêcherie qui désigne le recrutement des juvéniles dans le stock de pêche. Ici, il s'agit d'un recrutement au sens écologique : la larve est recrutée au sein du récif, et commence alors pour elle une phase juvénile récifale, succédant à la phase pélagique. Le recrutement recouvre en réalité deux processus : la "colonisation du récif" (*reef colonization*), c'est-à-dire le changement d'habitat entre l'océan et le récif, et la phase "d'installation" (*settlement*) dans un habitat approprié au sein du récif.

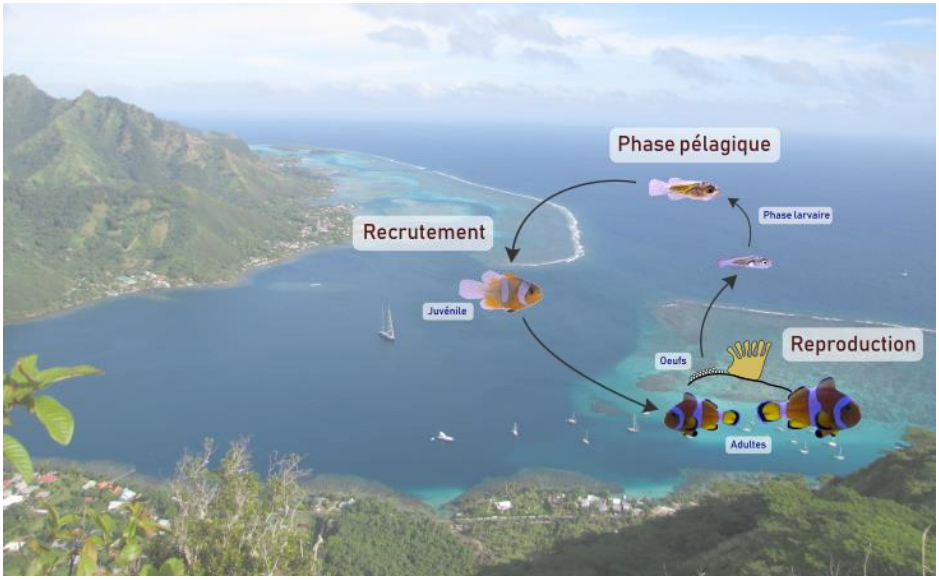


Figure 44 - Cycle de vie biphasique des poissons coralliens.

Les adultes se reproduisent au niveau du récif. Ensuite, les larves vivent une phase pélagique de durée variable avant de revenir au niveau du récif (recrutement larvaire). Cycle illustré : espèce *Amphiprion ocellaris* (ponte sur le substrat, près d'une anémone). Clichés *: Moorea et poissons adultes : T. Lorin ; larves et juvénile : N. Roux.*

Il ne faudrait toutefois pas imaginer que la phase pélagique caractérise une larve inepte, brassée par les courants et incapable d'avoir un déplacement horizontal propre (à l'instar du plancton) (Fisher *et al.*, 2000, 2005; Fisher, 2005). Les larves de poissons coralliens sont parfois capables de maintenir une nage active s'opposant au courant, voire même remontant le courant pendant plusieurs dizaines d'heures (Stobutzki and Bellwood, 1997), et peuvent rester près de leur récif d'origine (Cowen *et al.*, 2005; Irisson and Lecchini, 2008). Les signaux environnementaux (et comment ils sont perçus) permettant à la larve pélagique, "perdue" dans l'océan, de revenir vers un récif, commencent à être connus.

Une île corallienne, ou un récif, a quelques particularités pour un organisme pélagique, qui lui permettent de le distinguer de l'océan environnant. (i) Le récif est "odorant" : chimiquement, le récif est un lieu où de nombreuses molécules sont émises par les espèces qui y vivent, bien plus denses qu'en milieu pélagique ; de plus, la présence fréquente d'une île associe le récif à des composés spécifiques de feuilles mortes, de particules argileuses, qui ne sont pas présentes dans l'océan (Myrberg and Fuiman, 2002; Gerlach *et al.*, 2007) ; (ii) il est "bruyant" : mécaniquement, lorsque les vagues se brisent contre le récif, des ondes sont générées et signalent la présence d'un obstacle ; de plus, les organismes vivant dans le récif émettent de nombreux sons (Simpson, 2005; Simpson *et al.*, 2010).

Le récif est donc un "phare" qui attire et guide les larves… uniquement à condition qu'elles soient en mesure de le "voir" ! "Voir" un récif suppose que les structures sensorielles des larves, ainsi que les voies de transduction associées,

soient fonctionnelles. À quel moment du développement ces structures apparaissent-elles ? Comment sont-elles déclenchées par des signaux environnementaux ou endogènes ? Peu d'études ont été effectuées à ce propos. La première raison de cette méconnaissance physiologique de cette phase cruciale du cycle de vie est la rareté des larves en milieu pélagique, qui ne permet pas d'échantillonner assez d'individus synchrones pour une étude adaptée du développement.

Les larves capturées sont en effet souvent d'âge variable, et, pour celles de même âge, elles sont souvent à des étapes différentes de leur développement (Victor, 1986). Ainsi, la phase pélagique est de durée variable pour une même espèce, et certains individus peuvent se "retenir" de passer au stade développemental suivant, si les conditions environnementales ne sont pas favorables, indiquant une plasticité développementale dépendante de l'environnement. Par exemple, chez *Acanthurus triostegus*, des individus attrapés au moment du recrutement et remis dans des cages au large de l'île restent à l'état larvaire, à l'inverse des larves conservées dans le récif (McCormick, 1999).

Les processus physiologiques sous-jacents (expression génétique, hormones impliquées) et comment ils se traduisent phénotypiquement sont donc assez mal connus (Leis *et al.*, 2011). Des travaux récents suggèrent que la voie de signalisation des hormones thyroïdiennes pourrait être impliquée (je reviendrai par la suite sur cette voie de signalisation) et permettrait à la larve de (i) détecter les signaux environnementaux (ii) y répondre et (iii) adapter sa réponse en fonction des conditions environnementales (Holzer *et al.*, 2017).

2. *Amphiprion ocellaris* est un poisson-demoiselle du genre *Amphiprion*

a. Biologie des poissons-demoiselles, ou Pomacentridae

Après les Gobiidae (> 1 500 espèces) et les Labridae (> 600 espèces), les Pomacentridae sont la troisième famille de poissons coralliens la plus représentée, avec 392 espèces (Froese and Pauly, 2018). Leur nom latin provient de *Poma-* = opercule et de *kentron* = épine, par allusion à l'opercule qui est denté. Leur nom vernaculaire a quant à lui une origine moins claire, mais il est probablement lié au fait que ces poissons sont petits, souvent colorés, et d'apparence plaisante (Frédérich and Parmentier, 2016).

Presque exclusivement récifaux, les poissons de cette famille appartiennent à 5 sous-familles principales : les Stegastinae, les Lepidozyginae (une seule espèce), les Chrominae, les Abudefdufinae et les Pomacentrinae (Cooper and Westneat, 2009). Ces poissons sont soit des planctivores, soit des herbivores, soit des omnivores (Frédérich *et al.*, 2009).

Ce clade n'est présent que depuis 50 millions d'années dans le registre fossile, et le nombre d'espèces élevé dans ce groupe a pendant longtemps été associé à un événement unique de radiation (Schluter, 2000). Toutefois, en 2013, une étude a montré que cette diversité était en fait due à plusieurs petites radiations indépendantes associées à la colonisation de nouveaux milieux (émergence d'îles océaniques, fermeture de l'isthme de Panama…) (Frédérich *et al.*, 2013).

La distribution géographique (biogéographie) du groupe est vaste : l'essentiel des espèces (355) vivent dans la zone Indo-Pacifique et le reste dans l'Atlantique (Spalding *et al.*, 2007). Une seule espèce vit en Méditerranée : *Chromis viridis*, la "castagnole". Certaines espèces, telles que cette dernière, sont présentes dans toute l'aire de répartition de la famille ; à l'inverse, d'autres sont endémiques, à l'image de *Abudefduf abdominalis*, dans l'archipel d'Hawaï.

Cette famille a plusieurs caractéristiques propres. En premier lieu, de nombreuses espèces de Pomacentridae, telles que *Stegastes nigricans*, sont "cultivatrices" : solitaires, les individus défendent farouchement une zone contre les conspécifiques et les individus des autres espèces afin de s'assurer la récolte de tapis d'algues (Feitosa *et al.*, 2012). Ce comportement est unique chez les poissons coralliens. Par ailleurs, les Pomacentridae se distinguent par leurs capacités à émettre des signaux acoustiques ("bruits"), impliqués dans une diversité de contextes biologiques : parades nuptiales, reproduction, dissuasion, reconnaissance conspécifique et nettoyage du site de ponte (Parmentier *et al.*, 2010; Colleye and Parmentier, 2012).

Les espèces non-territoriales sont souvent grégaires et forment des groupes de 10 à 50 individus avec des structures sociales variables. Au sein du genre *Amphiprion,* les espèces sont strictement monogames jusqu'à la mort d'un des deux individus reproducteurs (Fautin and Allen, 1992). Chez *Dascyllus aruanus,* le groupe est structuré en harem avec un ou deux mâles reproducteurs et de nombreuses femelles (Fricke and Holzberg, 1974). Chez d'autres espèces du genre *Amblyglyphidodon*, les couples se forment et se défont au gré des saisons de reproduction. Ces structures sociales sont associées à différentes stratégies hermaphrodites. Chez les *Amphiprion*, l'hermaphrodisme est protandre, les mâles étant d'abord mâles puis femelles ; à l'inverse, chez les *Dascyllus*, l'hermaphrodisme est protogyne (femelle puis mâle). La spécificité de ce groupe est que les femelles pondent leurs œufs sur le substrat. Un seul mâle féconde alors les œufs et s'en occupe (soin parental mâle) jusqu'à l'éclosion en les oxygénant régulièrement et en nettoyant les œufs morts (Allen, 1991; DeAngelis and Rhodes, 2016).

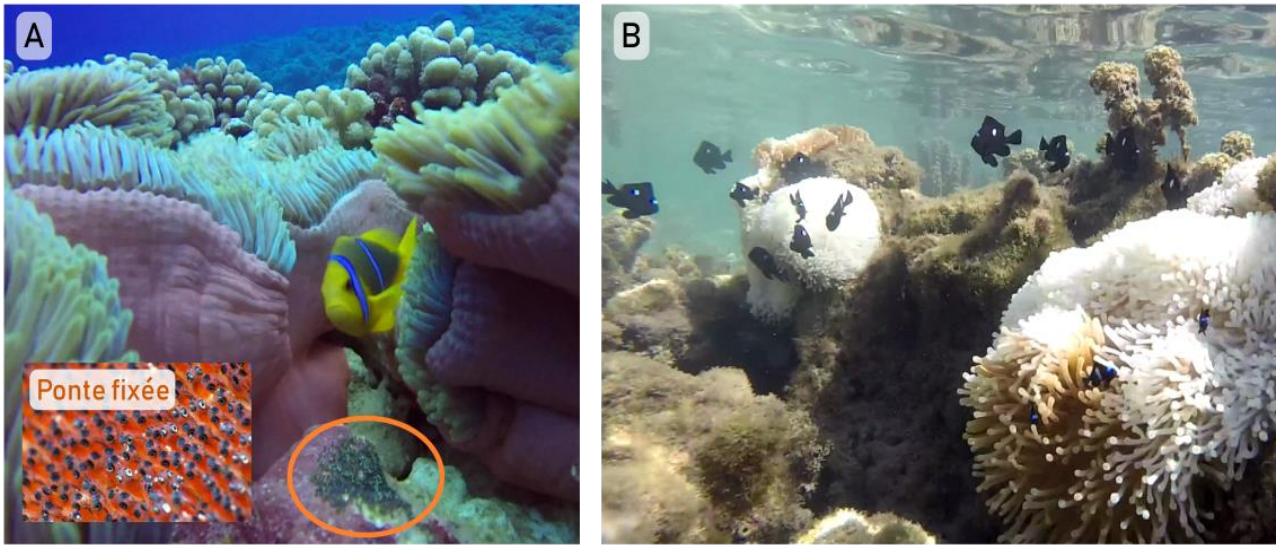


Figure 45 – Deux espèces de poissons demoiselles.

Chez les poissons demoiselles, la ponte est fixée sur le substrat. Certaines espèces qui vivent en symbiose avec des anémones fixent leur ponte non loin de celle-ci (exemple pour *Amphiprion chrysopterus*, à gauche). Chez les *Amphiprioninae*, il y a un seul couple reproducteur par anémone et l'hermaphrodisme est protandre. Chez *Dascyllus trimaculatus* (à droite), l'hermaphorisme est protogyne et les mâles se partagent un harem de femelles. Clichés *: organismes : T. Lorin ; détail de la ponte : Germain Boussarie.*

À l'instar de la quasi-majorité des poissons coralliens, le cycle de vie des Pomacentridae est biphasique (Leis and McCormick, 2002). Seules trois espèces de Pomacentridae ne présentent pas de phase pélagique, le soin parental durant jusqu'à ce que la descendance puisse échapper à la prédation (Bernardi, 2011; Bernardi *et al.*, 2017).

Chez les Pomacentridae, les adultes vivent près du fond (démersaux), et les larves, pélagiques, sont capables au moment du recrutement de changer directement d'habitat en allant investir l'habitat adulte (Leis and McCormick, 2002). Cette transition écologique rapide entre la larve pélagique et les juvéniles démersaux est associée à des faibles changements morphologiques mais à des modifications rapides de la pigmentation. Cette transition d'histoire de vie, "période critique" (sic) (Thorisson, 1994) du cycle de développement des poissons coralliens, est ainsi qualifiée de métamorphose (McCormick *et al.*, 2002a). Je la détaillerai plus loin, dans le cas du poisson-clown *Amphiprion ocellaris*.

b. Les poissons-demoiselles incluent les espèces du genre *Amphiprion*, ou "poissons-clowns"

i. le genre Amphiprion *contient une trentaine d'espèces…*

Au sein des Pomacentridae, les Amphiprioninae correspondent aux genres *Amphiprion* et *Premnas* et constituent les "poissons-clowns". Leur nom latin provient de *amphi-* (gr.) = des deux côtés et de *prion* = scie, par allusion à leurs opercules finement dentés. Le nom vernaculaire, "clown", provient de leurs couleurs bariolées, souvent constituées d'un fond orange, jaune ou noir et de bandes blanches verticales (Lieske and Myers, 2002). Des exceptions existent : *A. akallopisos* et *A. sandaracinos* ont une bande blanche dorsale horizontale.

Il existe une trentaine d'espèces au sein des *Amphiprioninae* réparties dans tout l'Indo-Pacifique (et absentes de l'Atlantique). La diversité maximale est présente au niveau du Triangle de Corail (Malaisie, Philippines, Indonésie) et diminue à mesure que l'on s'écarte géographiquement de cette région : il n'y a qu'une seule espèce en mer Rouge (*A. bicinctus*) et une en Polynésie (*A. chrysopterus*) (Allen, 1972).

Le nombre d'espèces précis n'est pas connu dans ce groupe puisque de nombreuses zones d'hybridation existent entre espèces. Par exemple, *A. leucokranos* pourrait être un hybride entre *A. chrysopterus* et *A. sandaracinos* (Ollerton *et al.*, 2007). La phylogénie des Pomacentridae montre que le genre *Amphiprion* est le groupe ayant eu le taux de diversification le plus élevé des Pomacentridae (Frédérich *et al.*, 2013). Cette diversification s'est effectuée au travers de deux radiations indépendantes ayant probablement eu lieu à la fin du miocène (entre 5 et 10 Ma) : après une première radiation ayant eu lieu dans la zone Indo-Australienne, une deuxième diversification a eu lieu dans l'Ouest de l'Océan Indien (Litsios *et al.*, 2012; Litsios, Pearman, *et al.*, 2014).

ii. … à l'écologie différente du reste des Pomacentridae

1. Les poissons-clowns sont tous associés à des anémones, avec une spécificité variable

La relation entre les poissons-clowns et les anémones est un exemple classique de mutualisme, dans lequel le poisson-clown est protégé des tentacules urticants de l'anémone. Ce cas est si emblématique que les poissons-clowns sont également appelés "anemonefish" en Anglais ; toutefois, ils ne sont pas les seuls à vivre dans des anémones et le terme de "poisson-clown" est plus adapté (Randall and Fautin, 2002). La mise en place de la symbiose repose sur des mécanismes de reconnaissance variables selon les espèces et assez mal caractérisés.

La protection contre l'anémone est vraisemblablement liée à des propriétés du mucus qui sont encore assez mal connues (Mebs, 2009). Par contre, enlever le mucus d'un poisson avant de le remettre dans l'anémone déclenche une réaction de libération des cnidocystes de l'hôte. Le poisson-clown tire un bénéfice évident de cette relation : au sein de l'anémone, il est protégé des prédateurs. D'ailleurs, dans la nature, les poissons-clowns, qui sont globalement d'assez mauvais nageurs, même si certaines espèces comme *A. clarkii* font exception, ne survivent pas s'ils ne sont pas associés à une anémone (Colleye *et al.*, 2016). Les poissons-clowns vivent d'ailleurs plus longtemps que les autres Pomacentridae : jusqu'à 30 ans pour *A. percula*, ce qui est deux fois plus que n'importe quel autre Pomacentridae et six fois plus que l'espérance de vie moyenne d'un poisson de cette taille (Buston and García, 2007). Au cours de cette longue vie, le poisson restera dans la même anémone. Des rares cas de mobilité de poissons entre anémones sont cependant décrits (individus "switchers"), et de mobilité même de l'anémone, poussant la colonie à suivre (Mitchell, 2003, 2005).

En retour, les avantages conférés à l'anémone par cette relation sont multiples. En premier lieu, les poisson-clowns, territoriaux, défendent l'anémone contre d'éventuels prédateurs (poissons-papillons notamment). De plus, les déjections ammoniaquées des poissons constituent une source d'azote pour les zooxanthelles mutualistes des anémones : la relation de mutualisme est donc tripartite. Enfin, certains auteurs ont suggéré que les anémones occupées par des poissons avaient une plus forte probabilité de survie, étant mieux oxygénées (Szczebak *et al.*, 2013).

Toutefois, les anémones sont souvent rencontrées seules dans la nature, donc la symbiose n'est pas obligatoire pour elles. Certains auteurs ont même été jusqu'à remettre en cause le bénéfice réciproque de la relation : les poissons-

clowns peuvent en effet consommer les gamètes des anémones, diminuant leur fitness (Scott and Francisco, 2006). La relation de mutualisme pourrait donc, dans certaines conditions, tendre vers le parasitisme.

La spécificité de la relation poisson-clown/anémone est variable selon les espèces. Certaines espèces sont généralistes (*A. clarkii*), pouvant coloniser les 10 espèces d'anémone colonisées par l'ensemble des espèces du genre *Amphiprion*. À l'inverse, certaines espèces (*A. sebae*) sont spécifiques d'une seule anémone hôte (Fautin and Allen, 1997). La répartition phylogénétique des types d'associations n'est pas claire, mais, souvent, des espèces sœurs de poissons-clowns ont des spécificités très différentes, l'une étant spécialisée, l'autre généraliste. La distribution des poissons-clowns étant intimement liée à celle des anémones hôtes, certains auteurs ont suggéré que cette relation de mutualisme était à l'origine de la (des) radiation(s) observée(s) au sein de ce groupe (Heard and Hauser, 1995; Elliott *et al.*, 1999; Litsios *et al.*, 2012). Toutefois, les espèces de poisson-clowns spécialistes par rapport à leurs hôtes sont en fait des généralistes du point de vue de la niche écologique : ils occupent souvent des aires de répartition plus grandes et vivent dans des gammes de température, salinité et pH plus importantes, ce qui suggère un "trade-off" entre spécialisation et généralisme entre l'hôte et l'environnement physique (Litsios, Kostikova, *et al.*, 2014). D'autres facteurs, tels que les hybridations fréquentes entre espèces de ce groupe, pourraient également expliquer la diversification du groupe (Litsios and Salamin, 2014).

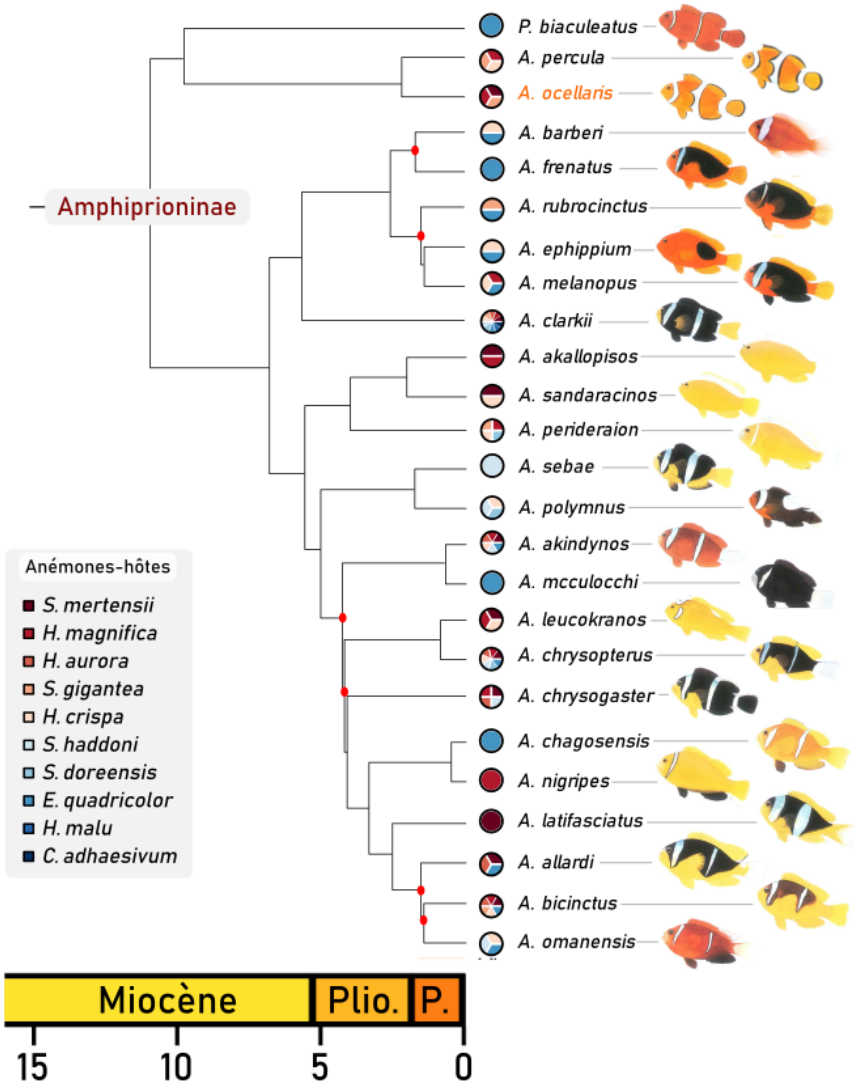


Figure 46 - Phylogénie des Amphiprioninae.

Les Amphiprioninae sont constitués du genre *Premnas* et du genre *Amphiprion*. Le genre *Premnas* ne compte qu'une espèce qui devrait en réalité s'appeler *Amphiprion*. La phylogénie est basée sur sept marqueurs nucléaires. Pour chaque espèce, le nombre d'anémones hôtes (cercle) est représenté. Le Poisson-clown *A. ocellaris* ("Némo") est indiqué (orange). Les points rouges indiquent les nœuds ayant une probabilité postérieure bayésienne inférieure à 0.95.

Arbre adapté d'après Litsios et al. (2012), Litsios, Pearman et al. (2014) et Litsios and Salamin (2014). Images de Lieske and Meyers (2009).

2. Les poissons-clowns ont une structure sociale très hiérarchisée

Au sein de chaque anémone, la colonie de poisson-clown ne contient qu'un seul couple reproducteur, qui est monogame strict, ainsi que plusieurs individus sexuellement immatures qui forment une "file d'attente" (Fricke and

Fricke, 1977; Moyer and Nakazano A, 1978; Mitchell, 2005). Le plus gros individu de l'anémone est la femelle ; le deuxième est le mâle, et tous les autres individus sont sexuellement immatures et sont de taille décroissante. Si la femelle meurt, le mâle devient une femelle, et le plus gros des juvéniles devient mâle. Si le mâle meurt, la femelle reste et le plus gros juvénile devient mâle. Plus l'anémone est grosse, plus le nombre d'individus qu'elle peut supporter est élevé (Mitchell, 2005).

Cette structure sociale est très hiérarchisée : un individu d'une taille donnée empêche les individus plus petits que lui de grossir pour atteindre sa taille (Fricke, 1979; Buston, 2003). Si tel est le cas, le belligérant est expulsé de l'anémone. De la même manière, un juvénile nouvel arrivant ne peut intégrer une nouvelle anémone que s'il est plus petit que tous les individus qui y sont déjà. Si un individu essaie d'intégrer une anémone alors qu'un individu de même taille est déjà présent, ce dernier l'expulsera (Fautin and Allen, 1992).

Chez les Amphiprioninae, la hiérarchie est maintenue par divers signaux conspécifiques, de menace et de soumission. Lorsqu'un individu secoue sa tête puis le reste du corps, cela correspond à une posture de soumission (Colleye and Parmentier, 2012). En retour, un individu agressif émet des signaux acoustiques à l'encontre d'un conspécifique. Chez *A. akallopisos*, plus l'individu est de taille importante, plus la fréquence des sons est faible et plus la durée des sons est importante (Colleye *et al.*, 2009). Il est même possible de prédire la taille d'un individu en fonction des signaux qu'il émet (Colleye *et al.*, 2011). De manière intéressante, un individu n'a pas la plus forte sensibilité pour sa propre fréquence, mais pour celle émise par les adultes, permettant probablement aux juvéniles de retrouver les adultes au sein du récif (Parmentier *et al.*, 2009). Cependant, ces signaux acoustiques n'interviennent pas au cours de la reproduction, à l'inverse des autres Pomacentridae (Colleye and Parmentier, 2012).

Les bases moléculaires du maintien de cette structure sociale sont mal connues. Si les changements de comportement après élimination de la femelle sont rapides (le mâle se met à avoir une attitude dominante et parade auprès des juvéniles), le vrai changement de sexe est lui plus long (Casas *et al.*, 2016). Chez *A. ocellaris*, le changement de sexe n'était pas visible 180 jours après enlèvement de la femelle (Iwata *et al.*, 2008), conduisant les auteurs à dire que ce sont les interactions sociales à long terme qui déterminent le changement de sexe (Iwata *et al.*, 2008). Ces interactions sociales agissent vraisemblablement sur le circuit de l'arginine vasotocine (un homologue de l'arginine vasopressine chez les Mammifères), impliqué dans le comportement reproducteur de nombreux Vertébrés : Oiseaux, Lépidosauriens, Amphibiens et poissons (De Vries and Panzica, 2006; Balment *et al.*, 2006; Iwata *et al.*, 2010). Cette voie, rapide, pourrait à terme induire l'expression de la voie de l'aromatase, enzyme transformant les androgènes en œstrogènes et exprimée dans le cerveau par le gène *cyp19a1* (Casas *et al.*, 2016). L'axe hypothalamo-hypophysaire prendrait donc ensuite le relais, provoquant une expression de l'aromatase dans les gonades, et provoquant la dégénérescence du tissu testiculaire (Abol-Munafi *et al.*, 2011; Gopurappilly *et al.*, 2013; Casas *et al.*, 2016).

3. La reproduction chez les poissons-clowns

Les espèces du genre *Amphiprion* ont une reproduction comparable à celle des autres Pomacentridae. La reconnaissance du partenaire repose sur des signaux visuels. Cela a été montré chez *A. bicinctus* : en peignant un des deux partenaires en vert (avec du vert de Bromo-crésol), le couple ne se reforme pas et l'individu est rejeté par le partenaire "non-peint" (Fricke, 1973). Les femelles pondent ensuite leurs œufs sur le substrat, non loin de l'anémone hôte. Le nombre d'œufs est variable : par exemple, chez *A. chrysopterus,* le nombre d'œufs varie de 400 à 3 500 selon la taille de la femelle, le site et la saison de ponte (Beldade *et al.*, 2012). Le développement dans l'œuf dure environ 6-8 jours, et, après éclosion, commence la phase pélagique. Au laboratoire, la photopériode influe sur le développement : en conditions de jour court (12h), la durée de la phase embryonnaire est augmentée de 2 à 3 jours par rapport à des conditions de jour long (16h) (Arvedlund, McCormick, *et al.*, 2000).

La phase pélagique dure environ entre 15 et 20 jours au cours desquels l'individu peut parcourir une distance jusqu'à 35 km (Wellington and Victor, 1989; Planes *et al.*, 2009) au terme desquels le recrutement larvaire a lieu (Thresher *et al.*, 1989). Les juvéniles sont alors attirés par une anémone, vraisemblablement par des signaux olfactifs encore mal

connus baptisés "amphikuemines" (Murata *et al.*, 1986; Arvedlund *et al.*, 1999; Arvedlund, Bundgaard, *et al.*, 2000) et captés par les récepteurs olfactifs très développés des larves (Arvedlund *et al.*, 2003). Il est également possible que la reconnaissance de l'anémone soit liée à "l'imprinting" par l'hôte au cours du développement dans l'œuf : la proximité d'une espèce d'anémone favoriserait le retour vers celle-ci (Miyagawa-Kohshima *et al.*, 2014). Enfin, les juvéniles naïfs peuvent détecter l'état de blanchissement de l'anémone et éviter les anémones blanchies (Scott and Dixson, 2016). Les juvéniles reconnaissent par ailleurs les conspécifiques par des signaux visuels (Brolund *et al.*, 2003).

Des études de connectivité marine ont montré que, bien souvent, en raison de la courte phase pélagique, les larves *d'Amphiprion* reviennent avec une forte probabilité (60 % chez *A. percula*) vers l'île dont elles sont issues, un processus appelé "auto-recrutement" (Almany *et al.*, 2007, 2017). Parfois, les larves reviennent à moins de 100 mètres du site de ponte (Jones *et al.*, 2005).

3. Spécificités de la biologie de *A. ocellaris*

a. Le poisson-clown *Amphiprion ocellaris* : une icône en aquariophilie et dans l'imaginaire collectif

Parmi toutes les espèces de poissons-clowns, une est en particulier devenue célèbre depuis une quinzaine d'années : *Amphiprion ocellaris*, ou "Némo". Dans le film *Le Monde de Némo* (Disney-Pixar), paru en 2003, un père poisson-clown appelé Marin part à la recherche de son fils Némo, capturé par des pêcheurs et revendu à un particulier, qui le met dans son aquarium. (On note à cet égard que, en réalité, Marin, le plus gros individu de l'anémone, devrait plutôt être une femelle, et que Némo ne devrait pas être un garçon mais un individu indifférencié sexuellement immature.)

Ce film pour jeunesse a eu un tel impact qu'il a été suivi d'un "effet Némo" : les ventes de ce poisson, qui était déjà populaire en aquariophilie, ont augmenté de 25% depuis la sortie du film (Militz and Foale, 2017). La pêche sauvage d'individus a de plus contribué à une réduction de la diversité génétique dans certains archipels indonésiens (Madduppa *et al.*, 2018).

b. La définition taxonomique d'*Amphiprion ocellaris* n'est pas claire

Amphiprion ocellaris est également appelé le "false percula clownfish" pour sa ressemblance phénotypique avec *A. percula*. Les deux espèces diffèrent par le nombre de rayons sur la nageoire dorsale : *A. percula* a 10 rayons épineux sur cette nageoire alors qu'il n'y en a que 11 pour *A. ocellaris* (Némo ayant 11 rayons, c'est donc un *A. ocellaris*). De plus, leurs spécificités d'hôtes sont légèrement différentes : *A. ocellaris* peut vivre dans les espèces d'anémones *Stichodactyla gigantea, S. mertensii* et *Heteractis magnifica*. À l'inverse, *A. percula* peut coloniser *S. gigantea, H. magnifica* et *H. crispa*. Enfin, leur distribution géographique est également légèrement différente : *A. ocellaris* est situé dans le triangle formé par le Nord-Ouest de l'Australie, la Thaïlande et le Japon alors que *A. percula* est restreint au Nord de l'Australie et à la Papouasie-Nouvelle-Guinée. Il existe toutefois une zone d'hybridation entre ces deux espèces, qui sont si proches qu'elles constituent un groupe parfois appelé "*percularis*" par les aquariophiles.

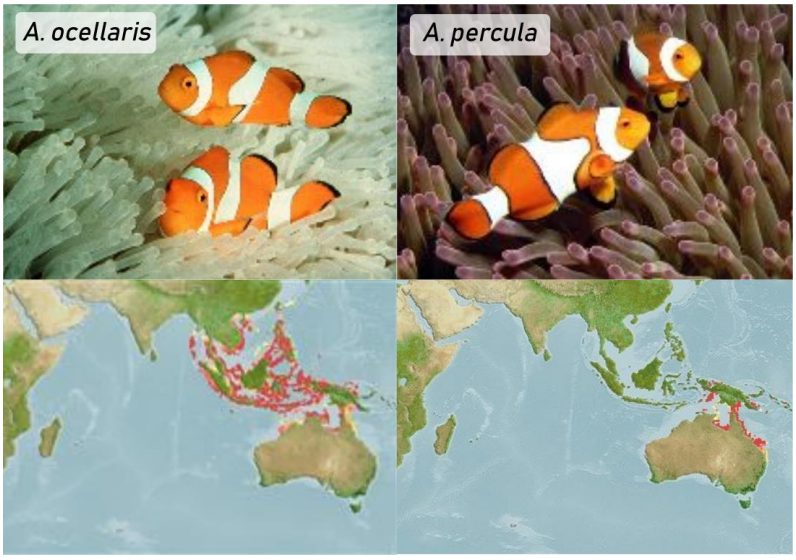


Figure 47 - Distributions comparées d'*Amphiprion ocellaris* et *A. percula*

A. ocellaris est présent du nord-ouest de l'Australie au sud du Japon et jusqu'en Thaïlande. *A. percula* est présent dans les environs du nord-est de l'Australie, des îles Salomon et des Vanuatu. Une zone d'hybridation existe entre les deux espèces qui sont toutefois considérées comme différentes. *Images : FishBase.*

c. La pigmentation adulte d'*Amphiprion ocellaris*

Les individus *Amphiprion ocellaris* sont, à l'état adulte (je présenterai plus loin l'état larvaire) constitués de trois bandes blanches, entourées de discrètes zones noires, dans un fond orangé. Cette pigmentation bariolée a parfois été considéré comme un signal d'aposématisme (Mariscal, 1966).

La pigmentation de cette espèce est un trait très populaire en aquariophilie, aussi de nombreux mutants ont été décrits. Ces mutants peuvent être issus de croisements *ocellaris* x *ocellaris*, *percula* x *percula* ou *percula* x *ocellaris*. Toutefois, aucune étude n'a à ce jour permis de comprendre la base génétique de cette variation pigmentaire, et mes demandes de résultats de croisement auprès de sociétés d'aquaculture sont restées lettre morte.

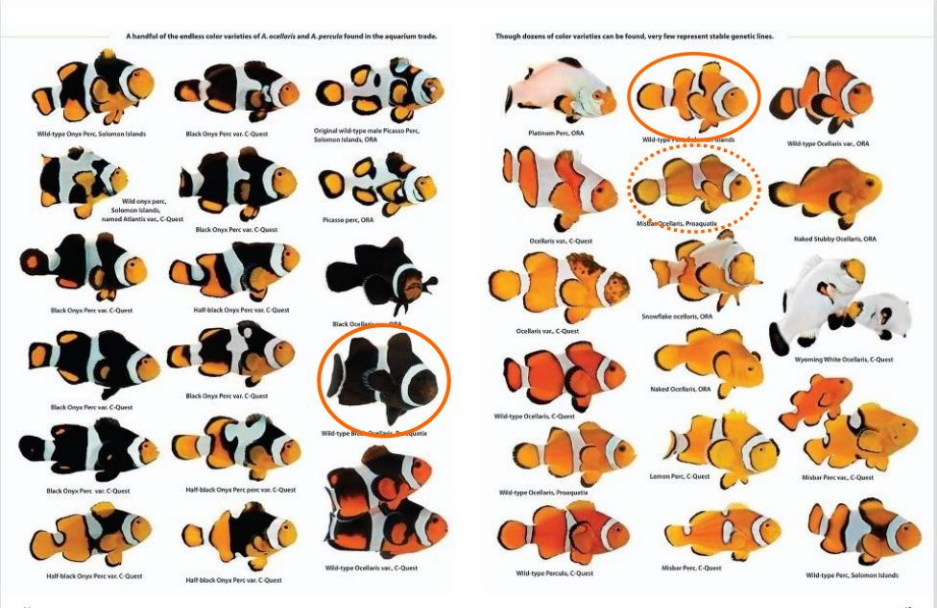


Figure 48 - Collection de morphes pigmentaires de Poissons-clowns.

Il existe en aquariophilie de nombreux morphes pigmentaires de Poissons-clowns, issus de croisements, mais dont le déterminisme génétique n'est pas connu. Dans la nature, on trouve un continuum entre les deux morphes sauvages orange et noir (trait plein). Le morphe *misbar* (pointillés) apparaît parfois. Enfin, un même poisson peut avoir une bande complète d'un côté et une bande interrompue de l'autre (non illustré). *Droits réservés : revhtree.*

La pigmentation est également un trait phénotypique variable chez cette espèce. À l'état adulte, au sein d'une même ponte, certains individus n'ont pas une bande blanche médiane continue : celle-ci peut être interrompue (on parle d'individus *misbar* ou *miss-band*), voire même réduite à un point. MM. Olivier Soulat et Patrick Masanet, aquariophiles de l'Aquarium de Canet en Roussillon et spécialistes du poisson-clown (ayant notamment fourni des individus au début de ce projet), m'ont indiqué qu'en réduisant la quantité de nourriture fournie aux larves, il était possible d'augmenter l'apparition de ces phénotypes. Cela est probablement dû à la teneur en acides gras du régime alimentaire des larves, mais le déterminisme de cette plasticité phénotypique n'est pas connu (Avella *et al.*, 2007).

Enfin, au sein-même de l'espèce *A. ocellaris*, on trouve un gradient de couleur de fond, depuis l'orange jusqu'au noir. À Darwin, dans le nord de l'Australie, il existe même une population naturelle d'*A. ocellaris* atteinte de mélanisme : les individus sont entièrement noirs – bandes blanches exceptées. Cette forme foncée est probablement liée à une hyperprolifération des mélanocytes et/ou à une surproduction de mélanine, même si, là encore, aucune étude ne permet d'étayer ces propos.

4. *A. ocellaris* est un intéressant modèle émergent

a. Qu'est-ce qu'un modèle ?

Plusieurs raisons, qui ne sont mutuellement pas exclusives, peuvent motiver le choix d'un organisme "modèle" : (i) l'utilité pour l'Homme, soit par sa proximité avec l'Homme dans une optique d'amélioration de la santé humaine (par exemple, le Chimpanzé), soit parce que le modèle provoque une maladie (le Virus de la grippe), soit parce que le modèle est une espèce économiquement importante (la Levure), soit parce qu'il est une source de nourriture pour l'Homme (le Maïs) ; (ii) une meilleure compréhension de processus biologiques ayant lieu chez l'Homme mais présents chez des organismes plus faciles d'accès (par exemple, la Souris pour l'étude de certaines tumeurs) ; (iii) l'étude pour elle-même du modèle, parce qu'un processus biologique particulier le caractérise (par exemple, la régénération des pattes chez l'arthropode *Parhyale*) ; (iv) la position phylogénétique intermédiaire entre deux taxons où des modèles sont déjà présents, afin de mieux comprendre la biologie du modèle en rapport avec celle des deux autres groupes (par exemple, le Lépisosté tacheté, Vertébré non-tétrapode non-Téléostéen) ; (v) plus récemment, l'étude du modèle pour une meilleure compréhension des impacts des activités humaines sur son cycle de vie. En cela, les organismes modèles ont une table de développement précise (Hopwood, 2007).

Dans tous les cas, un modèle est choisi pour représenter le nombre d'organismes le plus grand possible en fonction du processus à étudier (Leonelli and Ankeny, 2013). De manière pratique, un modèle est d'autant plus facilement choisi que son cycle de développement est court ; que les individus sont petits ; que la reproduction est facile en captivité ; que la descendance est nombreuse ; et que l'organisme est peu onéreux à maintenir. Une fois cela accompli, les études et premières cartographies génétiques peuvent être alors obtenues. On note qu'un cycle de vie court, un développement rapide et peu variable, avec une faible plasticité développementale, sont un biais notable associé au choix des modèles en biologie du développement (Bolker, 1995, 2012).

Enfin, l'aspect "communautaire" est un point parfois avancé pour caractériser une espèce modèle (Leonelli and Ankeny, 2013) : parce que les organismes modèles sont considérés comme représentatifs de processus qui les transcendent et concernent d'autres organismes, ils sont aussi étudiés pour d'autres aspects de leur biologie que le seul étudié dans un laboratoire donné. Ils rassemblent ainsi une communauté de spécialistes ayant des approches disciplinaires différentes, allant de l'écologie et l'évolution à la biologie moléculaire. Le partage de ressources scientifiques, humaines et matérielles à l'échelle mondiale autour d'un organisme, est, selon certains, une condition *sine qua none* à la définition d'un "modèle" (Leonelli and Ankeny, 2013). La Drosophile, *Caenorhabitis elegans*, la Levure ou la Bactérie *Escherichia coli* en sont peut-être les exemples les plus emblématiques. Tant que ces conditions ne sont pas remplies, on parlera plutôt "d'organisme expérimental", ou de "modèle émergent" (Gachelin, 2008; Cold Spring Harbor Laboratory Press, 2010).

b. Le poisson-clown *Amphiprion ocellaris*, un modèle émergent

Le poisson-clown *Amphiprion ocellaris* est étudié depuis plusieurs années par des entreprises d'aquaculture, qui, pour des raisons économiques, partagent très peu leurs résultats. Quelques laboratoires dans le monde travaillent occasionnellement sur cet organisme, notamment en Suisse, à Monaco, en Italie, en Israël, à Hawaï ou en Asie du Sud-Est, ce qui permet de mieux comprendre la biologie d'*A. ocellaris* (Yasir and Qin, 2010; Litsios, Pearman, *et al.*, 2014; Jacob *et al.*, 2017). Au sein du laboratoire de Vincent Laudet, co-encadrant de ma thèse, *A. ocellaris* est utilisé depuis quelques années comme objet d'étude en biologie du développement.

La principale justification de l'utilisation de cet organisme part d'un constat : dans le monde, les récifs coralliens, présentés plus haut pour leurs intérêts, sont dans une situation critique (Moberg and Folke, 1999; Bellwood *et al.*, 2004). Plus de 30% sont aujourd'hui dégradés, et 60% devraient l'être dans dix ans, avec des taux pouvant être plus élevés localement, notamment dans la zone caraïbe (Gardner *et al.*, 2003; Hughes *et al.*, 2003).

Les facteurs en cause sont multiples : la dégradation due à la déforestation en amont et à la surpêche (Moberg and Folke, 1999) ; l'élévation de la température du niveau marin et l'acidification des océans qui lui est associée et fragilise le squelette calcaire des cnidaires ou induit un blanchissement (Hoegh-Guldberg *et al.*, 2007; Hughes *et al.*, 2017; Eyre *et al.*, 2018) ; l'augmentation de la fréquence des ouragans (De'ath *et al.*, 2012) ; la pollution sonore, qui perturbe la réception des signaux acoustiques par les organismes récifaux (Simpson, 2005) ; et la pollution chimique, liée aux produits en tous genres déversés dans les océans (Negri *et al.*, 2011). Chez les poissons coralliens, espèces majeures de l'écosystème corallien, l'impact de ces paramètres anthropiques est assez mal caractérisé au cours du développement, et en particulier à une étape clef de leur cycle de vie : le recrutement larvaire (Jones *et al.*, 2004; Schunter *et al.*, 2018). En particulier, dans quelle mesure les facteurs sus-cités influent-ils sur le développement embryonnaire et post-embryonnaire de la larve, sur sa capacité à détecter le récif et décider de son retour, et sur sa capacité à coloniser un habitat favorable au sein du récif ?

Apporter des éléments de réponse à ces questions suppose d'étudier une espèce de poisson corallien dont on maîtrise la reproduction au laboratoire. Le poisson-clown *A. ocellaris* présente quelques spécificités qui en font un objet d'étude intéressant à ce titre : (i) les individus sont de petite taille et peuvent être élevés dans des bacs de taille moyenne (quelques dizaines de litres pour un couple) ; (ii) la reproduction peut avoir lieu en aquarium, et les œufs sont pondus sur un support fixe, à l'inverse de nombreux poissons coralliens donc les œufs sont souvent relargués dans la colonne d'eau ; (iii) la reproduction a lieu une fois toutes les trois semaines, indépendamment du cycle lunaire (en captivité), avec entre 300 et 500 œufs pondus environ ; (iv) la période d'incubation est de deux semaines environ, et le développement post-embryonnaire, après éclosion, dure moins de 20 jours avant que les larves n'aient des traits d'adulte.

Cet organisme apparaît donc comme un objet d'étude séduisant pour étudier directement sur un poisson corallien les effets de divers paramètres sur le développement. À défaut d'être déjà un "modèle", tous ces éléments font toutefois d'*A. ocellaris* un "modèle émergent" de poisson corallien. La mise en place de la symbiose avec l'anémone, très rapide (les larves peuvent instantanément coloniser une anémone, sans étape d'acclimatation), peut également être étudiée chez cet organisme. À l'inverse, *A. ocellaris* est peu adapté pour l'étude expérimentale des changements de sexe : on l'a vu, cela peut prendre jusqu'à 6 mois pour qu'un individu devienne sexuellement mature, même après que la hiérarchie de taille ait été établie (Iwata *et al.*, 2008).

c. État de l'art scientifique sur *A. ocellaris*

L'émergence du modèle va de pair avec un faible nombre de publications dans des revues à comité de lecture. En avril 2018, une recherche dans Pubmed avec les mots clefs "Amphiprion ocellaris" ne rendait que 32 résultats. Cette espèce étant populaire en aquariophilie depuis quelques dizaines d'années, certaines de ces études sont issues de laboratoires d'aquaculture et décrivent des maladies et des parasites de cette espèce avec des traitements associés (Bower *et al.*, 1887; Dhayanithi *et al.*, 2010; Pirarat *et al.*, 2011; Sharon *et al.*, 2015; Vorbach *et al.*, 2018). Les outils moléculaires étaient

toutefois inexistants chez cette espèce, malgré une méthode existante d'obtention de lignées cellulaires stables (Songwut *et al.*, 2014).

i. Structure sociale

La compréhension du déterminisme social au sein de cette espèce se développe. Des études ont ainsi mis en évidence le rôle de l'arginine vasotocine : des antagonistes de ce peptide diminuent l'agressivité des individus (Yaeger *et al.*, 2014) et exacerbent les soins parentaux mâles (DeAngelis *et al.*, 2017). Qui plus est, ce peptide est important dans la hiérarchie du groupe (Iwata *et al.*, 2010). Chez les individus dominants, la teneur en glucocorticoïdes ainsi qu'en aromatase cérébrale est plus élevée (Iwata *et al.*, 2012). Les polluants environnementaux, tels que les perturbateurs endocriniens, seraient susceptibles de perturber cette hiérarchie sociale : en présence d'ethinylestradiol, un œstrogène de synthèse trouvé dans les pilules contraceptives, la hiérarchie sociale est moins stable et les interactions agressives au sein du groupe durent plus longtemps que dans les groupes contrôles (Chen and Hsieh, 2017).

ii. Physiologie et développement

D'autres paramètres environnementaux ont été testés chez le poisson-clown *A. ocellaris* : l'acidification de l'eau n'a pas d'impact identifié sur l'assimilation des nutriments (Jacob *et al.*, 2017) et un réchauffement de quelques degrés de l'eau ne semble pas affecter sa physiologie (Madeira *et al.*, 2017).

L'étude du développement d'*A. ocellaris* n'en est qu'à ses débuts. Quelques équipes se sont intéressées au développement embryonnaire, avant éclosion (Liew *et al.*, 2006; Yasir and Qin, 2007), et d'autres ont étudié l'effet du régime alimentaire (teneur en acides gras, probiotiques) sur les stades développementaux (Avella *et al.*, 2010; Olivotto *et al.*, 2011). Un probable effet des hormones thyroïdiennes a également été documenté dans une étude relativement peu diffusée du Bulletin de l'Institut Océanographique de Monaco (Clement *et al.*, 2001).

L'étude de la pigmentation d'*A. ocellaris* est aujourd'hui cantonnée à quelques études provenant principalement d'une seule équipe d'aquaculteurs australo-indonésiens ayant pour objectif d'étudier les facteurs à même de faire varier la couleur des individus, dans une logique rentable. Les résultats montrent que les caroténoïdes alimentaires ont un effet sur les tons rouges des individus en augmentant probablement la teneur en carotènes des xanthophores (Yasir and Qin, 2010; Ho *et al.*, 2013). De même, des individus élevés dans des conditions de lumière faible sont plus "brillants" (sic), sans que ne soient avancées des hypothèses sur les bases cellulaires de cette "brillance", et des individus élevés sur un fond bleu ou vert sont plus orangés qu'un individu élevé sur fond blanc, ce qui peut être intéressant en aquariophilie mais ne révèle pas les mécanismes sous-jacents (Yasir and Qin, 2009a, 2009b).

Résumer en moins de deux pages la quasi-totalité de la littérature (hors écologie et évolution) disponible pour cette espèce souligne la rareté des informations disponibles, malgré l'importance culturelle de cette espèce emblématique des poissons coralliens. Toutefois, espérons que les années à venir soient associées à un regain d'intérêt pour *A. ocellaris* et les autres espèces de ce genre. La publication, au début de 2018, du génome d'*A. ocellaris,* va sans nul doute permettre de développer des outils moléculaires chez cette espèce (Tan *et al.*, 2018). (On note que ce génome, curieusement, a été obtenu à partir d'individus du morphe mélanique, pourtant minoritaire.) À un mois d'intervalle, le génome d'une autre espèce du même genre, *Amphiprion frenatus*, a été rendu disponible (Marcionetti *et al.*, 2018), puis, en mars 2018, le génome d'un troisième *Amphiprion* est paru – *A. bicinctus* (Casas *et al.*, 2018). La communauté semble donc se diriger vers une amélioration des outils moléculaires disponibles chez ces espèces.

II. Le développement d'*Amphiprion ocellaris* est caractérisé par une métamorphose

1. Changements morpho-anatomiques au cours du développement d'*A. ocellaris*

a. Les stades embryonnaires diffèrent des jours post-fécondation

Les études développementales précédentes chez *A. ocellaris* utilisaient les "jours post-fécondation" comme unité développementale (Yasir and Qin, 2007; Olivotto *et al.*, 2011). Cette mesure, quoique pratique, est en réalité peu adaptée à l'étude du développement. À l'instar de nombreuses autres espèces, le développement du poisson-clown varie selon la température, la nutrition, la durée du jour. Pour un "jour" de développement donné, plusieurs individus peuvent être à des "stades" plus ou moins avancés. Comparer des résultats d'une ponte à une autre pour un même couple reproducteur dans le même laboratoire, ou, à l'extrême, entre laboratoires différents, suppose d'avoir des repères communs et donc une table développementale aussi stricte que possible basée non pas sur des *jours* de développement, mais sur des *stades* (Hamburger and Hamilton, 1951; Ainsworth *et al.*, 2010). Cette table de développement est aujourd'hui disponible pour *A. ocellaris* (Roux *et al.*, in prep.)

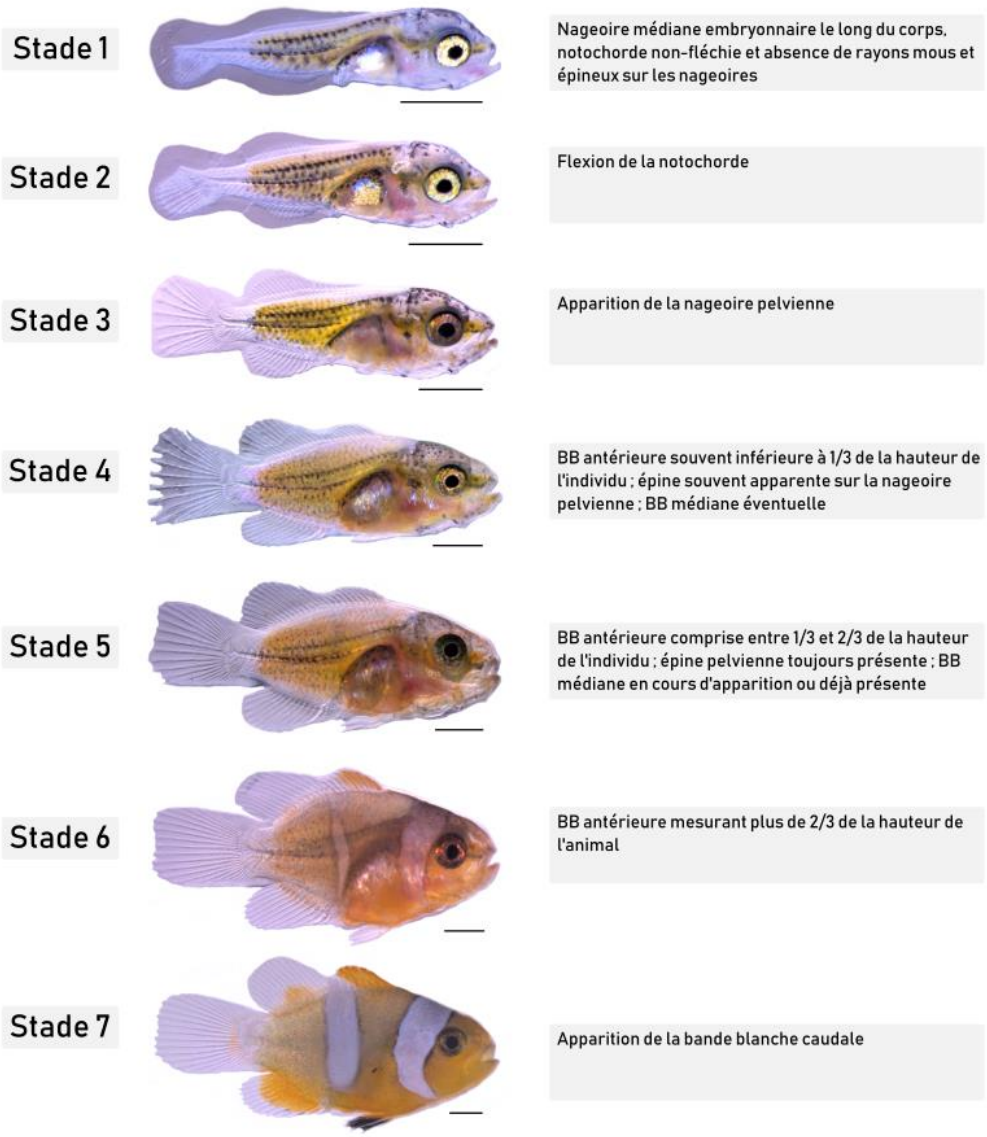


Figure 49 - Table de développement du poisson-clown *A. ocellaris*

Chaque stade développemental correspond à des critères anatomiques, morphologiques ou pigmentaires. BB : bande blanche. Barre d'échelle : 1 mm. *Clichés : N. Roux. Adapté de Roux et al. (in prep.)*

b. Acquisition de la pigmentation au cours du développement d'*Amphiprion ocellaris*

La définition des stades développementaux s'appuie sur ces critères morphologiques ou d'apparition de traits en lien avec la pigmentation, mais restant au niveau morphologique. Ils ne s'appuient pas sur une description cellulaire des cellules pigmentaires. Chez les poissons Téléostéens, comme nous l'avons vu dans le premier chapitre de cette introduction, il existe plusieurs types de cellules pigmentaires, ou chromatophores.

Chez *A. ocellaris*, les cellules noires et oranges sont, sans équivoque, respectivement, des mélanocytes et des xanthophores. Certains auteurs ont décrit des "érythrophores" chez cette espèce, qui apparaîtraient sur la nageoire caudale au bout de trois mois, mais cela est discutable (Yasir and Qin, 2007) : la publication ne montre pas de résultats réellement convaincants, et, par ailleurs, aucun érythrophore n'a été détecté dans notre laboratoire dans la peau de *A. ocellaris* (P. Salis, comm. pers.).

La couleur blanche peut être due à plusieurs types cellulaires : des leucophores, qui sont des cellules, on l'a vu, blanches (Kimura 2014), ou des iridophores, lorsque les plaquettes de guanine sont désorganisées (Rohrlich and Porter, 1972). La question de la nature de la bande blanche n'est pas résolue chez *A. ocellaris*. Les premières observations montrent que la peau blanche contient des iridophores et, dans une moindre mesure, des mélanocytes. Plusieurs autres arguments vont en faveur des iridophores. Jusqu'à présent, les leucophores n'ont été décrits que comme des cellules éparses dans un tissu, et pas comme des cellules très répandues. De plus, chez *A. chrysopterus*, une espèce polynésienne, la bande blanche peut parfois être bleutée, et ce bleu est variable selon l'intensité lumineuse (T. Lorin, obs. pers.). Ces deux éléments laissent donc penser qu'un phénomène d'iridescence est en jeu, et suggèrerait que les cellules qui donnent la couleur blanche chez *A. ocellaris* sont essentiellement des iridophores. Cependant, les iridophores semblent isolés dans le tissu et pourraient ne pas être seuls à l'origine de la couleur blanche.

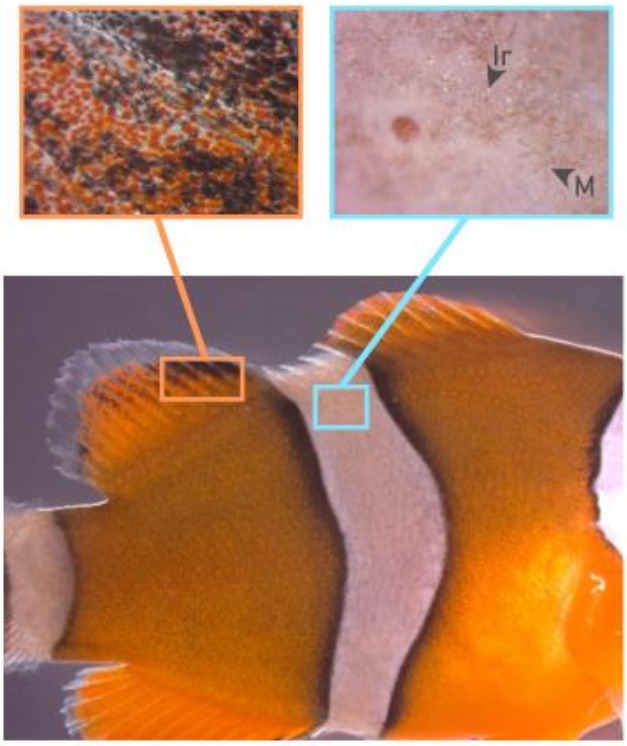
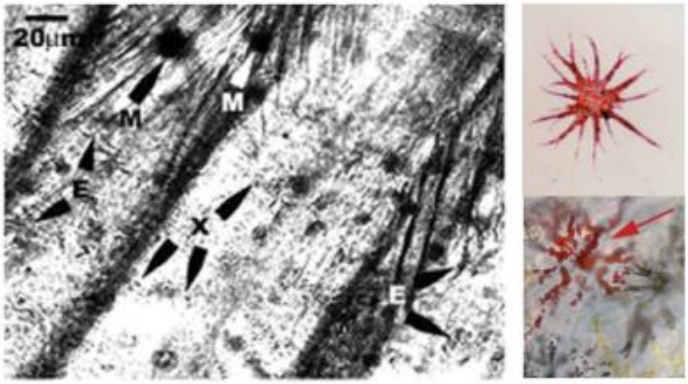


Figure 50 - Chromatophores chez *Amphiprion ocellaris*.
Au moins trois chromatophores co-existent chez *A. ocellaris* : mélanocytes, xanthophores et iridophores. La peau blanche contient (au moins) des iridophores (Ir) et des mélanocytes (M). Des érythrophores (E) ont également été décrits (à droite en haut). Toutefois, cela n'a pas été observé dans notre laboratoire et les clichés sont peu convaincants par rapport à des « vrais » érythrophores (en bas à droite). *Clichés : N. Roux (A. ocellaris) ; Yasir and Qin (2007) ; Ban et al. (2005) ; Cortesi et al. (2016).*



2. La métamorphose et son contrôle chez les Vertébrés

Now (…) their backs are green; their bellies, the largest part of their body, are white, and, as newly made frogs, they leap in their muddy pool. – Ovide, *Métamorphoses*, Livre VI

a. Aspects tissulaires, morphologiques et écologiques de la métamorphose chez les Vertébrés

Le développement d'*Amphiprion ocellaris* s'effectue selon des stades. Entre le S4 et le S5, de nombreux changements ont lieu, notamment pigmentaires, et, dans une moindre mesure, morphologiques. Chez d'autres espèces de poissons coralliens, ces derniers sont parfois plus marqués (McCormick *et al.*, 2002b). Qui plus est, chez *A. ocellaris,* c'est environ à cette période que les individus, à la fin de leur phase pélagique, sont recrutés dans le récif et colonisent une anémone. Enfin, cette étape du cycle de vie est associée, chez les poissons coralliens (ce n'est pas documenté pour *A. ocellaris*), à un remaniement tissulaire : par exemple, chez le Poisson-chirurgien bagnard *Acanthurus triostegus*, l'intestin s'allonge en rapport avec le passage à l'herbivorie (Holzer *et al.*, 2017).

Ce changement rapide et spectaculaire au niveau morphologique, pigmentaire, tissulaire, comportemental et écologique, correspond à la définition d'une "métamorphose" (en Grec, *meta-* = changement et *morphe* = forme). Le terme de "métamorphose" est utilisé pour qualifier le développement des poissons coralliens depuis au moins le début des années 1970 (Munro *et al.*, 1973). Dans le cas du poisson-clown *A. ocellaris*, les changements morphologiques sont assez peu spectaculaires. L'exemple de métamorphose le plus emblématique est probablement celui du Xénope : au cours de la métamorphose, des pattes apparaissent, la queue se résorbe et l'individu change de milieu de vie, depuis un mode de vie entièrement aquatique à un mode de vie amphibie (Brown, 1997). De même, le tube digestif est remanié en rapport avec un passage de l'herbivorie à la carnivorie (Heimeier *et al.*, 2010).

Chez les poissons Téléostéens, la métamorphose n'est pas uniquement présente chez les poissons coralliens : le Cardeau (*Paralichthys olivaceus*) ainsi que d'autres poissons-plats, telle que la Sole ou le Turbot, sont également concernés (Blaxter, 1988; Power *et al.*, 2001; Campinho *et al.*, 2010; Laudet, 2011). Au cours du développement, le plan d'organisation bilatéral de la larve est modifié, un œil migrant jusqu'à se retrouver du même côté que l'autre, et permettant un changement d'habitat (Schreiber and Specker, 1998; Schreiber *et al.*, 2005). La métamorphose concerne également la Dorade, le Congre et le Poisson-zèbre (Power *et al.*, 2001). Chez ce dernier, le patron de pigmentation adulte est mis en place au moment de la métamorphose (Brown, 1997; McMenamin *et al.*, 2014).

Dans le cas du Poisson-zèbre, certains auteurs suggèrent de ne pas utiliser le terme "métamorphose" pour qualifier le mode de développement général, puisqu'il n'y a pas de réel changement écologique ou morphologique (Parichy *et al.*, 2009). Le terme ne devrait être utilisé que pour évoquer le mode de développement du patron pigmentaire (Parichy *et al.*, 2009). Cette discussion souligne le côté arbitraire d'une définition de la métamorphose, puisque les changements "rapides" et "spectaculaires" peuvent ne s'appliquer qu'à un aspect du développement : écologie, physiologie, morphologie… Chez la majorité des poissons coralliens, le cycle de vie biphasique implique un changement de milieu. Par ailleurs, outre les poissons-clowns, de nombreux poissons coralliens ont des changements majeurs de la pigmentation au cours du recrutement (Baldwin, 2013). Le terme de "métamorphose" est donc effectivement consacré, même si son aspect "spectaculaire" peut être variable selon les espèces.

b. Aspects physiologiques

i. Un contrôle par les hormones thyroïdiennes

1. Mise en évidence historique

La métamorphose a fasciné les naturalistes et artistes depuis l'Antiquité, en inspirant notamment Aristote et Ovide. La mise en évidence de son déterminisme biologique remonte elle au début du XXème siècle. En 1912, Gudernatsch, scientifique américain de visite estivale en Europe, comme de coutume, donne des extraits de plusieurs organes de chevaux

à des têtards (Gudernatsch, 1912; Demeneix, 2017). Des extraits de thyroïde semblent en particulier induire la transformation d'un têtard en grenouille. C'est la première mise en évidence du rôle d'hormones dans le processus de métamorphose. En 1916, quelques années plus tard, Allen arrive à empêcher la métamorphose par une "extirpation" (sic) de la thyroïde (Allen, 1916).

2. Mécanismes moléculaires de la voie de signalisation thyroïdienne

Les hormones en question ont par la suite été mieux caractérisées : ce sont les hormones thyroïdiennes, appelées T3 et T4 chez les Vertébrés, et isolées dans les années 1950 (pour une revue historique, et quelque peu autobiographique, voir (Tata, 2008)). La T3, ou tri-iodothyronine, est l'hormone active et ne comporte que trois atomes d'iode. Elle dérive de la T4, ou tétra-iodothyronine, qui en comporte quatre et qui est produite au sein des follicules thyroïdiens. Des cellules spécialisées du follicule, les thyréocytes, sécrètent dans la lumière du follicule une protéine couplée à des tyrosines iodées : la thyroglobuline (Tg), longue protéine de 2700 acides aminés (Bergé-Lefranc *et al.*, 1981; Holzer, 2015). Des atomes d'iode sont alors ajoutés à certains acides aminés de la Tg, les tyrosines, par la thyropéroxydase (Tpo). Une fois la Tg iodinée, celle-ci est internalisée par le thyréocyte puis digérée dans des lysozymes avant export de la seule T4.

La transformation de T4 en T3 a lieu dans les organes périphériques et est assurée par une enzyme de la classe des déiodinases, la déiodinase de type 2 (Huang *et al.*, 2001; Cai and Brown, 2004; Bianco and Kim, 2006). Ces hormones activent l'expression de gènes cibles dans les tissus cibles en se fixant à des récepteurs nucléaires. La T3 se fixe à ses récepteurs, des facteurs de transcription, les TR (*Thyroid hormone Receptors*), qui s'hétérodimérisent avec d'autres facteurs de transcription, les RXR (*Retinoid X Receptors*) (Yaoita *et al.*, 1990). Ces derniers sont des récepteurs de la famille du récepteur de l'acide rétinoïque mais pour lesquels on ne connait pas de ligand (ils sont dits "orphelins") (Evans and Mangelsdorf, 2014).

Chez les Vertébrés, il y a deux TR : TRα et TRβ, aux propriétés différentes. Chez le Xénope, TRα permet d'empêcher l'entrée en métamorphose avant que les niveaux d'hormones thyroïdiennes ne soient assez élevés (Sachs *et al.*, 2002). À l'inverse, l'expression de TRβ est inhibée par TRα au début du développement. Elle est également contrôlée par les hormones thyroïdiennes et, chez les Anoures, son expression pique au moment du pic de concentration des hormones thyroïdiennes, correspondant au "climax" de la métamorphose (Machuca *et al.*, 1995). L'hétérodimère TRβ-RXR peut s'auto-activer et induire alors à son tour induire l'expression d'un faible nombre de gènes cibles, qui, à leur tour, induiront l'expression de gènes cibles secondaires à l'origine des remaniements tissulaires (Furlow and Neff, 2006; Brown and Cai, 2007).

Ces mécanismes moléculaires de transduction du signal thyroïdien sont variables chez les Vertébrés, et au sein-même des Amphibiens. Par exemple, chez les Urodèles, il n'y a pas d'auto-amplification de TRβ au moment du climax (Safi *et al.*, 2004). Chez les poissons Téléostéens, les mécanismes diffèrent également du Xénope. En premier lieu, suite à la Ts3R, il y a deux gènes pour TRα : TRα-A et TRα-B. De plus, les profils d'expression des différents TR et leur lien temporel avec le climax diffèrent et semblent espèce-spécifique (Yamano and Miwa, 1998; Laudet, 2011). Toutefois, des traitements aux hormones thyroïdiennes ou avec des antagonistes sont en mesure de perturber cette transition d'histoire de vie, chez le Poisson-zèbre *D. rerio* (McMenamin *et al.*, 2014) et chez diverses espèces de poissons coralliens (Clement *et al.*, 2001; Holzer *et al.*, 2017). De même, la voie de signalisation des hormones thyroïdiennes est impliquée dans la métamorphose de la Daurade *Sparus aurata* (Campinho *et al.*, 2010), du Thon (Kawakami *et al.*, 2008) et dans le processus de smoltification chez les saumons (acclimatation des jeunes saumons à l'eau de mer), parfois considéré comme une métamorphose "secondaire" (Rousseau and Dufour, 2012; Lorgen *et al.*, 2015; Holzer and Laudet, 2015). Ainsi, même si des variabilités de la voie de signalisation existent au sein des Vertébrés, le rôle de cette voie semble, lui, généralement conservé.

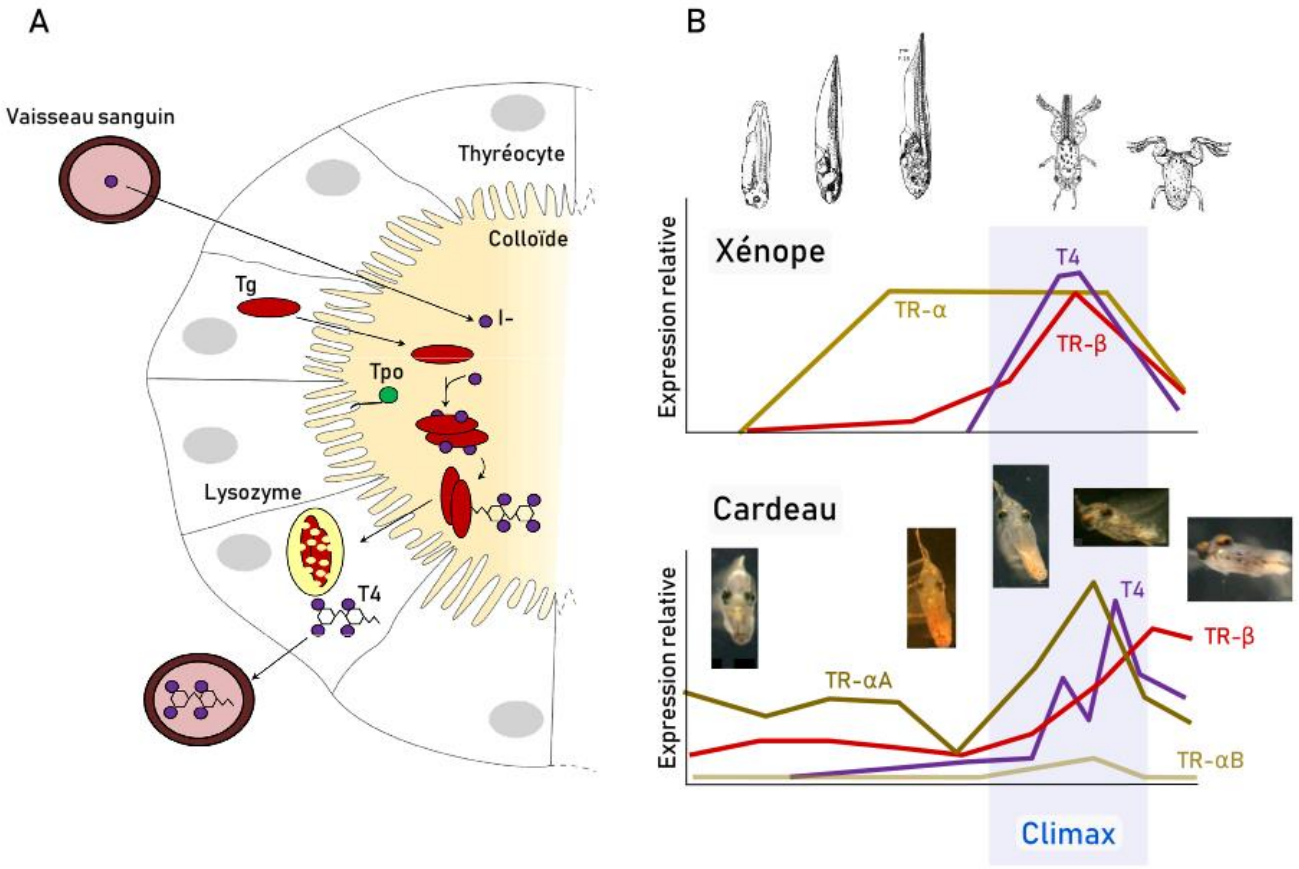


Figure 51 – Synthèse et niveaux d'expression de T4.
(A) Production de T4 au sein du follicule thyroïdien. La T4 est synthétisée à partir de Thyroglobuline (Tg), iodinée par la Thyropéroxydase (Tpo) dans la colloïde. *Adapté de Holzer (2015).* (B) Profils d'expression comparés de T4 et de ses récepteurs chez le Xénope et le Cardeau . Il y a un pic d'hormones au moment du climax qui coïncide avec un pic d'expression de TR-β chez le Xénope. Chez le Cardeau, en raison de la Ts3R, il y a deux TR-α dont l'un, TR-αA, pique en même temps que la T4. Ainsi, la métamorphose correspond à un pic de T4 qui se traduit par des expressions différentes des récepteurs selon les organismes. *Adapté de Yamano and Miwa (1997), Furlow and Neff (2006), Schreiber (2006) et Laudet (2011).*

ii. Interaction de la voie des hormones thyroïdiennes avec l'environnement : rôle des glucocorticoïdes et de l'axe hypothalamo-hypophysaire

Les mécanismes tels que décrits ci-dessus semblent immuables : un pic d'hormones induit, le plus souvent, un pic de récepteurs, qui active l'expression de gènes cibles. Toutefois, chez les poissons coralliens, la durée de la phase larvaire et donc le moment de la métamorphose, qui correspond au recrutement, peut varier. Des mesures d'âge au recrutement, effectuées par otolithométrie (mesure de la croissance des concrétions minérales de l'oreille interne), indiquent que les larves de *Thalassoma bifasciatum* ont des âges au recrutement allant de 38 à 78 jours (Victor, 1986). Les individus à 78 jours retardent donc leur métamorphose de plusieurs semaines. De plus, leur taux de croissance diminue pendant la durée du décalage de la métamorphose. De même, chez *A. triostegus*, il a été montré que des larves capturées au niveau du récif au moment du recrutement, et artificiellement conservées dans des cages au large, retardaient leur métamorphose par rapport à des larves gardées dans des cages au niveau du récif (McCormick, 1999).

Chez les Vertébrés, cette importante plasticité développementale est également caractéristique des Amphibiens (West-Eberhard, 2003). (*On note que les "records" sont observés chez les "InVertébrés", la métamorphose pouvant être retardée de l'ordre de plusieurs mois chez certains Gastéropodes ou de quelques années chez les Insectes (Jackson and Strathmann, 1981)*). Si le volume de l'eau du bac diminue, les têtards accélèrent leur développement et rentrent en

métamorphose plus tôt (Denver, 1988). De même, d'autres facteurs environnementaux influent sur l'entrée en métamorphose et le taux de croissance : la densité des conspécifiques et la nourriture disponible (Wilbur and Collins, 1973) ainsi que la présence de prédateurs (Relyea, 2007).

Ces résultats suggèrent que le déterminisme de la métamorphose est partiellement environnemental. Chez les Vertébrés, la production de T4 par la thyroïde est induite par la Thyroid-stimulating Hormone (TSH, ou thyrotropine, sécrétée par l'hypophyse). Chez les Mammifères, la TSH est libérée sous l'action de la Thyrotropin-releasing hormone (TRH) hypothalamique (Schally *et al.*, 1969; Burgus *et al.*, 1969). Or, chez les Vertébrés non-Mammifères, la libération de TSH n'est pas dépendante de TRH, mais de la corticotropin-releasing hormone (CRH, ou corticolibérine) hypothalamique (Denver, 1988; De Groef *et al.*, 2006). Cette dernière est par ailleurs une neurohormone induisant l'expression de l'hormone corticotrope ACTH, qui, en agissant au niveau des glandes surrénales, augmente la production de corticoïdes. Ainsi, chez les Amphibiens et également chez les Téléostéens, qui m'intéressent plus particulièrement dans le cadre de ma thèse, la voie de synthèse des hormones thyroïdiennes est hautement couplée à celle de la voie des glucocorticoïdes (Boorse and Denver, 2004).

Nous l'avons vu, la métamorphose est sous contrôle de multiples paramètres environnementaux. Cependant, au sein d'une cohorte d'individus tous élevés dans les mêmes conditions (même bac, même température, même fond génétique, même nourriture), il existe une variabilité de la métamorphose. Par exemple, chez *A. ocellaris*, deux individus de la même ponte ont un "taux développemental" (une vitesse de passage d'un stade à l'autre) variable. Le déterminisme de cette variabilité endogène n'est pas bien connu (Beck, 1997).

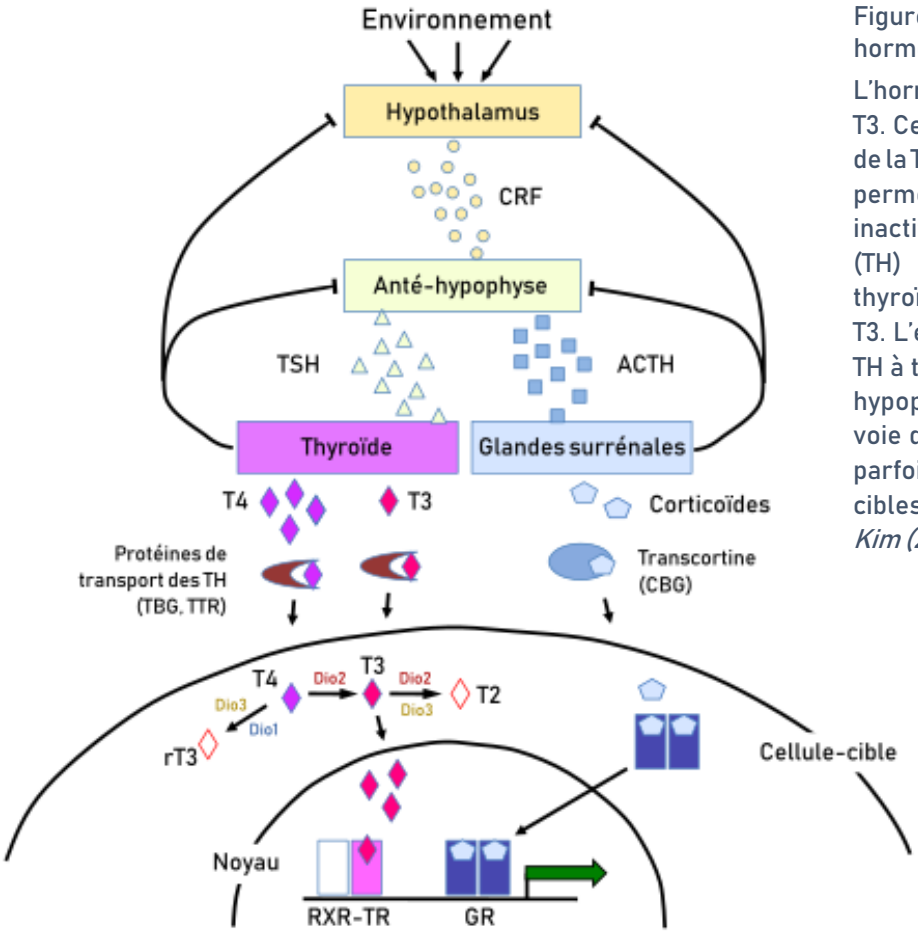


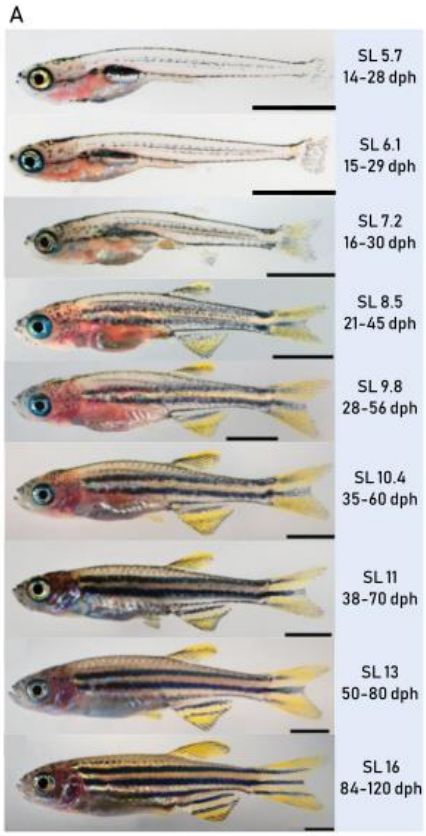
Figure 52 - La voie de signalisation des hormones thyroïdiennes.

L'hormone active se fixant à son récepteur est la T3. Celle-ci provient de la déiodination partielle de la T4 par la déiodinase 2. D'autres déiodinases permettent de transformer T3 et T4 en composés inactifs (T2 et rT3). Les hormones thyroïdiennes (TH) sont produites dans les follicules thyroïdiens, qui synthétisent 80% de T4 et 20% de T3. L'environnement influe sur la production de TH à travers le contrôle par l'axe hypothalamo-hypophysaire. Celui-ci contrôle également la voie de production des corticoïdes qui agissent parfois en synergie pour l'expression des gènes cibles. *Adapté de Laudet (2011) et de Bianco and Kim (2006).*

iii. Influence de la voie des hormones thyroïdiennes sur la pigmentation en particulier

La signalisation thyroïdienne est, on l'a vu, impliquée dans divers processus de la métamorphose. Chez le Poisson-zèbre *D. rerio*, la métamorphose est dans l'ensemble peu spectaculaire (Laudet, 2011), hormis concernant sa pigmentation (Parichy and Turner, 2003; Parichy *et al.*, 2009). Au cours de la métamorphose, une alternance de bandes sombres ("stripes") et claires ("interstripes") apparaît (cf. chapitre introductif n°1), à partir d'une larve quasiment transparente. À l'instar du poisson-clown *A. ocellaris*, mais avec une variabilité peut-être encore plus importante, le nombre de jours post-fécondation n'est pas le meilleur critère pour définir les changements ayant lieu au cours du développement, et l'équipe de D. Parichy a plutôt choisi de définir des stades en fonction de la longueur de l'individu (SL, pour "standard length").

Chez cette espèce modèle, les outils génétiques sont disponibles et l'influence de la voie des hormones thyroïdiennes a pu être expérimentalement testée, en particulier sur la mise en place du patron pigmentaire. Par exemple, lorsque le gène de la déiodinase 2 n'est pas fonctionnel, la pigmentation se met en place plus tardivement (Walpita *et al.*, 2009). Le phénotype sauvage peut être restauré par injection de T3, indiquant le rôle prépondérant de cette hormone (Walpita *et al.*, 2009). De plus, une hyperthyroïdie, accompagnée d'une plus forte expression de thyroglobuline et de T4, induit une hyper-prolifération des xanthophores et une répression des mélanocytes, qui se différencient plus mais meurent ensuite (McMenamin *et al.*, 2014). À l'inverse, une hypothyroïdie est accompagnée d'une absence de xanthophores et de mélanocytes ectopiques. La voie des TH semble donc être importante dans la mise en place du patron de pigmentation chez le poisson-zèbre. Enfin, si la thyroïde est enlevée à l'état adulte, une fois le patron pigmentaire mis en place, les mélanocytes prolifèrent, mais les xanthophores ne sont pas affectés. Ainsi, les TH semblent importantes dans le maintien de l'homéostasie des mélanocytes, mais pas dans celles des xanthophores, indiquant des sensibilités différentes à ces hormones dans ces deux lignées cellulaires (McMenamin *et al.*, 2014).



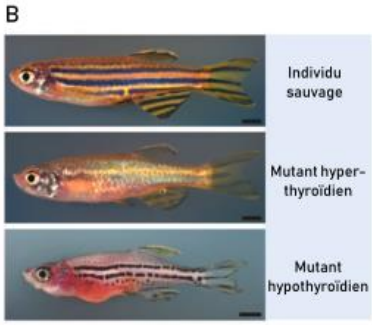


Figure 53 - Développement pigmentaire et hormones thyroïdiennes chez le Poisson-zèbre.

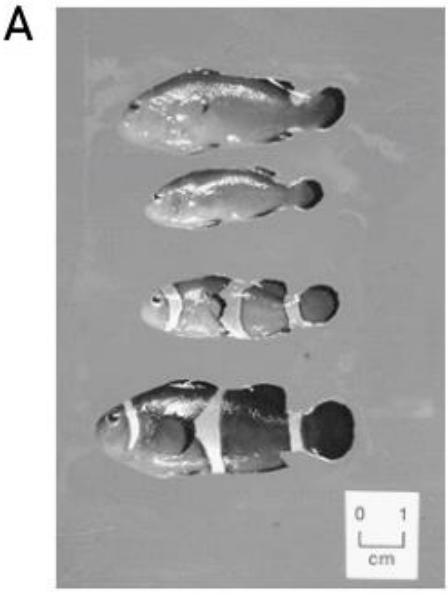
(A) La table du développement du Poisson-zèbre montre une métamorphose au niveau pigmentaire. On note l'importance de définir des stades sur d'autres critères que le seul nombre de jours post-éclosion (dph). Ici, c'est la longueur standard qui a été choisie (SL). Barre d'échelle : 2 mm. *Adapté de Parichy et al. (2009).*

(B) Le mutant hyperthyroïdien du Poisson-zèbre présente un excès de xanthophores et un défaut de mélanocytes. À l'inverse, les mutants hypothyroïdiens n'ont pas de xanthophores. *Adapté de McMenamin et al. (2014).*

Chez le poisson-clown *A. ocellaris*, le rôle des hormones thyroïdiennes a également déjà été documenté. Une étude menée en 2001, dont les méthodes sont décrites très succinctement, est celle de (Clement *et al.*, 2001). Dans cette étude les auteurs montrent qu'un traitement à la T3 exogène perturbe le développement du patron pigmentaire d'*A. ocellaris* : les bandes blanches ne se mettent pas en place. De plus, peu d'épines dorsales se mettent en place.

Au sein de l'équipe de V. Laudet, des traitements ont également été menés et montrent des résultats qui ne sont pas aussi catégoriques. Des individus traités par un mélange T3+IOP (3nM) à partir de 5 jours après éclosion ont, 2 jours plus tard (à 7 jours après éclosion), une bande blanche antérieure qui s'est formée plus rapidement. La deuxième blanche commence même à se former. À 10 jours, toutefois, les individus traités ont une bande blanche médiane moins développée : cette dernière se forme donc "plus vite mais moins bien". (*L'IOP est un inhibiteur des déiodinases qui convertissent la T4 en T3, permettant de s'assurer que les seules hormones actives sont celles ajoutées artificiellement, et donc de s'affranchir de la variabilité des niveaux de T4 entre les larves.)* On note que ces résultats sont préliminaires : les individus n'ont pas été stadés et une seule concentration d'hormone a été testée. La conclusion est donc qu'un traitement aux hormones thyroïdiennes a un effet sur le développement de la pigmentation chez *A. ocellaris*, mais on ne peut dire avec plus de précision à quel stade et selon quelles modalités.

Cette expérience préliminaire présente un résultat quelque peu différent de Clement et al. Cependant, les méthodes et concentrations employées ne sont pas les mêmes. Toutefois, la tendance est la même : les hormones thyroïdiennes semblent empêcher le développement des bandes (ou de la deuxième bande seulement) chez *A. ocellaris*.



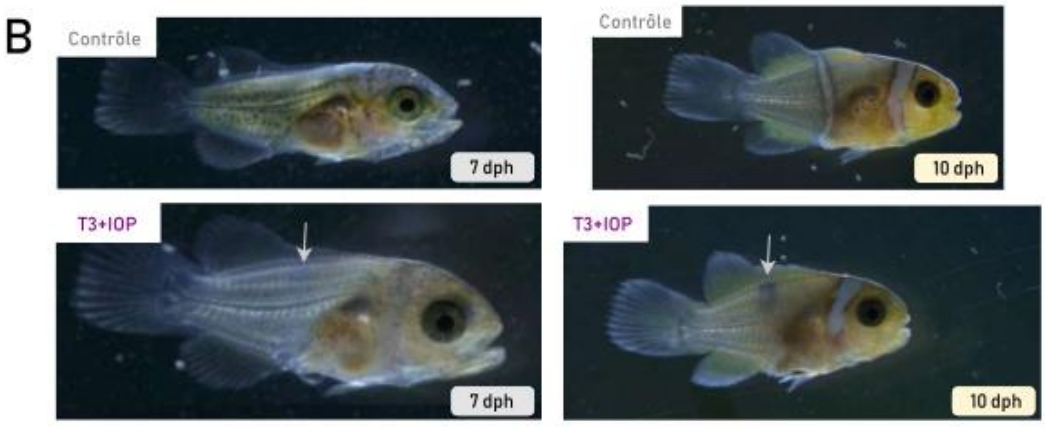


Figure 54 - Effet des hormones thyroïdiennes sur la pigmentation de *A. ocellaris.*

(A) Un traitement à la T3 exogène induirait une réduction des bandes blanches. *D'après Clement et al. (2011).*

(B) Au sein du laboratoire de V. laudet, des individus ont été traités par un mélange T3+IOP (3nM) à partir de 5 jours après éclosion ont, 2 jours plus tard (à 7 jours après éclosion), une bande blanche antérieure qui s'est formée plus rapidement. La deuxième bande blanche commence même à se former (flèche). À 10 jours, toutefois, les individus traités ont une bande blanche médiane moins développée (flèche). *Clichés : P. Salis et N. Roux.*

3. La métamorphose est un phénomène biologique pouvant être étudié par RNA-Seq

The power of sequencing RNA lies in the fact that the twin aspects of discovery and quantification can be combined in a single assay called RNA-seq." (Conesa *et al.*, 2016)

a. Le RNA-Seq est une méthode "omique" permettant d'étudier l'expression des gènes

i. Principe général du RNA-Seq

Le RNA-Seq est une technologie, qui, comme son nom l'indique, consiste à séquencer les ARN d'un organisme, d'un tissu ou d'une cellule. En réalité, ce n'est pas une molécule d'ARN qui est séquencée, mais l'ADN complémentaire issu de l'ARN initial par rétrotranscription. Le plus souvent, ce sont uniquement les ARN messagers qui sont séquencés. Les protocoles diffèrent selon les technologies, mais les grandes étapes sont conservées. Après une nécessaire extraction des ARN, les ARN polyA seuls sont conservés (souvent extraits par des séquences oligo-dT fixées sur des billes). Les ARN sont ensuite fragmentés afin d'atteindre une taille adaptée à la technologie de séquençage, puis rétro-transcrits en cDNA car, aujourd'hui, le séquençage de l'ADN est plus abouti que le séquençage direct des ARN. Les cDNA sont ensuite amplifiés par PCR afin d'amplifier le signal détecté au cours de l'étape de séquençage.

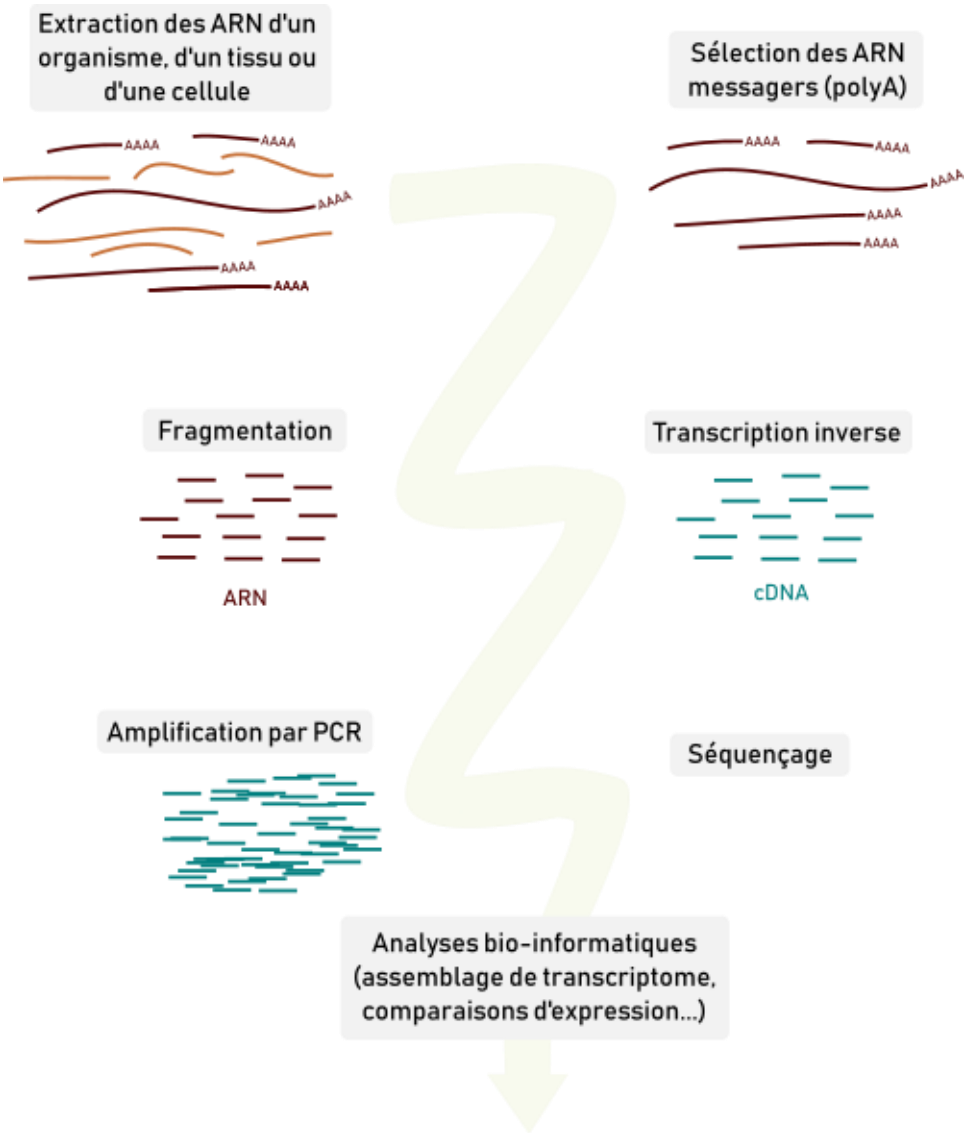


Figure 55 - Principe général du RNA-Seq.

Après extraction des ARN, les ARN polyA sont conservés. Les ARN sont ensuite fragmentés afin d'atteindre une taille adaptée à la technologie de séquençage, puis rétro-transcrits en cDNA. Les cDNA sont ensuite amplifiés par PCR afin d'amplifier le signal détecté au cours de l'étape de séquençage.

ii. Avantages du RNA-Seq et application à l'étude du poisson-clown A. ocellaris

Les débuts du RNA-Seq, qui suppose de séquencer indirectement les molécules d'ARN, ont été permis par la diminution drastique des coûts du séquençage au cours des années 2000 (Shendure *et al.*, 2004). Les nouvelles techniques de séquençage, ayant remplacé la méthode Sanger, parfois en s'inspirant de celle-ci, sont par ailleurs bien plus rapides (Sanger *et al.*, 1977). La concurrence fait rage entre les différentes "biotech" pour conquérir les parts de marché, et la pôle position est rude à conserver. Ainsi, alors que Roche-454 et son pyroséquençage était la technologie dominante au début des années 2000, la compagnie aujourd'hui appelée Illumina, à la technologie basée sur la fluorescence des nucléotides, est aujourd'hui la première "biotech" de séquençage.

Les méthodes de séquençage employées en RNA-Seq ne sont pas adaptées à des séquençages de longs reads (même si certaines technologies émergentes, comme celle développée par Pacific Biosciences, le permettent). Des jeux de données colossaux, contenant des centaines de millions de reads (entre 50 et 300 nucléotides, le plus souvent), sont alors générés. L'utilisation de ce nouvel outil est donc allée de pair avec une amélioration des capacités computationnelles, pour (i) aligner les reads sur le génome, quand celui-ci est connu ; (ii) assembler les reads en contigs, plus longs ; (iii) quantifier les différences d'expression entre les conditions (Garber *et al.*, 2011). Alors que les technologies de séquençage continuent de se développer, de nombreux outils bioinformatiques ont été développés, améliorant parfois de plusieurs ordres de grandeur le temps de calcul (Bray *et al.*, 2016). Des "revues" de l'ensemble des méthodes disponibles et des "manuels" d'utilisation, sorte de "pipelines" clés en main, sont parfois publiés pour aider la communauté à s'y retrouver (Conesa *et al.*, 2016).

Cette technologie offre des potentialités diverses en comparaison avec les micro-arrays, qui l'ont précédée. En 2001 des comparaisons entre micro-arrays, méthode alors très prometteuse, et les "vieillissantes" méthodes comme les Northern blot étaient effectuées (*Quantitative Assessment of DNA Micro-arrays - Comparison with Northern Blot Analyses* (Taniguchi *et al.*, 2001)). Sept ans plus tard, ironiquement, ces comparaisons avaient lieu entre le "prometteur" RNA-Seq et les "vieillissants" micro-arrays (*RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays* (Marioni *et al.*, 2008)). En s'affranchissant de l'hybridation avec des séquences déjà connues (comme c'était le cas pour les micro-arrays), le RNA-Seq a permis de découvrir une grande variété de nouveaux transcrits et de variants de splicing (Jacquier, 2009; Zhou *et al.*, 2010; Mortazavi *et al.*, 2010) ou des événements de fusions de gènes à l'origine de cancers (Berger *et al.*, 2010). Il a permis de plus d'améliorer des assemblages de génome déjà existants (Mortazavi *et al.*, 2010) et de comparer des jeux de données variés sans nécessiter de normalisation très sophistiquée, en utilisant des mesures comparables entre échantillons, telles que les RPKM (Reads par Kilobase par Million) (Mortazavi *et al.*, 2008), ou, aujourd'hui, plutôt, les TPM (Transcrits par Million).

Dans le cas du poisson-clown *A. ocellaris*, plusieurs propriétés du RNA-Seq en font un outil particulièrement intéressant. Contrairement aux micro-arrays, le RNA-Seq permet de travailler avec des organismes non-modèles, où les ressources bio-informatiques sont peu disponibles, et d'en obtenir un transcriptome sans génome (Grabherr *et al.*, 2011). Avant la publication du génome d'*A. ocellaris* en 2018, le génome publié de poisson-corallien le plus proche était celui de la demoiselle *Stegastes partitus*, ayant divergé de *A. ocellaris* il y a 50-70 millions d'années environ (Tan *et al.*, 2018). De plus, la profondeur de séquençage parfois très importante en RNA-Seq permet d'obtenir des reads qui correspondent à des gènes parfois très peu exprimés, tels que certains gènes de la voie des hormones thyroïdiennes, parfois exprimés dans des types cellulaires restreints (par exemple, les neurones de l'anté-hypophyse exprimant *tsh*).

Table 2 - Avantages du RNA-Seq par rapport aux micro-arrays.

(D'après Wang *et al.* (2009))

Technologie	Micro-arrays	RNA-Seq
Grands principes		
Méthode	Hybridation	Séquençage massif
Résolution	De quelques à 100 nt	1 nt
Dépendance au génome	Oui	Partielle
Bruit de fond	Elevé	Faible
Application		
Gamme de variation pour quantifier une différence d'expression	Jusqu'à quelques centaines de fois	Plus de 10 000 fois
En pratique		
Quantité d'ARN nécessaire	Elevée	Faible
Coût	Elevé	Plus faible

iii. Remarque conclusive

Outre l'aspect "technique" de cette technologie, l'aspect scientifique et le *pourquoi* de cette méthode doivent également être considérés. Cette technologie a permis d'aller plus loin que les micro-arrays dans la compréhension du transcriptome des organismes modèles, mais peut-être surtout dans la compréhension de ceux des organismes non-modèles qui ont ainsi été "démocratisés". Cela a sans aucun doute été un gain précieux, en particulier en Évo-Dévo (Roux *et al.*, 2015). Toutefois, je conclurai cette partie par les propres mots du généticien américain Jay Shendure : "*Although [RNA-Seq] may improve the quality of transcriptome profiling, we will continue to face what has probably been the larger challenge with micro-arrays: how best to generate biologically meaningful interpretations of complex datasets that are sufficiently interesting to drive follow-up experimentation."* (Shendure, 2008).

b. Le RNA-Seq est une méthode adaptée à l'étude du développement

Un des avantages principaux du RNA-Seq est qu'il peut aider à la capture de dynamiques transcriptomiques dans des jeux de données multiples, par exemple dans des stades développementaux successifs (Wang *et al.*, 2009).

i. Apports du RNA-Seq à la comparaison du développement entre espèces

Au cours du développement, nous l'avons vu, de nombreux changements ont lieu, à toutes les échelles : organe, cellule, tissu. Quels que soient les signaux déclencheurs, ces changements sont gouvernés, *in fine*, par des variations des niveaux d'expression de gènes impliqués dans divers processus, qui, ensemble, concourent à la définition du phénotype adulte (Pantalacci and Sémon, 2015). Le processus de développement peut donc être vu comme un complexe changement de modules de gènes utilisés à différents moments et dans différentes parties d'un organisme (Baron and Yanai, 2017).

De nombreuses différences au niveau morphologique apparaissent entre différents organismes au moment du développement, à toutes ses étapes. Toutefois, au niveau morphologique, il existe une période du développement, appelée "stade phylotypique", où plusieurs lignées d'un même phylum sont comparables : c'est le modèle du sablier (Duboule,

1994). Cette similarité morphologique correspond également à une similarité génétique, un plus grand nombre de gènes étant exprimé de la même manière au même moment du développement dans un phylum donné (Kalinka *et al.*, 2010). À l'inverse, entre différents phylums, l'expression génétique est plus similaire au début et à la fin du développement, où les morphologies sont très différentes, qu'au milieu du développement, où les morphologies sont comparables : c'est le modèle du "sablier inversé" (Levin *et al.*, 2016).

Les comparaisons multi-espèces de jeux de données RNA-Seq ont donc amélioré la compréhension du développement. Au XIXème siècle, c'était essentiellement l'anatomie comparée qui permettait d'établir des phylogénies (Gould 1977). Aujourd'hui, certains auteurs considèrent que la base de la diversité biologique est à rechercher dans les variations des profils d'expression des gènes au cours du développement. Levin et al. suggéraient d'ailleurs en 2016 qu'un phylum n'était rien d'autre qu'un "ensemble d'espèces qui partagent une expression génétique similaire à la transition mi-développementale (le "stade phylotypique") mais dont cette expression diffère par rapport aux espèces à l'extérieur du phylum." (Levin *et al.*, 2016). Les études "omiques" permettent donc une meilleure compréhension du développement.

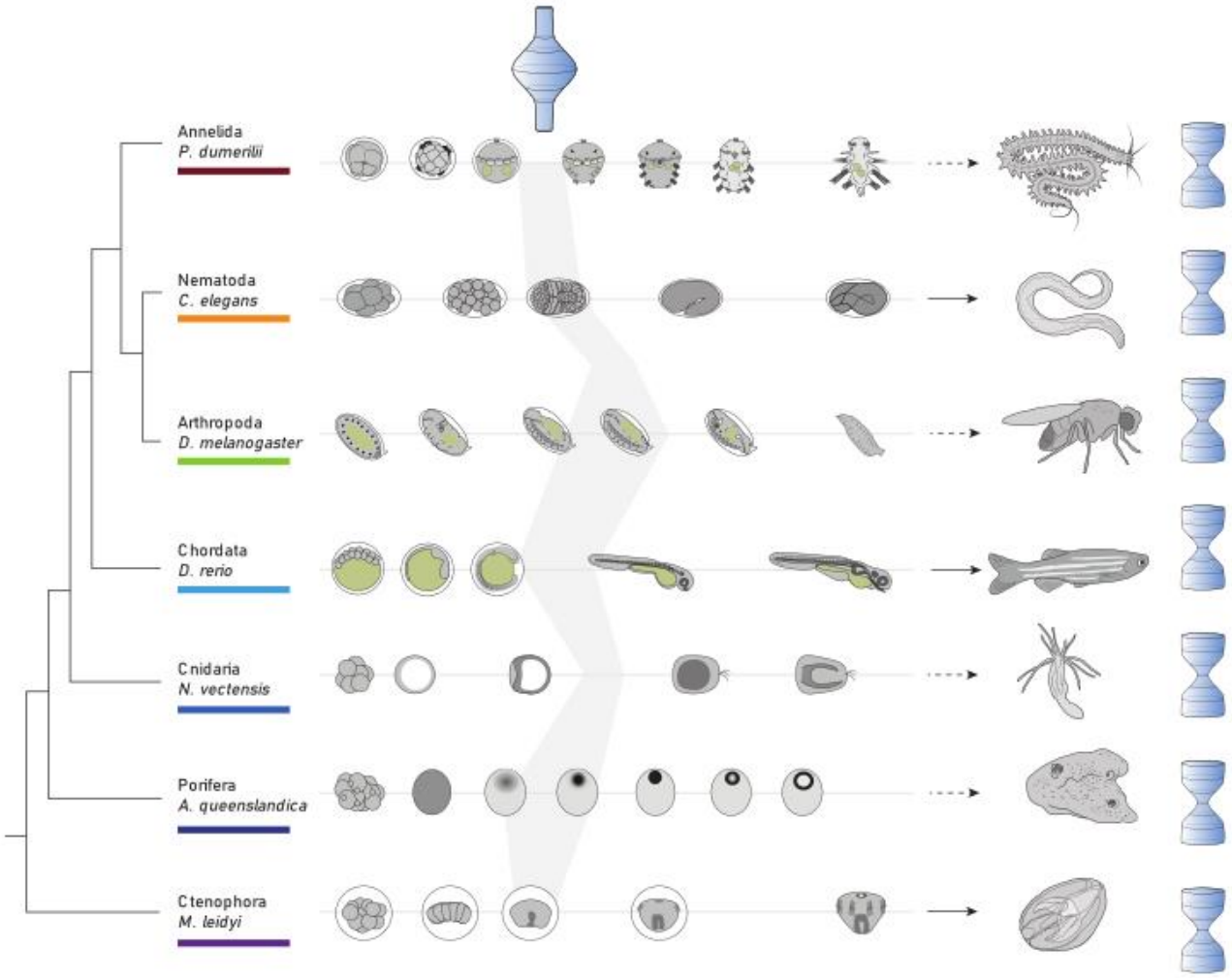


Figure 56 - Comparaison du développement au sein de différents phyla animaux.

Au sein d'un même phylum, les organismes passent par un stade appelé « transition mi-développementale » (en gris). Les espèces de ce phylum partagent alors une expression génétique moins divergente qu'avant et qu'après ce stade (modèle du sablier). À l'inverse, entre différents phylums, l'expression génétique est très divergente au moment de cette transition mi-développementale (modèle du sablier inversé).

ii. Apports du RNA-Seq à l'étude du développement au sein d'une espèce

L'utilisation de données de RNA-Seq pour des études comparatives, notamment en évo-dévo, a, on l'a vu, été amplement utilisée (Kalinka *et al.*, 2010; Levin *et al.*, 2016). Toutefois, les jeux de données RNA-Seq permettent également de répondre à des questions au sein d'une même espèce. Le développement d'un organisme peut être caractérisé par des stades, embryonnaires d'abord (par exemple, chez les Chordés : morula, blastula, gastrula, neurula, bourgeon caudal) puis post-embryonnaires. Est-ce que ces stades, définis morphologiquement par les embryologistes depuis de nombreux siècles, qui ont identifié des discontinuités de l'un à l'autre, reflètent également des discontinuités dans l'expression génétique qui les sous-tend ? Au contraire, y a-t-il, au sein d'un organisme donné, une expression génétique continue qui produit des variations discontinues du phénotype, "leurrant" ainsi les embryologistes ? Enfin, y a-t-il des découplages temporels entre des discontinuités de l'expression génétique et des discontinuités embryologiques ?

Les premières réponses à ces questions, non encore toutes résolues aujourd'hui, ont été permises par l'étude de jeux de données transcriptomiques. Par exemple, Levin 2012 ont montré chez *C. elegans* et des espèces proches que chaque stade correspondait à une pause de l'expression génétique et que chaque transition correspondait à un pic d'expression (Levin *et al.*, 2012). Une étude en 2011 a également montré que c'était le cas chez le Xénope (Yanai *et al.*, 2011).

Les échantillons développementaux, au sein d'une même espèce, sont généralement distincts dans une analyse en composantes principales (ACP), qui sont une représentation d'un jeu de données en *n* dimensions en quelques dimensions visualisables (Anavy *et al.*, 2014). Il est donc théoriquement possible de réaliser le procédé inverse, chez des espèces où les stades développementaux sont peu distincts les uns des autres. Par exemple, chez le Spongiaire *Amphimedon queenslandica*, aux stades développementaux difficilement discernables, les individus peuvent être stadés transcriptomiquement (Anavy *et al.*, 2014)

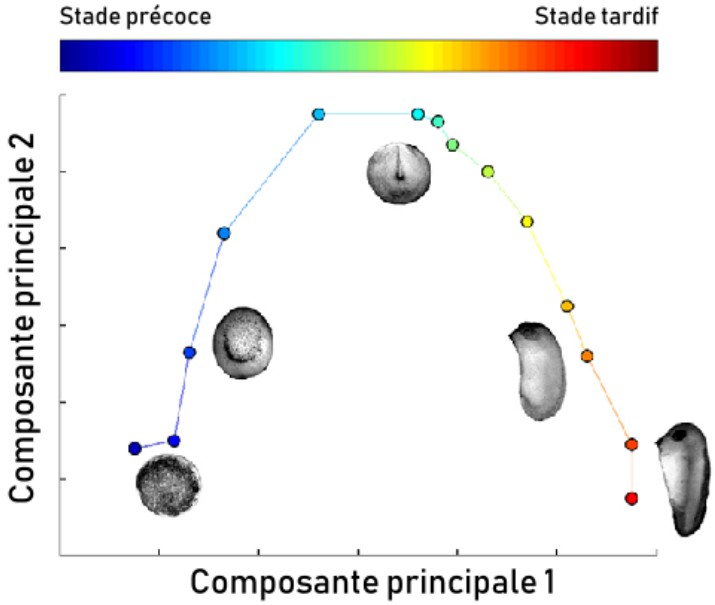


Figure 57 - Analyse en composantes principales de données de RNA-Seq développemental.

Chaque cercle correspond au transcriptome d'un embryon de Xénope d'un stade précoce ou tardif. La couleur indique l'âge de l'embryon. On peut voir que les échantillons développementaux s'ordonnent selon les deux premières composantes de l'ACP. *Adapté de Anavy et al. (2014).*

Une autre utilisation possible de jeux de données RNA-Seq est de suivre l'expression de très nombreux gènes au cours du développement. Ainsi, au lieu de suivre quelques gènes candidats (par exemple, les gènes de la voie des hormones thyroïdiennes et quelques gènes cibles), l'ensemble des gènes exprimés au cours du développement, ainsi que leur profil d'expression respectif, peut être obtenu. Ces résultats à grande échelle peuvent de fait permettre de répondre à une question biologique d'importance : quels sont les gènes dont le profil d'expression est corrélé à celui d'un gène d'intérêt, et qui appartiennent donc au même "cluster temporel d'expression" ? Que révèlent-ils sur les processus biologiques sous-jacents ? À quel moment du développement se mettent en place lesdits processus ?

Répondre à ces questions suppose d'utiliser des outils informatiques de "clustering" adaptés (Bar-Joseph *et al.*, 2012). Ceux-ci sont multiples et ont été amplement développés au cours des quinze dernières années, en rapport avec

l'augmentation des jeux de données associés (Tamayo *et al.*, 1999; Hashimshony *et al.*, 2015; Higuera *et al.*, 2015; Botía *et al.*, 2017; Pantalacci *et al.*, 2017). Un outil populaire est WGCNA (*Weighted Gene Correlation Network Analysis*) qui permet de réaliser des clusters de gènes aux profils d'expression similaires (Zhang and Horvath, 2005).

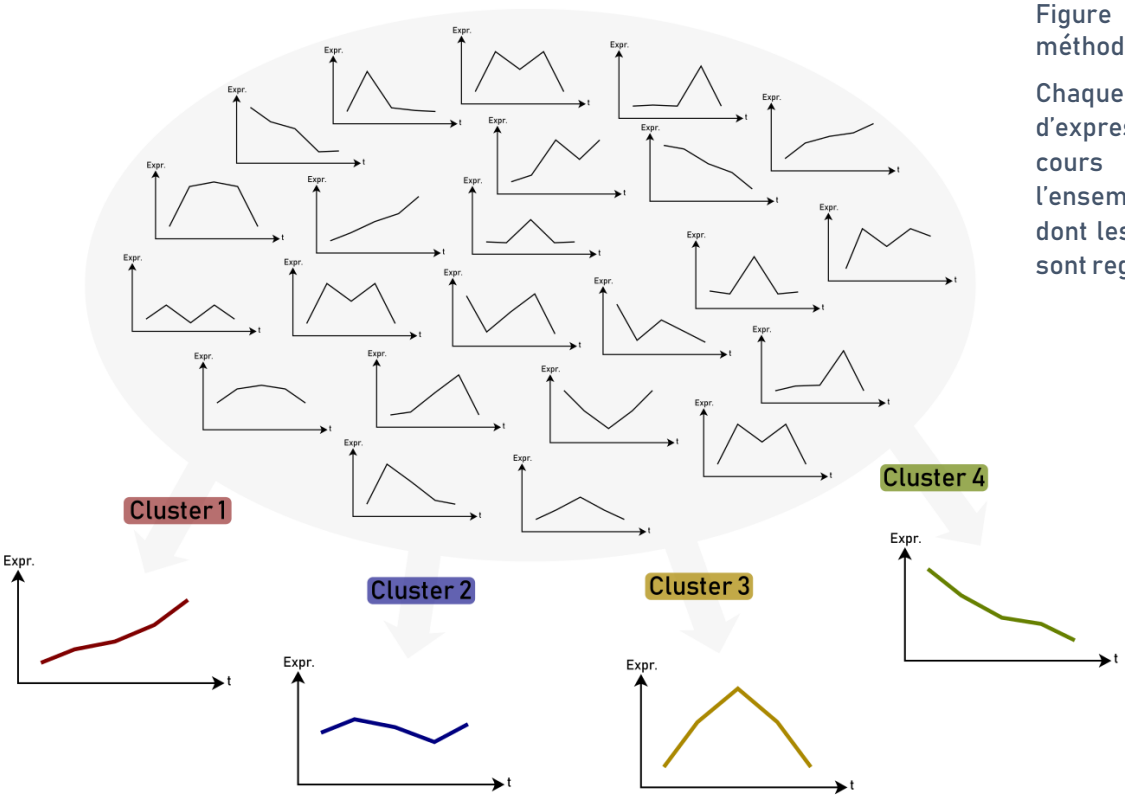


Figure 58 - Principe des méthodes de clustering.

Chaque gène a un profil d'expression (Expr.) propre au cours du temps (t). Parmi l'ensemble des gènes, ceux dont les profils sont similaires sont regroupés en clusters.

c. Le RNA-Seq est une méthode rarement utilisée pour étudier la métamorphose

L'utilisation de la transcriptomique pour étudier des processus temporels comme la différentiation de cellules souches, le développement d'organes spécifiques ou le développement embryonnaire a déjà été documentée à de nombreuses reprises (Cloonan *et al.*, 2008; Li *et al.*, 2010; Papatsenko *et al.*, 2011; Yanai *et al.*, 2011; Tan *et al.*, 2013; Macaulay *et al.*, 2016; Pantalacci *et al.*, 2017). Le développement embryonnaire a notamment été étudié chez les Téléostéens, comme la Morue (micro-arrays, (Kleppe *et al.*, 2014)), le Poisson-zèbre *D. rerio* (micro-arrays, (Mathavan *et al.*, 2005), RNA-Seq (Yang *et al.*, 2013)) et le killi *Austrofundulus limnaeus* (Wagner *et al.*, 2018).

L'utilisation de données transcriptomiques pour l'étude des stades larvaires (post-embryonnaires) des poissons marins est, elle, moins fréquente. Des études ont toutefois été réalisées chez la Daurade (Yúfera *et al.*, 2012), le Flétan et le Bar (Mazurais *et al.*, 2011). En 2011, Mazurais *et al.* alertaient déjà : "*Until now, panoramic time-course analyses of transcript expression during larval development have been carried out in only three marine fish species*". En 2018, sept ans plus tard, seule la Sole a été rajouté aux trois espèces précédentes (Ferraresso *et al.*, 2013). Alors que chez les Protostomiens les études transcriptomiques de la métamorphose abondent (Herranz *et al.*, 2013; Ou *et al.*, 2014; Niu *et al.*, 2016), cette transition d'histoire de vie est peu documentée en transcriptomique sur larve globale chez les Chordés.

POSITIONNEMENT SCIENTIFIQUE DE LA THÈSE

Contexte, problématique et objectifs

Le projet de ma thèse est d'étudier la coloration chez les Vertébrés sous plusieurs aspects.

En premier lieu, ce projet a commencé à un moment où le déterminisme génétique de la pigmentation arrivait à un état de connaissances certain : non seulement les gènes de la pigmentation étaient connus chez la Souris et quelques Mammifères, mais l'étude d'autres modèles de Vertébrés, en particulier poissons Téléostéens, avaient permis une meilleure compréhension des bases génétiques de la pigmentation et des interrelations entre les types de cellules pigmentaires. Par ailleurs, de nombreux génomes de Vertébrés sont aujourd'hui disponibles, et permettent de mieux comprendre l'évolution de l'ensemble des gènes de la pigmentation et des cellules pigmentaires.

L'évolution des cellules pigmentaires est allée de pair avec l'évolution des cellules des crêtes neurales, dont elles dérivent. Plusieurs auteurs ont suggéré que l'acquisition de ce type cellulaire a été permise par les grands réarrangements génomiques ayant eu lieu à la base des Vertébrés et ayant permis l'émergence de nouveaux réseaux de régulation de gènes.

Une première question de ma thèse a donc été : quelles ont été les conséquences des duplications de génome ayant eu lieu à la base des Vertébrés sur l'évolution des gènes de la pigmentation ? De plus, les Téléostéens constituent le groupe de Vertébrés le plus diversifié au niveau pigmentaire, et ont connu une duplication de génome surnuméraire. Une quatrième, au sein des Salmonidés, a également eu lieu plus récemment : quelles ont été les conséquences de ces duplications de génomes sur le répertoire des gènes de la pigmentation ? En quoi l'évolution de ce répertoire diffère-t-elle du reste du génome ?

Par ailleurs, le début de ma thèse a coïncidé avec l'émergence d'un nouvel organisme expérimental : le poisson-clown *Amphiprion ocellaris*. Chez cette espèce de poisson corallien, la pigmentation se met en place au cours du développement selon une métamorphose. Une deuxième partie de mon projet thèse a d'abord constitué en l'étude, à l'aune du travail sur les bases génétiques de la pigmentation réalisé auparavant, de la nature cellulaire et moléculaire de la bande blanche de cette espèce. En effet, la couleur blanche, chez les poissons Téléostéens, peut être due à deux types cellulaires : iridophores et/ou leucophores. Au cours de cette étude, un nouveau gène de la pigmentation, *roshi-a*, a été identifié. J'ai donc essayé d'identifier son origine évolutive ainsi que ses caractéristiques fonctionnelles.

Enfin, la mise en place de la bande blanche au cours du développement est progressive, et apparait liée à une étape clef de l'histoire de vie de cette espèce : la métamorphose. Cette étape dépend, chez de multiples Vertébrés, du métabolisme des hormones thyroïdiennes. J'ai donc essayé de mieux comprendre les liens entre le développement pigmentaire de *A. ocellaris* et le métabolisme des hormones thyroïdiennes, et ainsi de caractériser le déterminisme de la mise en place de la pigmentation.

Modèles d'études et méthodes employées

Une première partie de ce travail de thèse a consisté en une revue systématique de la littérature pour identifier des gènes de la pigmentation. La base de données disponible dans certaines études, datant d'il y a une dizaine d'années, a été enrichie par des gènes récemment caractérisés dans divers processus développementaux. Chaque gène a été associé à un processus développemental en lien avec la pigmentation.

Une fois cette base constituée, l'évolution de l'ensemble des gènes à travers l'histoire des Vertébrés a été réalisée. Cette étude a consisté à recenser l'ensemble des séquences disponibles pour les gènes de la pigmentation et leurs paralogues chez plus de 20 espèces de Deutérostomiens. Au sein des Vertébrés, nous avons échantillonné les grands groupes taxonomiques, avec un échantillonnage plus important des poissons Téléostéens, et deux espèces de Salmonidés au génome récemment séquencé. Les conséquences sur les gènes de la pigmentation, en termes de perte et de rétention, de plusieurs événements de duplication de génome, ont été identifiées.

Les études sur le modèle expérimental *A. ocellaris* ont été réalisées sur des échantillons de peau et sur des embryons à divers stades développementaux et correspondant à la table de développement récemment disponible. Les ARN de chacun de ces échantillons ont été séquencés, ce qui a permis des études transcriptomiques de l'évolution de l'expression génétique au cours du développement, et des différences entre échantillons de peau.

Enfin, la caractérisation du gène de pigmentation nouvellement identifié, *roshia*, a nécessité des approches bio-informatiques variées : caractérisation des domaines protéiques et études phylogénétiques.

RÉSULTATS

RÉSULTATS

CHAPITRE 1

Duplications de génomes et évolution des gènes de la pigmentation chez les Vertébrés

Avant-propos

En 2009, une étude sur l'évolution des gènes de la pigmentation chez les poissons Téléostéens avait montré que les gènes de la pigmentation étaient plus retenus en moyenne que l'ensemble du génome après la Ts3R (WGD spécifique des Téléostéens) (Braasch *et al.*, 2009).

Depuis presque 10 ans, de nombreux génomes de Téléostéens et d'autres Vertébrés ont été publiés. Qui plus est, de nouveaux gènes de la pigmentation ont été décrits. La première partie de ce travail a consisté à recenser l'ensemble de ces gènes à partir de diverses sources bibliographiques. Nous avons ensuite confirmé que les gènes de la pigmentation étaient plus retenus que le reste du génome après la Ts3R et également après une autre duplication de génome chez les Téléostéens, spécifique des salmonidés – la Ss4R. Enfin, nous avons voulu "prendre de la hauteur" par rapport aux Téléostéens et nous avons regardé l'évolution des familles de gènes contenant les gènes de la pigmentation à l'échelle des Vertébrés.

Ces résultats ont été acceptés pour publication dans le journal *G3 : Genes | Genomes | Genetics* le 26 mars 2018. Je vous prie de trouver ci-après le pdf de la version finale, disponible en ligne à l'adresse

http://www.g3journal.org/content/8/5/1795

Les figures supplémentaires, que je n'ai pas jointes ici, peuvent être téléchargées au lien suivant :

https://ndownloader.figshare.com/files/10912016

G3 Genes | Genomes | Genetics

INVESTIGATION

Teleost Fish-Specific Preferential Retention of Pigmentation Gene-Containing Families After Whole Genome Duplications in Vertebrates

Thibault Lorin,*[,1] Frédéric G. Brunet,* Vincent Laudet,† and Jean-Nicolas Volff*

*Institut de Génomique Fonctionnelle de Lyon, École Normale Supérieure de Lyon, UMR 5242 CNRS, Université Claude Bernard Lyon I, Université de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France and †Observatoire Océanologique de Banyuls-sur-Mer, UMR CNRS 7232 BIOM; Sorbonne Université; 1, Avenue Pierre Fabre, 66650 Banyuls-sur-Mer, France

ORCID IDs: 0000-0001-5145-8925 (F.G.B.); 0000-0003-4022-4175 (V.L.); 0000-0003-3406-892X (J.-N.V.)

ABSTRACT Vertebrate pigmentation is a highly diverse trait mainly determined by neural crest cell derivatives. It has been suggested that two rounds (1R/2R) of whole-genome duplications (WGDs) at the basis of vertebrates allowed changes in gene regulation associated with neural crest evolution. Subsequently, the teleost fish lineage experienced other WGDs, including the teleost-specific Ts3R before teleost radiation and the more recent Ss4R at the basis of salmonids. As the teleost lineage harbors the highest number of pigment cell types and pigmentation diversity in vertebrates, WGDs might have contributed to the evolution and diversification of the pigmentation gene repertoire in teleosts. We have compared the impact of the basal vertebrate 1R/2R duplications with that of the teleost-specific Ts3R and salmonid-specific Ss4R WGDs on 181 gene families containing genes involved in pigmentation. We show that pigmentation genes (PGs) have been globally more frequently retained as duplicates than other genes after Ts3R and Ss4R but not after the early 1R/2R. This is also true for non-pigmentary paralogs of PGs, suggesting that the function in pigmentation is not the sole key driver of gene retention after WGDs. On the long-term, specific categories of PGs have been repeatedly preferentially retained after ancient 1R/2R and Ts3R WGDs, possibly linked to the molecular nature of their proteins (*e.g.*, DNA binding transcriptional regulators) and their central position in protein-protein interaction networks. Taken together, our results support a major role of WGDs in the diversification of the pigmentation gene repertoire in the teleost lineage, with a possible link with the diversity of pigment cell lineages observed in these animals compared to other vertebrates.

KEYWORDS
pigmentation
chromatophores
vertebrates
teleost
whole-genome duplication

Pigmentation is a highly diverse trait in the animal kingdom, and its spectacular and intriguing features have always been an object of interest and fascination both for naturalists and the general public. Back then, Alfred R. Wallace has been among the first in the 19th century to understand pigmentation (called "coloration" at the time) in an adaptive context, and his classification of the different pigmentation functions remains largely accepted today (Wallace 1877; Caro 2017). What Wallace could not know at this time was the cellular and genetic bases of pigmentation, a topic that has been much studied since then, especially in vertebrates.

Except for structural colors such as in bird feathers, vertebrate pigmentation is mostly due to a single group of precursor cells that is a vertebrate-specific innovation: the neural crest cells (Green *et al.* 2015). While this transient group of cells emerging from the dorsal part of the neural tube has many derivatives during development (such as cranial cartilage, bone and connective tissue, adrenergic cells, sensory neurons), one category of neural crest-derived cells is of paramount importance in the context of pigmentation diversity: the pigment cells (Lapedriza *et al.* 2014). Among vertebrates, teleost fish harbor the highest number of pigment cell types. While

Copyright © 2018 Lorin *et al.*
doi: https://doi.org/10.1534/g3.118.200201
Manuscript received March 1, 2018; accepted for publication March 26, 2018; published Early Online March 29, 2018.
This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Supplemental material available at Figshare: https://doi.org/10.25387/g3.6061370.
[1]Author for correspondence: Institut de Génomique Fonctionnelle de Lyon, 32 Avenue Tony Garnier, 69007 Lyon, France. E-mail: thibault.lorin@ens-lyon.org

mammals only have melanocytes, teleosts have more than eight different cells (sub)types described: black melanocytes, silvery iridophores and possibly also iridophore-like cells, orange-yellow xanthophores, red erythrophores, whitish leucophores, blue cyanophores, dichromatic erythro-iridophores and cyano-erythrophores, and fluorescent chromatophores (Bagnara and Hadley 1973; Fujii 1993; Goda and Fujii 1995; Goda *et al.* 2011, 2013; Schartl *et al.* 2016). New pigment cell types are still being described in teleosts (Wucherer and Michiels 2012; Goda 2017) (see Figure S1) and the already large list of genes known to affect pigmentation in vertebrates is constantly increasing (see for instance (Mallarino *et al.* 2016)). Most pigmentation genes characterized so far in vertebrates come from mouse mutant studies (Hoekstra 2006), but others have been identified more recently in birds and teleost fish including zebrafish and medaka, which are particularly useful species to study pigment cell types absent from mammals (Roulin and Ducrest 2013; Lopes *et al.* 2016; Mundy *et al.* 2016; Frohnhöfer *et al.* 2016; Nüsslein-Volhard and Singh 2017). Most melanin pigmentation genes found in mammals are also involved in pigmentation in teleosts, suggesting conservation of an ancestral melanin production pathway in vertebrates (Rawls *et al.* 2001; Braasch *et al.* 2008).

The evolutionary origin of the neural crest, which has sometimes being referred to as the "4th germ layer in vertebrates" (Gilbert 2000), is not clear. It has been suggested that the evolutionary emergence of the neural crest and the diversification of its cellular derivatives including pigment cells has required many changes in gene regulatory networks. This has been proposed to be linked to the occurrence of two rounds of whole-genome duplications (WGDs) called 1R/2R early during vertebrate evolution after their split from urochordates (Davidson 2006; Green *et al.* 2015). Such events of genome duplications are revealed by the presence of homeologous chromosomes containing sets of paralogous genes known as "ohnologs" (Wolfe 2001; Turunen *et al.* 2009), in reference to Susumu Ohno and his seminal work on the genome duplication hypothesis (Ohno 1970, 1999). During subsequent evolution, additional WGDs occurred independently in different vertebrate lineages (Session *et al.* 2016). At the basis of teleost fish, a third round of WGD has taken place ca. 300 million years ago (teleost-specific genome duplication or Ts3R) (Meyer and Schartl 1999; Woods *et al.* 2000; Jaillon *et al.* 2004; Brunet *et al.* 2006). Within teleosts, Ts3R was followed by more recent lineage-specific events including a fourth round of WGD that occurred in salmonids ca. 100 million years ago (Salmonid-specific genome duplication or Ss4R; (Macqueen and Johnston 2014; Berthelot *et al.* 2014; Lien *et al.* 2016)) and genome tetraploidization in cyprinids (Li *et al.* 2015). The teleost lineage is thus considered to be prone to WGDs, even if such events also occurred in non-teleost ray-finned fishes, *e.g.*, sturgeons (Havelka *et al.* 2013).

Our aim is to investigate how different events of genome duplications have contributed to the evolution and diversification of the pigmentation gene repertoire in teleost fish and other vertebrates. Indeed, many pigmentation genes have been duplicated in teleosts by the Ts3R WGD (Braasch *et al.* 2008, 2009a, 2009b). Previous studies only focused on the impact of Ts3R on the evolution of this pathway (Braasch *et al.* 2008, 2009a). Here, we have compared the impact on the pigmentation gene repertoire of the ancestral 1R/2R vertebrate duplications and the teleost Ts3R and Ss4R WGDs. We show that pigmentation genes and their paralogs, if any, have been more frequently retained as duplicates than other genes after the fish-specific Ts3R and Ss4R WGDs, but not after the early vertebrate 1R/2R, with a possible link with the diversification of pigment cell types specifically found in teleost fish.

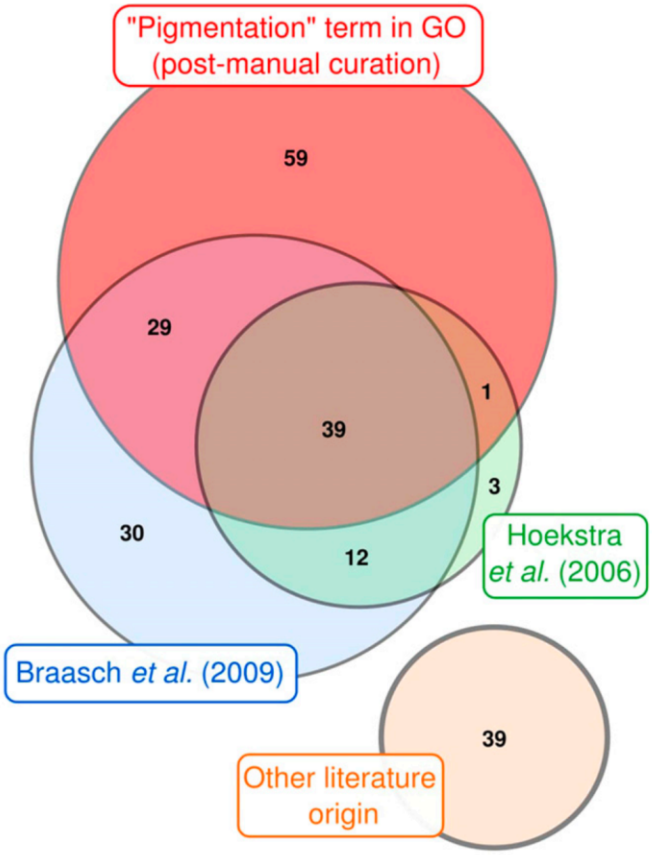


Figure 1 Origin of pigmentation genes studied in this work. Most PGs were either retrieved using the Gene Ontology term "pigmentation" (GO: 0043473) parsed for vertebrates (taxon id: 7742) or were previously cited in two studies (Hoekstra 2006; Braasch *et al.* 2009a). PGs described in more recent studies were also included ("other literature origin"). Numbers of genes are indicated. The Venn diagram was drawn using the BioVenn website (Hulsen *et al.* 2008).

MATERIAL AND METHODS

Pigmentation genes and their functional classification

Vertebrate *pigmentation genes* (PGs) are defined here as genes that are involved in the differentiation of a neural-crest derived pigment cell in at least one vertebrate species. This strict definition excludes in particular pigment cells present in the retinal pigmented epithelium, an anterior neural plate derivative (Ramón Martínez-Morales *et al.* 2004). According to this definition, one PG can be implicated in pigment cell differentiation in one species but not in others.

To establish the PG dataset used in this study (Figure 1; Table S1), we first retrieved genes annotated under the Gene Ontology (GO) term "pigmentation" (GO: 0043473) in vertebrates (taxon identifier: 7742) (http://www.geneontology.org/, last accessed on March 08th, 2017). We manually verified the correspondence of automatic GO annotations with our definition of a pigmentation gene and, when necessary, excluded some genes from the list. For instance, *sod2* was not kept in our dataset because its deficiency causes red blood cell damage (Friedman *et al.* 2004). As such, it does not fit the definition above. In addition, all genes previously described in two former papers on vertebrate pigmentation were included (Hoekstra 2006; Braasch *et al.* 2009a). PGs from these different sources were sometimes redundant (Figure 1). Finally, we conducted a review of the literature to include PGs from more recent studies (Figure 1; Table S1). Based on previous work,

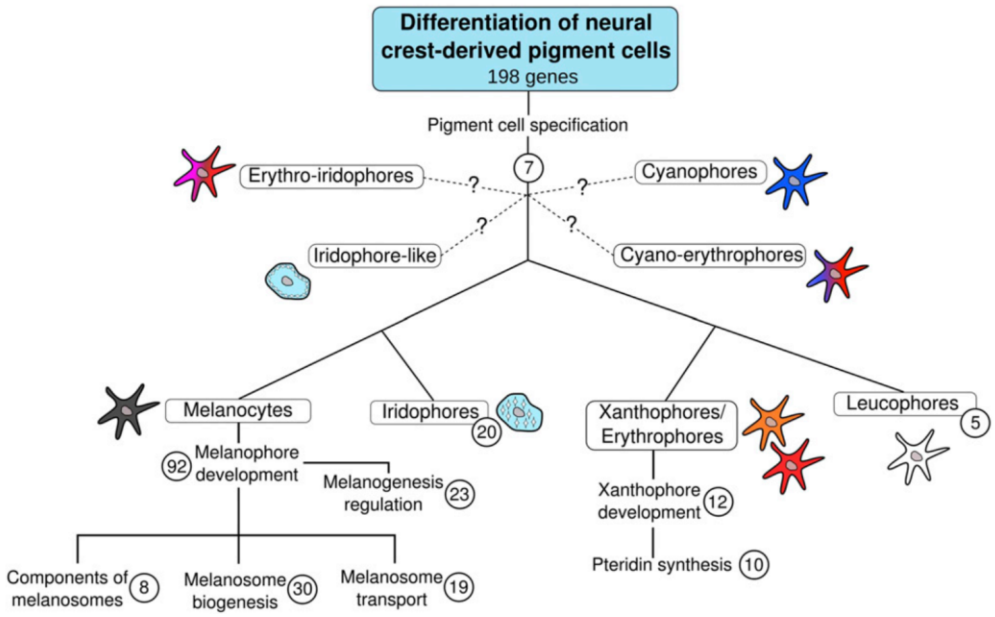


Figure 2 Classification of 198 vertebrate pigmentation genes according to functions and cell types. Numbers of pigmentation genes in (sub)-categories are indicated in circles (some genes can be present simultaneously in different subcategories). Ontogenetic relationships between melanocytes, iridophores, xanthophores/erythrophores and leucophores are as proposed by (Kimura *et al.* 2014). Functional classification within melanocytes and xanthophores is adapted from (Braasch *et al.* 2009a). Question marks indicate that these newly described pigment cell types may originate from neural crest (NC), but that there is no experimental evidence for that yet.

pigmentation genes were classified into categories depending on their functions (Table S1; Figure 2) (Braasch *et al.* 2009a).

Phylogenetic analysis

We define here as a *gene family* the genes that are derived by duplication and speciation from a pre-1R/2R ancestral deuterostome single-copy gene. After 1R/2R but before Ts3R, a gene family can be either multigenic if the ancestral gene has been kept as duplicates after 1R/2R, or monogenic if no duplicate has been maintained. We define as *pigmentation gene-containing family* (PGCF) a gene family containing at least one pigmentation gene. This implies that a PGCF can also contain genes that are not involved in pigmentation (experimental evidence), or for which no pigmentation function has been documented so far (lack of data) (see Figure S2 for a graphical representation of these definitions).

To assess the evolutionary history of each PGCF, and particularly how it was shaped by the four rounds of WGDs in vertebrates studied thereafter, sequences from 22 vertebrate genomes representing the major vertebrate lineages were analyzed. Eight sarcopterygian genomes including seven tetrapods were studied (*Latimeria chalumnae*, *Xenopus tropicalis*, *Anolis carolinensis*, *Gallus gallus* as well as the mammals *Ornithorhynchus anatinus*, *Monodelphis domestica*, *Mus musculus* and *Homo sapiens*). Thirteen actinopterygians (ray-finned fish) genomes were included in the study, encompassing a wide diversity of clades, including salmonids to study Ss4R (Near *et al.* 2012): the non-teleost spotted gar *Lepisosteus oculatus* (the closest outgroup to teleosts with sequenced genome that did not experience the Ts3R (Amores *et al.* 2011)) and the teleost species *Danio rerio*, *Astyanax mexicanus*, *Gadus morhua*, *Gasterosteus aculeatus*, *Tetraodon nigroviridis*, *Tak.fugu rubripes*, *Poecilia formosa*, *Xiphophorus maculatus*, *Oryzias latipes*, *Oreochromis niloticus*, *Oncorhynchus mykiss* and *Salmo salar*. To encompass the broad diversity of vertebrates we additionally included the genome of a chondrichthyan (cartilaginous fish), the elephant shark *Callorhinchus milii*. At the time of analysis, data from the lamprey genome were too incomplete to be included in the study. Sequences from five non-vertebrate deuterostomes that diverged from the vertebrate lineage before the 1R/2R WGDs were used as outgroups in the molecular phylogenies: the two urochordates *Ciona intestinalis* and *C. savignyi*, the cephalochordate *Branchiostoma floridae* (amphioxus) and the more distantly related ambulacrarians *Strongylocentrus purpuratus* (echinoderm) and *Saccoglossus kowalevskii* (hemichordate).

For each species, protein sequences of the longest isoform were retrieved from Ensembl v86 (www.ensembl.org; Oct. 2016) and NCBI (https://www.ncbi.nlm.nih.gov/, last accessed March 08th, 2017). When an Ensembl sequence was missing from the Ensembl annotation pipeline, we manually tested on the genome assembly (using Ensembl BLASTn method with the closest ortholog as query) whether this was due to gene loss or annotation skews. In the latter case, we used the FGENESH+ program of the Softberry suite (http://www.softberry.com/) to extract the gene sequence and predict the protein sequence.

Sequence alignments were performed using Clustal Omega v1.2.1 (Sievers and Higgins 2014) with default parameters and were individually manually curated for each PGCF (Blouin *et al.* 2009). Each alignment was then analyzed using the software prottest3 v.3.4.2 (Darriba *et al.* 2011) and the evolutionary model for tree-building was selected based on the Bayesian information criterion (Luo *et al.* 2010). Tree building was conducted with the best fitting model using the Maximum Likelihood method implemented in PhyML v3.1 (Guindon *et al.* 2010) with SH-aLRT support (Le and Gascuel 2008; Anisimova *et al.* 2011). The best-fitting model for each alignment is available (Table S2).

In order to assess if a duplication event was likely to be due to Ts3R and not to a small-scale duplication (SSD), and to identify large paralogons supporting Ts3R evidence, we conducted for each gene a synteny analysis using the Genomicus web browser (v86; http://genomicus.biologie.ens.fr/genomicus-86.01/cgi-bin/search.pl) (Louis *et al.* 2015).

Finally, as genes can be described in different organisms with different names, we used the unified nomenclature system for humans provided by the HGNC Database (Gray *et al.* 2015) to avoid nomenclature ambiguities.

Retention rate calculation

There are two main methods to calculate retention rates (RRs) of duplicated copies after WGD (Brunet *et al.* 2017). Hereafter, the RR

of duplicated copies was calculated as previously done by several authors (Figure S3) (Dehal and Boore 2005; Berthelot *et al.* 2014; Lien *et al.* 2016). When two ohnologs are present after one event of WGD, the duplicated copy has been maintained and its RR is 1. If only one copy is present after WGD, the duplicated copy has been lost and its RR is 0. We can calculate the average RR if several genes from a same family have undergone a same WGD (Figure S3).

Pathway enrichment analysis and protein-protein interaction networks

For each gene (PGs and their corresponding vertebrate 1R/2R-ohnologs, if any), we retrieved the UniProt ID of the human ortholog in the UniProtKB/Swiss-Prot database (www.uniprot.org/, The UniProt Consortium 2015, last accessed March 08th, 2017). Gene IDs were used to test for enrichment in Gene Ontology (GO) terms with g:Profiler (http://biit.cs.ut.ee/gprofiler/) (Reimand *et al.* 2016) using default parameters and the human genome as background. The significance threshold for enrichment was set at 0.05 (corrected p-value for multiple testing with Benjamini-Hochberg FDR) and the "Hierarchical filtering" was set at "Best per parent group" to obtain only hierarchically highest GO terms. The number of interactions for each protein was obtained from the manually curated "Binary interactions" tab in UniProt.

Data availability

All alignments and trees are available at FigShare; sequences were obtained from publicly available databases. All data necessary for confirming the conclusions presented in the article are represented fully within the article and its supplementary material. Supplemental material available at Figshare: https://doi.org/10.25387/g3.6061370.

RESULTS

Vertebrate pigmentation genes and pigmentation gene-containing families

We analyzed Gene Ontology databases and the existing literature to obtain a list of 198 vertebrate PGs that correspond to 181 PGCFs (the number of genes here is estimated after 1R/2R but before Ts3R; see Methods for details). Seventeen PGCFs contained two PGs, sometimes with different pigmentary functions (see Table S3). In contrast, most families (164/181) possessed only one PG, with its "non-PG" 1R/2R-ohnolog(s) (if any) having to the best of our knowledge no function described to date in the development of neural crest derived-pigment cells (either experimental evidence or absence of data).

The evolutionary history of nine gene families (*asip*, *foxd3*, *gja5*, *krt2a*, *rab8a*, *slc2a11*, *ugt1a8*, *xdh* and *zic2*) was too ambiguous to be included in our study because of the presence of many independent segmental duplication events and low phylogenetic resolution (this was previously reported for the *asip* gene family (Cal *et al.* 2017b); data not shown). We also excluded the *bcl2l11*, *dcfb103a* and *krtap21-1* families because these genes are either tetrapod-specific (*bcl2l11*) or eutherian-specific (*dcfb103a*, *krtap21-1*). The final analyzed dataset therefore consisted of 184 PGs and 144 non-PG paralogs for a total of 328 total genes within 169 PGCFs.

After phylogenetic and genomic analyses of 27 deuterostomes species (including 22 vertebrates), orthology and paralogy relationships were established for all the 328 genes. A schematic representation of the occurrence of all genes with copy numbers (corresponding to a total of 9,112 sequences) is provided for each studied species (see Figure S4A-J and an example case Figure S5).

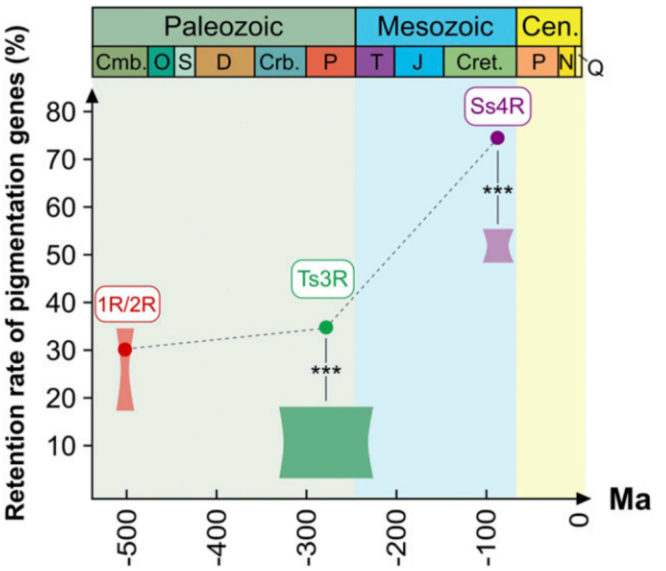


Figure 3 Retention rates for the whole pigmentation gene repertoire after successive WGDs in vertebrates. At each WGD event, heights of shaded areas represent the genome-wide retention rates estimates based on multiple studies (see text for detail); widths of shaded areas represent the temporal estimates available in the literature for each WGD event. Dots represent the retention rate for each WGD event when considering only PGs. Pigmentation repertoire retention rate is higher than genomic average after Ts3R- and Ss4R-WGDs. ***: $P < 10^{-7}$ (chi2 test, see text for values).

Retention rates of pigmentation genes after the fish-specific WGDs (Ts3R and Ss4R) are higher than genome average

We first observed that the global retention rate (RR) of PGs after Ts3R was 34.2% (= 63/184) (Figure 3; Table 1). Interestingly, this rate was significantly higher than the range of values reported for the whole genome, *i.e.*, 4–18% depending on the datasets used (Jaillon *et al.* 2004; Woods *et al.* 2005; Brunet *et al.* 2006; Howe *et al.* 2013; Inoue *et al.* 2015) ($P = 1.8 \times 10^{-8}$, chi2 test). Of note, values published in earlier studies were recalculated if needed before comparison to be consistent with our calculation method. This indicated that PGs have been globally more frequently retained as duplicates than other genes after the Ts3R WGD.

After the salmon-specific Ss4R WGD, the RR was 75.2% for PGs (182/242; Figure 3; Table 1). Again, this value was higher than the value reported for the whole genome (48–55%, $P = 4.3 \times 10^{-10}$, respectively, chi2 test) (Berthelot *et al.* 2014; Lien *et al.* 2016). Hence, PGs have also been preferentially retained after the Ss4R duplication compared to other genes.

We eventually assessed whether the retention rates after Ts3R and Ss4R would differ between genes from different functional categories in pigmentation (see Figure 2). There was no significant group effect after both Ts3R and Ss4R (p-value of Wald test = 0.30 and 0.34, respectively) (Figure S6A and 6B).

Pre-1R/2R pigmentation genes ancestors are not preferentially retained after 1R/2R

The ancestral vertebrate double event of WGD (1R/2R) is assumed to have played a major role in the evolution of vertebrate gene families (Dehal and Boore 2005; Xie *et al.* 2016). We therefore investigated the consequences of 1R/2R on PG evolution (note that the respective effects

Table 1 Retention rates after Ts3R and Ss4R of genes present in pigmentation gene-containing families

WGD event	Overall PG retention rate	PGs in monogenic PGCFs	PGs in multigenic PGCFs	Non-PGs in multigenic PGCFs
Ts3R	34.2% (63/184)	16.9% (13/77)	46.7% (50/107)	37.1% (52/140)
Ss4R	75.2% (182/242)	72.2% (65/90)	73.2% (112/153)	71.5% (133/186)

of 1R *vs.* 2R cannot be differentiated particularly due to the relatively short period of evolution between these two ancient WGD events). However, whether the ancestral pre-1R/2R genes had a function or not in pigmentation is not documented, hence the RR after 1R/2R could not be calculated for PGs *per se*. Instead, we calculated the post-1R/2R retention rates for pre-1R/2R pigmentation gene ancestors. We obtained a value of 30.4% (average RR for 169 PGCFs), which is well in the range of those previously reported for the complete genome (17.0–34.5% ; (Dehal and Boore 2005; Xie *et al.* 2016) (Figure 3).

If at least two post-1R/2R ohnologs are retained, the PGCF is "multigenic". Instead, if there is only one retained gene, the family is "monogenic" and only contains the pigmentation gene without any ohnolog. About half of PGCFs (92/169) (54.4%) were multigenic (Figure 4, case A). In these 92 multigenic families, 107 PGs were found (15 families harbored two PGs; Table S3) as well as 144 non-PGs. In contrast, 45.6% (77/169) of PGCFs were monogenic. When comparing the retention rates for the different genes in functional pigmentation categories (Figure 2), we found that the "melanosome biogenesis" group was enriched in monogenic PGCFs and therefore had a lower RR after 1R/2R compared to other groups (p-value of Wald test = 0.015) (Figure S6C).

Only PGs from 1R/2R multigenic families have higher retention rates after Ts3R compared to genomic average

We then assessed whether the pigmentation genes in post-1R/2R multigenic families had more chance to be retained as duplicates also after Ts3R than post-1R/2R singletons. While the overall RR for PGs was 34.2% after Ts3R (see above), we showed that genes in multigenic PGCFs had a higher retention rate after Ts3R (50/107 = 46.7%; Table 1; Figure 4, case C; $P = 2.6 \times 10^{-5}$, chi2 test) than PGs in monogenic PGCFs (13/77 = 16.9%; Table 1; Figure 4, case B). After Ts3R, the RR of PGs in monogenic 1R/2R PGCFs was within the reported genome range of retention (Figure 4, case B) while Ts3R RR for PGs in multigenic 1R/2R PGCFs was significantly higher (Figure 4, case C; $P = 2.4 \times 10^{-14}$, chi2 test). Thus, the global high RR for PGs after Ts3R is principally due to the high RR of PGs from 1R/2R multigenic families.

PGs from both monogenic and multigenic 1R/2R PGCFs have similar high retention rates after Ss4R

We subsequently compared the RR after Ss4R of PGs from 1R/2R monogenic *vs.* multigenic gene families. We showed that PGs that were in monogenic 1R/2R PGCFs (Table 1; Figure 4, cases D and E) had a RR higher than average similar to that of PGs in multigenic PGCFs (Table 1; Figure 4, cases F and G): [19+46]/[26+64] = 72.2% *vs.* [40+72]/[55+98] = 73.2%, respectively ($P = 0.98$, chi2 test). Thus, not only PGs from multigenic 1R/2R PGCFs are preferentially retained after Ss4R - as observed for Ts3R - but also PGs from 1R/2R monogenic families - in contrast to Ts3R.

Additionally, we assessed a possible correlation between Ts3R and Ss4r RRs for PGs in PGCFs. We showed that genes retained after Ts3R (Figure 4, cases D and F) were not more retained after Ss4R than genes not retained after Ts3R (Figure 4, cases E and G): [46+40]/[64+55] = 72.3% *vs.* [19+72]/[26+98] = 73.4%, respectively ($P = 0.96$, chi2 test). However, both have a higher RR than genomic average. Therefore, PGs have been retained at a higher rate than genomic average after Ss4R independently from the copy number of their gene family after both 1R/2R and Ts3R.

PGs and their non-pigmentary paralogs have similar high RRs after teleost-specific WGDs

We have then compared within 1R/2R multigenic gene families the RR of PGs with that of their non-pigmentary paralogs after both Ts3R and Ss4R. We first showed that the retention rate after Ts3R for non-pigmentary paralogs (37.1% = 52/140; Table 1; Figure 4, case C'; Table 1) was as high as the retention rates for PGs in general (34.2% = 63/184; $P = 0.67$, chi2 test) and particularly for PGs in multigenic PGCFs (46.7% = 50/107; Table 1; Figure 4, case C; $P = 0.17$, chi2 test). This rate was higher than the range of values reported for the whole genome ($P = 6.2 \times 10^{-9}$, chi2 test). Thus, PGs and their corresponding paralogs in multigenic PGCFs are both highly retained after Ts3R.

After Ss4R, as previously stated, PGs were retained at 73.2% (Table 1; Figure 4, cases F and G). Similarly, their paralogs were retained at [62+71]/[82+104] = 71.5% (Table 1; Figure 4, cases F' and G'). The values obtained were both higher than the global genomic range of retention (Figure 4, cases F' and G'). Hence, PGs and their paralogs in multigenic PGCFs are equally retained after both Ts3R and Ss4R.

Genes in multigenic PGCFs might be at the core of protein-protein interaction networks

Previous studies have shown that genes that are sensitive to dosage balance and those encoding proteins with a high number of interactions are more likely to be retained after WGD events (Freeling and Thomas 2006; McGrath *et al.* 2014; Bright *et al.* 2017). Hence, we hypothesized that genes in families with high RRs would have more "central" molecular functions (*e.g.*, transcription factors or intracellular signaling actors such as protein kinases) than PGs in families with low retention rates. Accordingly, enrichment analyses in Gene Ontology terms on genes in multigenic PGCFs showed that the most significant "Molecular Function" term was "transcription regulatory region DNA binding" (Table S4A). In contrast, PGs in monogenic PGCFs were mostly enriched in metabolic enzymes (Table S4B).

In addition, we estimated the number of interaction partners for each pigmentation protein using the UniProtKB/Swiss-Prot database (self-interactions being excluded). We showed that products of genes in multigenic PGCFs held more interactions on average than PGs in monogenic PGCFs (n = 5.5 *vs.* n = 2.2 on average, respectively; p-value of Welch *t*-test = 0.003). PGs and non-PGs in multigenic PGCFs, which have similar retention rates, also display similar average interaction number (n = 7.6 and n = 3.8, respectively; p-value of Welch *t*-test = 0.08).

DISCUSSION

In this study, we first showed that pigmentation genes were on average more retained than the genome average after both teleost-specific Ts3R and Ss4R WGDs, but not after the 1R/2R WGDs. Non-pigmentary post-1R/2R ohnologs of pigmentation genes are also more retained than the rest of the genome after WGDs in teleosts, suggesting that the function in pigmentation is not the only factor determining gene retention. After the

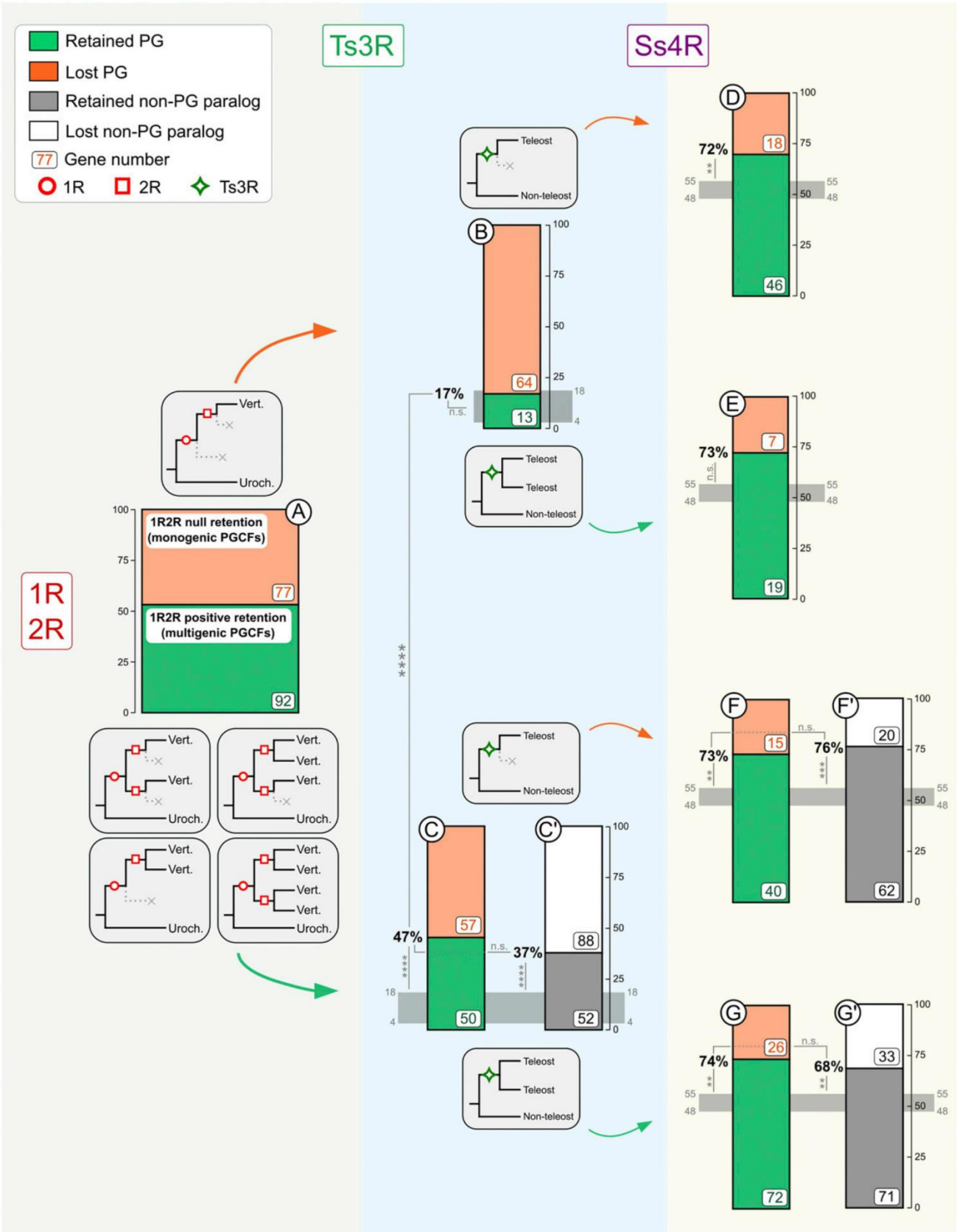


Figure 4 Retention rates for pigmentation genes from monogenic and multigenic families after successive WGDs in vertebrates. Green and orange boxes (respectively black and white) represent number of genes that are whether retained or not at the considered WGD for PGs (respectively non-PGs paralogs in multigenic PGCFs). Total numbers of genes per dataset are indicated. For 1R/2R, the bar does not represent retention rates but the 77 monogenic PGCFs *vs.* the 92 multigenic PGCFs. Retention rates are indicated (between 0 and 100%) for Ts3R and Ss4R.

evolutionary ancient 1R/2R and Ts3R WGDs, specific categories of genes have been repeatedly preferentially retained, which might have more central molecular functions compared to genes in monogenic PGCFs.

Pigmentation genes and retention rates

Pigmentation is a trait determined by many molecular actors and a strict definition of a "pigmentation gene" seems impossible. In particular, genes that are involved in the patterning of pigmentation such as *lvrn/taqpep* are not necessarily involved in the development of chromatophores, but rather influence the development of surrounding cell territories (Kaelin *et al.* 2012). In this work, we have studied genes that are implicated in the development of pigment cells as defined by (Schartl *et al.* 2016), *i.e.*, cells whose embryonic origin is the neural crest. Pigment cells of the retinal pigment epithelium, as they are not involved in body coloration and derive from the optic cup - a neuroepithelium derivative - were not studied here. We ascertained that genes qualified as "non-pigmentary paralogs" were not implicated in pigmentation by examining the annotations of the available mutants for these genes in the Mouse Genome Informatics (http://www.informatics.jax.org/allele) as well as in the Zebrafish Information network databases (last accessed February 9th, 2018), two of the most comprehensive mutant databases in vertebrates. Most presumed "non-pigmentary" post-1R/2R paralogs of PGs had indeed no "pigmentation" mutant described, confirming the robustness of the dichotomization "pigmentation"/"non-pigmentation" used here (data not shown). The lack of functional evidence can nevertheless either reflect a "real" absence of effect on pigmentation, or an artifactual lack of evidence due to a small number of studies.

A putative ancestral pigmentation gene repertoire in jawed vertebrates

Mammals only have one pigment cell type (melanocytes), while birds, lepidosaurians and amphibians have four to five. As already suggested by others, early gnathostomes (jawed vertebrates) were likely to possess most of these pigment cell types (Lapedriza *et al.* 2014). This implies that secondary losses of pigment cell lineages occurred in mammals, while new pigment cell types – *e.g.*, cyanophores – arose only in teleosts, with an important diversity of pigmentation patterns (Bagnara *et al.* 2007). In addition to the diversity in terms of pigmentation patterns, teleost fish also have the highest number of pigment cell types in vertebrates.

Our phylogenomic analysis of 198 PGs and their corresponding families confirmed the ancestry of the genetic control of the melanocyte/melanophore lineage in jawed vertebrates. This is supported by the conservation of many genes such as *mitf*, *mc1r*, *sox10*, *kit/kitlg* as well as genes from the Wnt pathway in both sarcopterygians and actinopterygians (Rawls *et al.* 2001; Braasch *et al.* 2009a) (Figure S7). The information on the functional conservation of underlying genes implicated in the development of other pigment cells absent from mammals is scarcer in the available literature. However, we have shown that genes involved in the formation of leucophores, erythrophores/xanthophores and iridophores are present in different vertebrate lineages, here again suggesting common ancestry (see corresponding PGCFs in Table S1 and Figure S4). These include the iridophore marker *foxd3* and the leucophore and xantho-/erythrophore genes *slc2a11* and *slc2a15* (Curran *et al.* 2010; Kimura *et al.* 2014). As previously shown by others, the xanthophore marker *somatolactin* (*smtl*) is present in actinopterygians and in the coelacanth but not in other tetrapods, suggesting a secondary loss in this lineage (Jiang *et al.* 2011).

Mutations of some pigmentation genes in non-mammalian species affect more pigment cell types than the sole melanocytes. This suggests that there are close ontogenetic relationships between pigment cell types. For instance, the transcription factor Mitf, which is involved in melanocyte formation in mammals, interacts with Foxd3 to determine the development of melanocytes and iridophores in zebrafish (Curran *et al.* 2010). In zebrafish and medaka, evidence of common progenitors has been found for leucophores and xanthophores on the one hand, and for iridophores and melanocytes on the other hand (Curran *et al.* 2010; Kimura *et al.* 2014). More functional studies are needed in other lineages including amphibians or lepidosaurians to assess the ancestry of ontogenetic relationships between pigment cell types in jawed vertebrates.

The presence of ancestral genetic networks driving the formation of pigment cell lineages early in vertebrate evolution is strongly supported by the conservation of pigmentary functions for genes in different vertebrate lineages. For instance, the *asip* locus has a conserved role in dorso-ventral pigment asymmetry in vertebrates, which results from a switch of melanin type in mammals and from the differential distribution of several types of chromatophores – including melanocytes – in both spotted gar and zebrafish, so presumably in the ancestor of ray-finned fish (Ceinos *et al.* 2015; Cal *et al.* 2017a). *Sox10* mutants present defects in melanocytes in mammals and in all pigment cell types in zebrafish (the corresponding mutant was incidentally coined *colorless*), suggesting a central function in the development of these neural crest derivatives (Rawls *et al.* 2001). Thus, the results provided here on the phylogenetic distribution within vertebrates of genes involved in pigmentation may provide a framework for the evolutionary and developmental understanding of the underlying genetic basis of vertebrate (and teleost) pigmentation.

Successive rounds of WGDs have extended the vertebrate pigmentation gene repertoire especially in teleost fish

While two first rounds of WGDs (1R/2R) have occurred early during the course of vertebrate evolution around 500 million years ago (Mya) (Putnam *et al.* 2008), a 3rd round of WGD specific to ray-finned fish (Ts3R) has taken place between 225 and 333 Mya before the teleost radiation (Hurley *et al.* 2007; Santini *et al.* 2009; Near *et al.* 2012). A 4th round of WGD (Ss4R) occurred in the salmonid fish lineage about 80-100 Mya (Macqueen and Johnston 2014; Berthelot *et al.* 2014; Lien *et al.* 2016). Such WGDs are drastic events leading to the doubling of genomes, which can strongly contribute to evolution at both gene and organismal levels (Conrad and Antonarakis 2007). As already suggested for Ts3R (Braasch *et al.* 2009a), we show that these WGDs have increased the evolvability of the pigmentation gene repertoire and therefore have contributed to the diversification of pigmentation cell types and patterns in vertebrates. This phenomenon might be especially true for teleost fish: for Ts3R and Ss4R, the retention rates obtained for the pigmentation repertoire are higher than genomic average, respectively), while they are in the genome-wide average for 1R/2R (Figure 3). The high pigmentation diversity observed in teleost fish in general, both in terms of chromatophore types and pigment patterns, is thus correlated to an extension of the repertoire of PGs after WGDs.

Horizontal gray bands represent the genome average retention rate observed in the literature for Ts3R (4–18%) and Ss4R (48–55%). *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$; ****: $P < 0.0001$; n.s.: not significant at a 0.05 level (chi2 test).

In salmonids, to the best of our knowledge, there is no obvious link between the Ss4R duplication and higher chromatophore diversity compared to other teleosts. This might be explained by the observed time-lag between WGD and evolutionary diversification in this family (Robertson *et al.* 2017) (but see (Colihueque 2010) for a discussion on the diversity of salmonid pigmentation). Of note, more salmonid genomes must be analyzed to avoid bias in the calculation of retention rates.

Pigmentation genes are predestinated to either loss or retention after ancient WGDs

After analyzing the long-term evolution of pigmentation genes after ancient WGDs, we observed that those that have been already retained after 1R/2R are also preferentially kept as duplicates after Ts3R. This indicated that after successive WGDs and long-term evolution, some pigmentation gene duplicates are recurrently retained, while others are preferentially lost. However, this phenomenon was not observed for the salmonid-specific Ss4R, after which PGs were retained at similar higher rates independently from their retention rate after previous WGDs. This is consistent with what has been previously found in salmon at the genome-wide level (Lien *et al.* 2016) but not in the rainbow trout (Berthelot *et al.* 2014). A potential reason to explain this apparent discrepancy could be the relatively recent occurrence of this event (ca. 96 Mya), with gene loss associated with the rediploidization process being still ongoing, at least in the Atlantic salmon (Robertson *et al.* 2017).

While general factors governing the destiny of genes after duplication events are still being debated (Rogozin 2014), several drivers of high retention after WGDs have been already described. Genes encoding transcription factors, signal transduction proteins and other proteins implicated in developmental processes are usually retained in excess compared to other genes (Conant and Wolfe 2008). This is not only valid for teleost fish (Steinke *et al.* 2006; Brunet *et al.* 2006; Hufton *et al.* 2008; Kassahn *et al.* 2009; Berthelot *et al.* 2014) but also for other vertebrates (Blomme *et al.* 2006; Session *et al.* 2016), yeast (Davis and Petrov 2005; Guan *et al.* 2007), plants (Seoighe and Gehring 2004; Maere *et al.* 2005; Jiang *et al.* 2013; Li *et al.* 2016) and paramecium (McGrath *et al.* 2014). At a protein family level, this is the case for receptor tyrosine kinase genes, which have been more frequently retained than other genes after 1R/2R and Ts3R (Brunet *et al.* 2016). The retention for transcription factor genes of the *Sox* family is also higher than genome average after Ts3R (Voldoire *et al.* 2017). Accordingly, we showed that pigmentation genes retained as duplicates after 1R/2R were enriched in transcription factor activity (Table S4A) compared to pigmentation genes in monogenic PGCFs (Table S4B).

It has also been proposed in the "dosage balance hypothesis", since all gene product quantities are doubled after a WGD, that the loss of one gene copy would be deleterious, especially if the protein has a central role in the cell with many interactors (Lemos *et al.* 2004; Birchler and Veitia 2007). As such, ohnologs are more often conserved than segmental duplicates because the latter immediately disrupt dosage balance (Makino and McLysaght 2010). In addition, at a genome-wide scale, duplicated genes tend to have more interactions than singleton genes (He and Zhang 2005) and proteins encoded by retained WGD ohnologs have significantly more protein interactions than proteins encoded by genes that are not duplicated (Wu and Qi 2010). Accordingly, we found that proteins encoded by PGs and in multigenic families held more interactions on average than PGs in monogenic families.

Finally, recurrent differential retention rates might be linked to pigmentation functions with different degrees of evolutionary constraints and evolvability: some genes involved in more conserved functions might be more refractory to retention than others. In our analysis, genes involved in "melanosome biogenesis" showed a lower retention rate after 1R/2R. In addition, although not significant, all genes involved in "pigment cell differentiation" (see Figure 2) belonged to multigenic PGCFs (Figure S6C), suggesting that genes that are involved in the cell differentiation process are highly retained.

Non-pigmentary paralogs of pigmentation genes also have high retention rates after teleost WGDs

We have reported here that non-pigmentary paralogs have similar high retention rates than the PGs from the same family. This was observed in multigenic PGCFs not only after the ancestral teleost-specific Ts3R but also after the more recent salmonid-specific Ss4R. This could suggest that the function in pigmentation is not the unique, or even not the main driver of high retention rates in PGCFs. As discussed above, the molecular nature and biochemical properties of a protein are certainly major factors determining the retention rate of a gene after genome duplication. This would explain why paralogous genes from a same gene family, which encode similar proteins, have similar RR independently from their biological function (involvement in pigmentation or not). However, one might expect, as observed for 1R/2R and Ts3R (see above), that a given gene family encoding related proteins will display a conserved bias of retention *vs.* loss after different WGDs. This is not the case after Ss4R, since PGs and their paralogs have been equally highly retained independently from their RR after the more ancient 1R/2R and Ts3R WGDs.

On the other hand, one might also consider that our category "non-pigmentation genes" might include PGs. Some post-1R/2R non-PG genes have been classified as such due to the absence of functional data, some genes might indeed play so far an unknown role in pigmentation. In addition, our functional annotation is mostly (but not exclusively) based on data from mouse mutants. It appears largely possible that some mammalian "non-PG genes" might have a pigmentation function in the ray-finned fish lineage. Such cases have been already documented, including the colony-stimulating factor 1 receptor gene *csf1ra* (aka *fms*) (Parichy *et al.* 2000) and the *erbb3b* gene, which encodes an EGFR-like receptor tyrosine kinase (Budi *et al.* 2008). Interestingly, both genes have been duplicated by the 1R/2R and 3R WGDs, and *csf1r* is a 1R/2R ohnolog of the well-studied pigmentation gene *kit* (Brunet *et al.* 2016). Such functional differences might be explained by a loss of function in pigmentation in the tetrapod/mammalian lineage, or alternatively by a gain of pigmentary function in the ray-finned/teleost fish lineage, as observed for several transcriptional regulators during the evolution of neural crest at the basis of vertebrates (Meulemans and Bronner-Fraser 2005). The activity of a transcription factor has indeed already been shown to be co-opted to control novel pigmentation patterns in fly, with the evolution of specific binding sites in a *cis*-regulatory element (Gompel *et al.* 2005). The recruitment to pigmentation of genes involved in other processes such as organ development has mainly been shown in arthropods but is likely to occur in vertebrates too (Werner *et al.* 2010; Monteiro 2012). In teleost fish, *fhl2b*, a gene duplicated after Ts3R, has been shown to be involved in the development of iridophores (more specifically in the formation of egg-spots in cichlids) consecutively to the insertion of a transposable element that induced changes in the regulation of its expression (Santos *et al.* 2014).

Finally, after a duplication event, gene conversion can "homogenize" the sequences of the two duplicates and thus counteracts duplicate loss through a pseudogenization process. This process can also contribute to their diversity, as reported for some gene families including cadherins (Noonan 2004; Ohta 2010; McGrath *et al.* 2014). Thus, even if they are not directly involved in pigmentation, the presence of non-pigmentary paralogous copies might favor the maintenance and evolution of pigmentation genes through gene conversion and ectopic recombination.

Conclusion

The present analysis supports a major role of WGDs in the diversification of the pigmentation gene repertoire, especially in the teleost fish lineage. This might have promoted the diversity of pigment cell lineages observed in these animals compared to other vertebrates. Besides, the functional basis of pigmentation is still being uncovered in non-mammal vertebrate lineages that harbor a high chromatophore diversity, including amphibians, lepidosaurians and birds. Thus, our genome-wide characterization of 1R/2R pigmentation gene-containing families provides an important framework to approach the ancestral pigmentation gene repertoire in vertebrates and better understand the evolution of pigmentation in relation to the gain or loss of specific pigment cell types.

ACKNOWLEDGMENTS

We thank M. Robinson-Rechavi, C. Dechaud, J. Barbier and C. Rey for helpful discussion and comments on earlier versions of the manuscript. This work was supported by grants from the École Normale Supérieure of Lyon. T. L. is recipient of an individual grant from the French Ministry of Research and Education.

LITERATURE CITED

Amores, A., J. Catchen, A. Ferrara, Q. Fontenot, and J. H. Postlethwait, 2011 Genome Evolution and Meiotic Maps by Massively Parallel DNA Sequencing: Spotted Gar, an Outgroup for the Teleost Genome Duplication. Genetics 188(4): 799–808. https://doi.org/10.1534/genetics.111.127324

Anisimova, M., M. Gil, J.-F. Dufayard, C. Dessimoz, and O. Gascuel, 2011 Survey of Branch Support Methods Demonstrates Accuracy, Power, and Robustness of Fast Likelihood-based Approximation Schemes. Syst. Biol. 60(5): 685–699. https://doi.org/10.1093/sysbio/syr041

Bagnara, J. T., P. J. Fernandez, and R. Fujii, 2007 On the blue coloration of vertebrates. Pigment Cell Res. 20(1): 14–26. https://doi.org/10.1111/j.1600-0749.2006.00360.x

Bagnara, J. T., and E. M. Hadley, 1973 Chromatophores and Color Change. The Comparative Physiology of Animal Pigmentation, Prentice-Hall, Englewood Cliffs.

Berthelot, C., F. Brunet, D. Chalopin, A. Juanchich, M. Bernard *et al.*, 2014 The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. Nat. Commun. 5: 3657.

Birchler, J. A., and R. A. Veitia, 2007 The Gene Balance Hypothesis: From Classical Genetics to Modern Genomics. PLANT CELL ONLINE 19(2): 395–402. https://doi.org/10.1105/tpc.106.049338

Blomme, T., K. Vandepoele, S. De Bodt, C. Simillion, S. Maere *et al.*, 2006 The gain and loss of genes during 600 million years of vertebrate evolution. Genome Biol. 7(5): R43. https://doi.org/10.1186/gb-2006-7-5-r43

Blouin, C., S. Perry, A. Lavell, E. Susko, and A. J. Roger, 2009 Reproducing the manual annotation of multiple sequence alignments using a SVM classifier. Bioinformatics 25(23): 3093–3098. https://doi.org/10.1093/bioinformatics/btp552

Braasch, I., F. Brunet, J.-N. Volff, and M. Schartl, 2009a Pigmentation pathway evolution after whole-genome duplication in fish. Genome Biol. Evol. 1(0): 479–493. https://doi.org/10.1093/gbe/evp050

Braasch, I., D. Liedtke, J.-N. Volff, and M. Schartl, 2009b Pigmentary function and evolution of tyrp1 gene duplicates in fish. Pigment Cell Melanoma Res. 22(6): 839–850. https://doi.org/10.1111/j.1755-148X.2009.00614.x

Braasch, I., J. N. Volff, and M. Schartl, 2008 The evolution of teleost pigmentation and the fish-specific genome duplication. J. Fish Biol. 73(8): 1891–1918. https://doi.org/10.1111/j.1095-8649.2008.02011.x

Bright, L. J., J.-F. Gout, and M. Lynch, 2017 Early stages of functional diversification in the Rab GTPase gene family revealed by genomic and localization studies in Paramecium species. Mol. Biol. Cell. 28: 1101–1110.

Brunet, F. G., T. Lorin, L. Bernard, Z. Haftek-Terreau, D. Galiana *et al.*, 2017 Case studies of seven gene families with unusual high retention rate since the vertebrate and teleost whole-genome duplications, pp. 369–396 in *Evolutionary Biology: Se.f/Nonse.f Evolution, Species and Complex Traits Evolution, Methods and Concepts*. Springer International Publishing, Champaign, IL.

Brunet, F. G., H. Roest Crollius, M. Paris, J. Aury, P. Gibert *et al.*, 2006 Gene Loss and Evolutionary Rates Following Whole-Genome Duplication in Teleost Fishes. Mol. Biol. Evol. 23(9): 1808–1816. https://doi.org/10.1093/molbev/msl049

Brunet, F. G., J.-N. Volff, and M. Schartl, 2016 Whole Genome Duplications Shaped the Receptor Tyrosine Kinase Repertoire of Jawed Vertebrates. Genome Biol. Evol. 8(5): 1600–1613. https://doi.org/10.1093/gbe/evw103

Budi, E. H., L. B. Patterson, and D. M. Parichy, 2008 Embryonic requirements for ErbB signaling in neural crest development and adult pigment pattern formation. Development 135(15): 2603–2614. https://doi.org/10.1242/dev.019299

Cal, L., M. Megías, J. M., Cerdá-Reverter, J. H., Postlethwait, I., Braasch *et al.*, 2017a BAC recombineering of the agouti loci from spotted gar and zebrafish reveals the evolutionary ancestry of dorsal-ventral pigment asymmetry in fish. J. Exp. Zool. B Mol. Dev. Evol. 328: 697–708.

Cal, L., P. Suarez-Bregua, J. M. Cerdá-Reverter, I. Braasch, and J. Rotllant, 2017b Fish pigmentation and the melanocortin system. Comp. Biochem. Physiol. A Mol. Integr. Physiol. 211: 26–33. https://doi.org/10.1016/j.cbpa.2017.06.001

Caro, T., 2017 Wallace on Coloration: Contemporary Perspective and Unresolved Insights. Trends Ecol. Evol. 32(1): 23–30. https://doi.org/10.1016/j.tree.2016.10.003

Ceinos, R. M., R. Guillot, R. N. Kelsh, J. M. Cerdá-Reverter, and J. Rotllant, 2015 Pigment patterns in adult fish result from superimposition of two largely independent pigmentation mechanisms. Pigment Cell Melanoma Res. 28(2): 196–209. https://doi.org/10.1111/pcmr.12335

Colihueque, N., 2010 Genetics of salmonid skin pigmentation: clues and prospects for improving the external appearance of farmed salmonids. Rev. Fish Biol. Fish. 20(1): 71–86. https://doi.org/10.1007/s11160-009-9121-6

Conant, G. C., and K. H. Wolfe, 2008 Turning a hobby into a job: How duplicated genes find new functions. Nat. Rev. Genet. 9(12): 938–950. https://doi.org/10.1038/nrg2482

Conrad, B., and S. E. Antonarakis, 2007 Gene Duplication: A Drive for Phenotypic Diversity and Cause of Human Disease. Annu. Rev. Genomics Hum. Genet. 8(1): 17–35. https://doi.org/10.1146/annurev.genom.8.021307.110233

Curran, K., J. A. Lister, G. R. Kunkel, A. Prendergast, D. M. Parichy *et al.*, 2010 Interplay between Foxd3 and Mitf regulates cell fate plasticity in the zebrafish neural crest. Dev. Biol. 344(1): 107–118. https://doi.org/10.1016/j.ydbio.2010.04.023

Darriba, D., G. L. Taboada, R. Doallo, and D. Posada, 2011 ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27(8): 1164–1165. https://doi.org/10.1093/bioinformatics/btr088

Davidson, E., 2006 The Regulatory Genome.

Davis, J., and D. Petrov, 2005 Do disparate mechanisms of duplication add similar genes to the genome? Trends Genet. 21(10): 548–551. https://doi.org/10.1016/j.tig.2005.07.008

Dehal, P., and J. L. Boore, 2005 Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol. 3: e314.

Freeling, M., and B. C. Thomas, 2006 Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. Genome Res. 16(7): 805–814. https://doi.org/10.1101/gr.3681406

Friedman, J. S., M. F. Lopez, M. D. Fleming, A. Rivera, F. M. Martin *et al.*, 2004 SOD2-deficiency anemia: protein oxidation and altered protein expression reveal targets of damage, stress response, and antioxidant responsiveness. Blood 104(8): 2565–2573. https://doi.org/10.1182/blood-2003-11-3858

Frohnhöfer, H. G., S. Geiger-Rudolph, M. Pattky, M. Meixner, C. Huhn *et al.*, 2016 Spermidine, but not spermine, is essential for pigment pattern formation in zebrafish. Biol. Open 5(6): 736–744. https://doi.org/10.1242/bio.018721

Fujii, R., 1993 Cytophysiology of fish chromatophores. Int. Rev. Cytol. 143: 191–255.

Gilbert, S. F., 2000 *Developmental Biology*. Sinauer Associates, Sunderland, MA.

Goda, M., 2017 Rapid integumental color changes due to novel iridophores in the chameleon sand tilefish Hoplolatilus chlupatyi. Pigment Cell Melanoma Res. 30(3): 368–371. https://doi.org/10.1111/pcmr.12581

Goda, M., and R. Fujii, 1995 Blue Chromatophores in Two Species of Callionymid Fish. Zool. Sci. 12(6): 811–813. https://doi.org/10.2108/zsj.12.811

Goda, M., Y. Fujiyoshi, M. Sugimoto, and R. Fujii, 2013 Novel Dichromatic Chromatophores in the Integument of the Mandarin Fish Synchiropus splendidus. Biol. Bull. 224(1): 14–17. https://doi.org/10.1086/BBLv224n1p14

Goda, M., M. Ohata, H. Ikoma, Y. Fujiyoshi, M. Sugimoto *et al.*, 2011 Integumental reddish-violet coloration owing to novel dichromatic chromatophores in the teleost fish, Pseudochromis diadema. Pigment Cell Melanoma Res. 24(4): 614–617. https://doi.org/10.1111/j.1755-148X.2011.00861.x

Gompel, N., B. Prud'homme, P. J. Wittkopp, V. A. Kassner, and S. B. Carroll, 2005 Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in Drosophila. Nature 433(7025): 481–487. https://doi.org/10.1038/nature03235

Gray, K. A., B. Yates, R. L. Seal, M. W. Wright, and E. A. Bruford, 2015 Genenames.org: the HGNC resources in 2015. Nucleic Acids Res. 43(D1): D1079–D1085. https://doi.org/10.1093/nar/gku1071

Green, S. A., M. Simoes-Costa, and M. E. Bronner, 2015 Evolution of vertebrates as viewed from the crest. Nature 520(7548): 474–482. https://doi.org/10.1038/nature14436

Guan, Y., M. J. Dunham, and O. G. Troyanskaya, 2007 Functional Analysis of Gene Duplications in Saccharomyces cerevisiae. Genetics 175(2): 933–943. https://doi.org/10.1534/genetics.106.064329

Guindon, S., J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk *et al.*, 2010 New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst. Biol. 59(3): 307–321. https://doi.org/10.1093/sysbio/syq010

Havelka, M., M. Hulák, D. A. Bailie, P. A. Prodöhl, and M. Flajšhans, 2013 Extensive genome duplications in sturgeons: new evidence from microsatellite data. J. Appl. Ichthyology 29(4): 704–708. https://doi.org/10.1111/jai.12224

He, X., and J. Zhang, 2005 Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. Genetics 169(2): 1157–1164. https://doi.org/10.1534/genetics.104.037051

Hoekstra, H. E., 2006 Genetics, development and evolution of adaptive pigmentation in vertebrates. Heredity (Edinb) 97(3): 222–234. https://doi.org/10.1038/sj.hdy.6800861

Howe, K., M. D. Clark, C. F. Torroja, J. Torrance, C. Berthelot *et al.*, 2013 The zebrafish reference genome sequence and its relationship to the human genome. Nature 496: 498–503 (erratum: Nature 505: 248). https://doi.org/10.1038/nature12111

Hufton, A. L., D. Groth, M. Vingron, H. Lehrach, A. J. Poustka *et al.*, 2008 Early vertebrate whole genome duplications were predated by a period of intense genome rearrangement. Genome Res. 18(10): 1582–1591. https://doi.org/10.1101/gr.080119.108

Hulsen, T., J. de Vlieg, and W. Alkema, 2008 BioVenn – a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC Genomics 9(1): 488. https://doi.org/10.1186/1471-2164-9-488

Huminiecki, L., and C. Heldin, 2010 2R and remodeling of vertebrate signal transduction engine. BMC Biol. 8(1): 146. https://doi.org/10.1186/1741-7007-8-146

Hurley, I. A., R. L. Mueller, K. A. Dunn, E. J. Schmidt, M. Friedman *et al.*, 2007 A new time-scale for ray-finned fish evolution. Proc. Biol. Sci. 274 (1609): 489–498. https://doi.org/10.1098/rspb.2006.3749

Inoue, J., Y. Sato, R. Sinclair, K. Tsukamoto, and M. Nishida, 2015 Rapid genome reshaping by multiple-gene loss after whole-genome duplication in teleost fish suggested by mathematical modeling. Proc. Natl. Acad. Sci. USA 112(48): 14918–14923. https://doi.org/10.1073/pnas.1507669112

Jaillon, O., J.-M. Aury, F. Brunet, J.-L. Petit, N. Stange-Thomann *et al.*, 2004 Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. Nature 431(7011): 946–957. https://doi.org/10.1038/nature03025

Jiang, Q., W. K. W. Ko, and A. O. L. Wong, 2011 Insulin-like growth factor as a novel stimulator for somatolactin secretion and synthesis in carp pituitary cells via activation of MAPK cascades. Am. J. Physiol. Metab. 301: E1208–E1219. https://doi.org/10.1152/ajpendo.00347.2011

Jiang, W., Y. Liu, E. Xia, and L. Gao, 2013 Prevalent Role of Gene Features in Determining Evolutionary Fates of Whole-Genome Duplication Duplicated Genes in Flowering Plants. Plant Physiol. 161(4): 1844–1861. https://doi.org/10.1104/pp.112.200147

Kaelin, C. B., X. Xu, L. Z. Hong, V. A. David, K. A. McGowan *et al.*, 2012 Specifying and sustaining pigmentation patterns in domestic and wild cats. Science 337: 1536–1541. https://doi.org/10.1126/science.1220893

Kassahn, K. S., V. T. Dang, S. J. Wilkins, A. C. Perkins, and M. A. Ragan, 2009 Evolution of gene function and regulatory control after whole-genome duplication: Comparative analyses in vertebrates. Genome Res. 19(8): 1404–1418. https://doi.org/10.1101/gr.086827.108

Kimura, T., Y. Nagao, H. Hashimoto, Y.-i. Yamamoto-Shiraishi, S. Yamamoto *et al.*, 2014 Leucophores are similar to xanthophores in their specification and differentiation processes in medaka. Proc. Natl. Acad. Sci. USA 111(20): 7343–7348. https://doi.org/10.1073/pnas.1311254111

Kuraku, S., 2010 Palaeophylogenomics of the Vertebrate Ancestor–Impact of Hidden Paralogy on Hagfish and Lamprey Gene Phylogeny. Integr. Comp. Biol. 50(1): 124–129. https://doi.org/10.1093/icb/icq044

Lapedriza, A., K. Petratou, and R. N. Kelsh, 2014 Neural crest cells and pigmentation, pp. 287–311 in *Neural Crest Cells*. Elsevier.

Le, S. Q., and O. Gascuel, 2008 An Improved General Amino Acid Replacement Matrix. Mol. Biol. Evol. 25(7): 1307–1320. https://doi.org/10.1093/molbev/msn067

Lemos, B., C. D. Meiklejohn, and D. L. Hartl, 2004 Regulatory evolution across the protein interaction network. Nat. Genet. 36(10): 1059–1060. https://doi.org/10.1038/ng1427

Li, Z., J. Defoort, S. Tasdighian, S. Maere, Y. Van de Peer *et al.*, 2016 Gene Duplicability of Core Genes Is Highly Consistent across All Angiosperms. Plant Cell 28(2): 326–344. https://doi.org/10.1105/tpc.15.00877

Li, J.-T., G.-Y. Hou, X.-F. Kong, C.-Y. Li, J.-M. Zeng *et al.*, 2015 The fate of recent duplicated genes following a fourth-round whole genome duplication in a tetraploid fish, common carp (Cyprinus carpio). Sci. Rep. 5(1): 8199. https://doi.org/10.1038/srep08199

Lien, S., B. F. Koop, S. R. Sandve, J. R. Miller, M. P. Kent *et al.*, 2016 The Atlantic salmon genome provides insights into rediploidization. Nature 533(7602): 200–205. https://doi.org/10.1038/nature17164

Lopes, R. J., J. D. Johnson, M. B. Toomey, M. S. Ferreira, P. M. Araujo *et al.*, 2016 Genetic Basis for Red Coloration in Birds. Curr. Biol. 26(11): 1427–1434. https://doi.org/10.1016/j.cub.2016.03.076

Louis, A., N. T. T. Nguyen, M. Muffato, and H. Roest Crollius, 2015 Genomicus update 2015: KaryoView and MatrixView provide a genome-wide perspective to multispecies comparative genomics. Nucleic Acids Res. 43(D1): D682–D689. https://doi.org/10.1093/nar/gku1112

Luo, A., H. Qiao, Y. Zhang, W. Shi, S. Y. Ho *et al.*, 2010 Performance of criteria for selecting evolutionary models in phylogenetics: a comprehensive study based on simulated datasets. BMC Evol. Biol. 10(1): 242. https://doi.org/10.1186/1471-2148-10-242

Macqueen, D. J., and I. A. Johnston, 2014 A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. Proc. Biol. Sci. 281(1778): 20132881. https://doi.org/10.1098/rspb.2013.2881

Maere, S., S. De Bodt, J. Raes, T. Casneuf, M. Van Montagu *et al.*, 2005 Modeling gene and genome duplications in eukaryotes. Proc. Natl. Acad. Sci. USA 102(15): 5454–5459. https://doi.org/10.1073/pnas.0501102102

Makino, T., and A. McLysaght, 2010 Ohnologs in the human genome are dosage balanced and frequently associated with disease. Proc. Natl. Acad. Sci. USA 107(20): 9270–9274. https://doi.org/10.1073/pnas.0914697107

Mallarino, R., C. Henegar, M. Mirasierra, M. Manceau, C. Schradin *et al.*, 2016 Developmental mechanisms of stripe patterns in rodents. Nature 539(7630): 518–523. https://doi.org/10.1038/nature20109

McGrath, C. L., J.-F. Gout, P. Johri, T. G. Doak, and M. Lynch, 2014 Differential retention and divergent resolution of duplicate genes following whole-genome duplication. Genome Res. 24(10): 1665–1675. https://doi.org/10.1101/gr.173740.114

Meulemans, D., and M. Bronner-Fraser, 2005 Central role of gene cooption in neural crest evolution. J. Exp. Zool. Part B Mol. Dev. Evol. 304B: 298–303.

Meyer, A., and M. Schartl, 1999 Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. Curr. Opin. Cell Biol. 11(6): 699–704. https://doi.org/10.1016/S0955-0674(99)00039-3

Monteiro, A., 2012 Gene regulatory networks reused to build novel traits. BioEssays 34(3): 181–186. https://doi.org/10.1002/bies.201100160

Mundy, N. I., J. Stapley, C. Bennison, R. Tucker, H. Twyman *et al.*, 2016 Red Carotenoid Coloration in the Zebra Finch Is Controlled by a Cytochrome P450 Gene Cluster. Curr. Biol. 26(11): 1435–1440. https://doi.org/10.1016/j.cub.2016.04.047

Near, T. J., R. I. Eytan, A. Dornburg, K. L. Kuhn, J. A. Moore *et al.*, 2012 Resolution of ray-finned fish phylogeny and timing of diversification. Proc. Natl. Acad. Sci. USA 109(34): 13698–13703. https://doi.org/10.1073/pnas.1206625109

Noonan, J. P., 2004 Gene Conversion and the Evolution of Protocadherin Gene Cluster Diversity. Genome Res. 14(3): 354–366. https://doi.org/10.1101/gr.2133704

Nüsslein-Volhard, C., and A. P. Singh, 2017 How fish color their skin: A paradigm for development and evolution of adult patterns. BioEssays 39 (3): 1600231. https://doi.org/10.1002/bies.201600231

Ohno, S., 1970 Evolution by gene duplication. https://doi.org/10.1007/978-3-642-86659-3

Ohno, S., 1999 The one-to-four rule and paralogues of sex-determining genes. Cell. Mol. Life Sci. 55(7): 824–830. https://doi.org/10.1007/s000180050336

Ohta, T., 2010 Gene Conversion and Evolution of Gene Families: An Overview. Genes (Basel) 1(3): 349–356. https://doi.org/10.3390/genes1030349

Parichy, D. M., D. G. Ransom, B. Paw, L. I. Zon, and S. L. Johnson, 2000 An orthologue of the kit-related gene fms is required for development of neural crest-derived xanthophores and a subpopulation of adult melanocytes in the zebrafish, Danio rerio. Development 127: 3031–3044.

Putnam, N. H., T. Butts, D. E. K. Ferrier, R. F. Furlong, U. Hellsten *et al.*, 2008 The amphioxus genome and the evolution of the chordate karyotype. Nature 453(7198): 1064–1071. https://doi.org/10.1038/nature06967

Ramón Martínez-Morales, J., I. Rodrigo, and P. Bovolenta, 2004 Eye development: a view from the retina pigmented epithelium. BioEssays 26(7): 766–777. https://doi.org/10.1002/bies.20064

Rawls, J. F., E. M. Mellgren, and S. L. Johnson, 2001 How the Zebrafish Gets Its Stripes. Dev. Biol. 240(2): 301–314. https://doi.org/10.1006/dbio.2001.0418

Reimand, J., T. Arak, P. Adler, L. Kolberg, S. Reisberg *et al.*, 2016 g:profiler — a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res. 44(W1): W83–W89. https://doi.org/10.1093/nar/gkw199

Robertson, F. M., M. K. Gundappa, F. Grammes, T. R. Hvidsten, A. K. Redmond *et al.*, 2017 Lineage-specific rediploidization is a mechanism to explain time-lags between genome duplication and evolutionary diversification. Genome Biol. 18(1): 111. https://doi.org/10.1186/s13059-017-1241-z

Rogozin, I. B., 2014 Complexity of Gene Expression Evolution after Duplication: Protein Dosage Rebalancing. Genet. Res. Int. 2014: 1–8. https://doi.org/10.1155/2014/516508

Roulin, A., and A.-L. Ducrest, 2013 Genetics of colouration in birds. Semin. Cell Dev. Biol. 24(6-7): 594–608. https://doi.org/10.1016/j.semcdb.2013.05.005

Santini, F., L. J. Harmon, G. Carnevale, and M. E. Alfaro, 2009 Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. BMC Evol. Biol. 9(1): 194. https://doi.org/10.1186/1471-2148-9-194

Santos, E., I. Braasch, N. Boileau, B. S. Meyer, A. Bo *et al.*, 2014 The evolution of cichlid fish egg-spots is linked with a cis-regulatory change. Nat. Commun. 5: 5149. https://doi.org/10.1038/ncomms6149

Schartl, M., L. Larue, M. Goda, M. W. Bosenberg, H. Hashimoto *et al.*, 2016 What is a vertebrate pigment cell? Pigment Cell Melanoma Res. 29(1): 8–14. https://doi.org/10.1111/pcmr.12409

Seoighe, C., and C. Gehring, 2004 Genome duplication led to highly selective expansion of the Arabidopsis thaliana proteome. Trends Genet. 20(10): 461–464. https://doi.org/10.1016/j.tig.2004.07.008

Session, A. M., Y. Uno, T. Kwon, J. A. Chapman, A. Toyoda *et al.*, 2016 Genome evolution in the allotetraploid frog Xenopus laevis. Nature 538(7625): 336–343. https://doi.org/10.1038/nature19840

Sievers, F., and D. G. Higgins, 2014 Clustal omega, accurate alignment of very large numbers of sequences. Methods Mol. Biol. 1079: 105–116. https://doi.org/10.1007/978-1-62703-646-7_6

Steinke, D., W. Salzburger, I. Braasch, and A. Meyer, 2006 Many genes in fish have species-specific asymmetric rates of molecular evolution. BMC Genomics 7(1): 20. https://doi.org/10.1186/1471-2164-7-20

Turunen, O., R. Seelke, and J. Macosko, 2009 In silico evidence for functional specialization after genome duplication in yeast. FEMS Yeast Res. 9(1): 16–31. https://doi.org/10.1111/j.1567-1364.2008.00451.x

UniProt Consortium, 2015 UniProt: a hub for protein information. Nucleic Acids Res. 43: D204–D212. https://doi.org/10.1093/nar/gku989

Voldoire, E., F. Brunet, M. Naville, J.-N. Volff, and D. Galiana, 2017 Expansion by whole genome duplication and evolution of the sox gene family in teleost fish. PLoS One 12: e0180936.

Wallace, A. R., 1877 The colours of animals and plants. Am. Nat. 11(12): 713–728. https://doi.org/10.1086/271996

Werner, T., S. Koshikawa, T. M. Williams, and S. B. Carroll, 2010 Generation of a novel wing colour pattern by the Wingless morphogen. Nature 464(7292): 1143–1148. https://doi.org/10.1038/nature08896

Wolfe, K. H., 2001 Yesterday's polyploids and the mystery of diploidization. Nat. Rev. Genet. 2(5): 333–341. https://doi.org/10.1038/35072009

Woods, I. G., P. D. Kelly, F. Chu, P. Ngo-Hazelett, Y. L. Yan *et al.*, 2000 A comparative map of the zebrafish genome. Genome Res. 10(12): 1903–1914. https://doi.org/10.1101/gr.10.12.1903

Woods, I. G., C. Wilson, B. Friedlander, P. Chang, D. K. Reyes *et al.*, 2005 The zebrafish gene map defines ancestral vertebrate chromosomes. Genome Res. 15(9): 1307–1314. https://doi.org/10.1101/gr.4134305

Wu, X., and X. Qi, 2010 Genes encoding hub and bottleneck enzymes of the Arabidopsis metabolic network preferentially retain homeologs through whole genome duplication. BMC Evol. Biol. 10(1): 145. https://doi.org/10.1186/1471-2148-10-145

Wucherer, M. F., and N. K. Michiels, 2012 A fluorescent chromatophore changes the level of fluorescence in a reef fish. PLoS One 7: e37913. https://doi.org/10.1371/journal.pone.0037913

Xie, T., Q.-Y. Yang, X.-T. Wang, A. McLysaght, and H.-Y. Zhang, 2016 Spatial Colocalization of Human Ohnolog Pairs Acts to Maintain Dosage-Balance. Mol. Biol. Evol. 33(9): 2368–2375. https://doi.org/10.1093/molbev/msw108

Communicating editor: J. Birchler

RÉSULTATS

CHAPITRE 2

Détermination de la nature de la bande blanche chez le poisson-clown *Amphiprion ocellaris*

Avant-propos

Ce travail a impliqué la réalisation conjointe d'analyses fonctionnelles et expérimentales, ainsi que bio-informatiques. Les résultats expérimentaux ont été essentiellement obtenus par Pauline Salis, chercheure post-doctorante à l'Observatoire Océanologique de Banyuls-sur-Mer, avec la contribution de Natacha Roux et Laurence Besseau. Suite à la découverte de marqueurs d'iridophores intéressants chez le poisson-clown *Amphiprion ocellaris*, nous avons initié une collaboration avec l'équipe du Pr. David Parichy, à l'Université de Virginie (Etats-Unis). Son doctorant, Victor Lewis, réalise actuellement les expériences de CRISPR/Cas9 chez le Poisson-zèbre. J'ai pour ma part mené les analyses bio-informatiques : assemblage et annotation du transcriptome d'*A. ocellaris*, analyses d'expression différentielles, analyses d'enrichissement de groupes de gènes et comparaisons avec des données déjà publiées. J'ai été en cela aidé par Marie Sémon et Carine Rey. Enfin, Jean-Nicolas Volff et Vincent Laudet ont tous deux co-supervisé ce travail.

Cet article est actuellement encore en cours d'écriture. Il manque quelques résultats, notamment expérimentaux, avant que la soumission ne soit possible. Je vous prie donc d'en trouver ci-dessous une version non-définitive, que j'ai en partie rédigée.

Color of white stripe in clownfish is due to iridophores, not leucophores

Pauline Salis*[1], Thibault Lorin*[2], Victor Lewis[3], Carine Rey[4,5], Natacha Roux[1], Laurence Besseau[1], Marie Sémon[4], David Parichy[3], Jean-Nicolas Volff[2], Vincent Laudet[1]

*: contributed equally to this work

1: Observatoire Océanologique de Banyuls-sur-Mer, UMR CNRS 7232 BIOM ; Sorbonne Université ; 1, Avenue Pierre Fabre, 66650 Banyuls-sur-Mer, France

2: Institut de Génomique Fonctionnelle de Lyon, École Normale Supérieure de Lyon, UMR 5242 CNRS, Université Claude Bernard Lyon I, Université de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France

3: Department of Biology and Department of Cell Biology, University of Virginia, Charlottesville VA 29903

4: UnivLyon, ENS de Lyon, Univ Claude Bernard, CNRS UMR 5239, INSERM U1210, Laboratoire de Biologie et Modélisation de la Cellule, 15 parvis Descartes, F-69007, Lyon, France;

5: Laboratoire de Biométrie et Biologie Evolutive (LBBE), Université de Lyon, Université Lyon 1, CNRS, Villeurbanne, France

ABSTRACT

In fish, the cellular basis of pigmentation is very diverse, as this lineage harbors more than eight pigment cell types. In particular, the white color can be due to either leucophores or iridophores. Among all animals, the clownfish *Amphiprion ocellaris*, popularized by the *Finding nemo* movie, represents a widely used species in aquarium trade, particularly due to its staggering striped pattern consisting of alternating white and orange stripes. In this work, we aimed at better understanding the nature of the white stripe in clownfish A. ocellaris. We first performed both stereomicroscopy and electron microcopy analyses that confirmed that iridophores provided the basis for white stripe color. In addition, RNA-Seq data corresponding to orange and white stripes were obtained using skin samples. Through comparisons with public iridophore specific RNA-Seq data, we showed that clownfish white tissue had a transcriptomic signature of iridophores and additionally identified new putative iridophore markers such as roshia, apod1a or alx4b. We finally injected zebrafish *Danio rerio* larvae with CRISPR/Cas9 targeting some of the candidate genes. Most of these mosaic F0 mutants showed indeed a reduction in iridophore number. Altogether, this work shows that color of clownfish white stripe is due to iridophores, not leucophores. This work also provides a list of 12 new putative iridophore markers conserved in both zebrafish and clownfish and thus probably conserved in most teleost fish.

INTRODUCTION

Colors in the surrounding living beings have for long been a subject of both attraction and fascination for mankind. Animal pigmentation, in particular, is of great interest to the general public and to naturalists, including animals breeders that have long been interested in obtaining color morphs (Shapiro *et al.*, 2013). The 19th century represents a milestone for the understanding of the adaptive value of such colors, with the first classification of animal "coloration" – as termed at the time – by Wallace (Wallace, 1877), and the work of Darwin on artificial selection taking advantage of pigeons tremendous variations in color morphs (Darwin, 1868).

More than 150 years later, we now have a more complete understanding of pigmentation in an evolutionary and adaptive context – even if more work is needed (Cuthill *et al.*, 2017). In particular, the field of genetics in general, and more particularly the genetic and cellular basis of pigmentation in vertebrates, have much benefited from work on pigmentation mutants in model animals, including mouse (Hoekstra, 2006). However, the most compelling pigmentation features in vertebrates are seen in fish, and especially in the teleost lineage. This lineage harbors more than 30,000 species, *i. e.* more than half of all extant vertebrates (Near *et al.*, 2012; Nelson *et al.*, 2016), and is characterized by the occurrence of the most diverse pigmentation features, in terms of both pigment cell types (or chromatophores) and pattern diversity. Indeed, while mammals only have one pigment cell type, the melanocytes, fish harbor more than eight types: melanophores, iridophores, xanthophores, erythrophores, leucophores, cyanophores, erythro-iridophores, cyano-erythrophores, and even iridophore-like cells or fluorescent chromatophores (Bagnara and Hadley, 1973; Fujii, 1993; Goda and Fujii, 1995; Goda *et al.*, 2011, 2013; Schartl *et al.*, 2016). New pigment cell types are still being described in fish (Wucherer and Michiels, 2012; Goda, 2017), and, if no species has them all, most species harbor at least three or four of them. Consequently, in the last decades, the field of vertebrate pigmentation has much benefited from the introduction of new teleost model species, such as zebrafish and medaka.

The zebrafish *Danio rerio* harbors three different types of pigment cells: carotenoid- and pteridine-containing xanthophores that provide the yellow/orange hue; purine-containing silvery or blue iridophores reflecting light; and melanin-containing black melanophores absorbing light across the spectrum (Hirata *et al.*, 2003). These chromatophores interact with each other to provide the characteristic striped pattern in *D. rerio* (Singh and Nüsslein-Volhard, 2015). Studies on zebrafish have greatly contributed to the characterization of the developmental mechanisms and of the genetic determinants underlying such patterning, and to the cellular and molecular characterization of these three chromatophores (Hirata *et al.*, 2003). The medaka *Oryzias latipes*, another fish model, is of high interest in the context of vertebrate pigmentation since it harbors an additional pigment cell type, the leucophores. Iridophores and leucophores are both purine-containing cells: iridophore reflecting platelets are constituted of guanine whereas the white hue of leucophore platelets is due to uric acid (Kimura *et al.*, 2014).

One of the most popular fish to date in aquaculture is the clownfish *Amphiprion ocellaris* (Cuvier) that has been widely popularized by the so-called "Nemo effect" consecutive to the movie *Finding Nemo* (Militz and Foale, 2017). The clownfish *A. ocellaris* belongs to a genus of ca. 30 clownfish – also known as anemonefish – species (Litsios *et al.*, 2012). Most of these species harbor a pigmentation pattern with one or multiple white stripes, hence their name of "clownfish" based on the resemblance between this stripe pattern and a clown's outfits. The adaptive function of these white stripes in clownfish has not been fully characterized yet but might be implied in aposematism (Mariscal, 1966). In nature but even more in the aquaculture field, plethora of variants in color morphs exist based on the relative widths of the white and orange stripes.

So far, there has been no characterization at the cellular and genetic level on the nature of the white stripe in this emblematic species, and especially whether this hue is mainly due to white leucophores, white iridophores, or both cell types. In this work, we first characterized at the cellular level, using microscopy, the skin organization in *A. ocellaris* orange and white stripes. We showed that chromatophores in these stripes were mostly xanthophores and iridophores, respectively. In addition, we performed transcriptomic analyses by comparing both types of pigmented skin tissue and showed that the white stripe harbored a transcriptomic signature comparable to zebrafish iridophores. Conserved markers between zebrafish and clownfish were identified. We finally used CRISPR/Cas9 targeting specific candidate genes in zebrafish larvae and thus functionally confirmed the major role played in iridophore differentiation by some of the newly characterized genes. We conclude that the white coloration of the iconic clownfish *A. ocellaris* is due to the

particular organization of purine platelets in iridophores, and that the study of new fish models enlightens the processes underlying iridophore differentiation.

MATERIAL AND METHODS

Electronic microscopy

3-4 cm *A. ocellaris* individuals were euthanized using MS-222 Tricaine-S (REF) and the white and orange skin were dissected directly in a mixture of 3% paraformaldehyde and 3% glutaraldehyde in 0.1 M cacodylate buffer (pH7.2), and fixed overnight at 4 °C, according to (Djurdjevič *et al.*, 2015). After 1% OsO4 post-fixation (1h at RT) individuals were dehydrated in a graded series of ethanol solutions and embedded in Epon 812 resin (Electron Microscopy Science).

Semi-thin (1µm) and ultra-thin (80nm) sections were cut using a Leica Ultracut R microtome. Semi-thin sections were contrasted with toluidine blue and examined under the light microscope to determine the position of pigment cells. Regions observed by TEM were selected on the basis of the preliminary analysis of the semi-thin sections by light microscope. Ultra-thin sections were contrasted with uranyl acetate and lead citrate and examined with a Hitachi H7500 TEM for cell ultrastructure analysis.

RNA extraction and purification

3-4 cm *A. ocellaris* individuals were euthanized using MS-222 Tricaine-S. Then, a shallow incision was quickly made with a razor in the white and orange skin. The skin was slowly pulled back with forceps while using the razor blade to make sure the skin separated cleanly from the underlying tissue. Tissue was excised from the body regions outlined in figures 1 and 2. The RNA was extracted using Trizol (Trizol Reagent 15596-026 kit, AMBION) followed by DNase treatment (DNA-free AM1906 kit, AMBION) and then purified with 0.025 µm dialysis membranes to remove reverse transcriptase inhibitors (Nitrocellulose Membrane Filters- VSWP01300 Merck Millipore).

RNA-Seq library preparation and sequencing

RNA-Seq libraries were generated with the TruSeq Stranded mRNA Sample Preparation Kit (Illumina, Part Number RS-122-2101/2102) from 400 ng of total RNA using samples from skin, muscle and several developmental stages. Briefly, following purification with poly-T oligonucleotide-attached magnetic beads, the mRNA was fragmented using divalent cations at 94°C for 2 minutes. The cleaved RNA fragments were copied into first strand cDNA using reverse transcriptase and random primers. Strand specificity was achieved by replacing dTTP with dUTP during second strand cDNA synthesis using DNA Polymerase I and RNase H. Following addition of a single 'A' base and subsequent ligation of the adapter on double stranded cDNA fragments, the products were purified and enriched with PCR (30 sec at 98°C; [10 sec at 98°C, 30 sec at 60°C, 30 sec at 72°C] x 12 cycles; 5 min at 72°C) to create the cDNA library. Surplus PCR primers were further removed by purification using AMPure XP beads (Beckman Coulter) and the final cDNA libraries were checked for quality and quantified using capillary electrophoresis. The libraries were loaded in the flow cell at 2 nM (skin samples) or 2.8 nM (developmental samples) concentration. Clusters were generated in the Cbot and sequenced on the Illumina Hiseq 4000 as Single-end 50 base reads (single-end skin data) or Paired-end 100 base reads (paired-end skin and developmental data) following Illumina's instructions.

Reference transcriptome assembly and annotation

Adapter dimer reads of skin, muscle and developmental paired-end data were removed *in silico* using DimerRemover (https://sourceforge.net/projects/dimerremover/) and reads were trimmed from adapters using BBDuk from the BBTools suite (Bushnell B. – https://sourceforge.net/projects/bbmap/). Reads corresponding to rRNAs, although mostly absent due to library preparation protocol, were further removed using sortMeRNA (Kopylova *et al.*, 2012) with *Oreochromis niloticus* rRNA sequences downloaded from Ensembl Biomart (Ensembl release 88 – March 2017) as a reference (Kinsella *et al.*, 2011). Reads were further normalized using the *insilico_read_normalization.pl* tool included in the Trinity package (Grabherr *et al.*, 2011).

In order to get a genome-guided transcriptome assembly, reads were mapped to the *Amphiprion ocellaris* genome sequence (courtesy of Marcionetti et al.) using Hisat2 v2.0.5 (Kim *et al.*, 2015) and mapping results were further processed using the SAMtools suite v1.3.1 (Li *et al.*, 2009). Contig assembly was performed using the Trinity program v2.3.2 with default parameters (Grabherr *et al.*, 2011). In addition to the genome-guided assembly, we obtained a *de novo* transcriptome assembly from the same set of reads and combined both assemblies before identification of

coding regions within transcript sequences using the TransDecoder program v3.0.1 (http://transdecoder.sf.net). We used information from the closely related species *Stegastes partitus* to annotate *Amphiprion ocellaris* contigs. A subset of the NCBI non-redundant nucleotide database containing only *S. partitus* sequences was obtained and we ran a local BLASTN comparison (Altschul *et al.*, 1990); each contig was annotated using best blast hit (e-value threshold: 10^{-5}).

Quantification and differential expression analyses

Single-end reads corresponding to skin samples were aligned to the transcriptome using the pseudo-alignment tool *kallisto* v0.43.0 (Bray *et al.*, 2016). This program was used for quantifying abundances of transcripts. Estimated counts were further imported in the R environment v3.3.2 (http://www.R-project.org/, R Core Team 2016) using the *tximport* package (Soneson *et al.*, 2015). Differential expression (DE) analysis was performed using the DEseq2 package (DESeq2, version 1.14.1), testing for color variations with the individual effect as control level (Love *et al.*, 2014).

Phylogenetic analyses

In this study, transcript annotation was based on a BLASTN comparison with a different species to *Amphiprion ocellaris*: *Stegastes partitus* (closest relative with an annotated genome at the time of this study). The teleost lineage is highly prone to duplications, especially considering genes related to pigmentation (Jaillon *et al.*, 2004; Lorin *et al.*, 2018). Thus, one should not consider only the name of the sequence found with the best BLASTN hit to ascertain gene identification. To ensure the identity of the 10 most differentially expressed genes in each condition (20 genes in total), we conducted a phylogenetic analysis for each of them. Using *Stegastes partitus* sequences as query, homologous sequences were obtained by a BLASTN analysis on multiple vertebrate genomes using Ensembl (Yates et al 2016). Sequences corresponding to the best BLASTN hits for each species were retrieved. After translation, alignment was run on corresponding protein sequences using Clustal Omega implemented in Seaview software v4.6.1 using default parameters (Gouy *et al.*, 2010; Sievers and Higgins, 2014). After manual curation, tree-building was then conducted using Maximum Likelihood LG model implemented in PhyML v3.1 with SH-aLRT support (Le and Gascuel, 2008; Guindon *et al.*, 2010; Anisimova *et al.*, 2011).

Gene Set Enrichment and functional analyses

Gene Set Enrichment Analyses (GSEA) were performed with the GSEA software available at http://software.broadinstitute.org/gsea/index.jsp using the Java GSEA implementation (Subramanian *et al.*, 2005). In addition, for each DEG, the closest zebrafish ortholog was obtained for further Gene Ontology enrichment analyses using local BLASTN comparison on coding sequences available at Ensembl FTP (ftp://ftp.ensembl.org/pub/release-88/fasta/danio_rerio/). Gene Ontology and KEGG pathway enrichment analyses were conducted with zebrafish IDs using KOBAS 3.0 (KEGG Orthology-Based Annotation System) software available online at http://kobas.cbi.pku.edu.cn/index.php (Xie *et al.*, 2011). Gene Ontology terms were slimmed using KOBAS.

qPCR assay and in situ hybridization

Work in progress (P. Salis).

Obtention of mosaic CRISPR/Cas9 F0 mutant larvae

Work in progress (V. Lewis). Here is a summary:

Two sgRNA targeting a 5' exon for each gene were designed. Embryos were initially injected with the two sgRNA and Cas9 protein at the 1-2 cell stage. If survival was low (<10%), embryos were injected again at the late 16-32 cell stage to increase mosaic mutagenesis as there is no longer cytoplasmic sharing between blastomeres at this point of development. Mutagenesis was confirmed by Sanger sequencing. Fish were screened for the number or dorsal iridophores at 4 days post fertilization.

RESULTS

Microscopy analysis of skin pigmentation in *Amphiprion ocellaris*

To characterize the pigment cells found in the white and orange skins and their ultrastructural organization, we carried out light microscopy and transmission electron microscopy (TEM) analyses.

Cellular basis for the orange skin color

In the orange skin region (**Fig. 1A**), stereomicroscope experiments revealed two types of pigmented cells: orange xanthophores and black-brown melanophores (**Fig. 1B**). Xanthophores can harbor two types of pigment organelles

such as pterinosomes/xanthosomes (storing pteridines) and carotenoid vesicles. Both structures can be differentiated by their size that correspond to 0.5 µm and 0.1 µm in diameter, respectively (Djurdjevič *et al.*, 2015). Here, xanthophores were composed of xanthosomes with a characteristic size of 0.5 µm (**Fig. 1C, 1D**). Xanthophores were mostly found in the epidermis and the dermis (**Fig. 1C, 1E**). Sometimes, isolated xanthophores were observed in the hypodermis (data not shown). Interestingly, black-brown cells were observed in the orange skin. Such cells were dendritic and TEM experiments revealed that they corresponded to melanophores composed of 0.5 µm melanosomes. They were found in the epidermis and dermis (data not shown).

Cellular basis for the white skin color

We explored, using a stereomicroscope, the type of chromatophores present in *A. ocellaris* white stripe (**Fig. 2A**). Interestingly, we identified three types of cells: iridescent cells, non-iridescent white cells (not visible with a stereomicroscope) and melanophores (**Fig. 2B**). We further used TEM to understand the nature of the white coloration as two types of cells, leucophores and iridophores, can confer the white coloration in teleost fish (Kimura *et al.*, 2014). , In fish, leucophores harbor spherical vesicles with a diameter of about 0.5 µm composed of uric acid. In contrast, iridophores contain a distribution of thin crystalline platelets. Here, we observed that white cells corresponded to iridophores with typical parallel organization of platelets (**Fig. 2C and D**). These cells could be observed in epidermis, dermis and some isolated cells in the hypodermis (**Fig. 2C and 2E**). Platelet size was about 0.5 µm in iridophores, and platelets were organized concentric to the nucleus (**Fig. 2D**). In addition, isolated melanophores were also found in white skin (**Fig. 2B, 2E**).

Transcriptomic analyses of skin tissues in *Amphiprion ocellaris*

The schematic pipeline of RNA-Seq analysis that we conducted in this study is represented in **Supplementary Figure 1**. We first obtained a reference transcriptome for *A. ocellaris* that was used for comparative analyses of transcriptomic data. White and orange skin samples were compared (**Fig. 3A**). First, we showed by a Principal Component Analysis (PCA) that color was the main factor contributing to the differences between samples (**Fig. 3B**). However, a strong individual effect was visible (second axis) and justified to correct for individual effect when testing for differences in expression. In total, out of 19,204 expressed genes, 169 of them were significantly differentially expressed in our dataset ($p < 0.05$; Benjamini-Hochberg adjusted p-value): 83 and 86 genes were overexpressed in orange and white tissue, respectively (**Fig. 3C**). The full list of DEGs is available in **Supplementary Table 1.** The top 10 differentially expressed genes (DEGs) in each condition are highlighted (**Fig. 3D**).

Gene identification, especially when multiple paralogs are known, was ascertained for the 10 most DEGs in each skin type using phylogenetic analyses (trees not shown). Within the three most overexpressed genes in white stripe, two genes had previously been functionally confirmed as iridophore markers: *fhl2a* and *fhl2b*, involved in cichlid egg-spot development (Santos *et al.*, 2014). Another gene, *gpnmb*, had previously been shown to be overexpressed in iridophores compared to other pigment cells (Higdon *et al.*, 2013). Overall, this made a total of three genes proven to be (or probably) related to iridophore differentiation, out of the four most overexpressed genes in white skin. Hence, this validated the sampling and RNA-Seq protocol.

In addition, we identified putative new iridophore markers. The 2nd most overexpressed gene in the white tissue was an ortholog to zebrafish ENSDARG00000055172 (Ensembl) / 795494 (NCBI) (**Fig. 3D)**. This gene is currently annotated as *si:ch211-256m1:8* with no known associated function in Ensembl, NCBI or ZFIN databases. Here, we refer to it as *roshia* (see Discussion). Other genes overexpressed in the white stripe, such as *apod1a*, *alk* and *drd5*, had no function related to iridophore differentiation, to our knowledge (see Discussion). In contrast, *gja5* (*cx41.8*) has previously been shown to be important for melanophore differentiation in fish (Watanabe, 2017) and *trpm1a* is another melanogenic marker (Oancea *et al.*, 2009). The overexpression of these genes in the white stripe could correspond to the larger number of melanophores in this stripe compared to the orange stripe.

In the orange stripe, we identified *plin1l* as the most overexpressed gene. This gene encodes a perilipin protein. Another gene in this family, *plin6*, has previously been shown to encode a protein important for concentrating carotenoids in zebrafish xanthophores (Granneman *et al.*, 2017). In addition, we identified *retsatl*, a paralog to *retsat*, that encodes a retinol saturase and that is involved in carotenoid metabolism (Moise *et al.*, 2007). We also identified *bscl2b,* a paralog to human *bscl2* (*seipin*) that plays a role in lipid droplet formation (Szymanski *et al.*,

2007). Since xanthophores contain carotenoid, which are of lipophilic nature, the finding of such genes with a function linked to lipid metabolism within the most overexpressed genes confirmed the quality of the orange skin sampling. Surprisingly, however, all other most overexpressed genes in the orange stripe had no known function in xanthophore differentiation or in lipid metabolism, to our knowledge. Hence, they could represent new putative xanthophore markers.

Comparison of white skin tissue in *A. ocellaris* with zebrafish iridophores

Altogether, these results suggested that there was an iridophore identity in the set of the most overexpressed genes in the white stripe. To further confirm that the white color was likely due to iridophores, we retrieved a publicly available list of genes that were identified in a purified cell type transcriptomic experiment on zebrafish iridophores (Higdon *et al.*, 2013). We compared that list of "specific zebrafish iridophore markers" to our list of 86 genes overexpressed in white skin. To do so, we first retrieved *S. partitus* orthologs of zebrafish iridophore markers to obtain a list of 237 genes. The comparison of 86 clownfish "white skin genes" with the list of 237 zebrafish iridophore markers showed that 12 genes were in common between both lists (**Fig. 4A**). These genes could thus correspond to conserved iridophore markers between both species. Within these 12 genes, we identified well-established iridophore markers such as *fhl2a* and *pnp4a* (Santos *et al.*, 2014; Kimura *et al.*, 2017). More interestingly, we also identified *roshia*, *gpnmb* and *apod1a*, confirming hereby their putative function in iridophore differentiation. We equally identified *tfec*, which was previously suggested to be a marker of early iridophore lineage in zebrafish (Lister *et al.*, 2011), and *prtfdc1*, an enzyme probably involved in purine metabolism (Welin *et al.*, 2010). *Slc2a9*, an uric acid – another purine compound – transporter, was also found within the 12 genes and could thus constitute an additional putative iridophore marker (Caulfield *et al.*, 2008). All other genes in the intersect between both lists had no clear function that could be related to iridophore differentiation, to our knowledge. Of note, there was only one gene in common between the list of 83 genes overexpressed in the orange skin and the "zebrafish iridophore list" **(Fig. 4A**). This gene was *fmn2*, encoding a formin protein, with no known function in iridophore nor xanthophore differentiation – to our knowledge.

The *S. partitus* genome has 22,655 annotated protein-coding genes and the probability to randomly obtain an "iridophore gene" is 237/22,655 = 1.1 %. Thus, by chance, one would expect to obtain in the white overexpressed genes 86 x 0.011 = 0.95 gene that would also be in the list of "iridophore genes". The obtained intersect of 12 was significantly higher than this value (p-value of chi2-test < 1e-16), which suggests that the list of genes overexpressed in the white stripe was enriched in iridophore markers. When considering the list of genes overexpressed in the orange skin, there was only one gene in the intersect, showing that, relatively to iridophore gene enrichment, this tissue was not different than an average genome sample.

In addition to these 237 genes, and taking profit of recent literature reviews on different chromatophore markers (Braasch *et al.*, 2009; Lorin *et al.*, 2018), we analyzed if several well-established literature-originating iridophore markers were enriched in the RNA-Seq expression results **(Fig. 4B**). Hence, we conducted a Gene Set Enrichment Analysis using the list of iridophore markers as query. We showed that such a gene set is indeed enriched in the white tissue (FDR < 25%; p < 0.01) (**Fig. 4B**).

Gene Ontology and KEGG Pathway enrichment analyses

A functional classification of the DEGs in white and orange tissue was conducted (only results with adjusted p-value < 0.05 were considered). When considering all orange and white DEGs, the enriched terms included pigmentation related processes (**Table 1A**). More specifically, genes overexpressed in the white stripe also corresponded to pigmentation related processes (**Table 1B**). In addition to these processes, transcription factors and responses to radiation stimulus were also enriched in orange skin (**Table 1C**). KEGG analysis showed that no pathway was statistically enriched (with adjusted p-value < 0.05). Altogether, these results confirmed that DEGs corresponded to pigmentation processes.

Interestingly, within the enriched terms in orange skin were "response to abiotic stimulus" that included "response to radiation". This term corresponded to the genes *opxn4*, *rlbp1*, *rgr* and *junb* that are involved in visual cycle and could suggest, as observed in other fish, a light-sensitive property of orange skin in *Amphiprion ocellaris* (see Discussion).

qPCR validation of RNA-Seq results and skin *in situ* hybridization of candidate genes

This work is still ongoing and is conducted by P. Salis.

Experimental evidence for the implication of newly identified genes in iridophore differentiation

This work is still ongoing and is conducted by Victor Lewis in Washington University (Seattle). I show here preliminary unpublished results (courtesy of V. Lewis and D. Parichy).

To further confirm the role in iridophore differentiation for some of the genes overexpressed in white skin samples, zebrafish embryos were injected with CRISPR/Cas9 targeting a 5' exon for each gene. In particular, we designed sgRNAs targeting *roshia*, *gpnmb* and *apod1a* as well as non-specifically both *fhl2a* and *fhl2b*. These mosaic F0 mutants were then compared to wild-type, non-injected larvae, and dorsal iridophores were numbered after 4 days post-fertilization (**Fig. 5**). On average, the iridophore number for both *roshia* and *gpnmb* as well as for *fhl2a/b* mosaic F0 fish was reduced compared to wild-type, non-injected larvae ($p < 0.01$, post-ANOVA Dunnett's test). In contrast, the *apod1a* F0 mutants were not different from non-injected fish. *Roshia* mosaic larvae had the most reduced iridophore number: from ca. 20 dorsal larval iridophores in non-injected fish, there were only ca. 12 in *roshia* F0 mutants, corresponding to a 40 % decrease in iridophore number. Overall, these results provided a functional confirmation for genes identified *in silico*.

DISCUSSION

In teleost fish, the white color can be due to either leucophores or iridophores. We first performed microscopy (including TEM) analyses and showed that white band in the clownfish *Amphiprion ocellaris* harbored numerous iridophores. In addition, comparisons of clownfish orange and white skin RNA-Seq data showed that the white stripe harbored a transcriptomic signature of iridophores. The specific comparison between clownfish white stripe overexpressed genes and specific zebrafish iridophore markers highlighted 12 conserved genes that could represent conserved iridophore markers in fish. Last, some of the newly characterized candidates for iridophore differentiation were experimentally knocked down in zebrafish CRISPR/Cas9 mosaic F0 embryos. In some cases, most injected individuals subsequently developed a reduced number of dorsal iridophores.

Multiples lines of evidence suggest that color of white stripes in clownfish is due to iridophores

The implication of iridophores in white color had not been characterized in the sole previous study – to our knowledge – on clownfish color cellular basis (Yasir and Qin, 2007). In this work, we showed by multiple approaches that the white hue in clownfish is due to iridophores.

First, we performed microscopy analyses and showed that iridophores were numerous in white stripe (**Fig. 2)**. The iridophores in clownfish white stripe would correspond to zebrafish S-type iridophores. These cells contain densely stacked purine platelets with a Small platelet size of ca. 0.5-2 µm compared to L-iridophores that instead have Larger purine platelets (ca. 10 µm) organized parallel to the dermis (Hirata *et al.*, 2003). The latter iridophore type was not observed in clownfish skin sections (**Fig. 2)**. Of note, white skin – as well as orange skin – also harbored some melanophore that likely also contributed to the visual appearance of the stripes. This was in line with previous evidence in zebrafish showing that the interaction between multiple cell types can explain the final observed color (Nüsslein-Volhard and Singh, 2017).

Second, using transcriptomic analysis, we independently confirmed that white color was due to iridophores. We indeed retrieved previously identified iridophore markers as most overexpressed genes in white skin (**Fig. 3**). In particular, top 1 and 3 genes corresponded to *fhl2b* and *fhl2a*. These genes are expressed in cichlid fish anal fin egg-spots, which are localized spots of high-density iridophores. They are thus considered as iridophore markers (Santos *et al.*, 2014). Both paralogs were functionally validated by injection of CRISPR/Cas9 targeting both genes in zebrafish larvae (**Fig. 5**). These F0 mosaic mutants showed on average a 25 % reduction in iridophore number compared to wild-type, non-injected fish ($p < 0.01$, post-ANOVA Dunnett's test; **Fig. 5)**. In addition, we measured the enrichment in white skin of a previously established literature-based list of iridophore markers (Braasch *et al.*, 2009; Lorin *et al.*, 2018). We showed that this gene set was statistically enriched in clownfish white skin. Of note, some 'classical' iridophore markers were not enriched in white skin and were rather overexpressed in orange skin, although not significant (**Fig. 4B**). These included *chm*, a crucial gene involved in iridophore distribution in *D. rerio* as well as *mpv17*, involved in iridophore survival in *D. rerio* and responsible, when mutated, for the *transparent* zebrafish mutant (Starr *et al.*, 2004; Krauss *et al.*, 2013). To our knowledge, both genes have not been shown to be involved in iridophore

differentiation in any other fish species, including medaka. Hence, this could suggest that *mpv17* and *chm* are specific to zebrafish iridophore differentiation. Alternatively, these genes could have lost their function in the clownfish lineage after divergence from zebrafish ca. 230 Ma (Betancur-R *et al.*, 2017). Finally, this could also indicate that, at adult stage, they are no longer required in iridophore metabolism but that they have another unknown function in orange skin.

Third, we compared a list of previously published zebrafish iridophore markers with overexpressed genes in clownfish white skin (Higdon *et al.*, 2013). The overlap between both of these lists was higher than at random, showing that white skin in clownfish had a transcriptomic signature of iridophores (**Fig. 4A**). In particular, 12 genes were common to both datasets and could thus constitute new iridophore markers (see below). Yet, although significant, the overlap was relatively small compared to the size of respective clownfish (86) and zebrafish candidates (237). This could indicate, in line with the expression pattern of *chm* and *mpv17* discussed above, that zebrafish and clownfish iridophore differentiation processes partially differ. Of note, the zebrafish RNA-Seq data corresponded to purified iridophores whereas clownfish RNA-Seq corresponds to whole skin samples. Thus, clownfish data represented a more diverse set of cells and tissues compared to a single-cell, purified, experiment. This biological difference in the data could alternatively explain the apparent discrepancy observed between both datasets.

Finally, the probable implication of iridophores in white stripe is consistent with previous observations that other clownfish species, such as *Amphiprion chrysopterus*, *A. frenatus* and *A. allardi* have one or multiple white stripes that can turn to blue depending on observation angle (T. Lorin, P. Salis, pers. obs.). In particular, the blue hue is more pronounced in *A. chrysopterus* that is hence commonly named the "blue striped clownfish". Thus, the cellular basis for *A. ocellaris* skin color could be conserved with other species in this genus.

Identification of new putative iridophore markers previously known to be involved in other pigmentation-related processes

The comparison of white and orange stripes skin samples provided a list of multiple genes overexpressed in each condition, some of which might constitute new iridophore and xanthophore markers (xanthophore markers are discussed below). Among the 10 most overexpressed genes in clownfish white stripe, three were previously associated with confirmed or putative iridophore differentiation processes: *fhl2a* and *fhl2b* (discusses above) and *gpnmb*.

The top 4 gene overexpressed in white stripe was *gpnmb*, encoding a glycoprotein. The first pigmentation-related function of this gene in pigmentation has been melanosome development in melanocytes (Zhang *et al.*, 2012). More recently, this gene was shown to be more expressed in zebrafish iridophores than melanophores (Higdon *et al.*, 2013). This gene was also expressed in cichlid egg-spots (Santos *et al.*, 2016), raising the possibility that melanosomes and purine platelets could share formation processes (Higdon *et al.*, 2013). We functionally confirmed that F0 *gpnmb* mosaic mutants had on average 30 % fewer iridophores at four days post-fertilization compared to non-injected fish ($p < 0.001$, post-ANOVA Dunnett's test; **Fig. 5**). Our results confirmed the previously hypothesized implication of *gpnmb* in iridophore differentiation.

Among the 10 most overexpressed genes in the white band, we additionally identified two other genes associated to iridophore-*independent* pigmentation processes: *gja5* and *trpm1a*. The former, also coined *cx41.8*, encodes a connexin that has already been shown to be involved in pattern formation in zebrafish: a defect in this gene gives the *leo* mutant, characterized by a spotted pigment pattern (Irion *et al.*, 2014). This gene is thus involved in fish general patterning as well as in melanophore and xanthophore – but not iridophore – differentiation (Irion *et al.*, 2014). In zebrafish, the striped pattern is mainly due to a reaction-diffusion mechanism and the presence of this gene among the most overexpressed genes in white tissue could indicate that this very mechanism is producing the striped pattern in clownfish (Watanabe and Kondo, 2015; Eom and Parichy, 2017). However, clownfish bands do not form through a reaction-diffusion mechanism (Salis *et al.*, submitted). Thus, interestingly, and even though gene expression does not necessarily translates into gene function, *gja5* might be involved in clownfish white band development through a reaction-diffusion *independent* mechanism.

In addition, *trpm1*, a melanogenic gene, was highly overexpressed in white skin (Oancea *et al.*, 2009). This melanogenic gene was retained after the teleost-specific whole-genome duplication, or TSGD (Braasch *et al.*, 2009; Lorin *et al.*, 2018). In white tissue *trpm1a* was the 10th most overexpressed gene (**Fig. 3**) and *trpm1b* was ranked 11th (**Supplementary Table 1**). This observation is quite

puzzling as these genes have not previously been associated with iridophore differentiation – to our knowledge. Their presence could suggest that white skin harbors more melanophores than orange skin, leading to the overexpression of these markers. However, in that case, we should also identify numerous other melanogenic genes in white skin, such as well-known candidates *mc1r* or *mitfa*. This was not the case and, among the numerous genes involved in melanogenesis, only one was also overexpressed in white skin – *slc24a5*, a cation exchanger involved in melanocyte development (Lamason *et al.*, 2005) (**Supplementary Table 1**). Thus, we suggest that the role of *trpm1* in fish iridophore differentiation should be scrutinized with more detail.

Finally, within the 12 genes present in both zebrafish and clownfish lists, we also identified *tfec* that has previously been suggested to be an early iridophore marker (Lister *et al.*, 2011; Higdon *et al.*, 2013). Our results confirmed its implication in clownfish, hereby adding a line of evidence to its conserved function in iridophore differentiation in fish.

Identification of new putative iridophore markers with no previously known pigmentary function

In addition to the pigmentation markers discussed above, we identified a range of other genes that have not been previously shown to be involved in pigmentation acquisition – to our knowledge. Yet, some of them could represent putative new markers.

First, among the identified genes, three were previously known to be involved in purine metabolism. Since iridophore coloration is based on purine platelets, this could reflect a putative function of these genes in platelet formation. The first such gene was *ada2* (or *cecr1*), encoding an adenosine deaminase which has previously been shown to be involved in purine metabolism (Pei *et al.*, 2016). This gene was ranked 6th within the most overexpressed genes in white skin (**Fig. 3)**. Within the 12 genes in the overlap between zebrafish and clownfish lists was present *prtfdc1*, an enzyme probably involved in purine metabolism but for which no precise function has been documented to date (Welin *et al.*, 2010) (**Fig. 4A**). In this very list intersect, *slc2a9* was another gene related to purine metabolism since it is involved in uric acid transport (Caulfield *et al.*, 2008) **(Fig. 4A)**. Though uric acid is normally restricted to leucophores, the current understanding of *slc2a9* function is limited and cannot exclude a more general role for this gene transport of other purine-derived compounds.

Furthermore, and very interestingly, we also identified in this study genes with no known function to date in any process related to pigmentation nor purine metabolism. Here, we only discuss such genes that are either genes in the 10 white most overexpressed genes or in the 12 genes intersect. Indeed, the plethora of "omics" studies is currently providing a wealth of candidate genes (San-Jose and Roulin, 2017). Finding "the needle in the haystack", *i. e.* identifying meaningful markers within numerous genes in a list, is a harassing process. Thus, one often first considers the genes with the highest fold-change value between conditions, as exemplified by the recently discovered pigmentation marker *alx3* in rodents (Mallarino *et al.*, 2016). As also performed in the present study, another way of restricting the number of potential candidates and to retrieve conserved genes in multiple organisms is to cross results between multiple studies on independent datasets and species.

Ranked 2nd within the most overexpressed genes in white stripe was *si:ch211-256m1:8,* hereafter coined *roshia* (**Fig. 3**). This gene was also identified in the comparison between both zebrafish and iridophore datasets (**Fig. 4A).** *Roshia* seemed to be one of both paralogs of an ancestor *roshi* gene that has been conserved after the TSGD genome duplication (Lorin *et al.*, in prep.). Of note, this name is after the eponym character that constantly wears sunglasses to be protected from the shining sun – in relation with the probable function of *roshia* in the differentiation of these silvery, light-reflecting cells. Most interestingly, *roshia* was indeed functionally confirmed as an iridophore marker (**Fig. 5**). The mosaic CRISPR/Cas9 F0 mutant for *roshia* harbored ca. 40% fewer iridophores compared to wild-type non-injected larvae ($p < 0.001$, post-ANOVA Dunnett's test; **Fig. 5**). Although not significant, this reduction was even more drastic than observed for "classical" iridophore markers *fhl2a/b*.

Also found to be conserved in both clownfish and zebrafish was *apod1a*, encoding an apolipoprotein involved in lipid binding and transport. The apolipoprotein gene *apod1* has, to date, never been identified in iridophore development (but see non-peer reviewed work by (Gu *et al.*, 2017)). Yet, the occurrence of this gene within both datasets encouraged us to test its function in iridophore differentiation. However, we could not detect any difference between injected and non-injected fish (**Fig. 5**).

Finally, only present in the 10 most overexpressed genes was also *drd5*, a dopamin receptor, and *alk*, encoding

a tyrosine kinase receptor that is conserved in vertebrates – including humans – with unknown biochemical function (Fadeev *et al.*, 2018). The finding of the latter is very puzzling as one would expect to find *ltk* instead – the *alk* paralog (Brunet *et al.*, 2016). Indeed, zebrafish deficient in *ltk* ("shady" mutants) do not harbor iridophores and are widely-used models to study vertebrate pigmentation (Lopes *et al.*, 2008; Fadeev *et al.*, 2016). Intriguingly, *ltk* was not identified within DEGs (**Supplementary Table 1)** but was still more expressed in white skin compared to orange skin, although not significant (**Fig. 4B**). Thus, in clownfish, a putative role for *alk* in iridophore differentiation is suggested for the first time and suggest that this gene could be interchangeable with its *ltk* paralog to accomplish iridophore differentiation. This new function would thus have been acquired after divergence between clownfish and zebrafish lineages ca. 230 Ma (Betancur-R *et al.*, 2017). Of note, this would not be the first occurrence of a pigmentation gene involved in some process in one fish species and not in another, as was recently demonstrated for the role of *sox5*, repressing xanthophores in zebrafish and activating them in medaka (Nagao *et al.*, 2018).

Orange skin in clownfish has a peculiar transcriptomic signature

As far as orange tissue is concerned, one would expect to retrieve well-established xanthophore markers such as *csf1ra* (aka *fms*) (Parichy and Turner, 2003). However, in this work, we did not identify any previously known xanthophore marker within genes overexpressed in the orange tissue– neither in the 10 most overexpressed genes in orange stripe nor in the whole DEG list (**Fig. 3, Supplementary Table 1)**. In contrast, surprisingly, *csf1ra* was identified as slightly overexpressed in the white tissue (**Supplementary Table 1).**

Ranked 1st in orange tissue was the perilipin-1-like *plin1l*. This gene is part of the PLIN multigene family, a major protein family involved in lipid droplet formation (Granneman *et al.*, 2017). This protein could thus play the same role in carotenoid-droplet formation than the xanthophore marker *plin6* (Granneman *et al.*, 2017). Two other genes involved in lipid metabolism, *retsat*, a retinol saturase, and *bscl2*, involved in lipid droplet formation, could constitute new xanthophore markers that would require further validation. Other most overexpressed genes in the orange skin had no clear functional relationship with specific xanthophore properties.

Interestingly, the first enriched term in orange DEGs corresponded to "response to abiotic stimulus" and includes all genes in the second term, "response to radiation": *opxn4*, *rlbp1*, *rgr* and *junb* (**Table 1**). These four genes have been found to be involved in the visual cycle and in opsin activities (Chen *et al.*, 2001; Gu *et al.*, 2007; Moraes *et al.*, 2014; Xue *et al.*, 2015). Skin chromatophores are known to be light-sensitive in a wide range of animal pigment cells, including in fish erythrophores (Ban *et al.*, 2005; Chen *et al.*, 2013). The enrichment of such term in *A. ocellaris* orange skin could reveal the ability of this tissue to change the distribution of pigments, or increase their synthesis, in response to light.

Leucophores are probably not involved in clownfish white stripe

In this work, we showed with multiples lines of evidence that white color of clownfish stripes was most probably due to iridophores. In addition, several other lines of evidence suggested that this color is likely not due to leucophores. First, at a histological level, leucophores that have been described to date are usually dispersed isolated cells (Kimura *et al.*, 2014). Thus, they do not constitute a dense, intermingled network of chromatophores that could provide the basis for the pigmentation of a whole stripe. Besides, current knowledge of iridophore markers is limited since only one gene, *slc2a15b*, has been documented as a specific leucophore marker (Kimura *et al.*, 2014). This gene was not found to be expressed in skin samples, and its Ts3R-paralog, *slc2a15a*, had comparable levels between white and orange skin (not shown).

CONCLUSION

In this study we have shown using multiple lines of evidence that the white color of an iconic fish species, clownfish *Amphiprion ocellaris*, is due to iridophores. In addition, we confirmed the role of previously hypothesized iridophore markers and identified new iridophore markers, some of which were functionally verified using CRISPR/Cas9.

ACKNOWLEDGEMENTS

The authors are much obliged to A. Marcionetti and N. Salamin for providing access to the *A. ocellaris* genome sequence. T. L. is recipient of an individual grant from the French Ministry of Research and Education.

LITERATURE CITED

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J Mol Biol* **215**: 403–410.

Anisimova M, Gil M, Dufayard J-F, Dessimoz C, Gascuel O (2011). Survey of Branch Support Methods Demonstrates Accuracy, Power, and Robustness of Fast Likelihood-based Approximation Schemes. *Syst Biol* **60**: 685–699.

Bagnara JT, Hadley EM (1973). *Chromatophores and Color Change. The Comparative Physiology of Animal Pigmentation*. Englewood Cliffs: Prentice-Hall.

Ban E, Kasai A, Sato M, Yokozeki A, Hisatomi O, Oshima N (2005). The signaling pathway in photoresponses that may be mediated by visual pigments in erythrophores of Nile tilapia. *Pigment Cell Res* **18**: 360–369.

Betancur-R R, Wiley EO, Arratia G, Acero A, Bailly N, Miya M, *et al.* (2017). Phylogenetic classification of bony fishes. *BMC Evol Biol* **17**: 162.

Braasch I, Brunet F, Volff J-N, Schartl M (2009). Pigmentation pathway evolution after whole-genome duplication in fish. *Genome Biol Evol* **1**: 479–493.

Bray NL, Pimentel H, Melsted P, Pachter L (2016). Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**: 525–527.

Brunet FG, Volff J-N, Schartl M (2016). Whole Genome Duplications Shaped the Receptor Tyrosine Kinase Repertoire of Jawed Vertebrates. *Genome Biol Evol* **8**: 1600–1613.

Caulfield MJ, Munroe PB, O'Neill D, Witkowska K, Charchar FJ, Doblado M, *et al.* (2008). SLC2A9 Is a High-Capacity Urate Transporter in Humans (A Hattersley, Ed.). *PLoS Med* **5**: e197.

Chen P, Hao W, Rife L, Wang XP, Shen D, Chen J, *et al.* (2001). A photic visual cycle of rhodopsin regeneration is dependent on Rgr. *Nat Genet* **28**: 256–260.

Chen S-C, Robertson RM, Hawryshyn CW (2013). Possible Involvement of Cone Opsins in Distinct Photoresponses of Intrinsically Photosensitive Dermal Chromatophores in Tilapia Oreochromis niloticus (K-W Koch, Ed.). *PLoS One* **8**: e70342.

Cuthill IC, Allen WL, Arbuckle K, Caspers B, Chaplin G, Hauber ME, *et al.* (2017). The biology of color. *Science (80-)* **357**: eaan0221.

Darwin C (1868). *The Variation of Animals and Plants Under Domestication*. John Murray, London.

Djurdjevič I, Kreft ME, Sušnik Bajec S (2015). Comparison of pigment cell ultrastructure and organisation in the dermis of marble trout and brown trout, and first description of erythrophore ultrastructure in salmonids. *J Anat* **227**: 583–595.

Eom DS, Parichy DM (2017). A macrophage relay for long-distance signaling during postembryonic tissue remodeling. *Science (80-)* **355**: 1317–1320.

Fadeev A, Krauss J, Singh AP, Nüsslein-Volhard C (2016). Zebrafish Leucocyte tyrosine kinase controls iridophore establishment, proliferation and survival. *Pigment Cell Melanoma Res* **29**: 284–296.

Fadeev A, Mendoza-Garcia P, Irion U, Guan J, Pfeifer K, Wiessner S, *et al.* (2018). ALKALs are in vivo ligands for ALK family receptor tyrosine kinases in the neural crest and derived cells. *Proc Natl Acad Sci* **115**: E630–E638.

Fujii R (1993). Cytophysiology of Fish Chromatophores. In: International Review of Cytology, pp 191–255.

Goda M (2017). Rapid integumental color changes due to novel iridophores in the chameleon sand tilefish Hoplolatilus chlupatyi. *Pigment Cell Melanoma Res* **30**: 368–371.

Goda M, Fujii R (1995). Blue Chromatophores in Two Species of Callionymid Fish. *Zoolog Sci* **12**: 811–813.

Goda M, Fujiyoshi Y, Sugimoto M, Fujii R (2013). Novel Dichromatic Chromatophores in the Integument of the Mandarin Fish Synchiropus splendidus. *Biol Bull* **224**: 14–17.

Goda M, Ohata M, Ikoma H, Fujiyoshi Y, Sugimoto M, Fujii R (2011). Integumental reddish-violet coloration owing to novel dichromatic chromatophores in the teleost fish, Pseudochromis diadema. *Pigment Cell Melanoma Res* **24**: 614–617.

Gouy M, Guindon S, Gascuel O (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**: 221–224.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, *et al.* (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **29**: 644–652.

Granneman JG, Kimler VA, Zhang H, Ye X, Luo X, Postlethwait JH, *et al.* (2017). Lipid droplet biology and evolution illuminated by the characterization of a novel perilipin in teleost fish. *Elife* **6**.

Gu D, Beltran WA, Li Z, Acland GM, Aguirre GD (2007). Clinical Light Exposure, Photoreceptor Degeneration, and AP-1 Activation: A Cell Death or Cell Survival Signal in the Rhodopsin Mutant Retina? *Investig Opthalmology Vis Sci* **48**: 4907.

Gu L, Zhang C, Xia C (2017). Revelation of the Genetic Basis for Convergent Innovative Anal Fin Pigmentation Patterns in Cichlid Fishes. *bioRxiv*.

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst Biol* **59**: 307–321.

Higdon CW, Mitra RD, Johnson SL (2013). Gene Expression Analysis of Zebrafish Melanocytes, Iridophores, and Retinal Pigmented Epithelium Reveals Indicators of Biological Function and Developmental Origin (YF Leung, Ed.). *PLoS One* **8**: e67801.

Hirata M, Nakamura K, Kanemaru T, Shibata Y, Kondo S (2003). Pigment cell organization in the hypodermis of zebrafish. *Dev Dyn* **227**: 497–503.

Hoekstra HE (2006). Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity (Edinb)* **97**: 222–234.

Irion U, Frohnhöfer HG, Krauss J, Çolak Champollion T, Maischein H-M, Geiger-Rudolph S, *et al.* (2014). Gap junctions composed of connexins 41.8 and 39.4 are essential for colour pattern formation in zebrafish. *Elife* **3**.

Jaillon O, Aury J-M, Brunet F, Petit J-L, Stange-Thomann N, Mauceli E, *et al.* (2004). Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature* **431**: 946–957.

Kim D, Langmead B, Salzberg SL (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**: 357–360.

Kimura T, Nagao Y, Hashimoto H, Yamamoto-Shiraishi Y -i., Yamamoto S, Yabe T, *et al.* (2014). Leucophores are similar to xanthophores in their specification and differentiation processes in medaka. *Proc Natl Acad Sci* **111**: 7343–7348.

Kimura T, Takehana Y, Naruse K (2017). pnp4a Is the Causal Gene of

the Medaka Iridophore Mutant guanineless. *G3 Genes|Genomes|Genetics Genes|Genomes|Genetics* **7**: 1357–1363.

Kinsella RJ, Kahari A, Haider S, Zamora J, Proctor G, Spudich G, *et al.* (2011). Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database* **2011**: bar030-bar030.

Kopylova E, Noé L, Touzet H (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**: 3211–3217.

Krauss J, Astrinides P, Frohnhofer HG, Walderich B, Nusslein-Volhard C (2013). transparent, a gene affecting stripe formation in Zebrafish, encodes the mitochondrial protein Mpv17 that is required for iridophore survival. *Biol Open* **2**: 703–710.

Lamason RL, Mohideen M-APK, Mest JR, Wong AC, Norton HL, Aros MC, *et al.* (2005). SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science* **310**: 1782–6.

Le SQ, Gascuel O (2008). An Improved General Amino Acid Replacement Matrix. *Mol Biol Evol* **25**: 1307–1320.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, *et al.* (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–9.

Lister JA, Lane BM, Nguyen A, Lunney K (2011). Embryonic expression of zebrafish MiT family genes tfe3b, tfeb, and tfec. *Dev Dyn* **240**: 2529–2538.

Litsios G, Sims CA, Wüest RO, Pearman PB, Zimmermann NE, Salamin N (2012). Mutualism with sea anemones triggered the adaptive radiation of clownfishes. *BMC Evol Biol* **12**: 212.

Lopes SS, Yang X, Müller J, Carney TJ, McAdow AR, Rauch G-J, *et al.* (2008). Leukocyte Tyrosine Kinase Functions in Pigment Cell Development (G Barsh, Ed.). *PLoS Genet* **4**: e1000026.

Lorin T, Brunet FG, Laudet V, Volff J-N (2018). Teleost Fish-Specific Preferential Retention of Pigmentation Gene-Containing Families After Whole Genome Duplications in Vertebrates. *G3 Genes|Genomes|Genetics*: g3.200201.2018.

Love MI, Huber W, Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Mallarino R, Henegar C, Mirasierra M, Manceau M, Schradin C, Vallejo M, *et al.* (2016). Developmental mechanisms of stripe patterns in rodents. *Nature* **539**: 518–523.

Mariscal RN (1966). The symbiosis between tropical sea anemones and fishes: a review. In: *The Galápagos: Proceedings of the Symposia of the Galápagos International Scientific Project*, University of California Press, p 318.

Militz TA, Foale S (2017). The 'Nemo Effect': Perception and reality of Finding Nemo 's impact on marine aquarium fisheries. *Fish Fish* **18**: 596–606.

Moise AR, Isken A, Domínguez M, de Lera AR, von Lintig J, Palczewski K (2007). Specificity of Zebrafish Retinol Saturase: Formation of All-trans -13,14-dihydroretinol and All- trans -7,8- dihydroretinol †. *Biochemistry* **46**: 1811–1820.

Moraes MN de CM, Lima LHRG de, Ramos BCR, Poletini M de O, Castrucci AM de L (2014). Endothelin modulates the circadian expression of non-visual opsins. *Gen Comp Endocrinol* **205**: 279–286.

Nagao Y, Takada H, Miyadai M, Adachi T, Seki R, Kamei Y, *et al.* (2018). Distinct interactions of Sox5 and Sox10 in fate specification of pigment cells in medaka and zebrafish (M Schartl, Ed.). *PLOS Genet* **14**: e1007260.

Near TJ, Eytan RI, Dornburg A, Kuhn KL, Moore JA, Davis MP, *et al.* (2012). Resolution of ray-finned fish phylogeny and timing of diversification. *Proc Natl Acad Sci* **109**: 13698–13703.

Nelson JS, Grande TC, Wilson MVH (2016). *Fishes of the World*, 5th edn. John Wiley & Sons, Inc: Hoboken, NJ, USA.

Nüsslein-Volhard C, Singh AP (2017). How fish color their skin: A paradigm for development and evolution of adult patterns. *BioEssays* **39**: 1600231.

Oancea E, Vriens J, Brauchi S, Jun J, Splawski I, Clapham DE (2009). TRPM1 Forms Ion Channels Associated with Melanin Content in Melanocytes. *Sci Signal* **2**: ra21-ra21.

Parichy DM, Turner JM (2003). Temporal and cellular requirements for Fms signaling during zebrafish adult pigment pattern development. *Development* **130**: 817–33.

Pei W, Xu L, Varshney GK, Carrington B, Bishop K, Jones M, *et al.* (2016). Additive reductions in zebrafish PRPS1 activity result in a spectrum of deficiencies modeling several human PRPS1-associated diseases. *Sci Rep* **6**: 29946.

San-Jose LM, Roulin A (2017). Genomics of coloration in natural animal populations. *Philos Trans R Soc B Biol Sci* **372**: 20160337.

Santos ME, Baldo L, Gu L, Boileau N, Musilova Z, Salzburger W (2016). Comparative transcriptomics of anal fin pigmentation patterns in cichlid fishes. *BMC Genomics* **17**: 712.

Santos E, Braasch I, Boileau N, Meyer BS, Bo A, Belting H, *et al.* (2014). The evolution of cichlid fish egg-spots is linked with a cis-regulatory change. *Nat Commun*.

Schartl M, Larue L, Goda M, Bosenberg MW, Hashimoto H, Kelsh RN (2016). What is a vertebrate pigment cell? *Pigment Cell Melanoma Res* **29**: 8–14.

Shapiro MD, Kronenberg Z, Li C, Domyan ET, Pan H, Campbell M, *et al.* (2013). Genomic Diversity and Evolution of the Head Crest in the Rock Pigeon. *Science (80-)* **339**: 1063–1067.

Sievers F, Higgins DG (2014). Clustal omega, accurate alignment of very large numbers of sequences. In: *Methods in Molecular Biology*, Vol 1079, pp 105–116.

Singh AP, Nüsslein-Volhard C (2015). Zebrafish Stripes as a Model for Vertebrate Colour Pattern Formation. *Curr Biol* **25**: R81–R92.

Soneson C, Love MI, Robinson MD (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* **4**: 1521.

Starr CJ, Kappler JA, Chan DK, Kollmar R, Hudspeth AJ (2004). From The Cover: Mutation of the zebrafish choroideremia gene encoding Rab escort protein 1 devastates hair cells. *Proc Natl Acad Sci* **101**: 2572–2577.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, *et al.* (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci* **102**: 15545–15550.

Szymanski KM, Binns D, Bartz R, Grishin N V., Li W-P, Agarwal AK, *et al.* (2007). The lipodystrophy protein seipin is found at endoplasmic reticulum lipid droplet junctions and is important for droplet morphology. *Proc Natl Acad Sci* **104**: 20890–20895.

Wallace AR (1877). The colours of animals and plants. *Am Nat* **11**: 713–728.

Watanabe M (2017). Gap Junction in the Teleost Fish Lineage: Duplicated Connexins May Contribute to Skin Pattern Formation and Body Shape Determination. *Front Cell Dev Biol* **5**.

Watanabe M, Kondo S (2015). Is pigment patterning in fish skin determined by the Turing mechanism? *Trends Genet* **31**: 88–96.

Welin M, Egeblad L, Johansson A, Stenmark P, Wang L, Flodin S, *et al.* (2010). Structural and functional studies of the human phosphoribosyltransferase domain containing protein 1. *FEBS J* **277**: 4920–4930.

Wucherer MF, Michiels NK (2012). A Fluorescent Chromatophore Changes the Level of Fluorescence in a Reef Fish (J Zeil, Ed.). *PLoS One* **7**: e37913.

Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, *et al.* (2011). KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res* **39**: W316–W322.

Xue Y, Shen SQ, Jui J, Rupp AC, Byrne LC, Hattar S, *et al.* (2015). CRALBP supports the mammalian retinal visual cycle and cone vision. *J Clin Invest* **125**: 727–738.

Yasir I, Qin JG (2007). Embryology and early ontogeny of an anemonefish Amphiprion ocellaris. *J Mar Biol Assoc UK* **87**: 1025.

Zhang P, Liu W, Zhu C, Yuan X, Li D, Gu W, *et al.* (2012). Silencing of GPNMB by siRNA Inhibits the Formation of Melanosomes in Melanocytes in a MITF-Independent Fashion (RA Veitia, Ed.). *PLoS One* **7**: e42955.

FIGURES

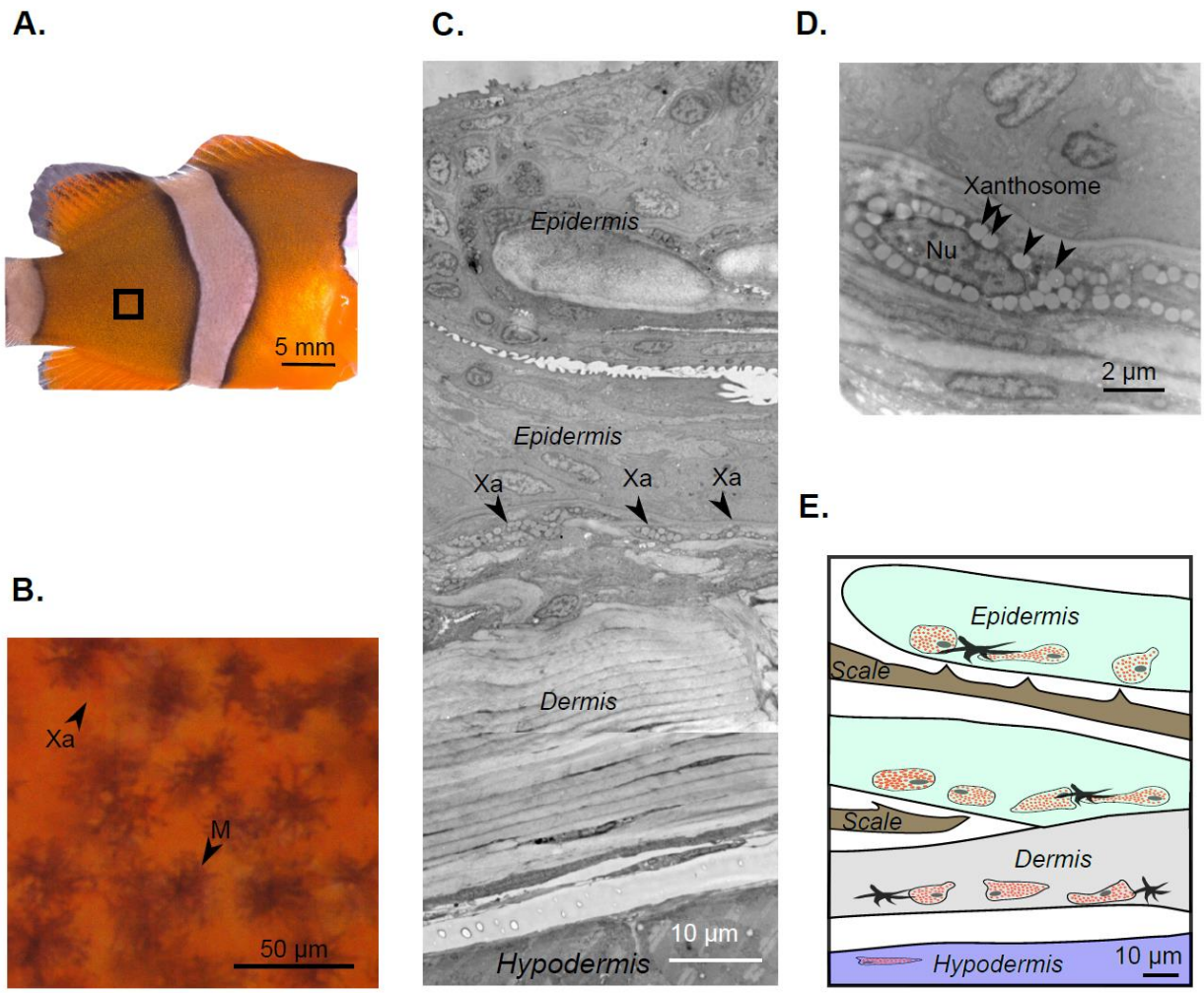


Figure 1: Pigment cells organization in orange skin of *A. ocellaris*. (A) Photography of the region used for stereomicroscope images and TEM experiments. (B) Stereomicroscope image of the orange skin showing xanthophores (Xa) and melanophores (M) cells. (C) TEM image of a transversal section of the orange skin. (D) TEM high magnification of a xanthophore showing xanthosome (arrowhead). (E) Cartoon representing a transversal section of the orange skin of *A. ocellaris* with xanthophores (represented by the orange cells) and melanophores (cells with stars shape). [M= melanophores, Xa= Xanthophores, Nu= Nucleus]

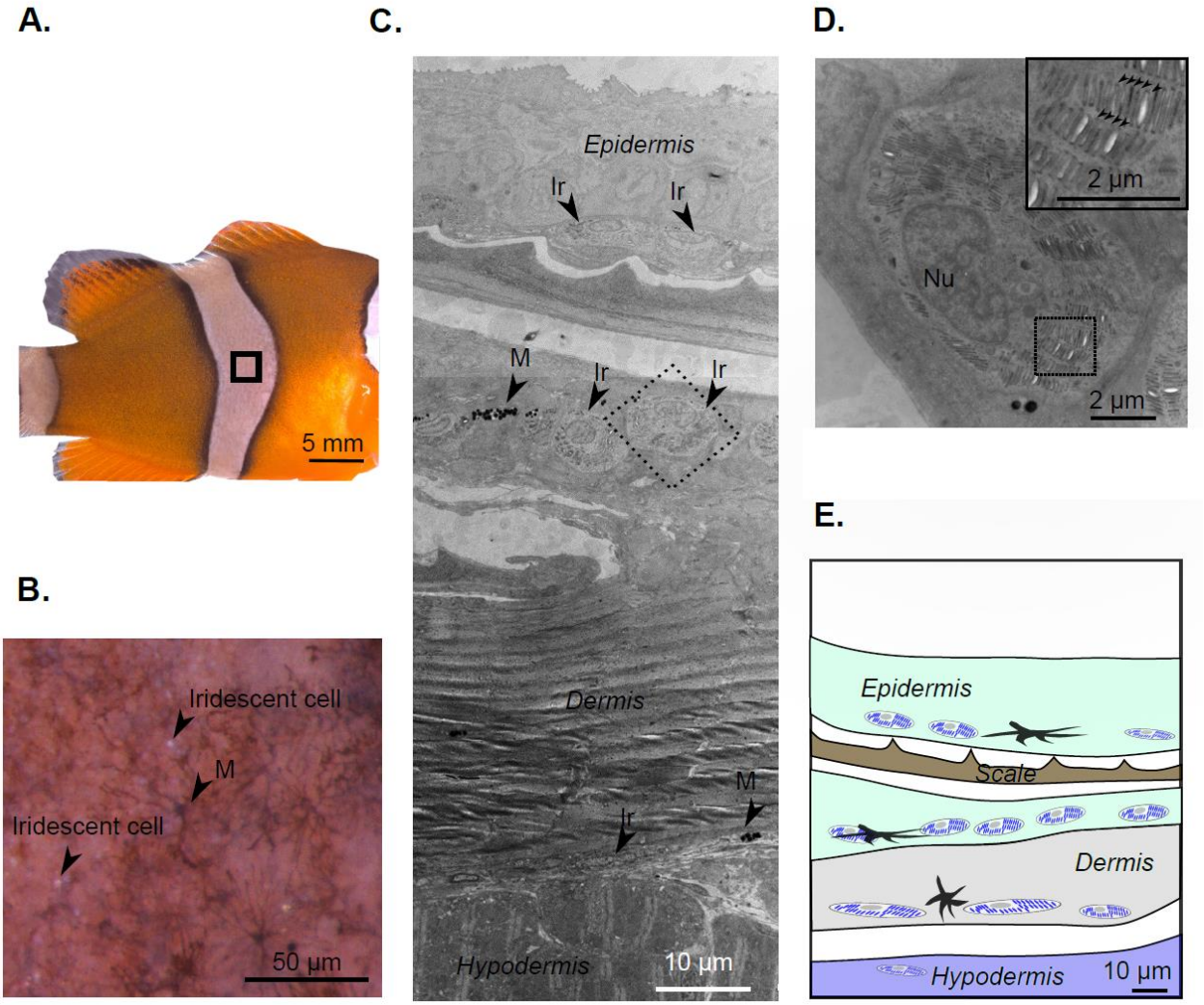
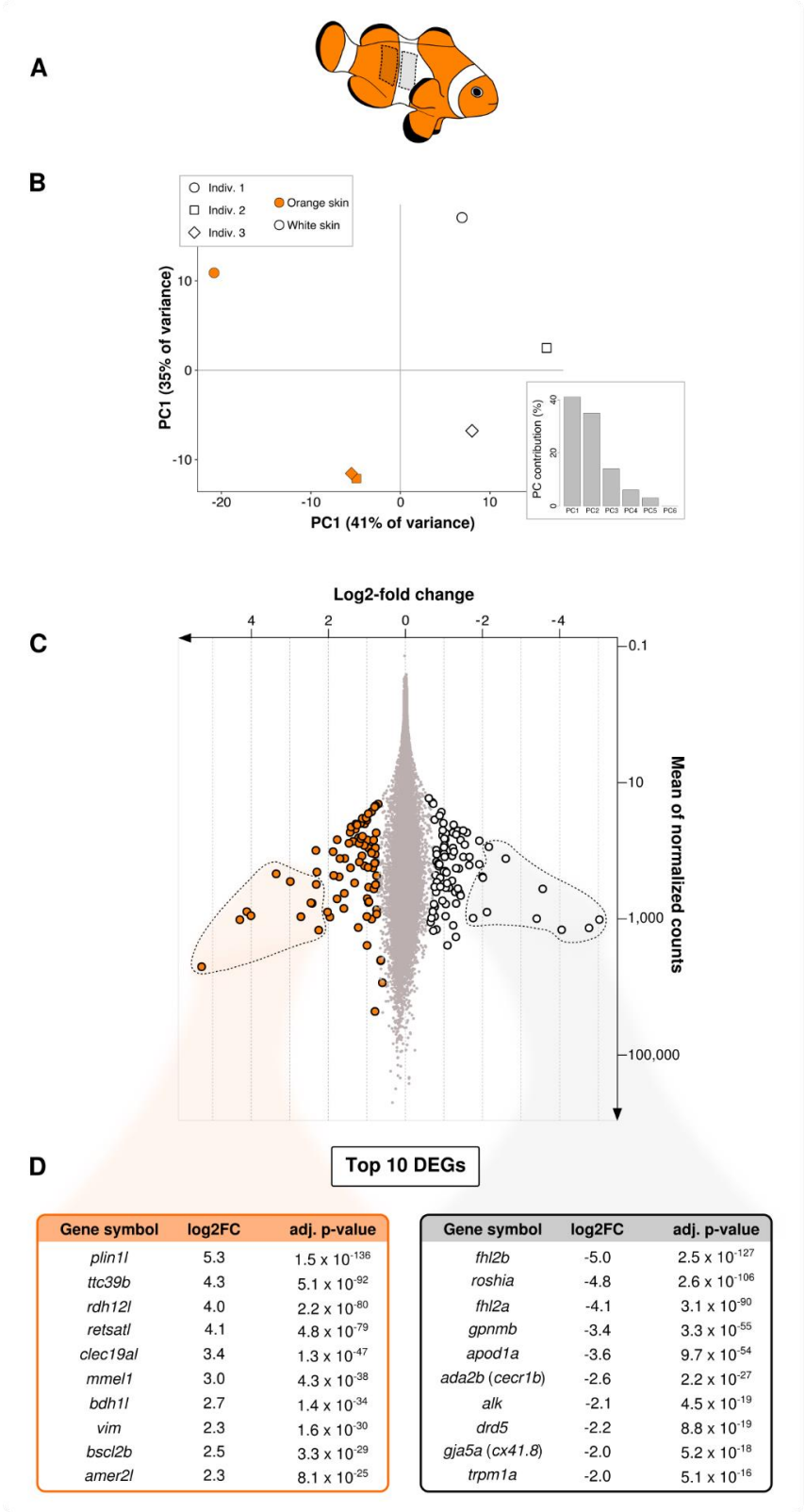


Figure 2: Pigment cells organization in white skin of *A. ocellaris*. (a) Photography of the region used for stereomicroscope images and TEM experiments. (b) Stereomicroscope image of the white skin showing Ir cells, white color and M cells. (c) TEM image of a transversal section of the white skin. (d) TEM high magnification of a Ir showing parallel platelets (arrowheads). (e) Cartoon representing a transversal section of the white skin of *A. ocellaris* with Ir (represented by the blue cells) and Melanophores (cells with stars shape). [M= melanophores, Ir = Iridophores]



Gene symbol	log2FC	adj. p-value
plin1l	5.3	1.5×10^{-136}
ttc39b	4.3	5.1×10^{-92}
rdh12l	4.0	2.2×10^{-80}
retsatl	4.1	4.8×10^{-79}
clec19al	3.4	1.3×10^{-47}
mmel1	3.0	4.3×10^{-38}
bdh1l	2.7	1.4×10^{-34}
vim	2.3	1.6×10^{-30}
bscl2b	2.5	3.3×10^{-29}
amer2l	2.3	8.1×10^{-25}

Gene symbol	log2FC	adj. p-value
fhl2b	-5.0	2.5×10^{-127}
roshia	-4.8	2.6×10^{-106}
fhl2a	-4.1	3.1×10^{-90}
gpnmb	-3.4	3.3×10^{-55}
apod1a	-3.6	9.7×10^{-54}
ada2b (*cecr1b*)	-2.6	2.2×10^{-27}
alk	-2.1	4.5×10^{-19}
drd5	-2.2	8.8×10^{-19}
gja5a (*cx41.8*)	-2.0	5.2×10^{-18}
trpm1a	-2.0	5.1×10^{-16}

Figure 3: Comparison of transcriptomic results between white and orange skin samples.
(A) A schematic representation of the dissected area is shown (shaded areas). **(B)** PCA of data shows a separation between white and orange samples. **(C)** On the MA plot, statistically DEGs are highlighted compared to background genes: orange (resp. white) dots represent 83 genes with positive log2-fold change (resp. 86 genes with negative log2-fold change) genes overexpressed in orange (resp. white) tissue. **(D)** The 10 most overexpressed genes in each condition (dashed line) are highlighted. DEGs: Differentially Expressed Genes.

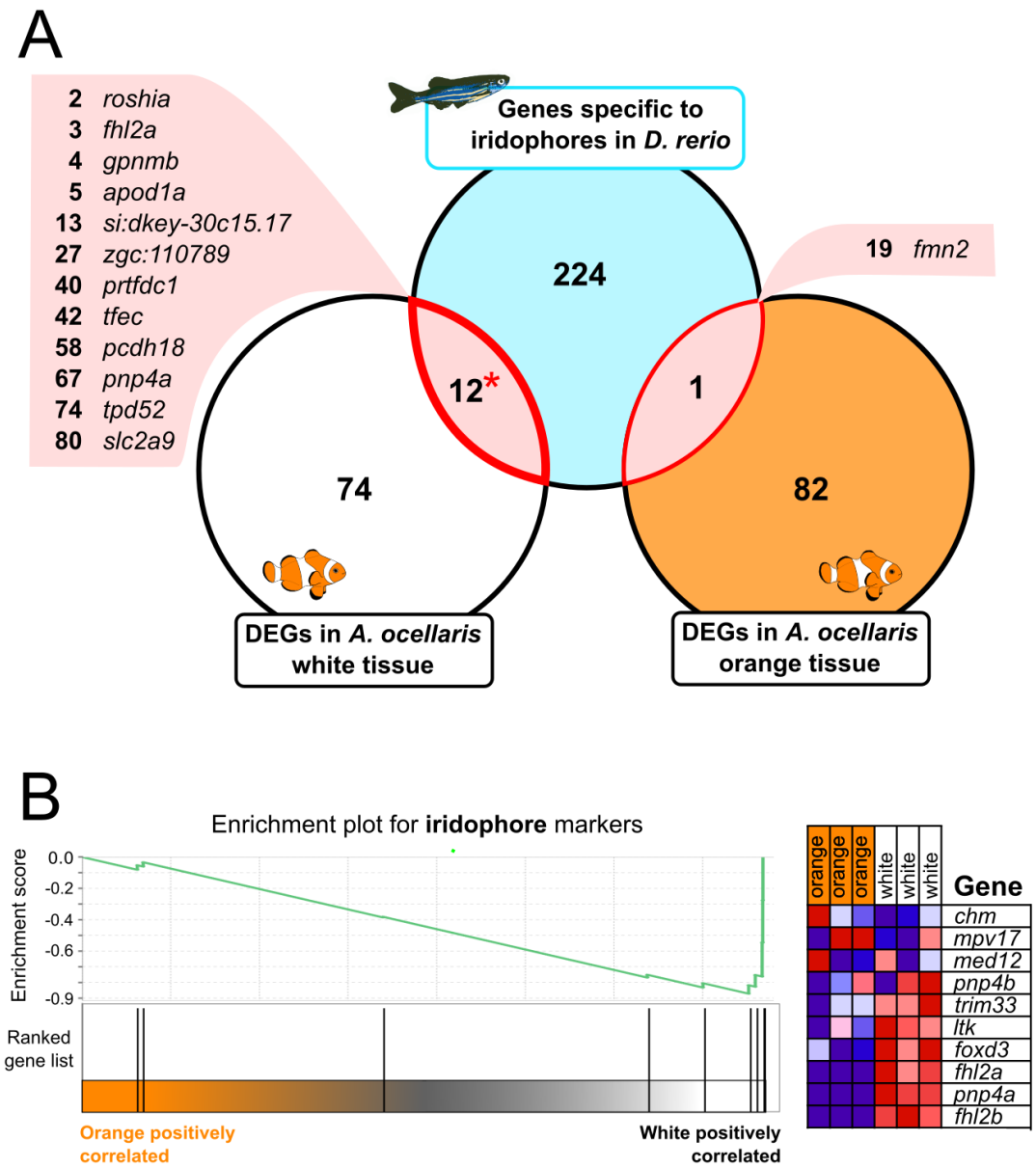


Figure 4: White band in clownfish is enriched in iridophore markers.
(A) 237 genes, markers of iridophores and recently characterized in zebrafish (Higdon et al 2013), were compared to the lists of 86 white and 83 orange DEGs obtained in this study. There were 12 genes in common between the iridophore gene list and the 86 clownfish white DEGs (intersection is significant, *p*-value < 1e-16, chi2 test), only 1 between the iridophore gene list and the orange DEGs (not significant) and no common gene between the white and orange overexpressed genes – by design of both lists. Numbers indicate the number of genes in each area. Gene symbols in common between datasets are indicated; numbers in bold indicate relative rank in the list of DEGs. For instance, *fhl2a* is the 3rd most significantly differentially expressed gene in orange tissue.
(B) GSEA results for literature-based iridophore markers. Horizontal axis of the enrichment plot represents the genes whose expression profiles are correlated to orange (left) or white (right) color. Colored squares indicate relatively high (red) or low (blue) expression in different samples. Enrichment score reflects the degree to which a gene set is overrepresented at the top or bottom of a ranked list of genes (http://software.broadinstitute.org/gsea/doc/GSEAUserGuideFrame.html). Most iridophore markers were enriched in white tissue but not in orange tissue.

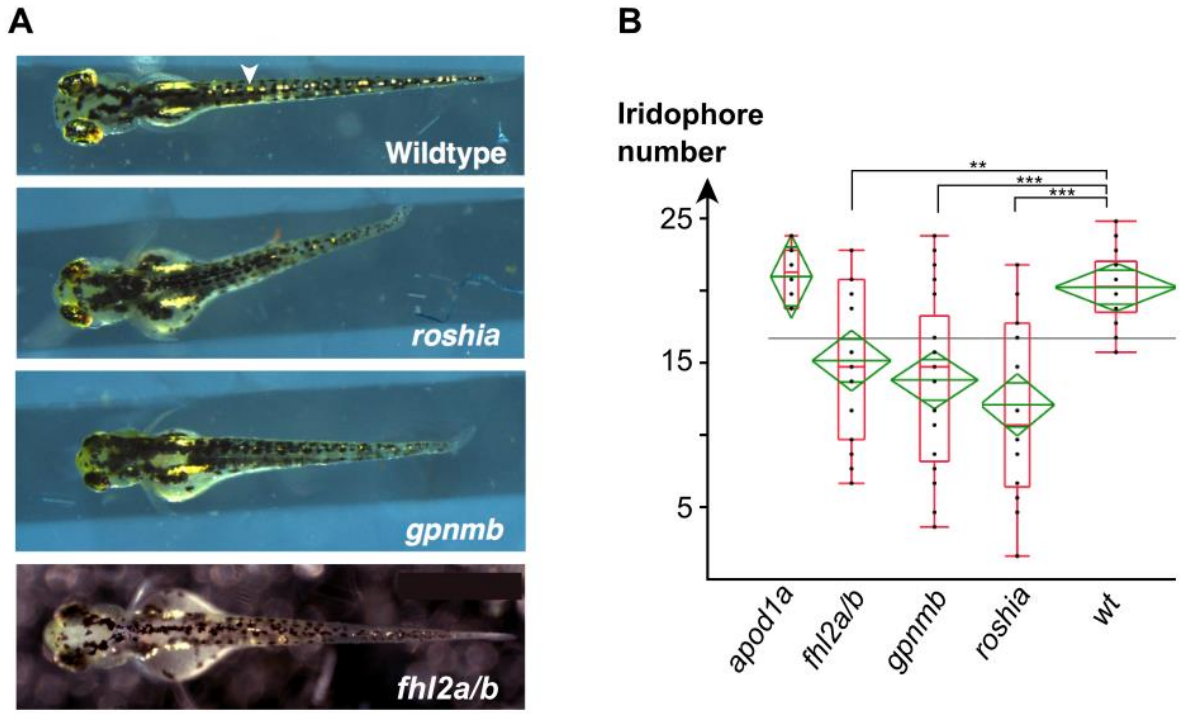


Figure 5: Reduced iridophore number in injected zebrafish larvae (4 dpf) with CRISPR/Cas9 targeting putative newly identified iridophore markers.

(A) Phenotypes of non-injected (wildtype) and injected individuals are shown. In wildtype individuals (top), numerous iridophores (located with the arrow head) were present along the body. In injected individuals, larval iridophores are less present.

(B) Quantification of dorsal iridophores in zebrafish 4 dpf larvae. Each dot corresponds to one individual. The whole distribution is represented. Stars indicate significant results after post-ANOVA Dunnett's test; **: $p < 0.01$; ***: $p < 0.001$. Pictures and experimental results: courtesy of Victor Lewis and David Parichy. Dpf: days post-fertilization.

Table 1A

Term	ID	Input number	Background number	Corrected P-Value
pigmentation	GO:0043473	6	67	0,0099
developmental pigmentation	GO:0048066	4	34	0,0428
response to abiotic stimulus	GO:0009628	7	158	0,0428

Table 1B

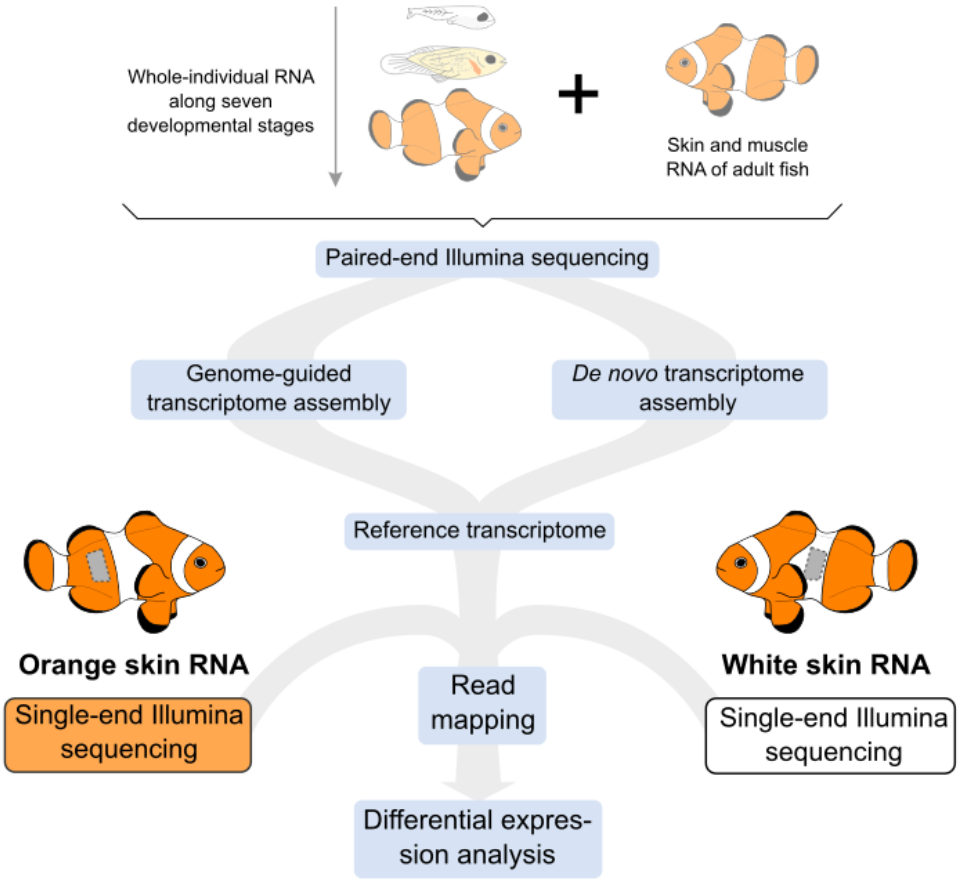
Term	ID	Input number	Background number	Corrected P-Value
developmental pigmentation	GO:0048066	4	34	0,0085
pigmentation	GO:0043473	4	67	0,0456
pigment cell differentiation	GO:0050931	3	27	0,0456

Table 1C

Term	ID	Input number	Background number	Corrected P-Value
response to abiotic stimulus	GO:0009628	5	158	0,0284
response to radiation	GO:0009314	4	88	0,0284
RNA polymerase II core promoter proximal region sequence-specific DNA binding	GO:0000978	3	33	0,0284
core promoter proximal region sequence-specific DNA binding	GO:0000987	3	35	0,0284
core promoter proximal region DNA binding	GO:0001159	3	35	0,0284
pigment granule	GO:0048770	2	9	0,0484
melanosome	GO:0042470	2	9	0,0484

Table 1: Gene Ontology enrichment terms on (A) all differentially expressed genes, (B) white skin DEGs and (C) orange skin DEGs.

SUPPLEMENTARY DATA



Supplementary Figure 1: RNA-Seq analysis pipeline used in this study. In order to get a representative reference transcriptome, RNA was first extracted from a collection of individuals at seven developmental stages as well as from muscle and skin adult tissues. Both genome-guided and *de novo* transcriptome assemblies were performed and combined. Subsequently, orange and white skin samples were collected and their RNA sequenced for differential expression analyses.

Supplementary Table 1 Next three pages

Gene list of differentially expressed genes ranked by adjusted p-value significance.

Blue (resp. orange) rows correspond to white (resp. orange) skin overexpressed genes. For top DEGs, phylogenetic analyses were performed to ascertain gene naming. Base mean corresponds to average expression over six samples in transcripts per million.

Gene ID	Gene Product	Base Mean	Log2 Fold Change	Adj. P-value	Manual curation of name using phylogeny
LOC103374546	perilipin-2-like	5046,9	5,3	1,5E-136	*plin1l*
LOC103370266	four and a half LIM domains protein 2-like	1024,7	-5,0	2,5E-127	*fhl2b*
LOC103358546	uncharacterized LOC103358546	1360,4	-4,8	2,6E-106	*roshia*
LOC103361331	tetratricopeptide repeat protein 39B-like	1028,2	4,3	5,1E-92	*ttc39b*
fhl2	four and a half LIM domains 2	1448,3	-4,1	3,1E-90	*fhl2a*
LOC103376104	retinol dehydrogenase 12-like	898,4	4,0	2,2E-80	*rdh12l-like*
LOC103364173	putative all-trans-retinol 13,14-reductase	782,9	4,1	4,8E-79	*retsat-like*
gpnmb	glycoprotein (transmembrane) nmb	992,2	-3,4	3,3E-55	*gpnmb*
LOC103353029	apolipoprotein D-like	362,9	-3,6	9,7E-54	*apod1a*
LOC103373034	galactose-specific lectin nattectin-like	218,3	3,4	1,3E-47	*clec19a-like*
mmel1	membrane metallo-endopeptidase-like 1	282,8	3,0	4,3E-38	*mmel1*
LOC103360024	D-beta-hydroxybutyrate dehydrogenase, mitochondrial-like	930,0	2,7	1,3E-34	*bdh1-like*
vim	vimentin	1465,6	2,3	1,6E-30	*vim*
bscl2	Berardinelli-Seip congenital lipodystrophy 2 (seipin)	582,5	2,5	3,3E-29	*bscl2b*
LOC103358107	adenosine deaminase CECR1-A-like	130,4	-2,6	2,2E-27	*ada2b = cecr1b*
LOC103368265	APC membrane recruitment protein 2-like	311,8	2,3	8,1E-25	
LOC103360395	patatin-like phospholipase domain-containing protein 2	589,6	2,4	2,3E-24	
LOC103361599	kinesin-1 heavy chain-like	205,2	2,3	1,1E-22	
LOC103368054	zona pellucida-like domain-containing protein 1	99,0	2,3	1,1E-21	
rasef	RAS and EF-hand domain containing	940,4	2,0	7,4E-21	
tmem159	transmembrane protein 159	795,5	2,0	1,3E-20	
LOC103372717	ALK tyrosine kinase receptor-like	795,5	-2,1	4,5E-19	*alk*
drd5	dopamine receptor D5	87,7	-2,2	8,8E-19	*drd5*
LOC103371017	gap junction alpha-5 protein-like	239,5	-2,0	5,2E-18	*gja5a = cx41.8*
prkar1b	protein kinase, cAMP-dependent, regulatory, type I, beta	507,9	1,8	3,7E-16	
LOC103374804	transient receptor potential cation channel subfamily M member 1-like	248,2	-2,0	5,1E-16	*trpm1a*
LOC103353689	transient receptor potential cation channel subfamily M member 1-like	973,2	-1,8	1,5E-15	*trpm1b*
LOC103354931	cAMP-regulated D2 protein-like	71,1	-1,9	9,5E-15	
LOC103369070	receptor-type tyrosine-protein phosphatase zeta-like	157,5	-1,9	9,5E-15	
LOC103375591	palmitoyltransferase ZDHHC2-like	230,2	1,9	3,0E-14	
LOC103356877	dehydrogenase/reductase SDR family member 12-like	103,1	1,9	7,3E-14	
LOC103356309	formin-2-like	241,1	1,7	7,1E-13	
mogat2	monoacylglycerol O-acyltransferase 2	68,9	1,8	2,3E-12	
LOC103355933	protein notum homolog	703,6	1,6	1,7E-11	
LOC103375397	ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase 1-like	130,2	1,7	3,1E-11	
LOC103367872	probable G-protein coupled receptor 61	162,8	-1,7	3,8E-11	
LOC103375795	ras-related protein Rab-27B-like	422,2	1,6	1,0E-10	*rab27bb*
LOC103360418	syntaxin-binding protein 5-like	123,4	-1,7	1,0E-10	
LOC103361807	sodium- and chloride-dependent betaine transporter-like	54,0	-1,6	1,9E-10	
plekhh1	pleckstrin homology domain containing, family H (with MyTH4 domain) member 1	457,0	-1,4	5,4E-10	
LOC103367178	dexamethasone-induced Ras-related protein 1-like	129,7	1,6	7,4E-10	
nkx6-2	NK6 homeobox 2	50,5	-1,5	2,5E-09	
proca1	protein interacting with cyclin A1	81,3	-1,5	3,7E-09	
LOC103354760	neuroblast differentiation-associated protein AHNAK-like	123,2	-1,5	5,0E-09	
pdzrn3	PDZ domain containing ring finger 3	594,6	-1,3	6,9E-09	
LOC103370010	creatine kinase B-type	358,6	-1,4	2,1E-08	
LOC103357637	mucin-19-like	53,6	1,4	2,3E-08	
LOC103361393	ensconsin-like	178,1	1,4	2,3E-08	
LOC103370819	forkhead box protein F2-like	69,6	-1,4	2,3E-08	
lpin2	lipin 2	1247,8	-1,2	2,5E-08	
tmem130	transmembrane protein 130	77,6	1,5	2,6E-08	
LOC103361237	protein tweety homolog 2-like	417,3	-1,4	3,3E-08	
LOC103366311	gap junction alpha-3 protein-like	44,8	1,4	3,6E-08	
LOC103366582	melanophilin-like	1340,5	1,2	4,5E-08	*mlpha*
akap12	A kinase (PRKA) anchor protein 12	1839,7	-1,3	1,1E-07	
LOC103365863	solute carrier family 23 member 1-like	50,7	-1,4	1,2E-07	
csrp2	cysteine and glycine-rich protein 2	296,9	1,3	1,4E-07	

arhgap20	Rho GTPase activating protein 20	404,2	-1,3	1,6E-07	
LOC103365076	paired box protein Pax-3-like	182,4	-1,4	1,9E-07	*pax3b*
LOC103367708	collagen alpha-1(XXVII) chain-like	40,4	1,3	3,3E-07	
LOC103374898	hibernation-associated plasma protein HP-20-like	72,9	1,4	3,6E-07	
apba2	amyloid beta (A4) precursor protein-binding, family A, member 2	48,4	1,3	7,4E-07	
LOC103369701	complement C1q-like protein 2	41,7	1,3	9,6E-07	
serping1	serpin peptidase inhibitor, clade G (C1 inhibitor), member 1	57,5	-1,3	1,1E-06	
LOC103361049	probable G-protein coupled receptor 21	41,5	-1,3	1,5E-06	
LOC103374947	actin-binding LIM protein 1-like, transcript variant X4	180,4	-1,3	1,8E-06	
LOC103359859	major facilitator superfamily domain-containing protein 8-like	54,1	-1,2	2,1E-06	
dner	delta/notch-like EGF repeat containing	50,1	-1,3	2,4E-06	
LOC103362443	synaptotagmin-like protein 1	2453,7	-1,1	2,4E-06	
ret	ret proto-oncogene	102,9	-1,3	7,8E-06	
alx4	ALX homeobox 4	348,6	-1,1	8,2E-06	
LOC103372544	immediate early response gene 2 protein-like	2456,0	1,0	1,2E-05	
LOC103367161	tubulin beta-3B chain-like	90,9	-1,2	1,3E-05	
LOC103359584	uncharacterized LOC103359584	65,0	1,2	1,4E-05	
LOC103369689	caprin-2-like	33,5	1,1	1,4E-05	
LOC103361106	seipin-like	924,0	1,0	1,5E-05	
LOC103361970	melanin-concentrating hormone receptor 2-like	39,6	1,2	1,6E-05	
slc24a4	solute carrier family 24, member 4	275,8	-1,2	1,9E-05	
prtfdc1	phosphoribosyl transferase domain containing 1	118,8	-1,2	1,9E-05	
LOC103374941	uncharacterized LOC103374941	146,3	1,2	2,2E-05	
npdc1	neural proliferation, differentiation and control, 1	78,0	1,2	2,2E-05	
rab38	RAB38, member RAS oncogene family	66,4	1,2	6,1E-05	
LOC103359796	androgen-dependent TFPI-regulating protein-like	33,8	1,1	6,3E-05	
LOC103362090	dihydropyrimidinase-related protein 2	607,1	-1,0	6,4E-05	
LOC103353204	retinaldehyde-binding protein 1-like	40,1	1,1	6,9E-05	
mgst2	microsomal glutathione S-transferase 2	32,8	1,0	7,9E-05	
LOC103356482	dixin-A-like	503,1	1,0	9,3E-05	
LOC103373000	cGMP-dependent protein kinase 1-like	118,9	1,1	1,2E-04	
rgr	retinal G protein coupled receptor	61,3	1,1	1,3E-04	
tfec	transcription factor EC	300,4	-1,1	1,4E-04	
nceh1	neutral cholesterol ester hydrolase 1	208,7	-1,1	1,6E-04	
LOC103356872	paired box protein Pax-3-like	98,7	-1,1	1,7E-04	*pax3a*
LOC103357625	fibronectin type III domain-containing protein 7-like	35,5	1,0	1,8E-04	
LOC103365588	voltage-dependent L-type calcium channel subunit alpha-1D-like	55,6	-1,0	2,0E-04	
LOC103372633	jun dimerization protein 2-like	573,8	1,0	2,0E-04	
zeb2	zinc finger E-box binding homeobox 2	914,5	-1,0	2,0E-04	
srp19	signal recognition particle 19kDa	28,6	1,0	2,1E-04	
LOC103354052	cysteine-rich protein 2-like	166,5	1,1	2,4E-04	
LOC103357875	transmembrane protein 254B-like	348,8	-1,0	2,7E-04	
mc5r	melanocortin 5 receptor	83,5	1,1	3,8E-04	
LOC103355756	serine/threonine-protein kinase MAK-like	27,0	-0,9	5,4E-04	
LOC103354955	tropomyosin alpha-1 chain-like	69,3	1,0	6,5E-04	
LOC103375584	phosphatidylinositol 4-phosphate 5-kinase type-1 gamma-like	82,9	-1,0	6,5E-04	
LOC103372631	proto-oncogene c-Fos-like	22880,7	0,8	7,7E-04	
bai3	brain-specific angiogenesis inhibitor 3	198,2	-1,0	7,7E-04	
LOC103362962	prickle-like protein 2	67,7	-1,0	8,1E-04	
LOC103366381	pinopsin-like	39,3	-1,0	8,4E-04	
pax7	paired box 7	122,9	-1,0	8,6E-04	*pax7a*
LOC103357149	alpha-2B adrenergic receptor-like	175,5	1,0	1,2E-03	
LOC103372478	VPS10 domain-containing receptor SorCS3-like	50,1	-1,0	1,2E-03	
LOC103356235	scavenger receptor cysteine-rich domain-containing group B protein-like	26,7	0,9	1,2E-03	
LOC103356924	integrin beta-1-like	91,6	1,0	1,3E-03	
LOC103366415	fasciculation and elongation protein zeta-2-like	66,3	-1,0	1,3E-03	
LOC103365931	phosphatase and actin regulator 3-like	111,3	-1,0	1,4E-03	
ncam2	neural cell adhesion molecule 2	343,3	1,0	1,4E-03	
slc7a4	solute carrier family 7, member 4	88,7	-1,0	1,4E-03	
mcf2l	MCF.2 cell line derived transforming sequence-like, transcript variant X4	336,6	-0,9	1,5E-03	

pcdh18	protocadherin 18	540,8	-1,0	1,5E-03	
LOC103361117	carbohydrate sulfotransferase 1-like	30,5	-0,9	1,5E-03	
LOC103374332	uncharacterized LOC103374332	201,1	-1,0	1,5E-03	
akap6	A kinase (PRKA) anchor protein 6	59,7	-1,0	2,9E-03	
LOC103369700	C1q-related factor-like	22,7	0,8	3,1E-03	
tsc22d1	TSC22 domain family, member 1	559,0	0,9	3,4E-03	
LOC103359170	caskin-1-like	248,9	-0,9	4,5E-03	
LOC103365919	pancreatic secretory granule membrane major glycoprotein GP2-like	21,9	0,8	4,5E-03	
LOC103359281	epithelial membrane protein 2-like	1007,0	0,9	4,6E-03	
LOC103361403	A-kinase anchor protein 12-like	922,7	-0,8	5,2E-03	
csf1r	colony stimulating factor 1 receptor	1431,0	-0,8	5,6E-03	*csf1ra*
ppfia2	protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 2, transcript variant X4	158,0	-0,9	7,8E-03	
tspan6	tetraspanin 6	108,1	0,9	9,5E-03	
LOC103356417	uncharacterized LOC103356417	92,9	0,9	9,7E-03	
LOC103362062	calpain clp-1-like	104,1	0,9	9,7E-03	
LOC103367714	purine nucleoside phosphorylase-like	20,4	-0,7	1,0E-02	*pnp4a*
LOC103374690	cholesterol 25-hydroxylase-like protein	849,9	0,7	1,1E-02	
LOC103376046	vitronectin-like	781,6	-0,7	1,1E-02	
npas3	neuronal PAS domain protein 3	362,9	0,8	1,5E-02	
LOC103366826	uncharacterized LOC103366826	54,5	0,8	1,5E-02	
LOC103371022	protein NDRG2	168,9	-0,9	1,5E-02	
LOC103370719	tyrosine-protein kinase receptor Tie-1-like	607,1	-0,7	1,5E-02	
kcnh7	potassium voltage-gated channel, subfamily H (eag-related), member 7	20,6	0,7	1,7E-02	
LOC103354333	vertebrate ancient opsin-like	19,6	-0,7	1,7E-02	
map1a	microtubule-associated protein 1A	154,1	-0,9	1,9E-02	
c1qtnf5	C1q and tumor necrosis factor related protein 5	735,2	0,7	2,2E-02	
gsto1	glutathione S-transferase omega 1	178,3	0,8	2,4E-02	
soat1	sterol O-acyltransferase 1	316,3	0,8	2,4E-02	
LOC103372150	secreted phosphoprotein 24-like	112,0	0,8	2,6E-02	
greb1	growth regulation by estrogen in breast cancer 1	35,0	-0,8	2,7E-02	
LOC103371747	tumor protein D52-like	338,7	-0,8	2,7E-02	*tpd52*
LOC103356903	rho guanine nucleotide exchange factor 25-like	228,7	-0,8	2,9E-02	
LOC103370645	ankyrin repeat domain-containing protein SOWAHC-like	289,0	-0,8	2,9E-02	
gpr149	G protein-coupled receptor 149	19,3	-0,7	2,9E-02	
LOC103360637	melanopsin-B-like	69,7	0,8	3,0E-02	
LOC103375404	protein FAM89B-like	152,0	0,8	3,0E-02	
LOC103355155	5'-3' exoribonuclease 1-like	97,9	-0,8	3,8E-02	
LOC103359885	cell death activator CIDE-3-like, transcript variant X4	112,4	-0,8	3,8E-02	
ankrd27	ankyrin repeat domain 27 (VPS9 domain)	403,3	-0,8	3,9E-02	
LOC103367959	solute carrier family 2, facilitated glucose transporter member 9-like	144,9	-0,8	3,9E-02	
LOC103359930	chromosome unknown open reading frame, human C10orf90	314,0	0,8	3,9E-02	
dpysl3	dihydropyrimidinase-like 3	1118,7	-0,7	3,9E-02	
LOC103355141	acyl-CoA desaturase-like	661,7	-0,8	3,9E-02	
hoxc11	homeobox C11	232,7	0,7	4,0E-02	
pnrc2	proline-rich nuclear receptor coactivator 2	90,7	0,8	4,0E-02	
LOC103367496	FCH domain only protein 1-like	970,2	-0,7	4,2E-02	
LOC103366495	melanoregulin-like	150,8	0,8	4,3E-02	
cebpb	CCAAT/enhancer binding protein (C/EBP), beta	4122,3	0,6	4,7E-02	
LOC103354647	early growth response protein 1-like	4040,3	0,6	4,7E-02	
wfikkn2	WAP, follistatin/kazal, immunoglobulin, kunitz and netrin domain containing 2	16,9	-0,6	4,7E-02	
fgl1	fibrinogen-like 1	1494,1	-0,7	5,0E-02	
junb	jun B proto-oncogene	8648,4	0,6	5,0E-02	

RÉSULTATS

CHAPITRE 3

Roshia, un nouveau marqueur d'iridophore acquis par transfert horizontal chez les poissons actinoptérygiens

Horizontal transfer of an iridophore gene into the ray-finned fish lineage

Thibault Lorin*[1], Frédéric Brunet[1], Vincent Laudet[2], Jean-Nicolas Volff[1]

1: Institut de Génomique Fonctionnelle de Lyon, École Normale Supérieure de Lyon, UMR 5242 CNRS, Université Claude Bernard Lyon I, Université de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France

2: Observatoire Océanologique de Banyuls-sur-Mer, UMR CNRS 7232 BIOM ; Sorbonne Université ; 1, Avenue Pierre Fabre, 66650 Banyuls-sur-Mer, France

ABSTRACT

Vertebrate coloration is a highly diverse trait determined by the interplay between physical and chemical colors and mainly based on pigment cells. Among vertebrates, teleost fish constitute the most diverse group of vertebrates in terms of coloration: they harbor very diverse pigment patterns and more than eight pigment cell types have been described in this lineage, including xanthophores, melanocytes and iridophores. In contrast, mammals and birds only have melanocytes, while other vertebrate lineages such as amphibians and lepidosaurians also have xanthophores and iridophores. In a recent study, a novel gene involved in iridophore formation in fish was identified, coined *roshi-a*. In the present work, we show that Roshi-a is a complex multidomain enzymatic protein likely to be localized in cytosol. Interestingly, within vertebrates, *roshi-a* is restricted to ray-finned fish (Actinopterygii) and absent from any other vertebrate lineage. In addition, *roshi* is sporadically found in a very restricted number of divergent non-vertebrate animal species, suggesting acquisition by horizontal transfer, and has been transferred from fish to viruses. The ancestral *roshi* gene might have first appeared in the fungal lineage before subsequent transmission to animals. The occurrence of a fish-specific gene necessary for the formation of iridophores might challenge our current understanding of pigment cell evolution in vertebrates. In particular, this finding questions the monophyly of iridophores, which could instead have originated independently in different vertebrate lineages.

INTRODUCTION

Color is a very diverse trait in vertebrates and is mainly due to the interplay between physical (structural) and chemical colors (pigments). Among vertebrates, teleost fish constitute the most diverse group in terms of coloration since, in addition to harboring both chemical and structural colors, these fish also have the highest diversity in terms of pigment cells within vertebrates (Parichy, 2006; Braasch *et al.*, 2008). Indeed, at least eight pigment cell types have been documented so far in teleosts (Schartl *et al.*, 2016). Some of them, such as blue cyanophores, appear to be specific to this group of animals and seem to be an evolutionary innovation of the teleost lineage (Goda and Fujii, 1995). However, some chromatophores are shared with other vertebrate lineages: black melanocytes (often coined melanophores, but see (Schartl *et al.*, 2016), yellow-orange xanthophores, red erythrophores and the silvery blue-white iridophores (Mellgren and Johnson, 2002). Present mammals and birds only harbor melanocytes in their integument (Braasch *et al.*, 2008), suggesting that secondary losses occurred in these lineages (Mellgren and Johnson, 2002).

Among all pigment cells, iridophores are of particular interest. These cells are called "pigment cells" and are indeed often categorized within pigment-containing cells like melanin-containing melanocytes. However, they actually contain specialized organelles called iridosomes that harbor light-reflecting purine platelets that, strictly speaking, are not pigments, *i. e.* light-absorbing molecules (Schartl *et al.*, 2016). Depending on physiological signals, the relative position and orientation of the platelets in the organelles can rapidly change; as a consequence, the fish color also changes (Nilsson Sköld *et al.*, 2013). Iridophores can thus be either white or silvery blue. In fish, the white color can also be due to leucophores and the determination of the cellular nature of a white-colored skin (iridophore-based vs. leucophore-based) can sometimes require the combination of both cellular and genetic approaches (Salis, Lorin et al., in prep.).

In a recent study, it has been shown that the color of the white band in the clownfish *Amphiprion ocellaris* is mainly due to dermal iridophores, not leucophores. We additionally identified a previously uncharacterized fish-specific iridophore marker, coined *roshi-a* (the "a" copy of *roshi*), which was shown through a CRISPR/Cas9 analysis to be required for iridophore differentiation (Salis, Lorin et al., in prep.). In this work, we aimed at better describing this gene. We first scrutinized its structure and the putative functions of the Roshi-a protein and showed that this gene harbors four main domains and that the protein might be of cytosolic localization with multiple enzymatic activities. We also conducted an analysis of the distribution of this gene within animals and identified that *roshi* is present in all ray-finned fish, but absent from other vertebrates. Using genomic and transcriptomic data, we identified partial or full *roshi* sequences in other animal species, including scorpions (Arthropoda), hydra (Cnidaria) and lingula (Brachiopoda). We were also able to retrieve *roshi* sequences in diverse fungal lineages. Last, we describe hereafter the transfer of one domain of *roshi* from a fish host to a virus of the Iridoviridae family that occurred less than 70 million years ago. Altogether, our results suggest that multiple events of horizontal gene transfer occurred throughout the evolution of *roshi* that could have subsequently been co-opted in the iridophore differentiation pathway. Iridophores having been described in other vertebrate taxa that do not share this crucial gene for iridophore differentiation, our results challenge the current hypothesis of the monophyly of this cell type (Mellgren and Johnson, 2002) and, together with other lines of evidence, could suggest that iridophores differ between multiple vertebrate lineages.

MATERIAL AND METHODS

Roshi-a structure

Structure prediction of the different domains of Roshi-a was performed using prediction tools based on Hidden Markov Models (HMMs). We first used Pfam sequence search tool (v31.0), based on a database of manually curated proteins and available at https://pfam.xfam.org/ (Finn *et al.*, 2016). The InterProScan tool was also used to predict domains that were not predicted by Pfam due to low e-value score (Jones *et al.*, 2014). Predictions of 3D structures for each protein domain of Roshi-a were obtained using the Phyre2 protein fold recognition web server available at http://www.sbg.bio.ic.ac.uk/phyre2/ (Kelley *et al.*, 2015).

Multiple tools were used for *in silico* prediction of cellular localization. We first tested whether Roshi-a would be a membrane-localized protein. Predictions of signal peptide sequences were done using Signal IP (v4.1) (Petersen *et al.*, 2011) and Phobius (Kall *et al.*, 2007). Transmembrane protein domains were predicted using the TMHMM Server 2.0 available at http://www.cbs.dtu.dk/services/TMHMM/ (Krogh *et al.*, 2001). An important driver of protein membrane

localization is lipidation (Casey, 1995). Diverse putative protein lipidation sites were tested using CSS Palm (v4.0) for palmitoylation (http://www.csspalm.biocuckoo.org/ ; Zhou *et al.*, 2006); GPS Lipid (v1.0) for other lipidations (http://lipid.biocuckoo.org/index.php ; Xie *et al.*, 2016) and GPI-SOM for GPI anchor prediction (http://gpi.unibe.ch/ ; Fankhauser and Maser, 2005).

In addition, we evaluated whether Roshi-a would be a nuclear protein using prediction tools for nuclear localization signals (NLSs). We used cNLS Mapper (http://nls-mapper.iab.keio.ac.jp/; v2012/11/7), SeqNLS (http://mleg.cse.sc.edu/seqNLS/), NLStradamus (http://www.moseslab.csb.utoronto.ca/NLStradamus/ ; v2015/02/24) and NucPred (https://nucpred.bioinfo.se/nucpred/) (Brameier *et al.*, 2007; Nguyen Ba *et al.*, 2009; Kosugi *et al.*, 2009; Lin *et al.*, 2012). We also tested if Roshi-a would harbor nuclear export signals using NetNES (http://www.cbs.dtu.dk/services/NetNES/; la Cour *et al.*, 2004). Putative DNA-binding domains were predicted using DNABIND (http://dnabind.szialab.org/ ; Szilágyi and Skolnick, 2006).

Roshi distribution in diverse lineages

We screened for *roshi* sequences in a wide diversity of lineages encompassing the diversity of Opisthokonta (ray-finned fish, lobe-finned fish, cartilaginous fish, cyclostomes, urochordates, cephalochordates, ambulacrarians, protostomians, cnidarians, sponges, ctenophores, choanoflagellates and eumycetes) as well as bacteria and seed plants. For each of these lineages, we retrieved, when available, genomic or transcriptomic data, and when different, the longest isoform sequence. We performed a combination of BLASTP and TBLASTN analyses on diverse genomic and transcriptomic databases (last accessed April 21st 2018) using *D. rerio* Roshi-a protein sequence as a query (Altschul *et al.*, 1990). We used general databases such as RefSeq (v87), Transcriptome Shotgun Assembly (TSA), Whole-Genome Shotgun contigs (WGS), Expressed Sequence Tags (ESTs), Ensembl (v91), Genbank and DNA Databank of Japan (Benson *et al.*, 2013; O'Leary *et al.*, 2016; Kodama *et al.*, 2018). We also screened for *roshi* in lineage-specific databases such as teleost-specific FISHT1K and fungi-specific MycoCosm (Grigoriev *et al.*, 2014; Sun *et al.*, 2016). We finally used specific genome databases such as the *Limulus polyphemus* genome portal (Battelle *et al.*, 2016), *Amphiprion frenatus* genome (Marcionetti *et al.*, 2018), sturgeon transcriptome database Anacariibase (Vidotto *et al.*, 2013) and available cnidarian transcriptome assemblies (Zapata *et al.*, 2015).

When a *roshi* sequence was identified in the genome but had partial or no corresponding annotation, we extracted the contig/scaffold sequence and used the FGENESH+ gene finding tool to retrieve the coding sequence (http://www.softberry.com/) (Solovyev, 2008).

Phylogenetic and synteny analyses

Sequence alignments were performed using Clustal Omega (v1.2.1; Sievers & Higgins 2014) with default parameters and were manually curated with the Seaview v4.6.5 software (Gouy *et al.*, 2010). Each alignment was then analyzed using either jModelTest or prottest3 for nucleotid and amino acid sequences, respectively (Posada, 2008; Darriba *et al.*, 2011). The evolutionary model for tree-building was selected based on the Bayesian information criterion (Luo *et al.*, 2010). Tree building was conducted with the best fitting model using the Maximum Likelihood method implemented in PhyML v3.1 (Guindon *et al.*, 2010) with SH-aLRT support (Le and Gascuel, 2008; Anisimova *et al.*, 2011).

Synteny analyses were conducted using the Genomicus web browser (v91; http://genomicus.biologie.ens.fr/genomicus-86.01/cgi-bin/search.pl) (Louis *et al.*, 2015). We also used synteny information available in the literature for the spotted gar and salmonids (Berthelot *et al.*, 2014; Braasch *et al.*, 2016; Lien *et al.*, 2016).

RESULTS

***roshi-a* encodes a protein with four conserved domains**

In the zebrafish *Danio rerio, roshi-a* corresponds to locus si:ch211-256m1:8 located on chromosome 3 (Ensembl identifier: ENSDARG00000055172 and NCBI gene ID: 795494). The predicted protein encoded by this gene is 1575 amino-acids long and the coding sequence is constituted of three different exons (**Fig. 1A and 1B**). Four functional domains were predicted in the Roshi-a protein based on multiple sequence alignments and domain prediction by hidden Markov models (HMM) (**Fig. 1C**).

The first predicted N-terminal domain (D1) corresponded to a Choline/Carnitine o-acyltransferase (Pfam ID: CL0149; e-value = 3.3e-58). Proteins harboring this domain can transfer acyl groups either to choline (forming the active neurotransmitter acetylcholine) or carnitine (necessary for the transport of fatty acids to the

mitochondria matrix). The typical corresponding genes are *chat* (choline acetyltransferase) and *crat* (carnitine O-acetyltransferase) that encode structurally similar but different enzymes (Cai *et al.*, 2004). We compared the protein sequence of domain 1 with those of human proteins ChAT and CRAT, for which a crystal structure and a detailed description of the function of critical amino-acids is available. The active site of both these enzymes is a histidine residue (His324 in ChAT and His343 in CRAT) that is conserved in fish Roshi-a (**Fig. S1A**). However, the binding of choline by ChAT requires an extra-tyrosine residue (Tyr552) that is not conserved in fish Roshi-a, suggesting that Roshi-a cannot bind choline. Domain 1 of Roshia could then be a putative functional carnitine o-acyltransferase domain.

The second domain (D2) corresponded to a predicted "asparaginase" domain (Pfam ID: CL0052; e-value = 9.2e-91). Asparaginase is an enzyme that breaks down the amino-acid asparagine and has a conserved distribution in animals, fungi, plants and bacteria (Borek *et al.*, 2014). This enzyme belongs more generally to the class of threonine peptidases, which are enzymes characterized by a nucleophile threonine at the N-terminus of the mature enzyme. Contrarily to Pfam analysis, the HMM analysis of *D. rerio* Roshi-a protein using InterProScan did not strictly predict any asparaginase (ID: IPR006034) but rather, less specifically, a threonine peptidase (ID: IPR006034). To further unravel the putative function of this second domain, we performed a BLASTP analysis, using domain 2 of zebrafish Roshi-a as query, on *Escherichia coli* protein sequences, since crystal structures of asparaginase and other threonine peptidases are available for this bacterial species. The best hit obtained corresponded to a predicted "beta-aspartyl-peptidase" (*iaaA* gene product), which aligned better with fish Roshi-a domain 2 than the *E. coli* asparaginase protein (not shown). This enzyme catalyzes the cleavage of an aspartate residue at the N-terminal part of a polypeptide. Most predicted functional amino-acid residues, including G10, N67, T179 (the residue with the most important catalytic activity), V180, T197, G199, R207, G209, D210, G231, G233, were conserved between *E. coli* beta-aspartyl-peptidase (Michalska *et al.*, 2008) and Roshi-a **(Fig. S1B**). Altogether, this suggests that domain 2 has a functional beta-aspartyl-peptidase enzymatic activity.

Roshi-a third domain (D3) is divided into two-subdomains (**Fig. 1C**). The first one has no predicted function by HMM prediction tools but is well conserved in fish and other organisms (see below). The second one is predicted to be a "HIT (Histidine Triad Motif) domain" (Pfam ID: CL0265; e-value = 4.5e-05). The HIT domain is usually shared by proteins with a nucleotide hydrolase and/or transferase (on the alpha-phosphate of ribonucleotides) activity (Brenner, 2002). Proteins in this family show the motif H-ϕ-H-ϕ-H-ϕ (where H is histidine and ϕ a hydrophobic amino-acid, hence the "histidine-triad" name) and the Hit domain is widely distributed in prokaryotes and eukaryotes (Séraphin, 1992). However, in Roshi-a, this catalytic triad is not conserved since fish Roshi sequences have only one (most often) or two (rarely) histidine residues, together with a lysine residue **(Fig. S1C**). The motif in Roshi-a is Y/H-ϕ-H-ϕ-K-ϕ. Thus, the catalytic domain may not be functional and we renamed this domain "HIT-like" in Roshi-a. This could also explain the relatively low e-value of this domain in BLAST sequence analyses compared to D1 and D2.

Finally, Roshi-a fourth and last C-terminal domain (D4) corresponds to a "tRNA synthetase class II (D, K and N)" domain (Pfam ID: CL0040; e-value = 5.5e-31). In general, amino-acyl tRNA synthetases of class II (aaRS II) are enzymes constituted of three different domains: the N-terminal domain recognizes the anticodon loop of tRNA; the middle domain binds the tRNA; the last domain, or 'core enzyme', is the C-terminal domain and binds the acceptor stem of the tRNA to perform the catalysis (Eriani *et al.*, 1995). These enzymes usually dimerize before catalyzing the reaction. However, Roshi-a only harbors the 'core enzyme' domain of class II aaRS and does not have the two N-terminal domains of classical aaRS II (**Fig. S1D**). The structure of the D4 domain corresponded to a seven-stranded anti-parallel beta-sheet, as observed for the yeast 'core-enzyme' of aaRS II (pdb ID: 1ASY) (**Fig. 1D**). In addition, most active residues are conserved compared to the yeast sequence, which has been functionally characterized. In particular, Pro273 and Phe324, important for both catalytic activity and dimerization, are well conserved in fish (these positions correspond to yeast aspartate aaRS – see (Eriani *et al.*, 1995)). Residues Glu327, Arg325, Phe338, Asp471, Glu478, important for binding the ATP molecule necessary to the catalysis, and residue His334, that binds the tRNA, are all well conserved in fish Roshi-a compared to yeast (not shown). Finally, the amino-acid binding site (position D342 in yeast aspartate aaRS) is conserved between asparagine aaRS, lysine aaRS and fish Roshi_a, but not with yeast aspartate aaRS (**Fig. S1D**). Altogether, this could suggest that Roshi-a has an aaRS activity II (probably lysine or asparagine) with putative dimerization properties.

Roshi-a is likely to be a cytosolic enzyme

We next investigated *in silico* the cellular localization of Roshi-a. We did not find any evidence of the presence of a N-terminus signal peptide, which would suggest targeting to either the secretory pathway or an intra-organelle localization (von Heijne, 1990) (**Fig. S2**). In addition, no transmembrane domain was predicted. Finally, we did not find any evidence for lipid covalent modifications that would enable membrane localization: there were no predicted sites for palmitoylation, myristoylation, farnesylation and geranylgeranylation, and no detected glycosylphosphatidylinositol (GPI) anchor signal sequence.

One of the programs we used for the detection of a nuclear localization signal (cNLS-Mapper) found a bipartite NLS with a score of 7.9 (**FIG. S3**), indicating that the protein could be partially located to the nucleus. However, none of the three other programs used detected any evidence for an NLS. Finally, there was also no nuclear export signal, and no DNA-binding domain was predicted in the protein (the probability of DNA-binding predicted by DNABIND was 0.196). Altogether, these results indicate that the Roshi-a protein is likely to be a free cytosolic protein enzyme.

***Roshi* is widely distributed in ray-finned fish and retained as partial duplicate after the teleost-specific whole genome duplication**

To better understand the evolution of *roshi* and its taxonomic distribution, we searched for homologous sequences in other organisms, including ray-finned fish, metazoans, opisthokonts, plants and bacteria (by increasing taxonomic order from a teleost fish perspective). No homologous sequence could be retrieved in both plants and bacteria.

Roshi-a is a gene present in teleost fish as well as in the spotted gar *L. oculatus* (non-teleost). This gene was detected in all teleost species with sequenced genomes that we tested. Compared to non-teleost ray-finned fish, a supplementary round of whole-genome duplication (WGD) has occurred at the basis of teleost fish: the teleost-specific genome duplication, or Ts3R ((Meyer and Schartl, 1999; Jaillon *et al.*, 2004). After this event, 10-20% of pre-Ts3R genes have been kept as duplicates – coined 'ohnologs' in reference to S. Ohno seminal work on WGDs; (Inoue *et al.*, 2015; Pasquier *et al.*, 2017). In all teleost species, a second *roshi* gene was found, called *roshi-b*. This gene is truncated, sharing with *roshi-a* only the sequence corresponding to both sub-domains of domain 3 including the HIT-like domain (**Fig. 1C**). *Roshi-b* is conserved in all teleost fish genomes, with a phylogeny compatible with a duplication through Ts3R (**Fig. 2A**). In addition, *roshi-a* and *roshi-b* were in conserved duplicated syntenic regions that correspond to paralogons from the Ts3R event (**Fig. 2B**). Altogether, these results show that *roshi-b* corresponds to a Ts3R-ohnolog of *roshi-a* that has lost domains 1, 2 and 4 after the divergence of the teleost species studied here (**Fig. 2**). Both *roshi-a* and *roshi-b* were also duplicated again by the salmonid-specific WGD called Ss4R (Berthelot *et al.*, 2014; Lien *et al.*, 2016) **(Fig. 2A**). Of note, one copy of *roshi-b* was apparently secondarily lost in the trout lineage.

Interestingly, two copies of the pre-Ts3R *roshi* gene were found in the gar genome: a gene containing all four domains of *roshi-a* and likely to represent the structure of the ancestral *roshi* gene (Ensembl ID: ENSLOCG00000011127), and a "*roshi*-like" (*roshi-l*) gene with only domain 3 (ENSLOCG00000017844) (**Fig. 2**). Both paralogs are on the same chromosome, chromosome 12, and are separated by ca. 80 kb and three annotated genes. This suggests that a small-scale duplication in the spotted gar genome conducted to a convergent loss of domains 1, 2 and 4 and gar *roshi-l*, as observed in the teleost *roshi-b* ohnolog after the Ts3R WGD.

Both spotted gar and teleost fish belong to Neopterygii, a sub-lineage of ray-finned fish (**Fig. 2A**). Using transcriptome data for more divergent lineages within ray-finned fish (there is no genome sequence available for species in this lineage), we retrieved in three sturgeon species including *Acipenser oxyrinchus* partial *roshi* sequences corresponding to D3 (**Fig. 2A**).A continuous sequence that would contain the other domains in addition to D3 could not be identified in any sturgeon species. Overall, these results indicated that *roshi* has been conserved in ray-finned fish for at least 350 million years at least under its truncated form.

Horizontal gene transfer of *roshi-b* to Megalocytiviruses (Iridoviridae) after 70 Ma

We have identified in public databases a viral sequence with very significant homologies to *roshi*. This sequence corresponded to ORF515L of the Red Seabream Iridovirus (RSIV), a double-stranded DNA virus from the *Megalocytivirus* genus within the Iridoviridae family (Kurita and Nakajima, 2012). The viral protein encoded by this gene has been qualified as "HIT-like" (Kurita and Nakajima, 2012). Since many viruses within Megalocytiviruses are fish-specific (Inouye *et al.*, 1992; Nakajima and Maeno, 1998), we hypothesized that this

gene, conserved in Megalocytiviruses, would have been acquired from a fish host through gene transfer. As initial phylogenetic reconstruction with both *roshi-a* and *roshi-b* suggested that the viral genes was closer to *roshi-b*, we retrieved sequences from multiple teleost genomes and transcriptomes to identify the teleost sequence that would the most closely related to the viral sequence (**Fig. 3**).

Among all fish sequences, HIT-like sequences of Red Seabream and Rock Bream Iridoviruses branched within Carangaria, a fish group including jacks. As this lineage is estimated to be 70 Ma old (Betancur-R *et al.*, 2017), this suggests a likely horizontal gene transfer from a carangarian host to an ancestral Megalocytivirus that occurred more recently (**Fig. 3**).

Roshi is a gene with a very patchy distribution in other animal lineages

In order to further analyze the distribution of *roshi,* we screened all available vertebrate genomes outside of ray-finned fish and could not identify any homologous sequence(**Fig. 4**). This was true for the whole *roshi* protein as well as for each of the four domains. In addition, *roshi* was absent from all deuterostomian lineages but present in cephalochordates. Indeed, in *Branchiostoma floridae* and *B. belcheri* genomes and transcriptomes, as well as in the *Asymmetron lucayanum* transcriptome (no genome is available for this species), genes homologous to *roshi* were identified – corresponding to domain 3 only (**Fig. 4).** No sequence was found in the *B. lanceolatum* transcriptome but this might be due to an absence of expression (Oulion *et al.*, 2012). This was not due to contamination since neighboring genes were conserved in cephalochordates.

Within protostomians, most other lineages do not harbor *roshi* sequences (nor any of its domains) (**Fig. 4**). However, we were able to identify sequences homologous to *roshi* in three different lineages. Using genomic data, we identified a *roshi* complete ortholog in the bark scorpion *Centruroides sculpturatus* (Buthidae, Chelicerata, Arthropoda) (NCBI ID: LOC111640366). On the same scaffold, another *roshi* ortholog was found (LOC111640371) with only domain 3 of *roshi*. Using publicly available transcriptomic data, we could identify that *roshi* was also present in another Buthidae species, *Tityus serrulatus*, whose lineage diverged from *C. sculpturatus* around 25-40 Ma (Ojanguren-Affilastro *et al.*, 2017). Within Arthropoda, we additionally identified a *roshi* D3 sequence in the oribatid mite *Nothrus* palustris, that diverged from scorpions ca. 450 Ma (Edgecombe and Legg, 2014). In particular, we could not find any sequence corresponding to any domain of *roshi* in all other tested arthropod genomes, including the house spider genome (Tetrapulmonata), the horseshoe crab genome (Xiphosura) and all Mandibulata genomes available to date, including hexapod genomes.

Finally, we identified *roshi* orthologs in another protostomian species, the brachiopod *Lingula anatina*. This species harbors two full length copies of *roshi* (LOC106171029 and LOC106154348, on two different scaffolds), two tandem duplicates of the sole D3 (LOC106175989 and LOC106175973), and one paralog of D4 (LOC106155739) (**Fig. 4**). Of note, we could not find any evidence for *roshi* in other brachiopod species, due to absence of both genomic and transcriptomic data.

Finally, we screened the genomes and transcriptomes of other taxa in early-diverging Metazoa (compared to Bilateria). *Roshi* was absent from most of these lineages but was present in *Hydra vulgaris* (Cnidaria, Hydrozoa) (**Fig. 4**). The *H. vulgaris* genome contained two orthologs of *roshi*: one copy contained all domains but D3 (gene ID: 100201936) while the other copy contained domain 1 and a fragment of domain 2 (gene ID: 100198269). We were not able to identify *roshi* sequences in other cnidarian lineages despite abundant genomic and transcriptomic data for more than 40 species (see for instance (Zapata *et al.*, 2015)).

Hence, *roshi* has a patchy distribution within metazoans. While *roshi* is found in all ray-finned fish lineages, it is not present in the vast majority of animal lineages despite extensive screening of public genomic and transcriptomic data. *Roshi* sequences could only be detected in very specific lineages.

Roshi gene tree reveals discrepancy with species tree in animals

We further built a phylogenetic tree for *roshi* protein sequences using the longest possible protein domain that was common diverse animal lineages (*i. e.* D1, D2 and D4). Interestingly, *roshi* protein tree revealed major discrepancies with the species tree (**Fig. 5**). First, whereas both scorpions and *Lingula* are within Protostomata and thus evolutionarily equally distant from fish, the scorpion sequences were closer to fish homologs than the *L. anatina* homologs (well supported by node support values). In addition, although *Hydra* is a cnidarian (hence, not a bilaterian, contrarily to scorpions and brachiopods), *H. vulgaris* homolog was more closely related to scorpion than the *L. anatina* sequences.

Altogether, patchy distribution of *roshi* in animals (**Fig. 4**) as well as tree topology (**Fig. 5**) strongly suggest that *roshi* has been transferred between multiple animal species through horizontal gene transfer.

Roshi has a patchy distribution in Eumycota and most species harbor partial _roshi_ sequences

As we identified a *roshi* ortholog in *Cenococcum geophilum*, an ectomycorrhizal Ascomycota (**Fig. 5**), we subsequently screened for *roshi* sequences in available fungal genomes and transcriptomes. We identified such sequences in diverse fungal lineages (**Fig. 6**). *Roshi* was not found in any of microsporidians, chytrids, class II zygomycetes, mucoromycetes nor glomeromycetes. Among basidiomycetes, only one species, *Hypsizygus marmoreus*, harbored a *roshi* sequence related to D1 (**Fig. 5**). The vast majority of lineages harboring a *roshi* sequences were ascomycetes. This lineage is characterized by a major polytomy with ambiguous nodes ca. 300-350 Ma due to conflicting phylogenies available in the literature (Schoch, Sung, *et al.*, 2009; Prieto and Wedin, 2013; Coelho *et al.*, 2013; Beimforde *et al.*, 2014). Within Ascomycota, *C. geophilum* (Dothideomycetes) was the only species with a *roshi* sequence with all four domains. Two other lineages within Dothideomycetes harbored three contiguous domains of *roshi* (D2, D3 and D4), as well as species in Eurotiomycetes (*Talaromyces* and *Penicillium* genera) and Lecanoromycetes (*Lobaria pulmonaria*). All other species in which a *roshi* sequence could be identified only harbored one or two complete or partial domains.

Hence, within each Ascomycota class, there was no observed clear phylogenetic evidence for conservation or loss of any of the *roshi* domains. Indeed, in many cases, species with *roshi* sequences were sister group to species with no *roshi* sequence. For instance, in Chaetotyriales, within the Eurotiomycetes class, only *Exophiala aquamarina* has retained an ortholog of *roshi* D1 while sister species have not.

Overall, these results show that (i) *roshi* is not restricted to animals; (ii) many fungi harbor different domains of *roshi*; (iii) only one fungal species seems to harbor all four domains; (iv) *roshi* taxonomic distribution is much patchy within Eumycota.

DISCUSSION

In this work we have characterized the structure and putative properties of *roshi-a* (si:ch211-256m1:8), a newly characterized gene in fish involved in iridophore formation that encodes a four-domains protein. We showed that this gene is conserved in ray-finned fish but absent from other vertebrates, with a very patchy distribution in other animals. *Roshi* was transferred from fish to iridoviruses. We finally detected *roshi* partial or complete sequences in diverse fungal species, indicating its modular evolution and suggesting a possible evolutionary origin.

Roshi-a is a long enzyme probably localized in cytosol

The predicted coding sequence of *roshi-a* encodes a protein composed of 1,575 amino acids (**Fig. 1**). This makes Roshi-a a protein with complex 3D folding, and one of the top 1% longest proteins in a eukaryotic rganism (Zhang, 2000). The different Roshi-a domains harbor different putative functions. They respectively encode a putative carnitine-acyl transferase domain, an aspartyl-peptidase domain, a HIT-like domain (within Roshi-a third domain) and a C-terminal aminoacyl tRNA synthetase (class II) domain. Apart from the HIT-like domain, all other domains share high similarity of active sites with crystallized, well-studied proteins (**Fig. S1**). They are thus possibly functional. However, it is difficult to bring a fully satisfactory function out for the whole protein based on these sole predictions. In addition, to our knowledge, there is not other eukaryotic protein that harbors any of these domains in association to any of the others. This makes *roshi-a* a unique protein in this aspect, with no possible direct comparison to other known proteins for an integrated functional prediction. However, we could not find clear relationship between the encoded functions and the biology of iridophores.

Roshi has been retained as complete or partial duplicates after fish-specific duplications

Whole-genome duplications (WGDs) can influence the evolution of a gene family (Ohno, 1970). After a WGD event, genes are either retained or lost. The retention pattern greatly varies with the function of the encoded protein, and genes that are retained in two copies often provide the raw material for the acquisition of new functions (Ohno, 1970; McGrath *et al.*, 2014). In this study, we showed that the ray-finned fish ancestral *roshi* gene, present in non-teleost fish, was retained after the teleost-specific Ts3R (**Fig. 2**). In addition, each of both Ts3R-ohnologs were retained after Ss4R – with, however, a possible loss of one *roshi-b* Ss4R-duplicate in the trout lineage. The retention of a gene over two successive whole-genome duplications in fish is not a frequent event. In a recent study on the evolution of the pigmentation gene repertoire, it has been shown that only

24.7% of the pre-Ts3R ancestral genes were likely to be retained over both successive fish-specific WGDs (Lorin *et al.*, 2018). Though retained after Ts3R, the retention pattern of *roshi* is notwithstanding peculiar. Indeed, *roshi-b* is shorter than the putative *roshi* pre-Ts3R ancestor, the latter being probably structurally similar to *roshi-a* (**Fig. 2**).

Finally, the spotted gar *Lepisosteus oculatus* equally harbors a "roshi-b-like" gene, above referred to as *roshi-l* (**Fig. 2**). This gene is a small-scale duplicate of the gar *roshi*, on the same chromosome (LG12). Interestingly, all domains of the ancestral *roshi* have been independently lost after duplication during the evolution of the gar *roshi-l* and of teleost *roshi-b*. We consider that independent loss is a more parsimonious scenario than considering an ancestral *roshil* that would have independently gained all three other domains in gar and in teleost fish evolution.

Acquisition of *roshi-b* by iridoviruses occurred through transfer from a fish host

In this work, we showed that *roshi-b* was transferred from a fish host to viruses of the Red Seabream Iridovirus (RSIV) species within the *Megalocytivirus* genus (Iridoviridae) (Kurita and Nakajima, 2012). In RSIV, it is known as *ORF515L*. This transfer likely occurred more recently than 70 Ma from a host in the Carangaria clade – the most resembling nucleic sequence to RSIV sequences being *C. ignobilis roshi-b*. Megalocytiviruses are large dsDNA viruses have a genome size of ca. 112,000 nucleotides encoding 116 ORFs and are 120-200 nm in diameter (Nakajima and Kunita, 2005). They have first been characterized in the early 1990s by studying diseased fish that were lethargic and showed severe anemia, gill petechiae (localized bleedings) and spleen enlargement that can lead to death in many cases (Inouye *et al.*, 1992). RSIV viruses are not specific to red seabream *Pagrus major*, where they were first isolated, and they have a broad host range within teleost fish. In particular, RSIV disease has already been identified in cultured Carangidae such as *Seriola dumerili* (Nakajima and Maeno, 1998). Thus, based on both phylogenetic inference and ecological considerations, the transfer of *roshi-b* from a carangarian fish host to a RSIV viral species is a likely event.

Horizontal gene transfer (HGT) can be defined as "the passage of DNA from one organism to another through a mechanism other than reproduction" and can occur through much diverse mechanisms (Soucy *et al.*, 2015; Gilbert and Cordaux, 2017). This process might result in the adaptive gain of novel genes and traits (Schönknecht *et al.*, 2014). HGT from vertebrate hosts to dsDNA viruses has not been reported to our knowledge in *Megalocytivirus* but has previously been shown in other dsDNA viruses such as Poxviridae (Bratke and McLysaght, 2008). In Poxviridae, HGT could proceed through reverse transcription of a host mRNA before integration of the cDNA into the viral genome (Bratke and McLysaght, 2008).

Interestingly, Iridoviridae owe their name to the iridescence observed in certain events of arthropod infection, such as those caused by Invertebrate Iridescent Iridoviruses (IIVs), that are iridoviruses different to Megalocytiviruses (Williams, 2008). In the case of IIVs, the iridescence is due to optical properties of virion accumulation in infected tissues that reflect light. It could be tempting to speculate that there is a direct link between this optical property and the presence of *ORF515L*, which is homolog to *roshi-b* – the latter being ohnolog to an iridophore marker. However, *ORF515L* is one of the rare genes to be RSIV-specific within the whole Iridoviridae family and is thus absent from IIVs (Kurita and Nakajima, 2012). In addition, RSIV viruses, though specifically harboring a *roshi-b* homolog, are Iridoviridae members that *do not* cause iridescence in infected fish – to our knowledge. Hence, the iridescence property of IIVs cannot be due to *roshi-b*. In conclusion, the link between a fish iridophore marker and the presence of its Ts3R-ohnolog in an Iridoviridae virus seems to be a fortuitous, yet curious, event.

Of note, the function of the viral gene *ORF515L* is not known in RSIV, where it is named "HIT-like" (Kurita and Nakajima, 2012). However, yet unknown, the conservation of a gene in a viral genome for such a long time in multiple strains of RSIV might reveal a useful function for the virus host cycle **(Fig. 3)**. In addition, it is not only the "HIT-like" domain of *roshi-b* that is conserved, but also the first subdomain of *roshi* D3 that has no predicted function (**Fig. 1**). Both of these subdomains were thus the only ones to be kept in both teleost *roshi-b* and gar *roshi-l* as well as in RSIV viruses. They were also conserved in other animal lineages (**Fig. 4**). Altogether, it seems that these two subdomains require to be closely linked to be fully functional and provide an adaptive, though unknown, advantage.

***Roshi* evolution in animals has probably been shaped by horizontal gene transfer**

In general, HGT is difficult to ascertain because two different methods can lead to opposite conclusions on the HGT origin – or not – of a given gene (Lander *et al.*, 2001; Stanhope *et al.*, 2001). To be rigorously inferred, one can

use a combination of several different criterions that independently are not sufficient proof (Bratke and McLysaght, 2008; Fitzpatrick, 2012): structural arguments, such as (i) a shared exon/intron structure in distantly related species, (ii) an intron-less gene in one of the species, indicative of a reverse transcription step before HGT or (iii) the presence, within a multispecies conserved synteny block, of an additional gene in one species that would not be present in closely related species indicating a new insertion; phylogenetic arguments, such as (iv) a patchy phyletic distribution of the gene and (v) a discrepancy between the species tree and the gene phylogenetic tree; ecological arguments, *i. e.* (vi) evolutionary distant sharing the same ecological niche; last, "parametric arguments" based on sequence composition, *e. g.* (vii) detection of codon usage variation between the potential newly acquired gene and the genome background of the host; (viii) variation in GC composition along the chromosome. However, methods (vii) and (viii) are not adapted to eukaryotic genomes as they are often large, more complex and present more intrinsic-variation than prokaryotic genomes (Mallet *et al.*, 2010). In addition, codon usage variation cannot be detected after long evolutionary times as newly transferred sequences rapidly homogenize with the host genome (Fitzpatrick, 2012).

In this work, we studied the evolution of *roshi* in animal lineages. We unambiguously demonstrated phylogenetic arguments (iv) and (v), both considered to be the most convincing evidence for a HGT detection (Fitzpatrick, 2012). In particular, *roshi* was found in very specific animal lineages (**Fig. 4**), suggesting several HGT events between animal taxa. We assume that multiple HGT within animals explains more parsimoniously the observed distribution rather than considering a gene ancestral to all Metazoa that would have been repeatedly lost. Of note, we exclude the hypothesis of multiple transfers between fungi and animals based on phylogenetic arguments, since fungal sequences always branch "outside" animals (not shown).

Besides, *roshi* gene tree revealed a major discrepancy with the species tree (**Fig. 5**). The position of *Hydra vulgaris* sequence in *roshi* gene tree was particularly unexpected. In addition, *Hydra vulgaris* harbored an intron-less *roshi* sequence, that would support argument (ii) if the donor sequence had introns – which we cannot affirm. The hydra lineage has equally been shown to have a high rate of gene acquisition through HGT in animals (Chapman *et al.*, 2010). Altogether this suggesting a putative HGT event between an ancestor hydra a chelicerate species.

Argument (iii) was also probably verified. Indeed, genes in the vicinity of *roshi* on gar LG12 (*mief1, rbs19bp1, atf4*) are also positioned in close proximity on human chromosome 22. However, in humans, there is no trace of *roshi* sequences in this syntenic block, which could indicate a fish-specific acquisition.

Roshi is present in all ray-finned fish tested here, a group that is at least 350 Ma old, and in no lobe-finned fish, that diverged from ray-finned fish ca. 440 Ma, nor in any other tested outgroup (cartilaginous fish and jawless fish). This could mean that this gene, if acquired in the ray-finned fish ancestor, was acquired between these two dates. The ecological argument for HGT of species sharing the same niche (vi) might also, surprisingly, be verified. Indeed, all fish, *L. anatina*, *H. vulgaris* are aquatic animals, and thus are susceptible of indirect contact – and thus transfer. However, mites and scorpions are not aquatic and are not in direct contact with aquatic animals. Yet, tree topology, with scorpions branching outside fish (**Fig. 5)** suggest that if there was a transfer between fish and scorpions, it must have occurred before the appearance of ray-finned fish lineage ca. 350 Ma. Scorpiones can be traced back to middle Silurian, ca. 430-440 Ma (Jeram, 1990). Whereas modern scorpions are terrestrial, scorpions were whether totally or at least partially aquatic at that time (Jeram, 1990; Braddy, 2001; Kenrick *et al.*, 2012). Their paleo-environment is thus compatible with an aquatic gene transfer from or to a fish ancestor. Of note, more scorpion genomes would be required to better estimate the timing of *roshi* acquisition in this lineage. This would also be needed in the Acari lineage since the presence of a *roshi* ortholog in *Nothrus palustris,* a soil mite, is very intriguing (Lebrun, 1969).

Thus, after a putative initial transfer from a fungal host to a brachiopod ancestor (which would explain the basal position of *L. anatina* in the gene tree), multiple HGT events between different aquatic species at around the same epoch is a possible scenario for *roshi* evolution. Different *roshi* copies within animals should thus actually be called "xenologs", *i. e.* shared sequences derived from HGT events (Gray and Fitch, 1984).

Potential mechanisms for multiple HGT events in animals

In the eventuality of multiple HGT events within animals, mechanisms for gene transfer need to be examined. Indeed, while HGT in bacteria and archea is a well-known evolutive force, it is much less present in eukaryotes and especially in animals (Crisp *et al.*, 2015; Soucy *et al.*, 2015). HGT between animals and other organisms could occur more frequently

than previously thought (Huang, 2013; Boto, 2014). In fish, a newly acquired gene from a bacterial donor has even previously been documented, with unknown function though (Sun *et al.*, 2015). Nevertheless, to date, direct HGT *between* animals are much rarely documented and the suggested mechanisms only hypothetical. Few such cases have nevertheless been unraveled, where arthropod parasites are confirmed or putative vectors of transposon HGT between tetrapods (Gilbert *et al.*, 2010, 2012). Thus, *roshi* transfer between different animal phyla could be mediated by parasites. As observed by others, this would not necessarily require a chromosomal integration of *roshi* within the parasite genome and could involve extrachromosomal vectors such as viruses (Piskurek and Okada, 2007; Gilbert *et al.*, 2010). The finding of a viral acquisition of a *roshi-b* paralog supports this hypothesis. In addition, a viral infection can enhance the activity of transposable elements and the frequency of DNA recombination (Liu *et al.*, 2010), hereby increasing the likelihood of gene transfer from a virus to a new animal host. Besides, transfer from viral material to eukaryotic host genome has already been documented (Liu *et al.*, 2010). Finally, many acarines are parasites and it has previously been suggested that mites phoretic and parasitic feeding behavior, as well as their frequent association with viruses, made them good vectors for HGT (Houck, 1994). The presence of a *roshi* copy in the mite *N. palustris* would be consistent with this scenario.

<u>*Roshi* evolutionary origin could be rooted in Fungi</u>

In this study, we identified multiple *roshi* orthologs in the Fungi lineage, with a puzzling patchy phylogenetic distribution, that, again, is parsimoniously more supported by the hypothesis of multiple HGT events rather than by multiple losses consecutive to an ancestral acquisition event (**Fig. 6**). *Roshi* homologs are thus not only found in Ascomycota but also in *Hypsizygus marmoreus* (Basidiomycota). HGT events, though thought to be generally rare within eukaryotes, are starting to be considered as a relatively major evolutionary force within the fungal lineage (Rosewich and Kistler, 2000; Fitzpatrick, 2012). This group is among the most densely represented in terms of genomic data within eukaryotes, facilitating comparative genomics analysis and the identification of HGT events, such as the transfer of pathogenicity chromosomes in *Fusarium* (Sordariomycetes, Ascomycota) (Ma *et al.*, 2010). As discussed above for animals, here again the mechanisms underlying the numerous observed HGTs are not well identified, but could involve numerous cell fusions, sometimes between distantly related fungal species (Fitzpatrick, 2012). The presence of numerous copies of *roshi* between distantly related species could thus be due to widespread occurrence of HGT within this lineage (**Fig. 6)**.

Phylogeny topologies indicated that fungal sequences are well separated to animal sequences (**Fig. 5**). This was verified for the ascomycete *Cenococcum geophilum* as well as for other fungal species (not shown). This could either suggest (i) a transfer from an animal ancestor to a fungal ancestor or (ii) an HGT from a fungal ancestor to an animal, prior to further independent HGT events in each of these lineages. Node dating in phylogenies is compatible with an HGT event between both Opisthokonta lineages. Indeed, Ascomycota, where *roshi* acquisition could have occurred, can be dated back to 350-400 Ma (**Fig. 6**). Most animal groups where *roshi* is present are at least this age (when known). We discuss here the likelihood of both scenarios: fungi to animals, or animals to fungi.

Gene transfers from fungi to arthropods have already been documented, for example in the case of the acquisition of carotenoid synthesis genes in aphids (Moran and Jarvik, 2010), or, and independently, in at least other arthropod lineages (Altincicek *et al.*, 2012; Cobbs *et al.*, 2013). Fungal gene transfer to chordates, including humans, has also been documented for hyaluronan synthase genes (Crisp *et al.*, 2015). Underlying mechanisms for such transfers, have, however, not been unraveled yet and are rarely discussed in corresponding studies. Nevertheless, the close ecological vicinity is sometimes evoked as a prerequisite to such transfers. Many marine fungi do exist, in various clades including Ascomycota and Basidiomycota (Jones *et al.*, 2015), and marine as well as intertidal species of Dothideomycetes, close relatives to *C. geophilum*, have been documented (Suetrong *et al.*, 2009).

On the other hand, gene transfers from animals to other lineages are much rarer. One such case has been documented for *set8*, a histone modifier, from human to *Plasmodium* (Kishore *et al.*, 2013), but this is the only documented case – to our knowledge.

Altogether, this suggests that the most likely scenario for *roshi* evolution is a transfer from a fungal ancestor that would have first acquired the ancestral *roshi* before HGT to an animal species, such as a brachiopod ancestor, before subsequent HGT both within animals and within fungi (**Fig. 7**).

Roshi evolution could reveal a new putative scenario for the evolution of vertebrate pigment cells

Roshi was first identified because one of its Ts3R-ohnologs, *roshi-a*, is necessary for iridophore differentiation (Salis, Lorin et al., *in prep.*). Phylogenetic distribution of *roshi* shows that, within vertebrates, this gene is restricted to ray-finned fish (**Fig. 4**). However, iridophores are not only present in the ray-finned fish lineage. Indeed, in tetrapods, amphibians as well as lepidosaurians harbor iridophores (Bagnara *et al.*, 1968; Kuriyama *et al.*, 2006). While only melanocytes have been described to date in cartilaginous fish, iridophores have also been described in lampreys (Wright and McBurney, 1992).

At the genetic level, it was recently shown that a teleost iridophore marker, *pnp4a*, is conserved in teleost fish, coelacanth and lepidosaurians, but not in other vertebrate lineages (Kimura *et al.*, 2017; Lorin *et al.*, 2018). In addition, two other iridophore markers that have been described in teleosts are not conserved throughout vertebrates: *csf1* is not present in amphibians and *mpv17* is not present in lizards and amphibians (Patterson and Parichy, 2013; Krauss *et al.*, 2013). Thus, the taxonomic distribution of *roshi-a* as well as of other teleost iridophore markers suggests that iridophore differentiation processes in ray-finned fish might partially differ from those within other vertebrate lineages.

A generally accepted hypothesis is that the ancestor of vertebrates already possessed melanocytes, iridophores and xanthophores (Mellgren and Johnson, 2002; Braasch *et al.*, 2009; Lapedriza *et al.*, 2014). Here, we suggest that all vertebrate iridophores are not homologous and that different lineages might have evolved specific differentiation pathways, including through the co-option of HGT-acquired genes.

ACKNOWLEDGEMENTS

The authors wish to thank C. Dechaud, J. Goedert, F. Lallemand, M. Naville and C. Rey for helpful discussion. T. L. is recipient of an individual grant from the French Ministry of Research and Education.

REFERENCES

Altincicek B, Kovacs JL, Gerardo NM (2012). Horizontally transferred fungal carotenoid genes in the two-spotted spider mite Tetranychus urticae. *Biol Lett* **8**: 253–257.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J Mol Biol* **215**: 403–410.

Anisimova M, Gil M, Dufayard J-F, Dessimoz C, Gascuel O (2011). Survey of Branch Support Methods Demonstrates Accuracy, Power, and Robustness of Fast Likelihood-based Approximation Schemes. *Syst Biol* **60**: 685–699.

Bagnara JT, Taylor JD, Hadley ME (1968). The dermal chromatophore unit. *J Cell Biol* **38**: 67–79.

Battelle B-A, Ryan JF, Kempler KE, Saraf SR, Marten CE, Warren WC, *et al.* (2016). Opsin Repertoire and Expression Patterns in Horseshoe Crabs: Evidence from the Genome of Limulus polyphemus (Arthropoda: Chelicerata). *Genome Biol Evol* **8**: 1571–1589.

Beimforde C, Feldberg K, Nylinder S, Rikkinen J, Tuovila H, Dörfelt H, *et al.* (2014). Estimating the Phanerozoic history of the Ascomycota lineages: Combining fossil and molecular data. *Mol Phylogenet Evol* **78**: 386–398.

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, *et al.* (2013). GenBank. *Nucleic Acids Res* **41**: D36–D42.

Berthelot C, Brunet F, Chalopin D, Juanchich A, Bernard M, Noël B, *et al.* (2014). The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nat Commun* **5**.

Betancur-R R, Wiley EO, Arratia G, Acero A, Bailly N, Miya M, *et al.* (2017). Phylogenetic classification of bony fishes. *BMC Evol Biol* **17**: 162.

Borek D, Kozak M, Pei J, Jaskolski M (2014). Crystal structure of active site mutant of antileukemic <scp>l</scp> -asparaginase reveals conserved zinc-binding site. *FEBS J* **281**: 4097–4111.

Boto L (2014). Horizontal gene transfer in the acquisition of novel traits by metazoans. *Proc R Soc B Biol Sci* **281**: 20132450.

Braasch I, Brunet F, Volff J-N, Schartl M (2009). Pigmentation pathway evolution after whole-genome duplication in fish. *Genome Biol Evol* **1**: 479–493.

Braasch I, Gehrke AR, Smith JJ, Kawasaki K, Manousaki T, Pasquier J, *et al.* (2016). The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. *Nat Genet* **48**: 427–437.

Braasch I, Volff JN, Schartl M (2008). The evolution of teleost pigmentation and the fish-specific genome duplication. *J Fish Biol* **73**: 1891–1918.

Braddy SJ (2001). Eurypterid palaeoecology: palaeobiological, ichnological and comparative evidence for a 'mass–moult–mate' hypothesis. *Palaeogeogr Palaeoclimatol Palaeoecol* **172**: 115–132.

Brameier M, Krings A, MacCallum RM (2007). NucPred Predicting nuclear localization of proteins. *Bioinformatics* **23**: 1159–1160.

Bratke KA, McLysaght A (2008). Identification of multiple independent horizontal gene transfers into poxviruses using a comparative genomics approach. *BMC Evol Biol* **8**: 67.

Brenner C (2002). Hint, Fhit, and GalT: function, structure, evolution, and mechanism of three branches of the histidine triad superfamily of nucleotide hydrolases and transferases. *Biochemistry* **41**: 9003–14.

Cai Y, Cronin CN, Engel AG, Ohno K, Hersh LB, Rodgers DW (2004). Choline acetyltransferase structure reveals distribution of mutations that cause motor disorders. *EMBO J* **23**: 2047–2058.

Casey PJ (1995). Protein lipidation in cell signaling. *Science* **268**: 221–5.

Chang Y, Wang S, Sekimoto S, Aerts AL, Choi C, Clum A, *et al.* (2015). Phylogenomic Analyses Indicate that Early Fungi Evolved Digesting Cell Walls of Algal Ancestors of Land Plants. *Genome Biol Evol* **7**: 1590–1601.

Chapman JA, Kirkness EF, Simakov O, Hampson SE, Mitros T, Weinmaier T, *et al.* (2010). The dynamic genome of Hydra. *Nature* **464**: 592–596.

Chen K-H, Miadlikowska J, Molnár K, Arnold AE, U'Ren JM, Gaya E, *et al.*

(2015). Phylogenetic analyses of eurotiomycetous endophytes reveal their close affinities to Chaetothyriales, Eurotiales, and a new order – Phaeomoniellales. *Mol Phylogenet Evol* **85**: 117–130.

Cobbs C, Heath J, Stireman JO, Abbot P (2013). Carotenoids in unexpected places: Gall midges, lateral gene transfer, and carotenoid biosynthesis in animals. *Mol Phylogenet Evol* **68**: 221–228.

Coelho MA, Gonçalves C, Sampaio JP, Gonçalves P (2013). Extensive Intra-Kingdom Horizontal Gene Transfer Converging on a Fungal Fructose Transporter Gene (J Heitman, Ed.). *PLoS Genet* **9**: e1003587.

la Cour T, Kiemer L, Mølgaard A, Gupta R, Skriver K, Brunak S (2004). Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng Des Sel* **17**: 527–536.

Crisp A, Boschetti C, Perry M, Tunnacliffe A, Micklem G (2015). Expression of multiple horizontally acquired genes is a hallmark of both vertebrate and invertebrate genomes. *Genome Biol* **16**: 50.

Darriba D, Taboada GL, Doallo R, Posada D (2011). ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27**: 1164–1165.

Edgecombe GD, Legg DA (2014). Origins and early evolution of arthropods (A Smith, Ed.). *Palaeontology* **57**: 457–468.

Eriani G, Cavarelli J, Martin F, Ador L, Rees B, Thierry J-C, *et al.* (1995). The class II aminoacyl-tRNA synthetases and their active site: Evolutionary conservation of an ATP binding site. *J Mol Evol* **40**.

Fankhauser N, Maser P (2005). Identification of GPI anchor attachment signals by a Kohonen self-organizing map. *Bioinformatics* **21**: 1846–1852.

Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, *et al.* (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* **44**: D279–D285.

Fitzpatrick DA (2012). Horizontal gene transfer in fungi. *FEMS Microbiol Lett* **329**: 1–8.

Gazis R, Kuo A, Riley R, LaButti K, Lipzen A, Lin J, *et al.* (2015). The genome of Xylona heveae provides a window into fungal endophytism. *Fungal Biol* **120**: 26–42.

Gilbert C, Cordaux R (2017). Viruses as vectors of horizontal transfer of genetic material in eukaryotes. *Curr Opin Virol* **25**: 16–22.

Gilbert C, Hernandez SS, Flores-Benabib J, Smith EN, Feschotte C (2012). Rampant Horizontal Transfer of SPIN Transposons in Squamate Reptiles. *Mol Biol Evol* **29**: 503–515.

Gilbert C, Schaack S, Pace II JK, Brindley PJ, Feschotte C (2010). A role for host–parasite interactions in the horizontal transfer of transposons across phyla. *Nature* **464**: 1347–1350.

Goda M, Fujii R (1995). Blue Chromatophores in Two Species of Callionymid Fish. *Zoolog Sci* **12**: 811–813.

Gouy M, Guindon S, Gascuel O (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**: 221–224.

Gray G, Fitch W (1984). Evolution of antibiotic resistance genes: the DNA sequence of a kanamycin resistance gene from Staphylococcus aureus. *Mol Biol Evol* **1**: 57–66.

Grigoriev I V., Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, *et al.* (2014). MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res* **42**: D699–D704.

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst Biol* **59**: 307–321.

Hansen K, Perry BA, Dranginis AW, Pfister DH (2013). A phylogeny of the highly diverse cup-fungus family Pyronemataceae (Pezizomycetes, Ascomycota) clarifies relationships and evolution of selected life history traits. *Mol Phylogenet Evol* **67**: 311–335.

von Heijne G (1990). The signal peptide. *J Membr Biol* **115**: 195–201.

Hongsanan S, Maharachchikumbura SSN, Hyde KD, Samarakoon MC, Jeewon R, Zhao Q, *et al.* (2017). An updated phylogeny of Sordariomycetes based on phylogenetic and molecular clock evidence. *Fungal Divers* **84**: 25–41.

Houck MA (1994). Mites as Potential Horizontal Transfer Vectors of Eukaryotic Mobile Genes: Proctolaelaps regalis as a Model. In: *Mites*, Springer US: Boston, MA, pp 45–69.

Huang J (2013). Horizontal gene transfer in eukaryotes: The weak-link model. *BioEssays*: n/a-n/a.

Hustad V, Miller A (2011). Phylogenetic placement of four genera within the Leotiomycetes (Ascomycota). *North Am Fungi* **6**: 1–13.

Igawa T, Nozawa M, Suzuki DG, Reimer JD, Morov AR, Wang Y, *et al.* (2017). Evolutionary history of the extant amphioxus lineage with shallow-branching diversification. *Sci Rep* **7**: 1157.

Inoue J, Sato Y, Sinclair R, Tsukamoto K, Nishida M (2015). Rapid genome reshaping by multiple-gene loss after whole-genome duplication in teleost fish suggested by mathematical modeling. *Proc Natl Acad Sci* **112**: 14918–14923.

Inouye K, Yamano K, Maeno Y, Nakajima K, Matsuoka M, Wada Y, *et al.* (1992). Iridovirus Infection of Cultured Red Sea Bream, Pagrus major. *Fish Pathol* **27**: 19–27.

Jaillon O, Aury J-M, Brunet F, Petit J-L, Stange-Thomann N, Mauceli E, *et al.* (2004). Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature* **431**: 946–957.

Jeram AJ (1990). Book-lungs in a Lower Carboniferous scorpion. *Nature* **343**: 360–361.

Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, *et al.* (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**: 1236–1240.

Jones EBG, Suetrong S, Sakayaroj J, Bahkali AH, Abdel-Wahab MA, Boekhout T, *et al.* (2015). Classification of marine Ascomycota, Basidiomycota, Blastocladiomycota and Chytridiomycota. *Fungal Divers* **73**: 1–72.

Kall L, Krogh A, Sonnhammer ELL (2007). Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. *Nucleic Acids Res* **35**: W429–W432.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**: 845–858.

Kenrick P, Wellman CH, Schneider H, Edgecombe GD (2012). A timeline for terrestrialization: consequences for the carbon cycle in the Palaeozoic. *Philos Trans R Soc B Biol Sci* **367**: 519–536.

Kimura T, Takehana Y, Naruse K (2017). pnp4a Is the Causal Gene of the Medaka Iridophore Mutant guanineless. *G3 Genes|Genomes|Genetics Genes|Genomes|Genetics* **7**: 1357–1363.

Kishore SP, Stiller JW, Deitsch KW (2013). Horizontal gene transfer of epigenetic machinery and evolution of parasitism in the malaria parasite Plasmodium falciparum and other apicomplexans. *BMC Evol Biol* **13**: 37.

Kodama Y, Mashima J, Kosuge T, Kaminuma E, Ogasawara O, Okubo K, *et al.* (2018). DNA Data Bank of Japan: 30th anniversary. *Nucleic Acids*

Res **46**: D30–D35.

Kosugi S, Hasebe M, Tomita M, Yanagawa H (2009). Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc Natl Acad Sci* **106**: 10171–10176.

Krauss J, Astrinides P, Frohnhofer HG, Walderich B, Nusslein-Volhard C (2013). transparent, a gene affecting stripe formation in Zebrafish, encodes the mitochondrial protein Mpv17 that is required for iridophore survival. *Biol Open* **2**: 703–710.

Krogh A, Larsson B, von Heijne G, Sonnhammer EL. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes11Edited by F. Cohen. *J Mol Biol* **305**: 567–580.

Kumar S, Stecher G, Suleski M, Hedges SB (2017). TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol Biol Evol* **34**: 1812–1819.

Kurita J, Nakajima K (2012). Megalocytiviruses. *Viruses* **4**: 521–538.

Kuriyama T, Miyaji K, Sugimoto M, Hasegawa M (2006). Ultrastructure of the Dermal Chromatophores in a Lizard (Scincidae: Plestiodon latiscutatus) with Conspicuous Body and Tail Coloration. *Zoolog Sci* **23**: 793–799.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature* **409**: 860–921.

Lapedriza A, Petratou K, Kelsh RN (2014). Neural Crest Cells and Pigmentation. In: *Neural Crest Cells*, Elsevier, pp 287–311.

Le SQ, Gascuel O (2008). An Improved General Amino Acid Replacement Matrix. *Mol Biol Evol* **25**: 1307–1320.

Lebrun P (1969). Écologie et biologie de Nothrus palustris (C. L. Koch) (Acari Oribatei) Densité et structure de la population. *Oikos* **20**: 34.

Lien S, Koop BF, Sandve SR, Miller JR, Kent MP, Nome T, *et al.* (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature* **533**: 200–205.

Lin J-R, Mondal A, Liu R, Hu J (2012). Minimalist ensemble algorithms for genome-wide protein localization prediction. *BMC Bioinformatics* **13**: 157.

Liu H, Fu Y, Jiang D, Li G, Xie J, Cheng J, *et al.* (2010). Widespread Horizontal Gene Transfer from Double-Stranded RNA Viruses to Eukaryotic Nuclear Genomes. *J Virol* **84**: 11876–11887.

Lorin T, Brunet FG, Laudet V, Volff J-N (2018). Teleost Fish-Specific Preferential Retention of Pigmentation Gene-Containing Families After Whole Genome Duplications in Vertebrates. *G3 Genes|Genomes|Genetics*: g3.200201.2018.

Louis A, Nguyen NTT, Muffato M, Roest Crollius H (2015). Genomicus update 2015: KaryoView and MatrixView provide a genome-wide perspective to multispecies comparative genomics. *Nucleic Acids Res* **43**: D682–D689.

Luo A, Qiao H, Zhang Y, Shi W, Ho SY, Xu W, *et al.* (2010). Performance of criteria for selecting evolutionary models in phylogenetics: a comprehensive study based on simulated datasets. *BMC Evol Biol* **10**: 242.

Ma L-J, van der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, Di Pietro A, *et al.* (2010). Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. *Nature* **464**: 367–373.

Mallet L V, Becq J, Deschavanne P (2010). Whole genome evaluation of horizontal transfers in the pathogenic fungus Aspergillus fumigatus. *BMC Genomics* **11**: 171.

Marcionetti A, Rossier V, Bertrand JAM, Litsios G, Salamin N (2018). First draft genome of an iconic clownfish species (Amphiprion frenatus). *Mol Ecol Resour*.

McGrath CL, Gout J-F, Johri P, Doak TG, Lynch M (2014). Differential retention and divergent resolution of duplicate genes following whole-genome duplication. *Genome Res* **24**: 1665–1675.

Mellgren EM, Johnson SL (2002). The evolution of morphological complexity in zebrafish stripes. *Trends Genet* **18**: 128–34.

Meyer A, Schartl M (1999). Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol* **11**: 699–704.

Miadlikowska J, Kauff F, Högnabba F, Oliver JC, Molnár K, Fraker E, *et al.* (2014). A multigene phylogenetic synthesis for the class Lecanoromycetes (Ascomycota): 1307 fungi representing 1139 infrageneric taxa, 317 genera and 66 families. *Mol Phylogenet Evol* **79**: 132–168.

Michalska K, Hernandez-Santoyo A, Jaskolski M (2008). The Mechanism of Autocatalytic Activation of Plant-type L-Asparaginases. *J Biol Chem* **283**: 13388–13397.

Moran NA, Jarvik T (2010). Lateral Transfer of Genes from Fungi Underlies Carotenoid Production in Aphids. *Science (80-)* **328**: 624–627.

Nakajima K, Kunita J (2005). Red sea bream iridoviral disease. *Uirusu* **55**: 115–25.

Nakajima K, Maeno Y (1998). Pathogenicity of Red Sea Bream Iridovirus and Other Fish Iridoviruses to Red Sea Bream. *Fish Pathol* **33**: 143–144.

Nguyen Ba AN, Pogoutse A, Provart N, Moses AM (2009). NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. *BMC Bioinformatics* **10**: 202.

Nilsson Sköld H, Aspengren S, Wallin M (2013). Rapid color change in fish and amphibians - function, regulation, and emerging applications. *Pigment Cell Melanoma Res* **26**: 29–38.

O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, *et al.* (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* **44**: D733–D745.

Ohno S (1970). *Evolution by gene duplication*. Springer.

Ojanguren-Affilastro AA, Adilardi RS, Mattoni CI, Ramírez MJ, Ceccarelli FS (2017). Dated phylogenetic studies of the southernmost American buthids (Scorpiones; Buthidae). *Mol Phylogenet Evol* **110**: 39–49.

Oulion S, Bertrand S, Belgacem MR, Le Petillon Y, Escriva H (2012). Sequencing and Analysis of the Mediterranean Amphioxus (Branchiostoma lanceolatum) Transcriptome (V Laudet, Ed.). *PLoS One* **7**: e36554.

Parichy DM (2006). Evolution of danio pigment pattern development. *Heredity (Edinb)* **97**: 200–210.

Park E, Hwang D-S, Lee J-S, Song J-I, Seo T-K, Won Y-J (2012). Estimation of divergence times in cnidarian evolution based on mitochondrial protein-coding genes and the fossil record. *Mol Phylogenet Evol* **62**: 329–345.

Pasquier J, Braasch I, Batzel P, Cabau C, Montfort J, Nguyen T, *et al.* (2017). Evolution of gene expression after whole-genome duplication: New insights from the spotted gar genome. *J Exp Zool Part B Mol Dev Evol* **328**: 709–721.

Patterson LB, Parichy DM (2013). Interactions with Iridophores and the Tissue Environment Required for Patterning Melanophores and Xanthophores during Zebrafish Adult Pigment Stripe Formation (GS Barsh, Ed.). *PLoS Genet* **9**: e1003561.

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* **8**: 785–786.

Piskurek O, Okada N (2007). Poxviruses as possible vectors for horizontal transfer of retroposons from reptiles to mammals. *Proc Natl Acad Sci* **104**: 12046–12051.

Posada D (2008). jModelTest: Phylogenetic Model Averaging. *Mol Biol Evol* **25**: 1253–1256.

Prieto M, Wedin M (2013). Dating the Diversification of the Major Lineages of Ascomycota (Fungi) (S Ho, Ed.). *PLoS One* **8**: e65576.

Regier JC, Shultz JW, Zwick A, Hussey A, Ball B, Wetzer R, *et al.* (2010). Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* **463**: 1079–1083.

Rosewich UL, Kistler HC (2000). Role of Horizontal Gene Transfer in the Evolution of Fungi. *Annu Rev Phytopathol* **38**: 325–363.

Schartl M, Larue L, Goda M, Bosenberg MW, Hashimoto H, Kelsh RN (2016). What is a vertebrate pigment cell? *Pigment Cell Melanoma Res* **29**: 8–14.

Schoch CL, Crous PW, Groenewald JZ, Boehm EWA, Burgess TI, de Gruyter J, *et al.* (2009). A class-wide phylogenetic assessment of Dothideomycetes. *Stud Mycol* **64**: 1–15.

Schoch CL, Sung G-H, López-Giráldez F, Townsend JP, Miadlikowska J, Hofstetter V, *et al.* (2009). The Ascomycota Tree of Life: A Phylum-wide Phylogeny Clarifies the Origin and Evolution of Fundamental Reproductive and Ecological Traits. *Syst Biol* **58**: 224–239.

Schönknecht G, Weber APM, Lercher MJ (2014). Horizontal gene acquisitions by eukaryotes as drivers of adaptive evolution. *BioEssays* **36**: 9–20.

Schwentner M, Bosch TCG (2015). Revisiting the age, evolutionary history and species level diversity of the genus Hydra (Cnidaria: Hydrozoa). *Mol Phylogenet Evol* **91**: 41–55.

Séraphin B (1992). The HIT protein family: a new family of proteins present in prokaryotes, yeast and mammals. *DNA Seq* **3**: 177–9.

Sharma PP, Baker CM, Cosgrove JG, Johnson JE, Oberski JT, Raven RJ, *et al.* (2018). A revised dated phylogeny of scorpions: Phylogenomic support for ancient divergence of the temperate Gondwanan family Bothriuridae. *Mol Phylogenet Evol* **122**: 37–45.

Sievers F, Higgins DG (2014). Clustal omega, accurate alignment of very large numbers of sequences. In: *Methods in Molecular Biology*, Vol 1079, pp 105–116.

Silva DN, Duplessis S, Talhinhas P, Azinheira H, Paulo OS, Batista D (2015). Genomic Patterns of Positive Selection at the Origin of Rust Fungi (M Robinson-Rechavi, Ed.). *PLoS One* **10**: e0143959.

Solovyev V (2008). Statistical Approaches in Eukaryotic Gene Prediction. In: Balding D, Cannings C, Bishop M (eds) *Handbook of Statistical Genetics*, John Wiley & Sons, Ltd: Chichester, UK, pp 97–159.

Soucy SM, Huang J, Gogarten JP (2015). Horizontal gene transfer: building the web of life. *Nat Rev Genet* **16**: 472–482.

Sperling EA, Pisani D, Peterson KJ (2011). Molecular paleobiological insights into the origin of the Brachiopoda. *Evol Dev* **13**: 290–303.

Stanhope MJ, Lupas A, Italia MJ, Koretke KK, Volker C, Brown JR (2001). Phylogenetic analyses do not support horizontal gene transfers from bacteria to vertebrates. *Nature* **411**: 940–944.

Suetrong S, Schoch CL, Spatafora JW, Kohlmeyer J, Volkmann-Kohlmeyer B, Sakayaroj J, *et al.* (2009). Molecular systematics of the marine Dothideomycetes. *Stud Mycol* **64**: 155–173.

Sun Y, Huang Y, Li X, Baldwin CC, Zhou Z, Yan Z, *et al.* (2016). Fish-T1K (Transcriptomes of 1,000 Fishes) Project: large-scale transcriptome data for fish evolution studies. *Gigascience* **5**: 18.

Sun B-F, Li T, Xiao J-H, Jia L-Y, Liu L, Zhang P, *et al.* (2015). Horizontal functional gene transfer from bacteria to fishes. *Sci Rep* **5**: 18676.

Szilágyi A, Skolnick J (2006). Efficient Prediction of Nucleic Acid Binding Function from Low-resolution Protein Structures. *J Mol Biol* **358**: 922–933.

Vasse M, Voglmayr H, Mayer V, Gueidan C, Nepel M, Moreno L, *et al.* (2017). A phylogenetic perspective on the association between ants (Hymenoptera: Formicidae) and black yeasts (Ascomycota: Chaetothyriales). *Proc R Soc B Biol Sci* **284**: 20162519.

Vidotto M, Grapputo A, Boscari E, Barbisan F, Coppe A, Grandi G, *et al.* (2013). Transcriptome sequencing and de novo annotation of the critically endangered Adriatic sturgeon. *BMC Genomics* **14**: 407.

Wang H, Xu Z, Gao L, Hao B (2009). A fungal phylogeny based on 82 complete genomes using the composition vector method. *BMC Evol Biol* **9**: 195.

Williams T (2008). Natural invertebrate hosts of iridoviruses (Iridoviridae). *Neotrop Entomol* **37**: 615–32.

Wright GM, McBurney KM (1992). Changes in the ventral dermis and development of iridophores in the anadromous sea lamprey, Petromyzon marinus, during metamorphosis: an ultrastructural study. *Histol Histopathol* **7**: 237–50.

Xie Y, Zheng Y, Li H, Luo X, He Z, Cao S, *et al.* (2016). GPS-Lipid: a robust tool for the prediction of multiple lipid modification sites. *Sci Rep* **6**: 28249.

Zapata F, Goetz FE, Smith SA, Howison M, Siebert S, Church SH, *et al.* (2015). Phylogenomic Analyses Support Traditional Relationships within Cnidaria (RE Steele, Ed.). *PLoS One* **10**: e0139068.

Zhang J (2000). Protein-length distributions for the three domains of life. *Trends Genet* **16**: 107–9.

Zhou L-W, Kõljalg U (2013). A new species of Lenzitopsis (Thelephorales, Basidiomycota) and its phylogenetic placement. *Mycoscience* **54**: 87–92.

Zhou F, Xue Y, Yao X, Xu Y (2006). CSS-Palm: palmitoylation site prediction with a clustering and scoring strategy (CSS). *Bioinformatics* **22**: 894–896.

FIGURES

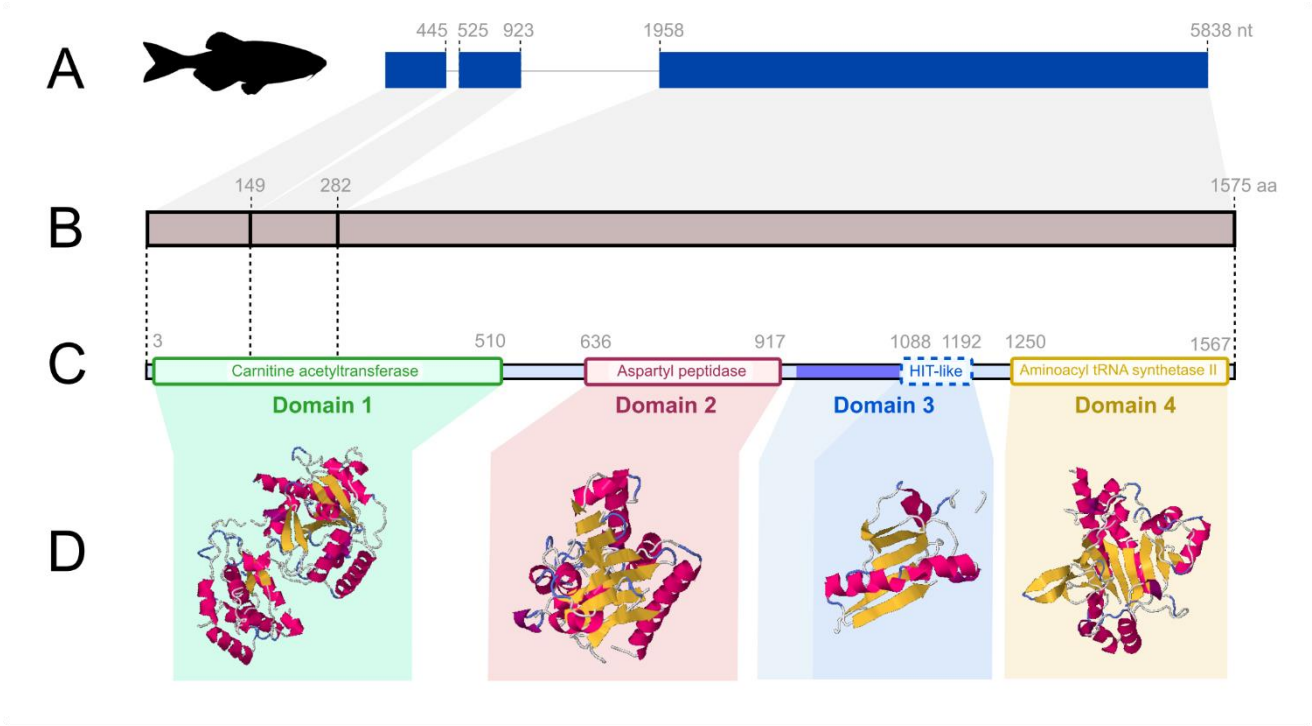


Figure 1: Structure of the zebrafish *roshi-a* gene (si:ch211-256m1:8) and of the Roshi-a protein. In the zebrafish *Danio rerio,* Roshi-a corresponds to locus si:ch211-256m1:8 located on chromosome 3. The coding sequence is constituted of three exons of respective length of 445, 398 and 3880 nt (A). The predicted zebrafish Roshi-a protein is 1575 amino acids long (B). Four domains structure the protein (C) and their 3D folding is represented (D). In particular, domain 4 has the canonical seven-stranded anti-parallel beta-sheet observed for yeast 'core-enzyme' of amino-acyl tRNA synthetases (class II).

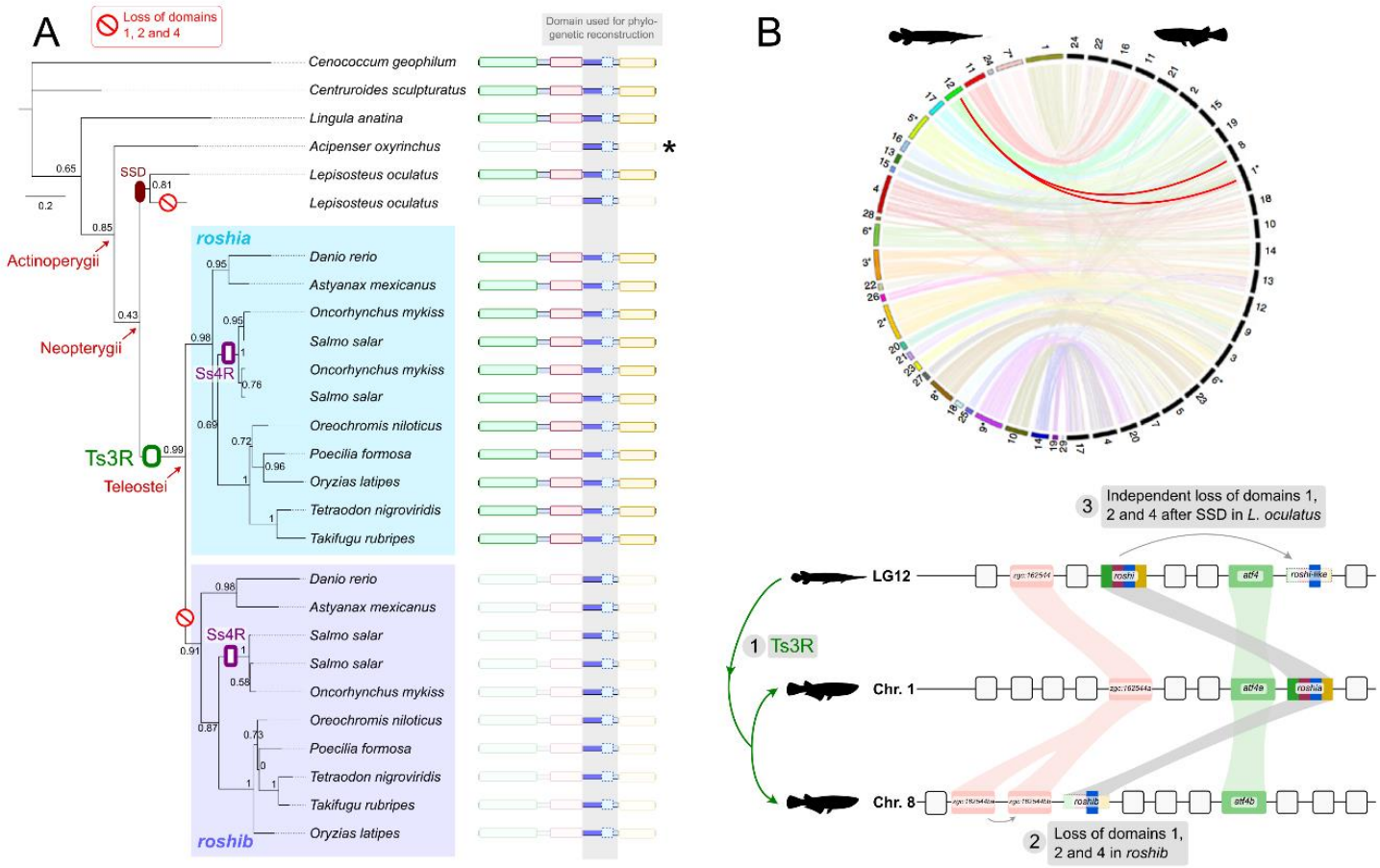


Figure 2: *Roshi-a* originates from an ancestral ray-finned fish *roshi* gene.

(A) Maximum-likelihood phylogenetic reconstruction of Roshi-a and Roshi-b protein sequences of D3 domain (highlighted in grey) supporting a Ts3R-mediated duplication of roshi in teleost fish. In addition, both *roshi-a* and *roshi-b* have been duplicated in the salmonid lineage after Ss4R. Of note, one Ss4R-ohnolog of *roshi-b* has been lost in the rainbow trout (*O. mykiss*) lineage. Gene structure is shown on the right of the tree. In *roshi-b*, all domains but D3 have ancestrally been lost in the teleost lineage after the Ts3R event. In the spotted gar *L. oculatus* lineage, an additional SSD has occurred. One of the paralogs lost D1, D2 and D4. Colors correspond to the different domains of *roshi-a* (see Fig. 1). SSD: Small-Scale Duplication. Ts3R: Teleost-Specific Whole-Genome duplication (WGD). Ss4R: Salmonid-Specific WGD. *C. geophilum* (Fungi), *C. sculpturatus* (Scorpiones) and *L. anatina* (Brachiopoda) are used as outgroups. SH-aLRT node support values are indicated. Star indicates sequence originating from publicly available transcriptomic data.

(B) Synteny analyses of teleost paralogons support a Ts3R-mediated duplication of roshi. *Top*: In medaka, chromosomes 1 and 8 harbor paralogons corresponding to linkage group 12 in *L. oculatus* (red curves). Adapted from (Braasch *et al.*, 2016). *Bottom*: Comparison of paralogons on spotted gar LG12 and medaka chromosomes 1 and 8 shows that *roshi-a* and *roshi-b* are in overall conserved regions in medaka. These paralogons have however been rearranged during evolution. In spotted gar, the ancestral *roshi* gene has been locally duplicated on LG12 into a "*roshi-like*" gene. The gene *zgc:162544* has been segmentally duplicated in medaka chromosome 8 (arrow). Blank boxes correspond to annotated genes that are present on each chromosome but for which no ohnolog is found.

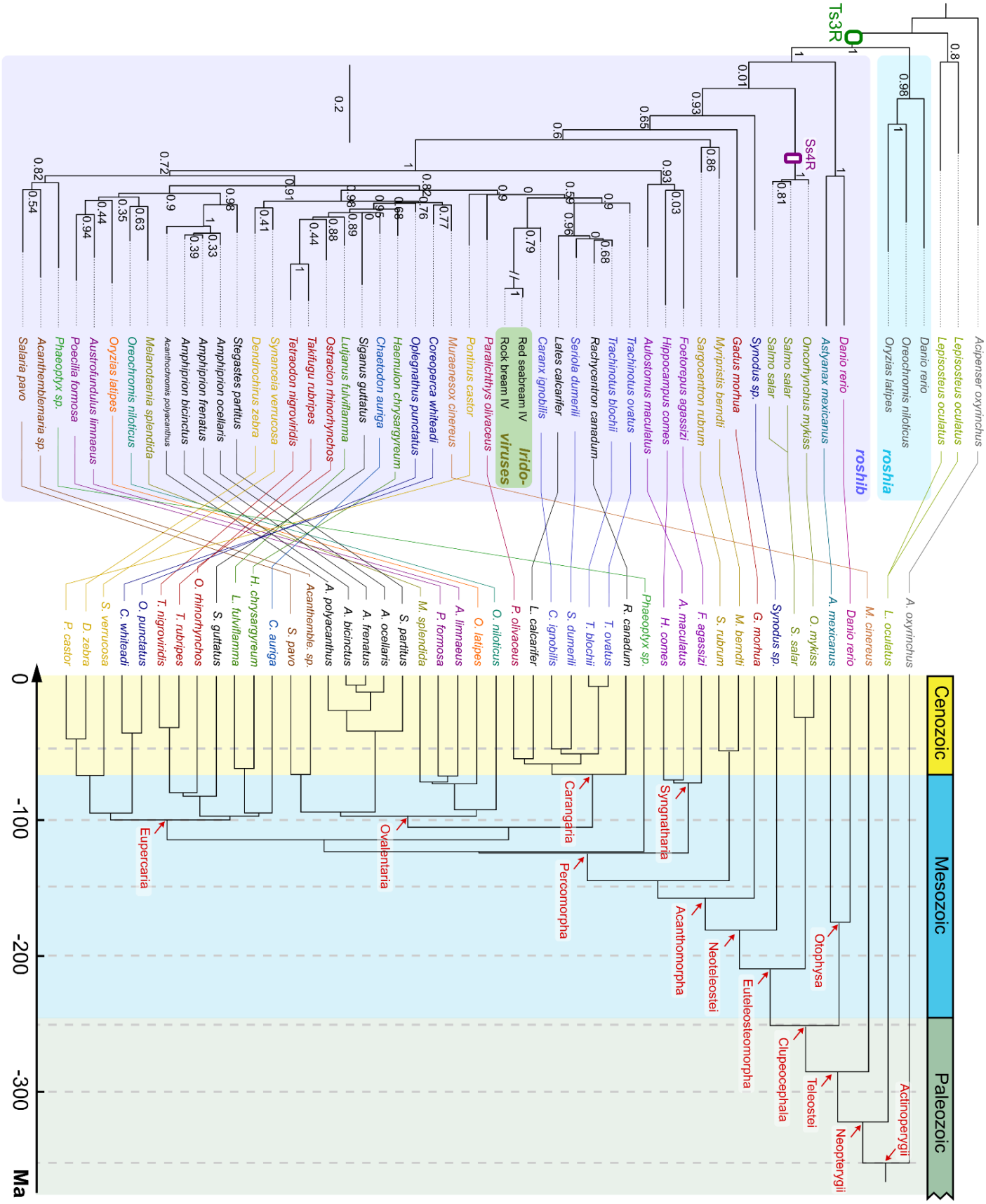

Acipenser oxyrinchus
Lepisosteus oculatus
Lepisosteus oculatus
Danio rerio
Oreochromis niloticus
Oryzias latipes
roshia
roshib
Ts3R
Ss4R
Danio rerio
Astyanax mexicanus
Oncorhynchus mykiss
Salmo salar
Salmo salar
Synodus sp.
Gadus morhua
Myripristis berndti
Sargocentron rubrum
Foetorepus agassizi
Hippocampus comes
Aulostomus maculatus
Trachinotus ovatus
Trachinotus blochii
Rachycentron canadum
Seriola dumerili
Lates calcarifer
Caranx ignobilis
Red seabream IV
Rock bream IV
Irido-viruses
Paralichthys olivaceus
Pontinus castor
Muraenesox cinereus
Coreoperca whiteadi
Oplegnathus punctatus
Haemulon chrysargyreum
Chaetodon auriga
Siganus guttatus
Lutjanus fulviflamma
Ostracion rhinorhynchos
Takifugu rubripes
Tetraodon nigroviridis
Synanceia verrucosa
Dendrochirus zebra
Stegastes partitus
Amphiprion ocellaris
Amphiprion frenatus
Amphiprion bicinctus
Acanthochromis polyacanthus
Melanotaenia splendida
Oreochromis niloticus
Oryzias latipes
Austrofundulus limnaeus
Poecilia formosa
Phaeoptyx sp.
Acanthemblemaria sp.
Salaria pavo
0.2
A. oxyrinchus
L. oculatus
M. cinereus
Danio rerio
A. mexicanus
O. mykiss
S. salar
Synodus sp.
G. morhua
M. berndti
S. rubrum
F. agassizi
A. maculatus
H. comes
Phaeoptyx sp.
R. canadum
T. ovatus
T. blochii
S. dumerili
C. ignobilis
L. calcarifer
P. olivaceus
O. niloticus
O. latipes
A. limnaeus
P. formosa
M. splendida
S. partitus
A. ocellaris
A. frenatus
A. bicinctus
A. polyacanthus
Acanthemble. sp.
S. pavo
C. auriga
H. chrysargyreum
L. fulviflamma
S. guttatus
O. rhinorhynchos
T. rubripes
T. nigroviridis
O. punctatus
C. whiteadi
S. verrucosa
D. zebra
P. castor
Cenozoic
Mesozoic
Paleozoic
Actinopterygii
Neopterygii
Teleostei
Clupeocephala
Euteleosteomorpha
Neoteleostei
Acanthomorpha
Otophysa
Syngnatharia
Carangaria
Percomorpha
Ovalentaria
Eupercaria
0
-100
-200
-300
Ma

Figure 3: Evidence for gene transfer from *roshi-b* to iridoviruses more recently than 70 million years ago. PREVIOUS PAGE

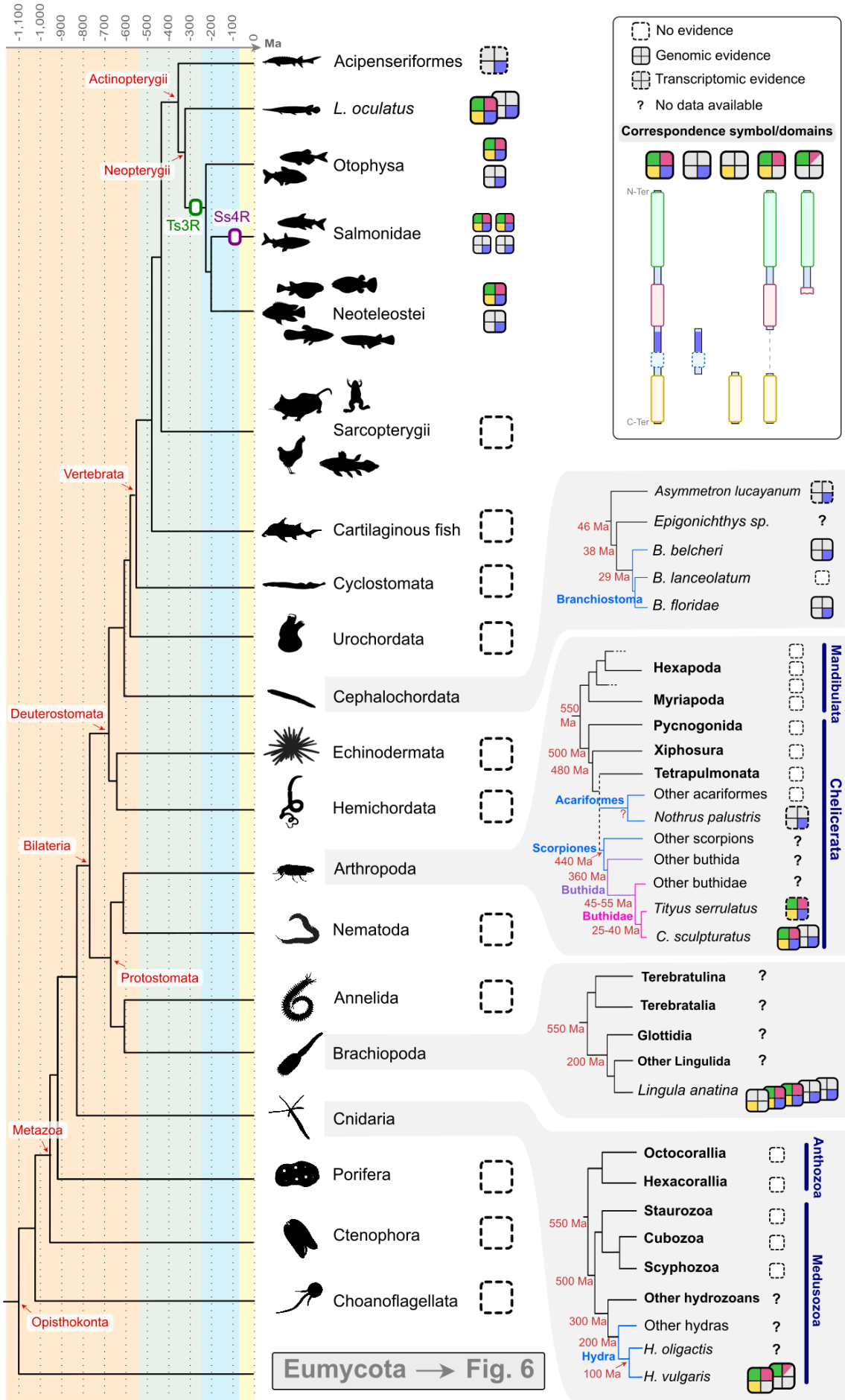
Left: Maximum-likelihood phylogenetic reconstruction (nucleic sequences) of *roshi-b* and *roshi-a* D3 domain using a dataset of fish genomic and transcriptomic data. Viral sequences are highlighted and branch in the Carangaria clade.

Right: phylogeny of the species used for phylogenetic reconstruction (topology, clade names and divergence times are based on (Betancur-R *et al.*, 2017)).

Overall, the *roshi-b* phylogeny reflects the species phylogeny with some discrepancy among Percomorpha linked to low bootstrap values. Gap length on iridovirus (IV) branch: 0.4. SH-aLRT node support values are indicated.

Figure 4: *roshi* taxonomic distribution in animals is very patchy. NEXT PAGE

Ray-finned fish left aside, *roshi* (or any of its domains) is not present in most animal lineages. However, sequences of specific domains or the whole sequence could be identified in some lineages: cephalochordates, the mite *N. palustris*, buthidae scorpions, the brachiopod *L. anatina* and the cnidarian *H. vulgaris*. Species tree topology and divergence times of major tree are based on http://timetree.org/ (Kumar *et al.*, 2017). Lineage trees topologies and divergence times are indicated – when known – for specific taxa. Phylogeny of Cephalochordata: see (Igawa *et al.*, 2017); phylogeny within Arthropoda is a synthesis based on (Regier *et al.*, 2010; Edgecombe and Legg, 2014; Ojanguren-Affilastro *et al.*, 2017; Sharma *et al.*, 2018); phylogeny within Brachiopoda: see (Sperling *et al.*, 2011); phylogeny within Cnidaria is a synthesis based on (Park *et al.*, 2012; Schwentner and Bosch, 2015; Zapata *et al.*, 2015).


-1,100
-1,000
-900
-800
-700
-600
-500
-400
-300
-200
-100
0
Ma
Actinopterygii
Neopterygii
Ts3R
Ss4R
Vertebrata
Deuterostomata
Bilateria
Protostomata
Metazoa
Opisthokonta
Acipenseriformes
L. oculatus
Otophysa
Salmonidae
Neoteleostei
Sarcopterygii
Cartilaginous fish
Cyclostomata
Urochordata
Cephalochordata
Echinodermata
Hemichordata
Arthropoda
Nematoda
Annelida
Brachiopoda
Cnidaria
Porifera
Ctenophora
Choanoflagellata
Eumycota → Fig. 6
No evidence
Genomic evidence
Transcriptomic evidence
? No data available
Correspondence symbol/domains
N-Ter
C-Ter
Asymmetron lucayanum
Epigonichthys sp.
B. belcheri
B. lanceolatum
B. floridae
46 Ma
38 Ma
29 Ma
Branchiostoma
Hexapoda
Myriapoda
Pycnogonida
Xiphosura
Tetrapulmonata
Other acariformes
Nothrus palustris
Other scorpions
Other buthida
Other buthidae
Tityus serrulatus
C. sculpturatus
550 Ma
500 Ma
480 Ma
Acariformes
Scorpiones
440 Ma
360 Ma
Buthida
45-55 Ma
Buthidae
25-40 Ma
Mandibulata
Chelicerata
Terebratulina
Terebratalia
Glottidia
Other Lingulida
Lingula anatina
550 Ma
200 Ma
Octocorallia
Hexacorallia
Staurozoa
Cubozoa
Scyphozoa
Other hydrozoans
Other hydras
H. oligactis
H. vulgaris
550 Ma
500 Ma
300 Ma
200 Ma
Hydra
100 Ma
Anthozoa
Medusozoa

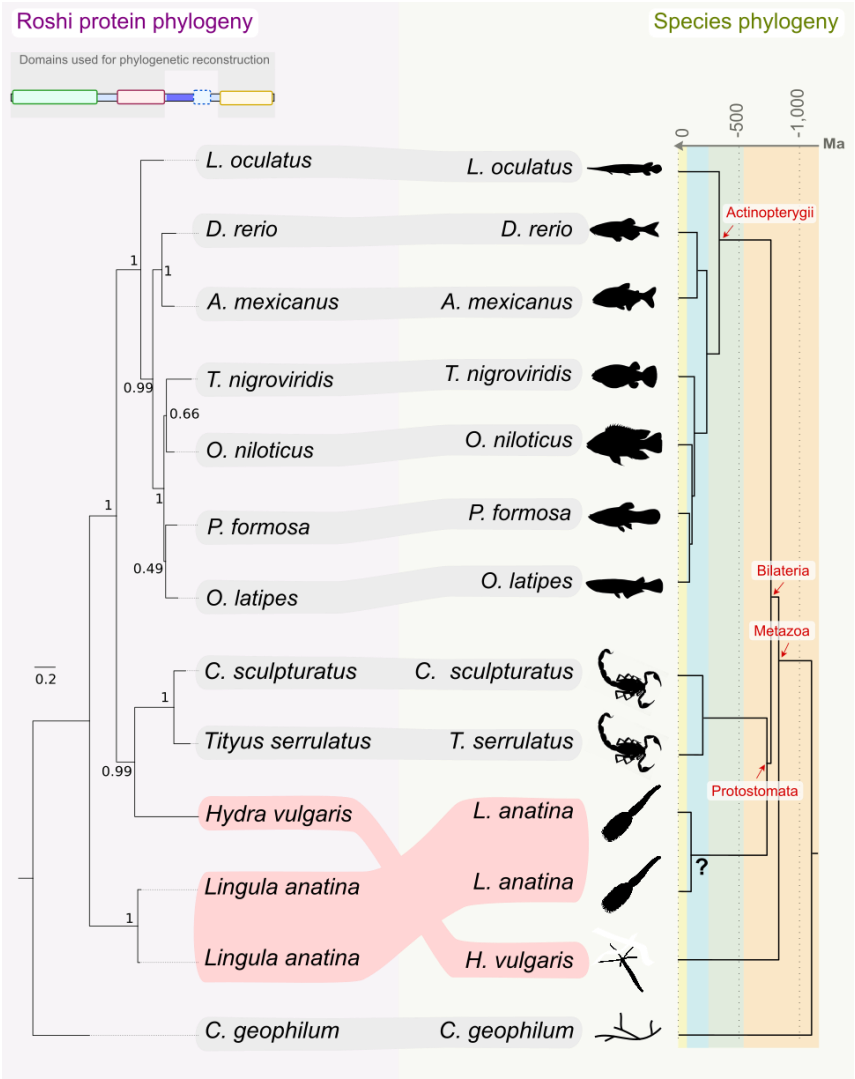


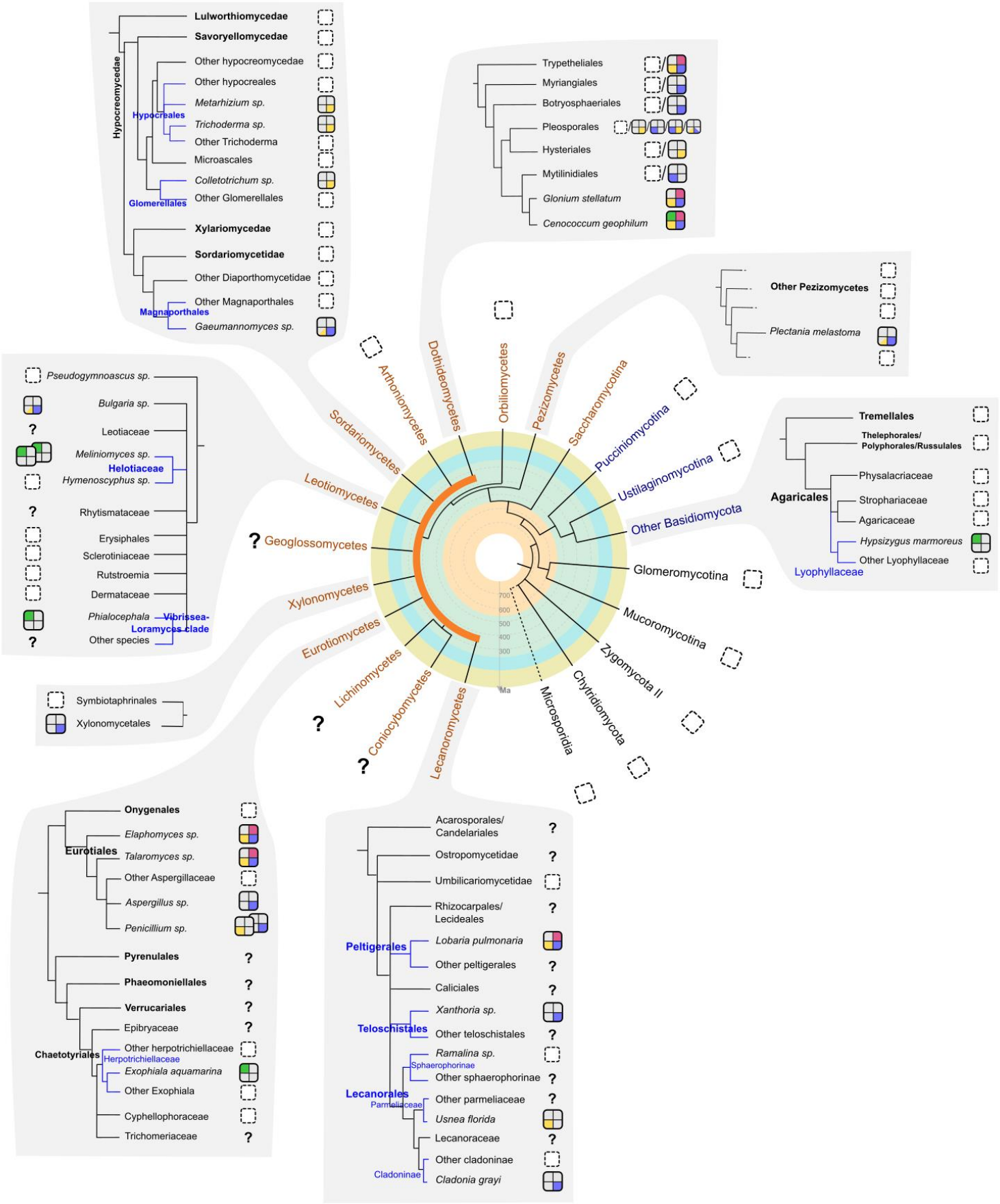
Figure 5: Roshi evolution in animals reveals discrepancies with species evolution.

Maximum-likelihood phylogenetic reconstruction based on amino acid sequences of Roshi and teleost Roshia-a reveals a discrepancy – highlighted in pink – between Roshi tree (left) and species tree (right). The regions used for phylogenetic reconstruction are indicated. *C. geophilum* (Fungi) is used as outgroup. SH-aLRT node support values are indicated.

Figure 6: Roshi and of each of its domains have a patchy taxonomic distribution in different fungal lineages. NEXT PAGE

The taxonomic distribution of *roshi* sequences is represented (same legend as in Fig. 4). Colored circles correspond to ages in million years. Taxa in orange (respectively, blue) correspond to classes of Ascomycota (respectively, Basiodiomycota). Orange represents a major polytomy due to node low supports and conflicting class resolution in the Ascomycota lineage.

Global fungi phylogeny: (Wang *et al.*, 2009; Chang *et al.*, 2015). Phylogenies within each lineage are based on a synthesis of multiple studies: (Schoch, Crous, *et al.*, 2009; Schoch, Sung, *et al.*, 2009; Hustad and Miller, 2011; Zhou and Kõljalg, 2013; Hansen *et al.*, 2013; Miadlikowska *et al.*, 2014; Gazis *et al.*, 2015; Chen *et al.*, 2015; Silva *et al.*, 2015; Vasse *et al.*, 2017; Hongsanan *et al.*, 2017)Schoch 2009 and Beimforde 2014 (Ascomycota); Hustad (2°11° (Leotiomycetes); Miadlikowska (2014) (Lecanoromycota); Chen 2015 and Vasse 2017 (Eurotiomycetes); Hongsanan 2017 (Sordariomycetes); Hansen 2013 (Pezizomycetes); Gazis 2015 (Xylonomycetes); Schoch 2009 (Dothideomycetes); Silva 2015 and 2hou 2013 (Basiodiomycota).


Hypocreomycedae
Lulworthiomycedae
Savoryellomycedae
Other hypocreomycedae
Other hypocreales
Metarhizium sp.
Hypocreales
Trichoderma sp.
Other Trichoderma
Microascales
Colletotrichum sp.
Glomerellales
Other Glomerellales
Xylariomycedae
Sordariomycetidae
Other Diaporthomycetidae
Other Magnaporthales
Magnaporthales
Gaeumannomyces sp.
Trypetheliales
Myriangiales
Botryosphaeriales
Pleosporales
Hysteriales
Mytilinidiales
Glonium stellatum
Cenococcum geophilum
Other Pezizomycetes
Plectania melastoma
Pseudogymnoascus sp.
Bulgaria sp.
Leotiaceae
Meliniomyces sp.
Helotiaceae
Hymenoscyphus sp.
Rhytismataceae
Erysiphales
Sclerotiniaceae
Rutstroemia
Dermataceae
Phialocephala
Vibrissea-Loramyces clade
Other species
Symbiotaphrinales
Xylonomycetales
Arthoniomycetes
Dothideomycetes
Orbiliomycetes
Pezizomycetes
Saccharomycotina
Pucciniomycotina
Ustilaginomycotina
Other Basidiomycota
Glomeromycotina
Mucoromycotina
Zygomycota II
Chytridiomycota
Microsporidia
Lecanoromycetes
Coniocybomycetes
Lichinomycetes
Eurotiomycetes
Xylonomycetes
Geoglossomycetes
Leotiomycetes
Sordariomycetes
Tremellales
Thelephorales/ Polyphorales/Russulales
Agaricales
Physalacriaceae
Strophariaceae
Agaricaceae
Hypsizygus marmoreus
Other Lyophyllaceae
Lyophyllaceae
Onygenales
Eurotiales
Elaphomyces sp.
Talaromyces sp.
Other Aspergillaceae
Aspergillus sp.
Penicillium sp.
Pyrenulales
Phaeomoniellales
Verrucariales
Epibryaceae
Chaetotyriales
Other herpotrichiellaceae
Herpotrichiellaceae
Exophiala aquamarina
Other Exophiala
Cyphellophoraceae
Trichomeriaceae
Acarosporales/ Candelariales
Ostropomycetidae
Umbilicariomycetidae
Rhizocarpales/ Lecideales
Peltigerales
Lobaria pulmonaria
Other peltigerales
Caliciales
Xanthoria sp.
Teloschistales
Other teloschistales
Ramalina sp.
Sphaerophorinae
Other sphaerophorinae
Lecanorales
Parmeliaceae
Other parmeliaceae
Usnea florida
Lecanoraceae
Other cladoninae
Cladoninae
Cladonia grayi
700
600
500
400
300
Ma

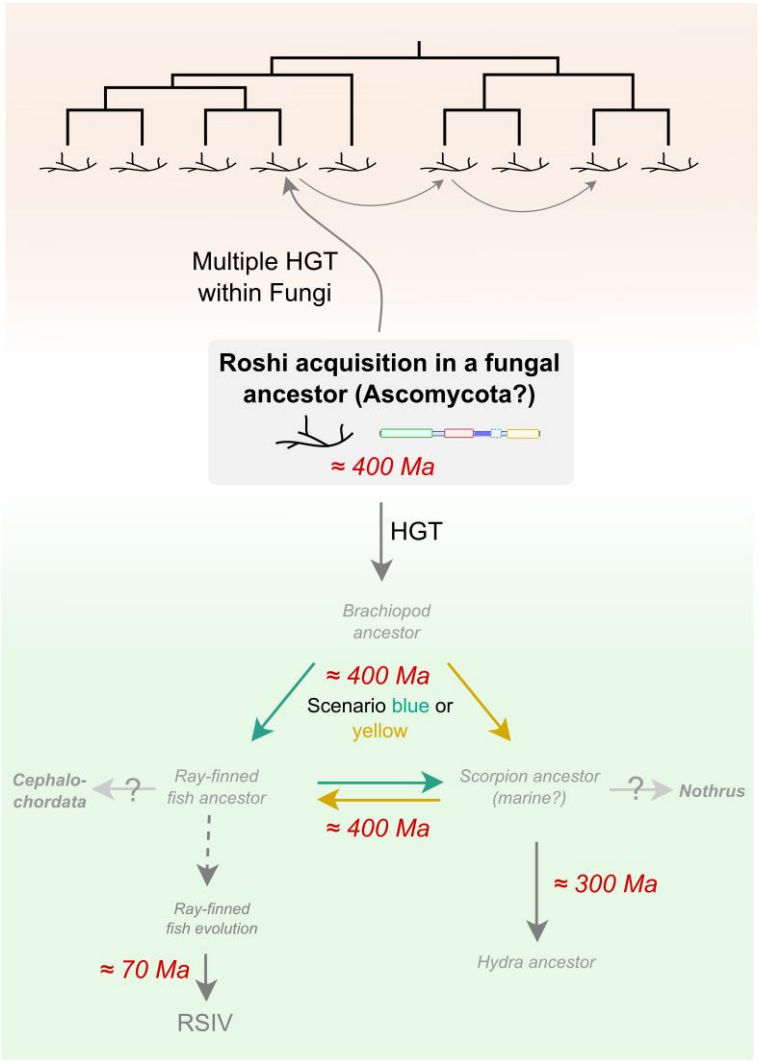


Figure 7. Hypothetical simplified scenario for the evolution of *roshi* within Opisthokonta (fungi on top, animals on bottom). Lineage-specific duplications or domain losses are not indicated. Putative ages for each event are indicated based on ages of the different taxa. Filled arrows correspond to HGTs. Dashed arrow corresponds to evolution of ray-finned fish. Question marks indicate HGTs that would need further analyses. See text for discussion on each event. RSIV: Red seabream iridovirus.

SUPPLEMENTARY FIGURES

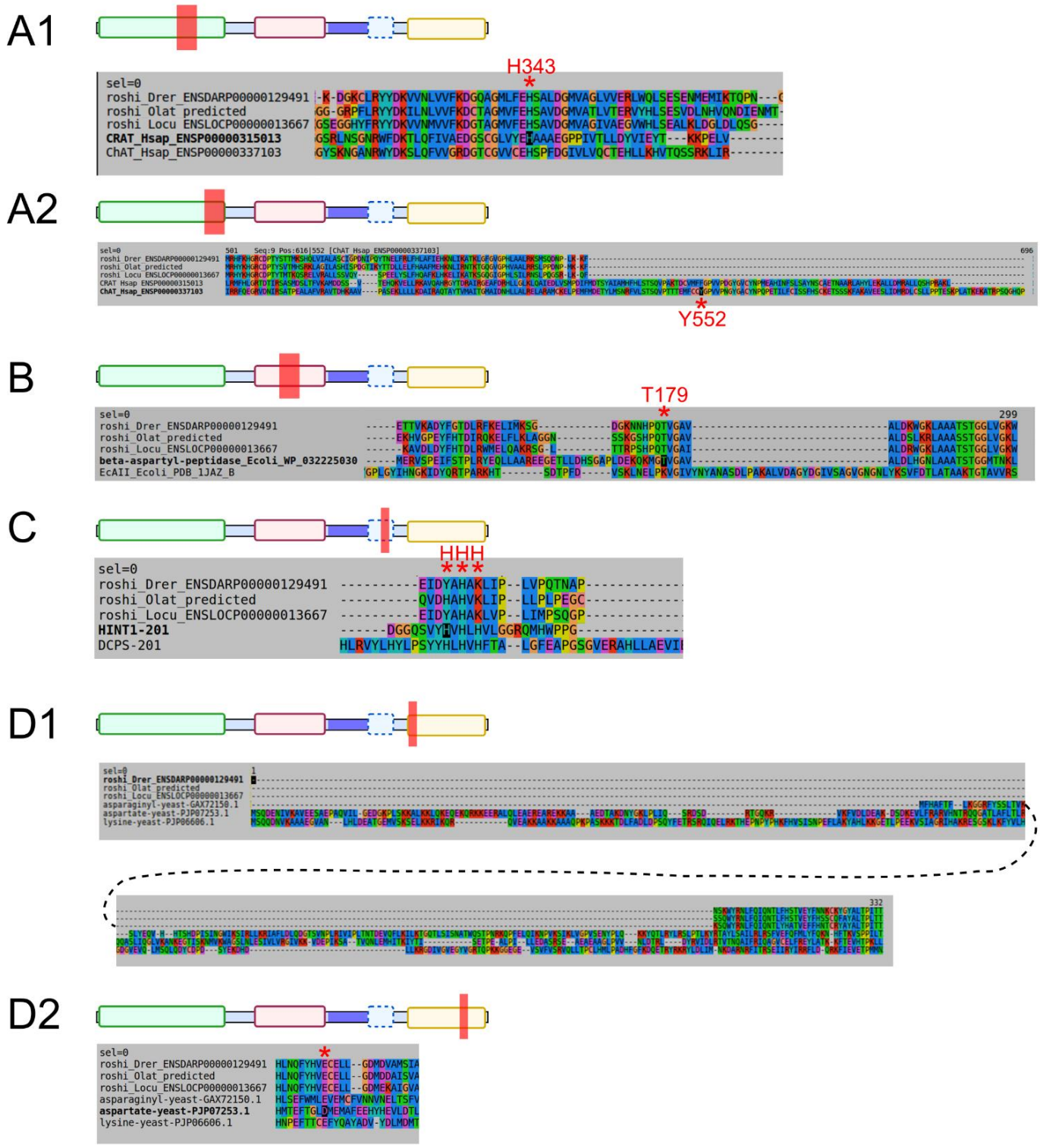


Figure S1. Amino acid alignments of multiple *roshi-a* fish sequences with other proteins corresponding to each *roshi-a* domain. Positions of the shown sequence are indicated compared to the whole *roshi-a* sequence. Drer: *Danio rerio*; Olat*: Oryzias latipes*; Locu: *Lepisosteus oculatus*.

(A) Protein alignment of Roshi-a with human choline acyltransferase (ChAT) and carnitine acyltransferase (CRAT). The active site of both these enzymes is a histidine residue corresponding to His343 in CRAT (A1). This site is conserved in

Roshi-a. However, the binding of choline by ChAT requires an extra-tyrosine residue (Tyr552) that is not conserved in *roshi-a* domain 1, that ends before (A2).

(B) Amino acid alignment of *roshi-a* with *E. coli* asparaginase and beta-aspartyl peptidase proteins. Overall, the beta-aspartyl peptidase sequence aligns better than the asparaginase. Most predicted functional amino acids, especially Thr179, were conserved between *E. coli* beta-aspartyl-peptidase and Roshi-a.

(C) Amino acid alignment of Roshi-a with humans HINT and DCPS proteins. Both these enzymes harbor a canonical HIT (histidine triad) domain with the motif H-ϕ-H-ϕ-H-ϕ (where H is histidine and ϕ a hydrophobic amino acid) (https://www.ebi.ac.uk/interpro/entry/IPR011146). However, in Roshi-a, this catalytic triad is not conserved since fish sequences only have one (most often) or two (rarely) histidine residues, together with a lysine residue. The motif in Roshi-a is thus Y/H-ϕ-H-ϕ-K-ϕ.

(D) Amino acid alignment of Roshi-a with yeast *S. cerevisiae* amino-acyl tRNA synthetases (class II) for asparagine, aspartate and lysine. In general, tRNA synthetases of class II (aaRS II) are enzymes constituted of three different domains. However, Roshi-a does only harbor the 'core enzyme' domain of class II aaRS and does not have both N-ter domains of canonical aaRS II (D1). The amino acid binding site (red star, D2) is conserved between asparagine aaRS, lysine aaRS and fish Roshi-a, but not with yeast aspartate aaRS.

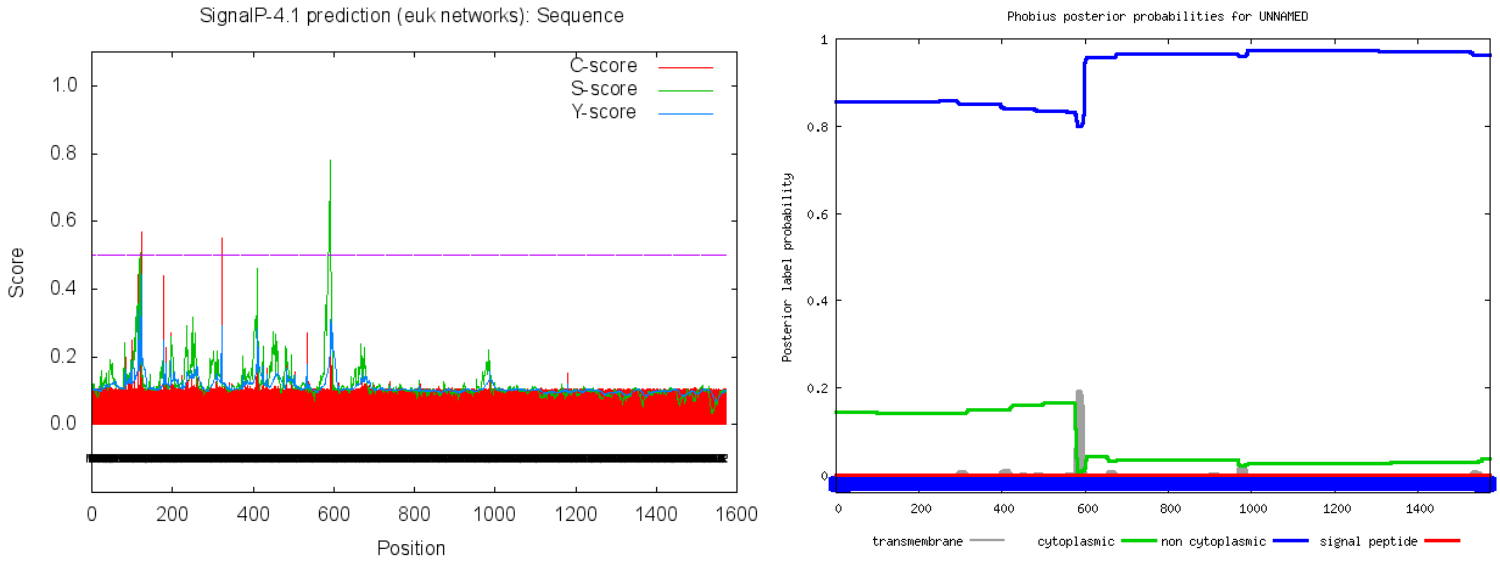


Figure S2. Results of signal peptide prediction wihin *Danio rerio* Roshi-a protein sequence. Left: SignalP result. A score above 0.5 on a sequence of ca. 30 amino acids, usually in N-Terminal extremity, is usually considered for prediction of potential NLS. This is not observed in Roshi-a. Right: Phobius result. No transmembrane segment nor signal peptide is predicted in Roshi-a. Here, the whole sequence is modelled as "non-cytoplasmic" but this is due to improper prediction: since there is no transmembrane domain, the program cannot properly predict if the whole protein is cytoplasmic or extra-cytoplasmic.

Predicted bipartite NLS		
Pos.	**Sequence**	**Score**
1202	PTYKGYVTSVDGPPANEEDLSEIHSKITKTI	7.9

Figure S3. Results of nuclear localization signal (NLS) prediction using cNLS-Mapper. A NLS was predicted with a score of 7.9, that indicates a protein "partially localized to the nucleus". However, no NLS was found by other detection tools: NucPred, NLStradamus and SeqNLS. Though seemingly positive, this result indicate that the Roshi-a protein is likely not to be localized in nucleus.

RÉSULTATS

CHAPITRE 4

Étude transcriptomique du métabolisme thyroïdien et des gènes de la pigmentation au cours du développement du poisson-clown *Amphiprion ocellaris*

Avant-propos

Au cours du premier chapitre, nous avons recensé et étudié l'évolution du répertoire de gènes pigmentaires. Cette liste de gènes a été en partie utilisée pour étudier les bases histologiques et cellulaires de la bande blanche chez le poisson-clown *A. ocellaris* (chapitre 2). Au cours de cette étude, nous avons identifié un nouveau marqueur d'iridophores, *roshi-a*, dont nous avons étudié la fonction potentielle et l'évolution au cours du chapitre 3.

Nous avions terminé le chapitre 2 sur la conclusion que la bande blanche du poisson-clown adulte est principalement constituée d'iridophores. Cette bande blanche, ou plutôt *ces* bandes blanches, ne sont pas présentes au début du développement et se mettent en place essentiellement au cours d'une courte période développementale, correspondant à une métamorphose. Dans le chapitre qui suit, je décris les résultats obtenus au cours d'une étude transcriptomique menée sur les différents stades développementaux du poisson-clown.

Ces résultats sont une contribution "transcriptomique" à une étude qui inclura également des aspects pharmacologiques encore en cours de réalisation, menés par Natacha Roux, doctorant à l'OOB. Cette étude n'étant pas, à ce stade, assez avancée pour la rédaction d'un article intégratif sur le développement du poisson-clown, je n'ai pas écrit la suite sous format d'article, mais comme une partie qui est, autant que possible, "chrono – logique".

FONDEMENT SCIENTIFIQUE DU PROJET

Ce projet a été initié suite à plusieurs constats, plus ou moins indépendants mais reliés au développement du poisson-clown, à partir desquels ont émergé de multiples questions.

Constat 1 : Au cours du développement d'un organisme, de nombreux changements anatomiques, histologiques et physiologiques ont lieu. Ceux-ci sont gouvernés par une interaction entre des facteurs internes et externes à l'organisme, qui se traduisent *in fine* par des modifications de l'expression génétique. Ainsi, il est possible de retrouver une information temporelle dans des analyses d'expression transcriptomiques. Chez certaines espèces dont la détermination des stades est difficile morphologiquement, car trop semblables, l'outil transcriptomique peut être une alternative adaptée pour "stader" les individus (Anavy *et al.*, 2014)

➔ Question 1 : La définition des stades telle qu'effectuée sur la base de critères morphologiques et anatomiques chez le poisson-clown *A. ocellaris* se traduit-elle par une signature transcriptomique ? Est-il possible, en utilisant les résultats de RNA-Seq, d'inférer des marqueurs de stades de développement ?

Constat 2 : Chez de nombreux Vertébrés, comme évoqué en introduction, le développement est caractérisé par une métamorphose (Laudet, 2011). Celle-ci correspond le plus souvent à un pic d'hormones thyroïdiennes (T3 ou T4), ainsi que, parfois, à un pic d'expression des récepteurs aux hormones thyroïdiennes, les TR. Les TR sont multiples et ont des patrons d'expression variables au cours du développement d'une même espèce et entre différentes espèces. Par exemple, chez le Xénope, il y a un pic de TR-β au moment du pic de T4 et de la métamorphose. Chez le Cardeau, le pic de T4 correspond à un pic de TR-αA (**Introduction – Fig. 51**). Chez les poissons coralliens, il y a également, le plus souvent, une métamorphose. Cependant, le rôle des hormones thyroïdiennes au cours de cette importante transition d'histoire de vie chez ces organismes clefs des écosystèmes coralliens n'est aujourd'hui pas encore bien connu. En conditions naturelles, une seule étude a, à ma connaissance, pointé le rôle des hormones thyroïdiennes au cours du développement d'un poisson corallien, le poisson-chirurgien *Acanthurus triostegus* (Holzer *et al.*, 2017). Quelques travaux préliminaires menés en laboratoire sur des poissons-clowns *Amphiprion ocellaris* ont également montré l'influence de ces hormones sur la métamorphose (**Introduction – Fig. 54**).

➔ Question 2 : Dans quelle mesure le métabolisme thyroïdien est-il impliqué au cours du développement du poisson-clown *A. ocellaris* ? De manière générale, comment les différents processus développementaux se mettent-ils en place au cours du développement ?

Constat 3 : Au cours du développement du poisson-clown *Amphiprion ocellaris*, de nombreux changements anatomiques interviennent. En particulier, les changements pigmentaires permettent de définir une véritable métamorphose ayant lieu entre les stades développementaux 4 à 6 (**Introduction – Fig. 49**). Cette métamorphose correspond au développement de la première et de la deuxième bande blanche, constituées d'iridophores.

➔ Question 3 : Comment la voie génétique des marqueurs d'iridophores se déploit-elle au cours du développement du poisson-clown *A. ocellaris* ?

Constat 4 : Chez de nombreuses espèces, notamment de poissons Téléostéens, les changements pigmentaires sont perturbés par des traitements pharmacologiques ciblant la voie des hormones thyroïdiennes. C'est le cas par exemple du poisson-zèbre *Danio rerio* (McMenamin *et al.*, 2014) ou du poisson rouge *Carassius auratus* (Reddy and Lam, 1992).

➔ Question 4 : Les voies génétiques de la pigmentation et des hormones thyroïdiennes sont-elles reliées chez le poisson-clown *A. ocellaris* ?

PLAN DÉTAILLÉ DU CHAPITRE

APPROCHE EXPÉRIMENTALE

RÉSULTATS ET DISCUSSION

1. Correspondance partielle entre les stades morphologiques et les stades transcriptomiques

2. Utilisation des données de RNA-Seq pour l'obtention de marqueurs de stades

3. Caractérisation des processus développementaux globaux par les données de RNA-Seq

4. L'étude de la voie des TH indique une forte variabilité de l'expression des gènes et définit des possibles déterminants endogènes de la métamorphose

5. Les marqueurs d'iridophores semblent être plus exprimés au cours de la transition S4/post-S4

CONCLUSION

APPROCHE EXPÉRIMENTALE

Afin de répondre aux questions ci-dessus, nous avons utilisé une approche de RNA-Seq. Une fois les stades définis et la table de développement complétée (Roux et *al.*, in prep.), trois individus de chaque stade ont été prélevés et les ARN messagers ont été extraits sur larve totale. Les sept stades ainsi définis correspondent au premier stade après l'éclosion (premier stade larvaire) jusqu'au stade juvénile (**Introduction – Fig. 49**). Au total, nous avons donc obtenu 21 échantillons développementaux. Il est à noter que ces échantillons proviennent de deux pontes différentes. En effet, étant donnée la mortalité massive qui peut souvent être rencontrée, les individus ont été échantillonnés sur une première ponte jusqu'à ce qu'il n'y ait plus de survivant. Au cours d'une deuxième ponte, les individus n'ont pas été échantillonnés au début du développement mais plutôt à la fin, afin de maximiser l'obtention d'individus à des stades tardifs.

En biologie du développement, l'effet de la ponte, ou de la portée, peut rajouter un facteur de variation notable lorsqu'on souhaite comparer des échantillons. Nous avons toutefois échantillonné, pour trois stades différents, des individus des deux pontes. Des analyses postérieures sur des individus nous ont permis de montrer que ce biais était ici inexistant.

Certains des gènes de la voie des hormones thyroïdiennes ont une expression très localisée ; par exemple, la TSH est uniquement produite par les cellules de l'anté-hypophyse. Ainsi, afin de maximiser la probabilité de détecter une expression pour ces gènes faiblement exprimés, nous avons opté pour une importante profondeur de séquençage après extraction d'ARN sur larve totale.

Le séquençage ainsi que la préparation des banques ont été réalisés à l'IGBMC, à Strasbourg, sur un séquenceur Illumina HiSeq4000. Les ARNm ont été extraits (polyA uniquement) et la profondeur de séquençage était, par échantillon, d'environ 50-100 millions de reads paired-end de 100 nt **(Fig. 1**).

Sample name	Total reads	Sample name	Total reads
1S1	203,537,012	2S1	166,894,468
3S1	109,933,108	1S2	116,235,526
2S2	140,516,392	3S2	108,700,362
1S3	127,839,634	2S3	140,593,322
3S3	90,568,190	1S4	130,144,502
2S4	148,389,392	3S4	154,644,736
1S5	170,894,926	2S5	181,108,352
3S5	160,413,804	1S6	95,232,040
2S6	100,581,684	3S6	108,849,184
1S7	145,021,974	2S7	142,620,292
3S7	59,779,386		
Total reads 2,802M			

Sample name	% of bases above Q30	Sample name	% of bases above Q30
1S1	93.94	2S1	93.41
3S1	91.25	1S2	92.97
2S2	93.80	3S2	92.99
1S3	92.49	2S3	92.72
3S3	93.34	1S4	92.91
2S4	93.00	3S4	93.11
1S5	93.86	2S5	93.89
3S5	93.61	1S6	91.72
2S6	93.80	3S6	93.36
1S7	93.57	2S7	93.36
3S7	91.30		

Fig. 1. Profondeur et qualité du séquençage des échantillons développementaux *d'A. ocellaris*. Gauche : Nombre de reads par échantillon (reads paired-end, 100 nt). Droite : qualité du séquençage. Par exemple, les échantillons 1S1, 2S1 et 3S1 correspondent respectivement au séquençage des ARN des individus 1, 2 et 3 du stade 1. Trois individus ont été séquencés pour chacun des sept stades. Un score supérieur à Q30 signifie que la probabilité d'une erreur de séquençage est de moins de 0.1 %.

Nous avons ensuite utilisé ces reads pour assembler un transcriptome de référence pour cette espèce. À l'époque de l'étude, le génome d'*Amphiprion ocellaris* n'était pas publié et nous avons donc bénéficié de l'aide d'Anna Marcionetti et de Nicolas Salamin (UNIL), qui nous ont très tôt donné accès à leur assemblage en cours. Cela nous a donc permis de réaliser un assemblage de transcriptome "genome-guided". Nous avons également assemblé un transcriptome "*de novo*", puis combiné les deux transcriptomes pour obtenir un transcriptome de référence. Ce transcriptome a été annoté par BLASTN en utilisant le meilleur hit blast de l'annotation de la demoiselle *Stegastes partitus*, l'espèce séquencée la plus proche du poisson-clown au moment de l'étude (seuil utilisé : 1e-5). L'utilisation de multiples stades développementaux, couvrant l'essentiel du développement post-embryonnaire, ainsi qu'une bonne profondeur de séquençage, nous a permis d'obtenir un transcriptome contenant 22 713 gènes sur les 23 678 gènes de *S. partitus*. Cette valeur est peut-être légèrement sur-estimée et diminue si on augmente la stringence du BLAST pour l'annotation. On peut toutefois estimer que le transcriptome de référence ainsi obtenu correspond à la quasi-totalité des gènes du génome.

Cet outil nous a ensuite servi pour quantifier l'expression des différents gènes. Cela a été réalisé avec l'outil de pseudo-alignement *kallisto* ainsi qu'avec le package DESeq2, implémenté dans R (Love *et al.*, 2014; Bray *et al.*, 2016).

RÉSULTATS ET DISCUSSION

1. Correspondance partielle entre les stades morphologiques et les stades transcriptomiques

Pour chacun des 21 échantillons, nous avons obtenu les valeurs d'expression, en transcrits par million, pour chaque gène. Ainsi, chaque échantillon correspond à un vecteur de plus de 20 000 valeurs d'expression (une par gène). Afin de représenter l'ensemble de ces données, correspondant à plus de 21 x 20 000 dimensions, dans un seul graphique, on utilise une représentation dite en "composantes principales". Cette transformation mathématique rend les données lisibles dans un graphe en 2 ou 3 dimensions et réduit chacun des échantillons à un seul point. Deux points proches correspondent donc à deux échantillons qui partagent le même profil d'expression, c'est-à-dire des valeurs d'expression comparables pour l'ensemble des gènes.

L'ACP ainsi réalisée sur les 500 gènes les plus variables montre une séparation des stades le long de l'axe 1, majoritaire, expliquant à lui seul 62 % de la variance des données. Dans l'ensemble, les individus de S1 sont différents des individus de S2, eux-mêmes différents de S3, etc. (**Figure 2**). Cependant, pour certains stades, quelques points semblent

être différents des autres. Il en va ainsi au stade 1, où un des trois individus (correspondant au 3ème jour post-éclosion) est presque, d'un point de vue transcriptomique, un individu du stade 3. Au stade 4, de même, l'individu du jour 11 est proche des individus des stades 5 et 6. Enfin, au stade 6, un des trois individus, du jour 11 également, est regroupé avec les individus du stade 5 (**Figure 2**). Ainsi, afin d'avoir des échantillons aussi représentatifs de leurs stades que possible, nous avons par la suite exclu ces trois échantillons des analyses.

Une autre façon de représenter les données est d'utiliser non pas une ACP mais un "clustering hiérarchique" (**Figure 3**). Chaque échantillon est ainsi comparé à tous les autres, ce qui permet de calculer une distance, et les échantillons les plus proches sont regroupés ensemble dans la représentation correspondante. Une telle représentation nous indique qu'au niveau transcriptomique, le développement d'*A. ocellaris* correspond à trois phases (qu'on devinait déjà sur l'ACP) : (i) une phase avant le 4ème stade, qui regroupe de manière indistincte les trois premiers stades qui forment un ensemble assez homogène ; (ii) une courte phase, correspondant au stade 4, où les échantillons sont distincts de la phase antérieure ; (iii) une phase terminale, regroupant les stades 5 à 7, d'une homogénéité équivalente à celle de la phase (i). Cette subdivision visible est confirmée statistiquement par la méthode ($p < 0.05$, calcul par pvclust) (Suzuki and Shimodaira, 2006).

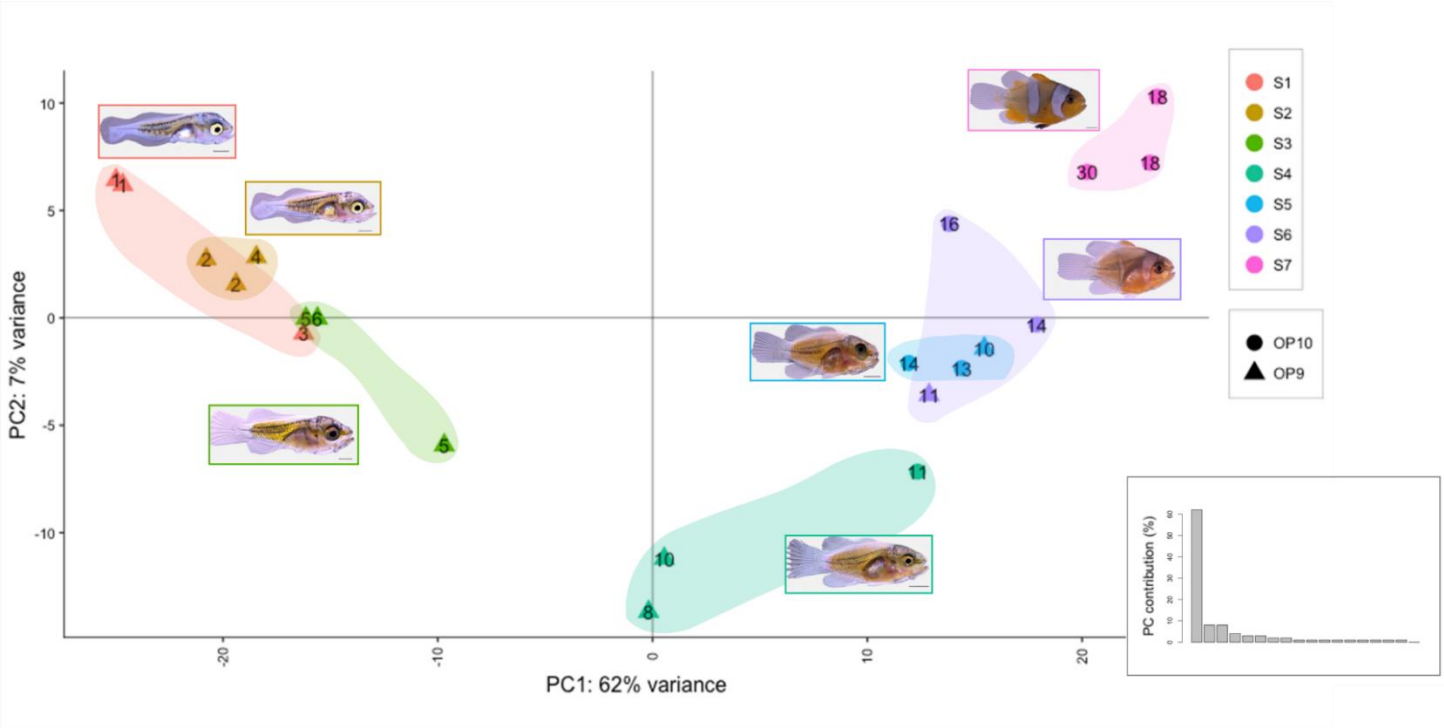


Figure 2. Analyse en composantes principales de l'expression des échantillons développementaux de poisson-clown *A. ocellaris*. Les points de même couleur correspondent aux individus d'un même stade. Les points de même forme correspondent aux individus issus de la même ponte (nommées OP9 ou OP10, pour "ocellaris ponte 9" ou "10"). Les nombres sur les points correspondent aux âges post-éclosion des individus (en jours). Le cartouche indique la relative contribution des différents axes de l'ACP. L'axe 1 est le principal axe expliquant les données (62%) et les autres axes n'expliquent qu'une faible proportion de la variance (moins de 7% pour chaque). Pour certains stades, quelques points semblent être différents des autres : au stade 1, un des trois individus (correspondant au 3ème jour post-éclosion) est presque, d'un point de vue transcriptomique, un individu du stade 3 ; au stade 4, l'individu de J11 est proche des individus des stades 5 et 6 ; au stade 6, un des trois individus, à J11 également, est regroupé avec les individus du stade 5.

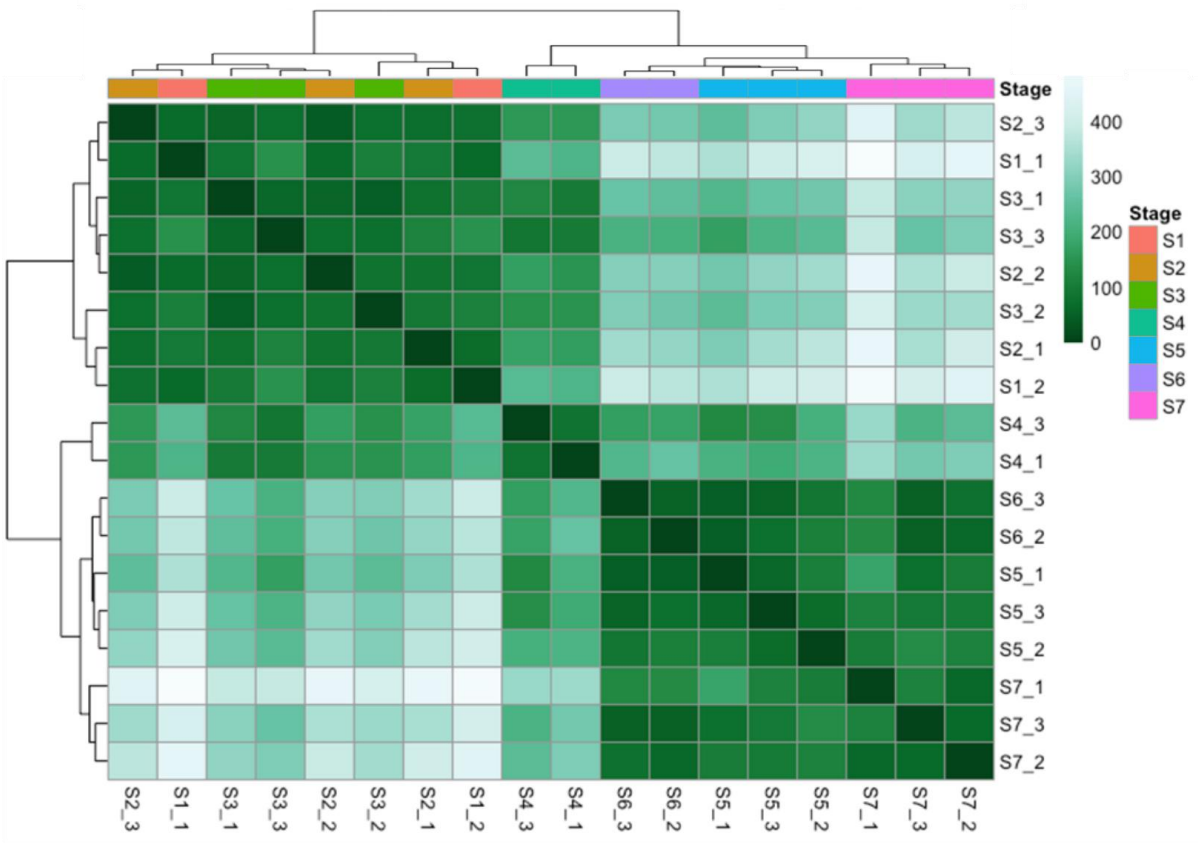


Figure 3. Clustering hiérarchique des échantillons développementaux d'*A. ocellaris*. Plus la distance est importante entre les échantillons (distance euclidienne), plus le carré est clair. Chaque ligne/colonne correspond à un échantillon comparé à tous les autres, dont lui-même. Par exemple, S2_3 correspond à l'individu 3 du stade 2.

2. Utilisation des données de RNA-Seq pour l'obtention de marqueurs de stades

Nous l'avons vu, il semblerait qu'au niveau des processus génétiques *globaux*, il soit plus pertinent de raisonner par groupe de stades (pré-S4 / S4 / post-S4) que par stade individuel, puisque, en particulier, les limites entre S1, S2 et S3 d'une part, ainsi qu'entre S4, S5 et S6 d'autre part, sont ténues.

Toutefois, la détermination de marqueurs génétiques *individuels* spécifiques de chaque stade peut s'avérer être intéressante. En particulier, dans des cas litigieux où la détermination du stade sur la base de seuls critères anatomiques est délicate, l'utilisation d'un marqueur génétique de stade fiable pourrait permettre de s'assurer du stade de l'individu testé. Cela pourrait en particulier être intéressant lors de traitements pharmacologiques perturbant le développement des individus traités, de telle sorte que leurs caractéristiques ne correspondent plus exactement à la table de développement. Par exemple, en utilisant un marqueur génétique et en mesurant son expression par qPCR, il serait possible de caractériser le stade d'un individu traité aux hormones thyroïdiennes, qui aurait un développement modifié et des défauts phénotypiques visibles.

Dans l'idéal, un marqueur du stade *i* correspondrait à un gène surexprimé au stade *i* par rapport à tous les autres stades. Par "le gène *g* est surexprimé", j'entends "toutes les comparaisons 2 à 2 entre le stade *i* et n'importe quel stade $j \neq i$ donnent le gène *g* surexprimé dans le stade *i* avec un log-fold change > 2 et une p-value < 0. 05". Par exemple, le gène *c21orf33,* un gène assez mal caractérisé mais probablement impliqué dans la différentiation des érythrocytes, est un marqueur spécifique du stade 4 (Stachorski *et al.*, 2014). De même, *rpl3l*, un gène codant pour une protéine ribosomale, est un marqueur spécifique du stade 7 (**Figure 4**). Pour le stade 4, cinq gènes ont un profil similaire ; pour le stade 7, quatre.

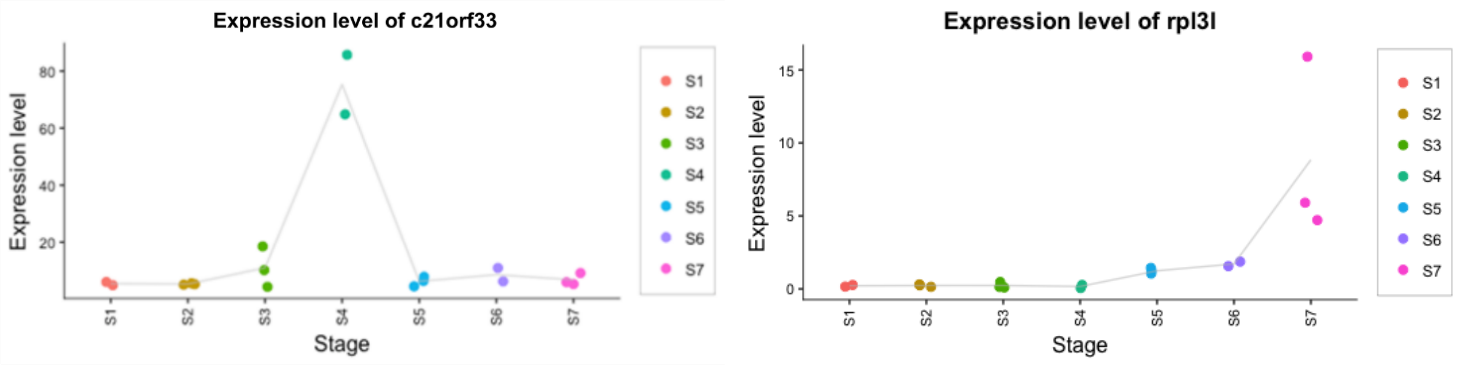


Figure 4. *c21orf33* et *rpl3l* sont des marqueurs respectifs des stades 4 et 7. Le niveau d'expression est en transcrits par million (TPM) et diffère en valeur absolue selon les graphes. La courbe grise relie les valeurs moyennes pour chaque stade.

Cependant, les stades 4 et 7 sont les seuls pour lesquels on observe des gènes avec un tel profil d'expression. Nous n'avons pas trouvé de gène qui aurait un profil comparable dans les stades 1 à 3 et aux stades 5 et 6.

En l'absence de gène ayant le profil "idéal", les marqueurs de stades doivent être choisis différemment. Nous proposons ici d'utiliser une combinaison de quelques gènes marqueurs qui permettraient, mesurés en simultané, de distinguer, d'une part, tous les individus des stades 1 à 4, et, d'autre part, tous les individus des stades 4 à 7.

Par exemple, mesurer en simultané les niveaux d'expression des trois gènes présentés sur la figure 5, dont le marqueur *c21orf33*, permettrait d'assigner un individu à un stade (**Figure 5**). En effet, avec le premier gène codant pour la Testisine, protéine impliquée dans la spermatogenèse (Scarman *et al.*, 2001), il est possible de distinguer les stades [1 et 3] (puisque ce marqueur est plus exprimé dans le stade 1 que dans le stade 3), [1 et 4], [2 et 3] et [2 et 4]. Le profil du gène *cpa1*, une exopeptidase pancréatique, permet, quant à lui de distinguer [1 et 2]. Enfin, *c21orf33* permet de distinguer [1 et 4]. La mesure du niveau d'expression de ces trois gènes permet donc de caractériser le stade d'un individu.

De la même manière, nous avons pu montrer que les trois gènes codant pour la Troponine T, ATP4B et la Carbonyl-réductase NADPH (type I) peuvent permettre d'assigner un individu aux stades 4, 5, 6 ou 7 (non montré ici). La validité et la reproductibilité des différents candidats devra être testée en qRT-PCR.

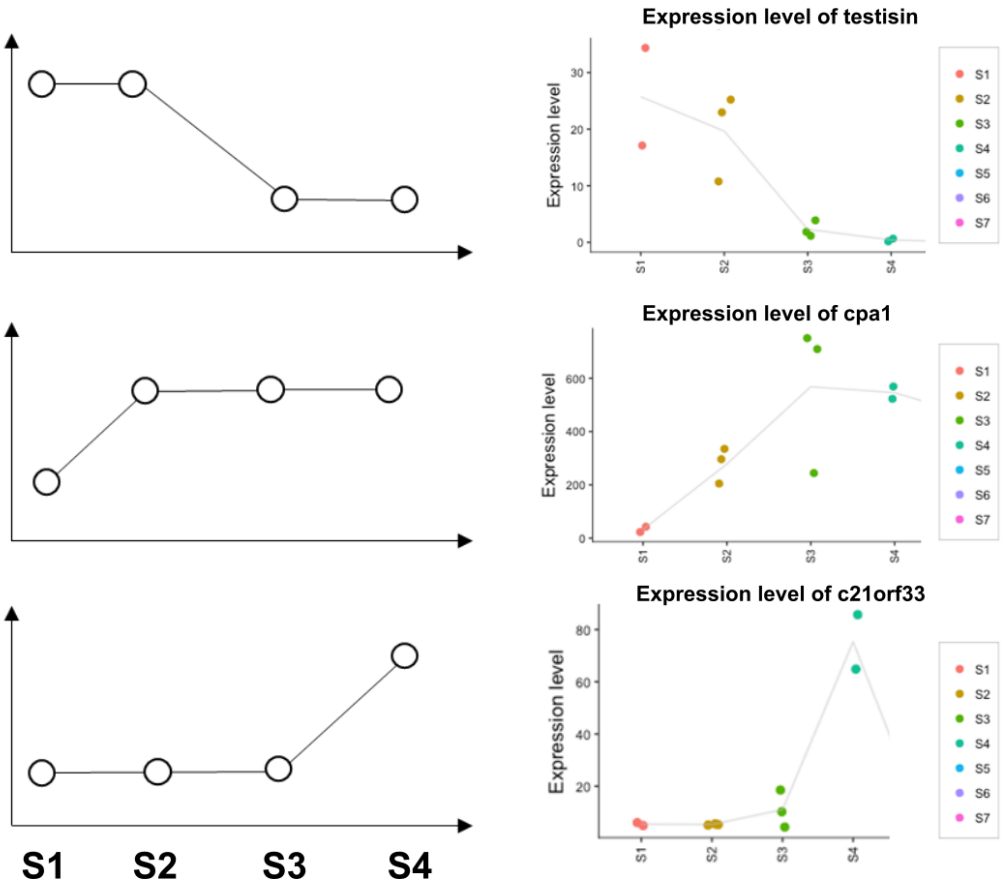


Figure 5. La mesure simultanée de trois gènes en qPCR pourrait permettre d'assigner les individus à un stade. Avec le premier gène codant pour la Testisine, il est possible de distinguer les stades [1 et 3] (puisque ce gène est surexprimé dans le stade 1), [1 et 4], [2 et 3] et [2 et 4]. Le profil du gène *cpa1*, une exopeptidase pancréatique, permet, quant à lui de distinguer [1 et 2]. Enfin, *c21orf33* permet de distinguer [1 et 4].

3. Caractérisation des processus développementaux globaux par les données de RNA-Seq

Les données de RNA-Seq permettent de diviser le développement du poisson-clown *A. ocellaris* en trois grands ensembles : avant le stade S4, le stade S4, et après le stade S4. L'étude de processus développementaux majeurs, impliquant des gènes multiples et des réseaux de régulation complexes, passe donc par des comparaisons de ces différents groupes de stades. Les stades *per se* sont trop similaires pour être comparés les uns aux autres.

Nous avons donc comparé deux à deux les trois différents groupes ainsi définis. Nous avons obtenu des listes de gènes différentiellement exprimés et avons tâché d'identifier les processus globaux à l'œuvre à chaque transition (pré-S4/S4 et S4/post-S4) et entre le début et la fin du développement post-embryonnaire (pré-S4/post-S4) (**Figure 6**). Ces processus correspondent aux enrichissements en termes de Gene Ontology et de KEGG pathways des listes de gènes obtenues. Ils ont été déterminés avec l'outil g:Profiler (Reimand *et al.*, 2016).

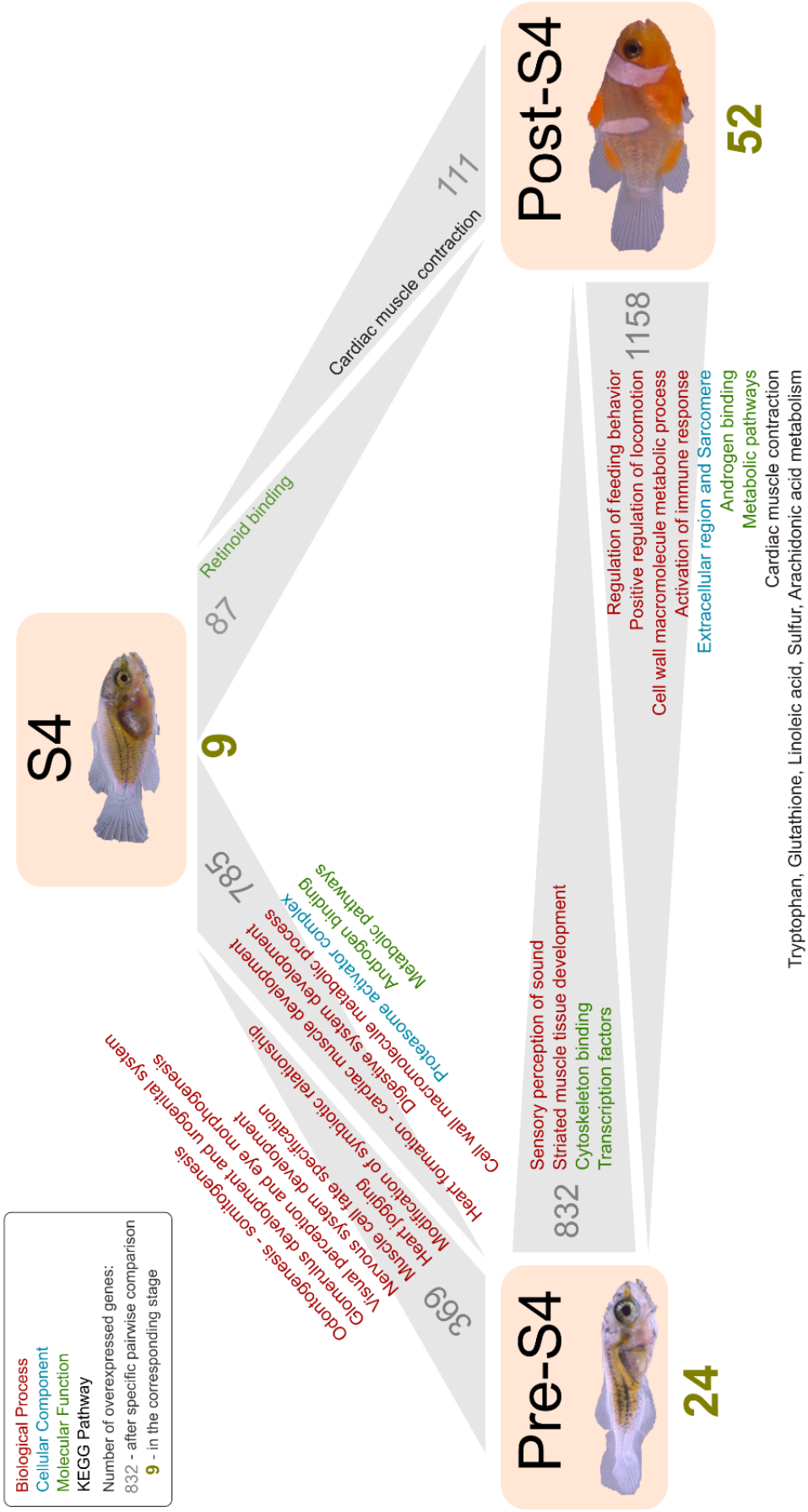
Il ressort en premier lieu que les groupes pré-S4 et S4 sont plus différents entre eux, au niveau du nombre de gènes différentiellement exprimés lors d'une comparaison 2 à 2, que les stades S4 et post-S4. Le plus grand nombre de gènes différentiellement exprimés est observé, comme attendu, entre les groupes pré-S4 et post-S4. De la même manière, de nombreux processus biologiques changent entre pré-S4 et S4, et beaucoup moins entre S4 et post-S4. Ainsi, il

semblerait qu'au niveau transcriptomique le développement d'*A. ocellaris* soit marqué par une distinction majeure entre pré-S4 et S4, et par une distinction plus faible entre S4 et post-S4.

Les processus enrichis impliqués sont très divers (**Figure 6**). Les gènes impliqués dans la liaison des androgènes, probablement importants pour la future différentiation du juvénile en mâle, sont exprimés dès le stade S4, donc relativement tôt, alors que la maturité sexuelle a lieu au bout d'un an ou deux chez un poisson-clown. Alors que le système visuel et urogénital se met en place dès avant le stade S4, le système digestif est lui plus tardivement mis en place, plutôt vers le stade S4 (**Figure 6**). Enfin, le terme "modification of symbiotic relationship" apparaît également enrichi en pré-S4, ce qui pourrait être intéressant au regard des relations symbiotiques entre poisson-clown et anémone. Ce terme correspond toutefois à deux gènes seulement, *actinoporin-like*, et *toll-like receptor 22*. À ma connaissance, aucun des deux n'a jusqu'ici été mis en évidence dans la relation symbiotique entre poisson-clown et anémone (même si le premier est un gène ayant une forte identité avec les gènes d'actinoporines qui constituent les pores lytiques des anémones (Črnigoj Kristan *et al.*, 2009)). De plus, la relation symbiotique se met en place bien plus tard, vers les stades 4 à 6. Si toutefois ces gènes étaient impliqués dans la mise en place de la symbiose, il serait donc étonnant de les voir exprimés si tôt au cours du développement.

Dans les listes de gènes obtenues, nous n'avons en particulier pas identifié de gène de la voie des hormones thyroïdiennes ou de la pigmentation.

Figure 6. PAGE SUIVANTE Analyses des différences d'expression entre paires de stades ou de groupes de stades. Le nombre de gènes différentiellement exprimés dans chacune des comparaisons 2 à 2 est indiqué (en gris). Par exemple, il y a 369 gènes surexprimés dans le groupe pré-S4 par rapport au stade 4. Le nombre de gènes spécifiques de chaque groupe de stades est également indiqué (en vert). Par exemple, il y a 9 gènes surexprimés dans S4 par rapport à pré-S4 et par rapport à post-S4. Les processus biologiques enrichis dans les gènes différentiellement exprimés sont représentés et la couleur du texte correspond au différent type de terme de Gene Ontology (ou KEGG) indiqué dans la légende.


Biological Process
Cellular Component
Molecular Function
KEGG Pathway
Number of overexpressed genes:
832 - after specific pairwise comparison
9 - in the corresponding stage
S4
9
785
Odontogenesis - somitogenesis
Glomerulus development and eye and morphogenesis
Visual perception and urogenital system
Nervous system development
Muscle cell fate specification
Heart jogging
369
Modification of symbiotic relationship
Digestive system development
Cell wall macromolecule metabolic process
Heart formation - cardiac muscle development
Androgen binding
Proteasome activator complex
Metabolic pathways
87
Retinoid binding
111
Cardiac muscle contraction
Pre-S4
24
832
Sensory perception of sound
Striated muscle tissue development
Cytoskeleton binding
Transcription factors
Post-S4
52
1158
Regulation of feeding behavior
Positive regulation of locomotion
Cell wall macromolecule metabolic process
Activation of immune response
Extracellular region and Sarcomere
Androgen binding
Metabolic pathways
Cardiac muscle contraction
Tryptophan, Glutathione, Linoleic acid, Sulfur, Arachidonic acid metabolism

4. L'étude de la voie des TH indique une forte variabilité de l'expression des gènes et définit des possibles déterminants endogènes de la métamorphose

Après des études "génome-wide", non-orientées vers un groupe de gènes particulier, nous avons ensuite étudié l'expression de gènes spécifiques de certaines voies biologiques. En particulier, nous avons tout d'abord étudié la voie des hormones thyroïdiennes.

Cette voie biologique est complexe et constituée d'un ensemble de gènes impliqués dans diverses étapes de la voie (**Introduction – Fig. 52**). On peut, pour résumer, distinguer (i) les gènes de différentiation de la glande thyroïdienne (qui correspond plutôt à un ensemble de follicules qu'à une vraie glande chez les poissons Téléostéens), (ii) les gènes régulant la synthèse des TH, (iii) les gènes de biosynthèse des TH, dont *tg* et *tpo* (**Introduction – Fig. 51**), (iv) les gènes de la voie en "aval" des TH, et (v) les gènes précoces de réponse.

Il n'est pas pertinent de montrer ici les profils d'expression individuels de chacun des gènes de la voie (ils sont environ une trentaine). Je choisis d'en montrer seulement certains, pour différentes raisons que j'explicite.

En premier lieu, lors d'une étude sur la métamorphose chez les Vertébrés, l'expression des récepteurs aux hormones thyroïdiennes (TR) est souvent mesurée. Parmi les trois TR, les deux TRα ont des profils comparables : une expression variable entre S1 et S3, une augmentation à partir de S4 jusqu'à un pic à S6, et une baisse à S7 (**Figure 7**). L'augmentation n'est pas très importante et il faudrait un plus grand nombre d'échantillons pour s'en assurer, mais elle est notable. On remarque en particulier que l'ordre relatif des échantillons est conservé entre TRα-A et pour TRα-B. Par exemple, l'individu à J14 du stade 5 est celui qui a la plus forte expression à la fois pour TRα-A et pour TRα-B. De même, l'individu à J13 a, pour les deux gènes, un niveau d'expression intermédiaire entre l'individu à J14 et à J10. Au stade 6, l'ordre des deux échantillons est également conservé pour les deux TRα. Ainsi, un individu qui a une forte expression de TRα-A a également une forte expression de TRα-B.

TR-β a un profil légèrement différent : de stable jusqu'au stade S4, son niveau d'expression augmente sans discontinuer dans les stades suivants (**Figure 7**).

On note que les taux *d'hormones* n'ont, à l'heure où j'écris ces lignes, pas encore été mesurés chez le poisson-clown, et il n'est donc pas possible de les corréler à l'expression des TR.

Chez le poisson-clown *A. ocellaris,* le climax de la métamorphose a lieu entre les stade S4 et S6 (**Introduction – Fig. 49**), ce qui pourrait correspondre au léger "pic" observé pour TRα-A et pour TRα-B ainsi qu'à l'augmentation de TR-β. Chez le poisson-chirurgien *Acanthurus triostegus*, un autre poisson corallien, le pic d'expression des TR a lieu au large, avant le recrutement larvaire et avant le climax de la métamorphose (Holzer *et al.*, 2017). Chez le poisson-clown, si le pic devait avoir lieu avant le recrutement, il devrait être avant les stades S4 à S6, puisque les individus recrutent environ à ce moment-là. Cela n'est pas observé dans les profils des TR, suggérant une variabilité de ce processus au sein des poissons coralliens.

Je finirai en précisant que tout cela reste à confirmer par qPCR sur un plus grand nombre d'échantillons, puisque ces résultats de RNA-Seq correspondent à des populations de n = 2 à n = 3 échantillons.

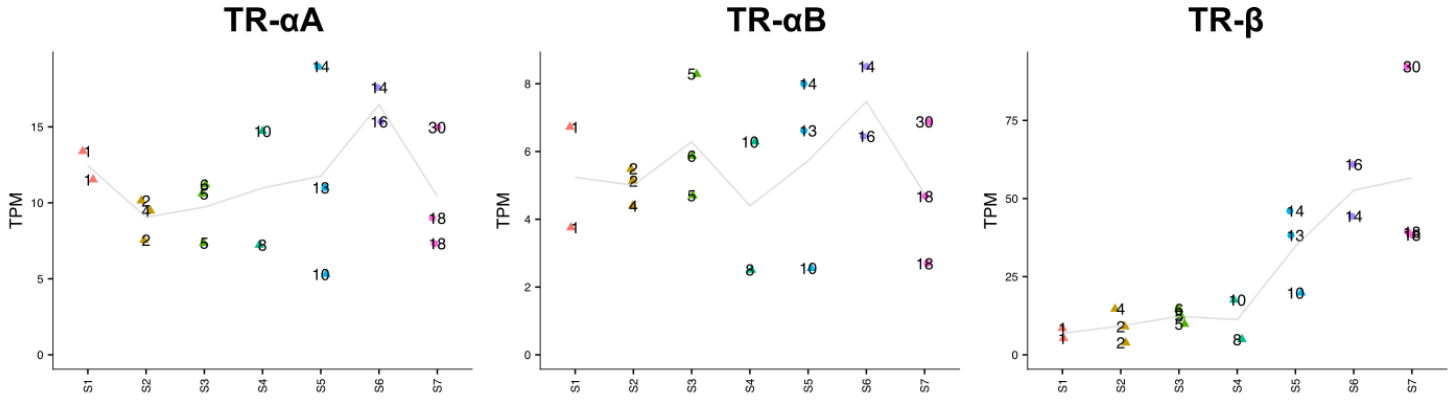


Figure 7. Profils d'expression des récepteurs aux hormones thyroïdiennes au cours du développement. Les jours de développement sont indiqués. Les deux TRα ont des profils comparables : une expression variable entre S1 et S3, une augmentation partir de S4 jusqu'à un pic à S6, et une baisse à S7. TR-β a un profil légèrement différent : de stable jusqu'au stade S4, son niveau d'expression augmente sans discontinuer dans les stades suivants.

Parmi les gènes de réponse aux hormones thyroïdiennes, *klf9* (codant pour un facteur de transcription impliqué dans la régulation de la différentiation cellulaire) est conservé chez les Vertébrés (Denver and Williamson, 2009; Holzer *et al.*, 2017). Cependant, son profil d'expression ne correspond pas, chez le poisson-clown, à un profil "typique" d'un gène de réponse, c'est-à-dire à un gène qui a un profil d'expression corrélé à celui des récepteurs (**Figure 8**). En revanche, *opn1lw*, autre gène de réponse aux hormones thyroïdiennes, codant pour une opsine à longue longueur d'onde, a un profil d'expression suggérant une activation potentielle par TR-β (Liu *et al.*, 2007). Cette expression tardive pourrait correspondre au changement d'habitat qui a lieu entre la larve pélagique, où la lumière bleue (courtes longueurs d'ondes) domine, et l'environnement récifal, où les longueurs d'ondes plus longues sont également présentes (**Introduction – Fig. 6**). De manière intéressante, l'opsine de courte longueur d'ondes *opn1sw* est elle exprimée selon une tendance opposée au cours du développement (**Figure 8**). Il a de plus déjà été montré que l'expression de cette opsine est inhibée par TR-β, suggérant une fois de plus l'implication de ce dernier (Bedolla and Torre, 2011).

Enfin, parmi l'ensemble des gènes de la voie des hormones thyroïdiennes, deux en particulier nous ont intrigué par leur profil d'expression : la thyroglobuline *tg* et la thyroïde péroxydase *tpo*. Ces deux gènes sont directement impliqués dans la synthèse d'hormone thyroïdienne au sein des follicules thyroïdiens (**Introduction – Fig. 51**). De manière intéressante, au sein de la même ponte, deux individus avaient des patrons opposés au stade 4 et un individu se démarquait clairement des deux autres au stade 5 **(Figure 9**). Ainsi, alors que l'individu S4 (échantillonné au jour 10) a une expression des deux gènes qui est maximale, l'autre individu a au contraire une expression minimale. (On note à cet égard, au vu des faibles valeurs de TPM pour *tpo*, l'importance d'avoir initialement choisi un séquençage avec une forte profondeur, sans lequel nous n'aurions pas pu détecter l'expression de *tpo*.)

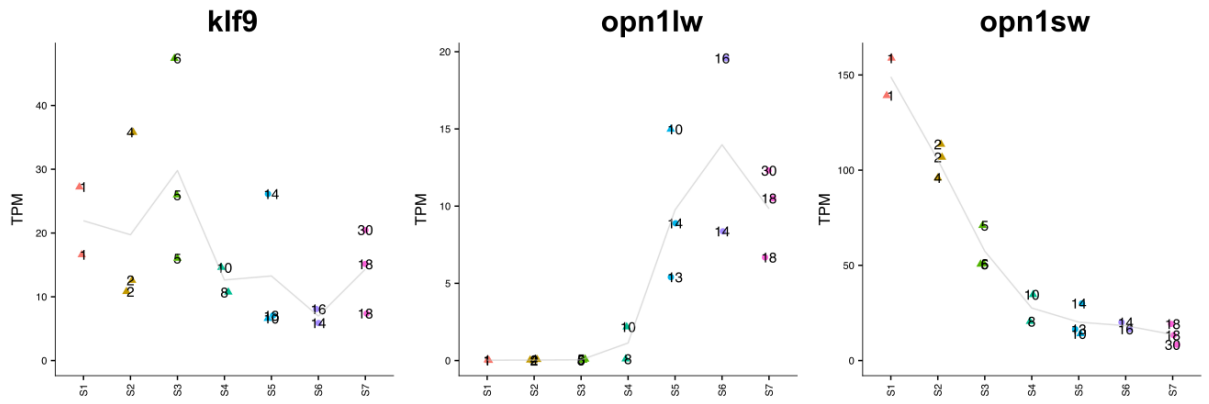


Figure 8. Profils d'expression de gènes de réponse aux hormones thyroïdiennes : *klf9* et *opn1lw*, deux gènes de réponse "positive", et *opn1sw*, un gène de réponse négative. *Klf9* n'a pas de manière évidente de réponse positive à la voie chez le poisson-clown. En revanche, les profils des deux opsines adaptées à des longueurs d'ondes différentes indiquent un contrôle par TR-β, positif pour l'opsine de grande longueur d'onde, et négatif pour celle de faible longueur d'onde. Cette corrélation reste cependant à tester, par exemple par des traitements pharmacologiques.

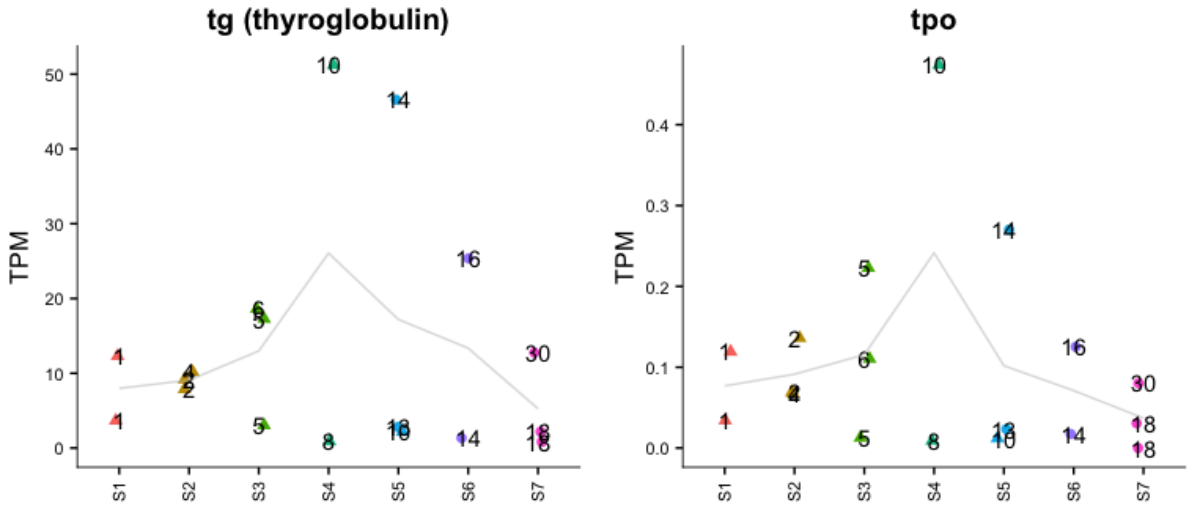


Figure 9. Profils d'expression de *tg* et de *tpo* au cours du développement. Ces deux gènes ont un profil d'expression comparable et montre curieusement une très forte variance au stade 4, indiquant que les deux individus du même stade sont pourtant à des états d'avancée différents dans la voie des TH.

Il semblerait donc que certains des individus soient, au S4, plus avancés que d'autres dans la voie des TH. Le rôle des facteurs environnementaux dans l'entrée ou non en métamorphose a été étudié à de multiples reprises. On sait également que des facteurs tels que la masse ou la taille peuvent influer sur l'entrée ou non en métamorphose. Cependant, le présent jeu de données pourrait permettre de répondre à une question qui est rarement considérée : quels sont les facteurs *endogènes* du déterminisme de la métamorphose ? En d'autres termes, quels sont les signaux internes qui induisent ou non l'entrée en métamorphose, indépendamment de l'environnement ? Une façon de répondre à cette question serait d'obtenir les gènes qui ont un patron d'expression similaire au patron d'expression de *tg* et *tpo*, c'est-à-dire l'ensemble des gènes qui ont une expression très forte dans un des individus et faible dans l'autre individu de S4. Trouver un sens à une telle liste de gène pourrait donc permettre de déterminer des facteurs génétiques en jeu dans le contrôle de l'entrée en métamorphose.

Afin d'identifier ces gènes, nous avons donc opté pour deux approches différentes. La première est une approche "stringente" : nous avons sélectionné tous les gènes pour lesquels l'expression est maximale dans l'individu à J10 du S4, et minimale dans celui à J8. La deuxième approche, plus permissive, a consisté à récupérer tous les gènes qui avaient la variance la plus élevée au stade 4 *et* une expression supérieure pour l'individu à J10 par rapport à celui à J8.

La première approche nous a fourni une liste de 21 gènes (**Table 1**). Parmi ceux-ci, il ne ressort pas clairement de gènes, qui, à ma connaissance, soient impliqués dans un processus développemental ou une voie métabolique particulière qui ait un sens. Cette liste a été envoyée à un expert de la voie des hormones thyroïdiennes, Peter Arvan, Professeur à l'Université du Michigan. Ce dernier a suggéré que ces gènes pourraient être des candidats plausibles de morphogenèse de la glande thyroïde. Cela reste toutefois à mieux caractériser.

Gene symbol	Gene annotation
ccdc181	coiled-coil domain containing 181
dmd	dystrophin (dmd), partial mRNA
efhc1	EF-hand domain (C-terminal) containing 1
foxj1	forkhead box J1
gnl1	guanine nucleotide binding protein-like 1
ikzf4	IKAROS family zinc finger 4 (Eos)
klhl3	kelch-like family member 3
LOC103358044	uncharacterized LOC103358044
LOC103360707	uncharacterized LOC103360707
LOC103361125	cytosolic carboxypeptidase 2-like
LOC103368470	dihydropyridine-sensitive L-type skeletal muscle calcium channel subunit alpha-1-like
LOC103368825	mucin-17-like
LOC103370561	myosin heavy chain, fast skeletal muscle-like
LOC103371215	dystrophin-like
LOC103375577	protocadherin beta-1-like
lrrk2	leucine-rich repeat kinase 2
nfkb1	nuclear factor of kappa light polypeptide gene enhancer in B-cells 1
ppp1r42	protein phosphatase 1, regulatory subunit 42
ppp2r3a	protein phosphatase 2, regulatory subunit B", alpha
usp28	ubiquitin specific peptidase 28
wfikkn1	WAP, follistatin/kazal, immunoglobulin, kunitz and netrin domain containing 1

Table 1. Liste des gènes qui ont un profil d'expression extrémal au stade 4, un individu (à jour 10) ayant la plus forte expression du gène, et l'autre (à jour 8) ayant la plus faible.

Afin d'élargir le spectre de gènes au profil comparable à *tg* et *tpo* et d'identifier une tendance globale (difficile à déceler avec les 21 gènes ci-dessus), nous avons recherché l'ensemble des gènes ayant la variance maximale au stade 4. Nous nous sommes également assurés que ces gènes avaient une expression de l'individu à J10 plus importante que celle de celui à J8. Nous avons ainsi obtenu une liste de 364 gènes (qui incluait les 21 gènes précédents). Cette liste étant trop longue pour être retranscrite ici, nous avons analysé les termes de Gene Ontology qui ressortaient enrichis dans cette liste de gènes (**Table 2**). Si de nombreux termes sont statistiquement enrichis, il est toutefois difficile de trouver un sens biologique à ceux-ci. Les ubiquitine ligases semblent être particulièrement représentées parmi ces gènes. Ces protéines n'ont, à ma connaissance, pas particulièrement été identifiées au cours de la métamorphose des Vertébrés. Elles jouent cependant un rôle dans la modification des réseaux synaptiques au cours de la métamorphose de la drosophile (Kuo *et al.*, 2006).

	Process	p-value
BP	protein modification by small protein conjugation or removal	0,00308
BP	cellular process	0,00983
BP	regulation of metal ion transport	0,0154
BP	somatic muscle development	0,0192
BP	intracellular signal transduction	0,0192
BP	calcium-mediated signaling	0,0192
BP	organelle organization	0,0236
BP	regulation of receptor activity	0,0374
BP	rhodopsin mediated signaling pathway	0,0374
BP	regulation of release of sequestered calcium ion into cytosol by sarcoplasmic reticulum	0,0374
BP	biological regulation	0,0376
BP	regulation of cellular localization	0,0376
BP	retina homeostasis	0,0406
CC	SCF ubiquitin ligase complex	0,0376
CC	ubiquitin ligase complex	0,0383
CC	intracellular	0,048
MF	ubiquitin-like protein ligase activity	0,00983
MF	ion channel binding	0,0295
MF	transcription regulatory region DNA binding	0,0395
MF	enzyme binding	0,0479
keg	Ubiquitin mediated proteolysis	0,00167

Table 2. Analyses des termes de Gene Ontology et KEGG enrichis dans la liste de gènes ayant la plus forte variance d'expression au stade 4. BP: Biological Process. CC: Cellular Component. MF: Molecular Function. Keg: KEGG pathway.

5. Les marqueurs d'iridophores semblent être plus exprimés au cours de la transition S4/post-S4

Au cours du développement du poisson-clown, les bandes blanches, constituées d'iridophores, se mettent en place à partir du stade 4 (**Introduction – Fig. 49**). Nous avons donc regardé si l'expression des gènes marqueurs d'iridophores identifiés dans la littérature suivait également cette tendance. Ces gènes sont au nombre de 27 chez les poissons et il est possible de regarder individuellement leurs profils d'expression (**Figure 10**). Au premier abord, il semble que leur expression soit très variable d'un stade à l'autre et qu'aucune tendance claire ne se dégage. Toutefois, en regardant les graphes plus en détail, on voit que de nombreux gènes (15 sur les 27) ont un profil d'expression qui correspond à une augmentation de l'expression entre le stade 4 et les stades 5 et 6 (**Figure 10**).

Notons que cette augmentation n'est pas nécessairement significative (elle ne l'est d'ailleurs, le plus souvent, pas), mais révèle une tendance de l'ensemble des gènes de cette voie.

Nous avons donc voulu tester si, par rapport à l'ensemble des gènes dont l'expression varie au cours du développement, les gènes considérés comme des marqueurs "classiques" d'iridophores sont statistiquement plus souvent "montant" entre les stades 4 et 6 qu'attendu par hasard pour un gène "moyen" du génome. (*Je considère ici uniquement les marqueurs "classiques" et pas les gènes néo-caractérisés au cours du chapitre 2, dont* roshia. *Je traite de ce dernier plus loin.*)

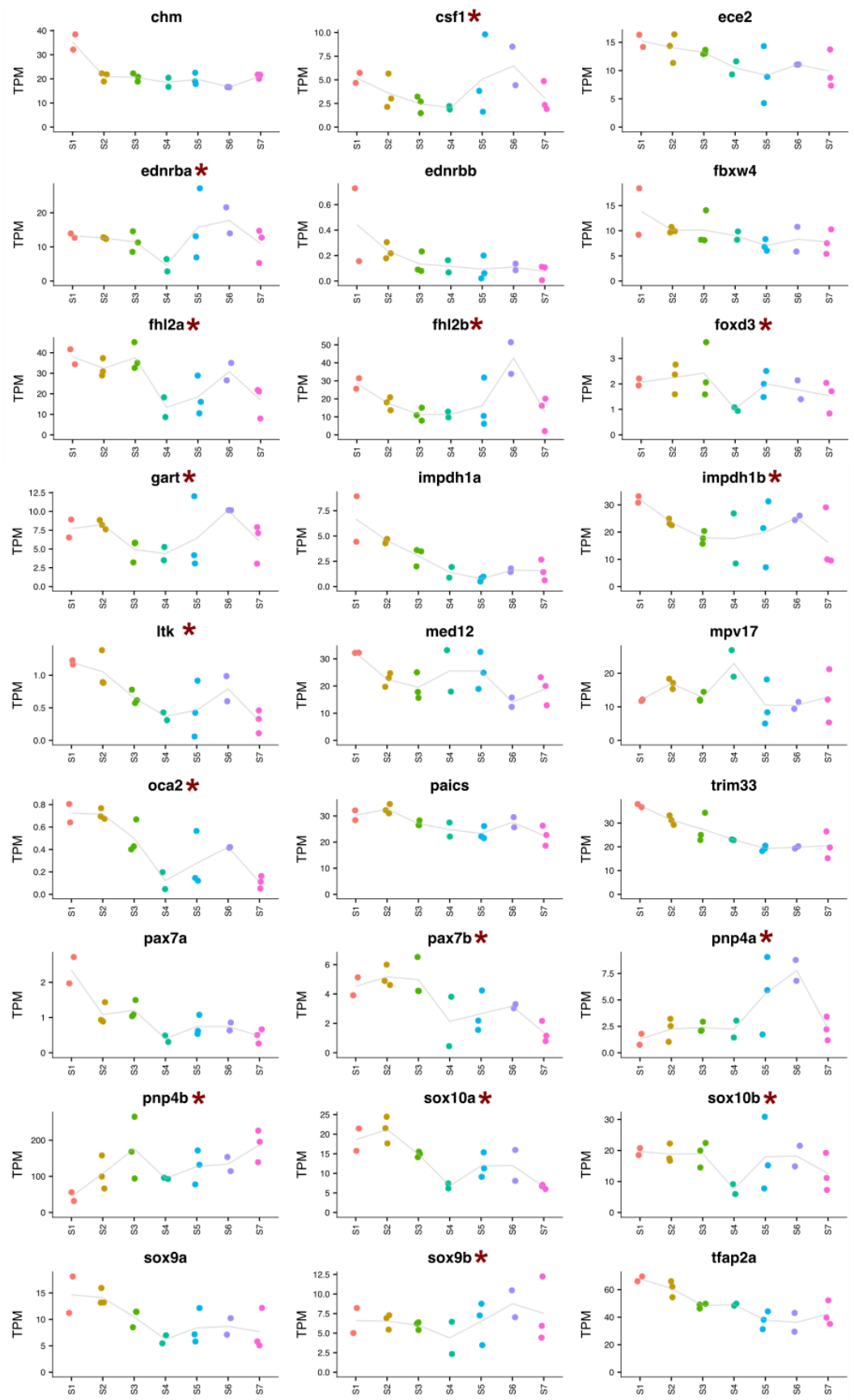


Figure 10. Profils d'expression des gènes marqueurs d'iridophores au cours du développement de *A. ocellaris*. Les gènes indiqués par une étoile (15 au total) sont ceux dont l'expression augmente en moyenne entre les stades 4 et 6.

Afin de répondre à cette question, nous avons utilisé une approche de clustering utilisant la méthode WGCNA, pour Weighted Gene Correlation Network Analysis (Langfelder and Horvath, 2008). Ce clustering se fonde sur plusieurs hypothèses, qui sont que les gènes ayant des profils comparables ont plus de chance d'être co-régulés, impliqués dans la réalisation d'une même fonction ou dans la même voie de signalisation. Le principe est le suivant :

- Chaque gène a un profil d'expression qui lui est propre au cours du développement
- Les gènes qui ont des profils d'expression comparables sont regroupés au sein d'un même cluster
- Chaque cluster correspond donc à un profil d'expression moyen dont chacun des gènes du cluster peut s'écarter légèrement, mais pas trop, sinon il le gène est considéré comme étant dans un autre cluster. Ce profil moyen théorique est représenté par le "eigengene" du cluster.
- Il est alors possible de visualiser graphiquement les relations entre les clusters. Les gènes appartenant à des clusters proches seront plus proches entre eux que des gènes aux profils d'expression opposés. Par exemple, si A veut dire "l'expression **A**ugmente entre le stade *i* et le stade *i+1*", et D veut dire que l'expression **D**iminue, les gènes ayant un profil qui correspond à A-A-A-A-A-A seront plus proches de ceux qui ont un profil A-A-A-A-A-D que de ceux qui ont un profil D-D-D-D-D-D.

On note que ce processus comporte une part importante d'arbitraire. À partir de quand deux profils d'expression sont-ils assez différents pour être considérés comme appartenant à deux clusters différents ? À partir de quelle taille un groupe de gènes aux profils similaires peut-il être considéré comme un cluster ?

Ici, de manière arbitraire, nous avons considéré qu'il fallait qu'un ensemble de gènes soit au moins constitué de 200 gènes pour être considéré comme un cluster. De même, nous avons arbitrairement choisi un seul de dissimilarité dans R de 0.3 : au-dessus de ce seuil, les clusters étaient considérés comme différents.

Nous avons donc extrait les 10 000 gènes ayant les profils les plus variables au cours du développement, en se basant sur la mesure de la variance de l'ensemble des valeurs pour chacun des gènes. Nous avons ensuite utilisé le package WGCNA dans R. Avec de tels choix, nous avons obtenu un total de 10 clusters de gènes (**Figure 11**). Pour que la représentation soit lisible, nous n'avons pas représenté les 10 000 gènes mais seulement 500, qui sont représentatifs en proportion des 10 000 initiaux. Nous avons ensuite tracé un graphe (méthode ForceAtlas2) qui représente le réseau constitué par ces 500 gènes. Plus des nœuds sont proches, plus les profils d'expression sont comparables. Nous avons également réalisé ces analyses sur l'ensemble des gènes du génome : hormis un temps de calcul très long (plusieurs heures sur une machine de bureau), le profil du graphe était quasiment identique.

À l'opposé du graphe se trouvent les gènes qui appartiennent à des clusters dont les "eigengenes" (le gène "moyen" théorique du cluster) ont des profils opposés : soit à l'expression qui diminue tout le long, soit à l'expression qui augmente tout au long du développement. Entre les deux se situent des gènes dont l'expression varie de manière moins linéaire. C'est le cas des gènes des clusters sur la gauche de la figure, qui ont plutôt un profil en cloche : faible expression au cours du développement, pic au milieu, et faible expression à la fin du développement (Figure 9). Sur la droite de la figure, on retrouve plutôt les gènes qui ont un profil en "creux" : faible expression autour du stade 4 et forte expression avant et après. On note que dans l'ensemble l'essentiel des gènes (79 %) sont soit dans le cluster "descendant" ou dans le cluster "montant". Cela a été vérifié par l'obtention du graphe utilisant non pas les 10 000 gènes les plus variables mais tous les gènes : le même profil est observé, avec deux clusters principaux opposés et des clusters intermédiaires.

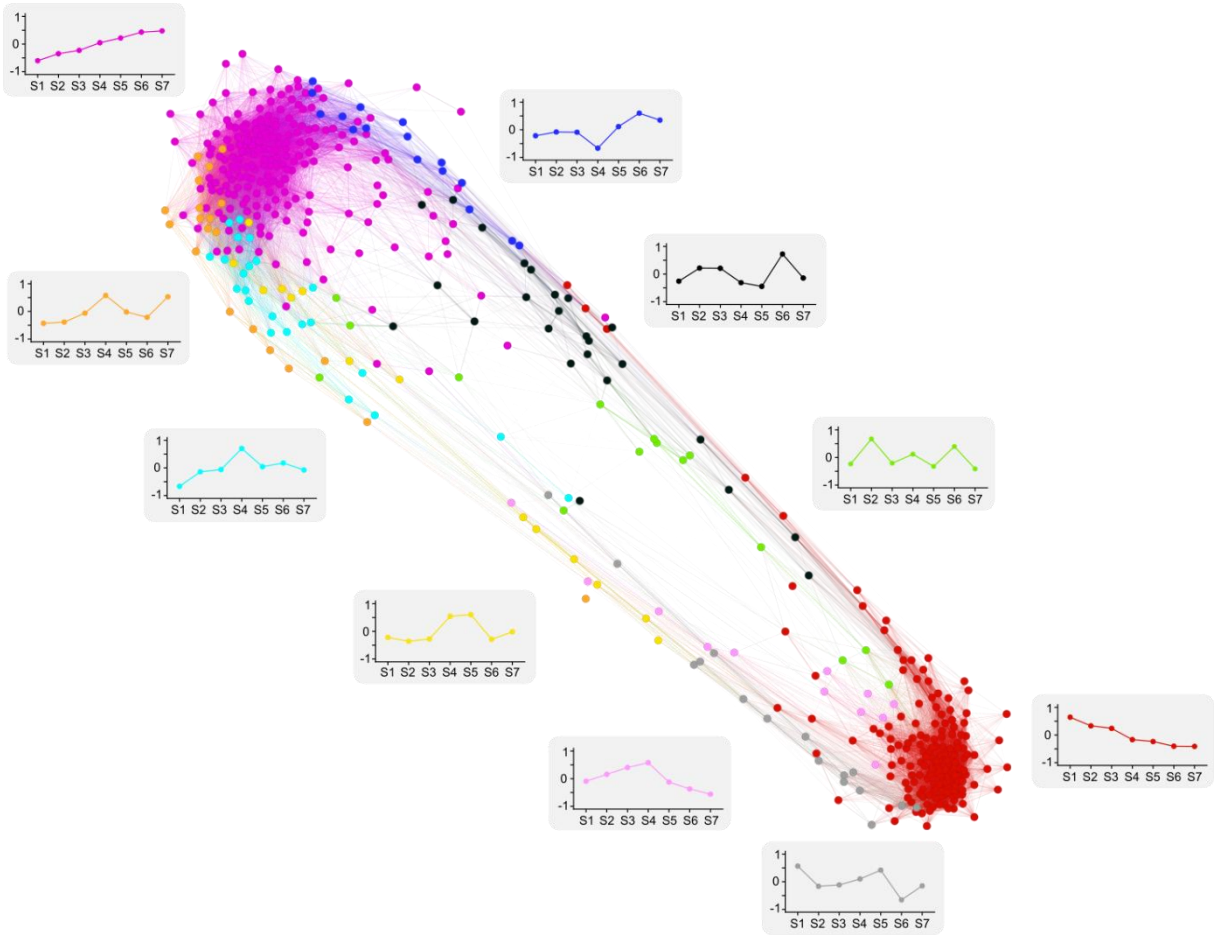


Figure 11. Réseau représentant les profils d'expression des gènes au cours du développement d'*A. ocellaris*. Chaque couleur correspond à un cluster. Pour chaque cluster, le profil d'expression de l'eigengene est indiqué par la couleur correspondante dans les cartouches en gris.

À ce réseau de 500 gènes nous avons ensuite ajouté les gènes marqueurs des iridophores. Après la sélection préalable des 10 000 gènes les plus variables, seulement 17 des 27 marqueurs initiaux ont été retenus. Certains, à l'instar de *fbxw4*, ont une expression quasi-linéaire au cours du développement et ne sont pas parmi les 10 000 les plus variables. Ces 17 marqueurs sont représentés sur le graphe de la **figure 12** (points blancs cerclés de noir).

De manière intéressante, 6 de ces marqueurs, correspondant à ce que nous observions sur la **figure 10**, sont souvent positionnés dans les clusters noir ou vert. On peut même en compter 7 si l'on inclut *fhl2a* – voire 11 si l'on est plus permissif et qu'on inclut également *pnp4a*, *sox9b, csf1* et *oca2* d'après leurs profils d'expression individuels de la **figure 10**.

Nous avons vu que 79 % des 10 000 gènes utilisés ici sont soit dans le cluster "descendant", soit dans le cluster "montant". Par hasard, on s'attendrait donc à ce que 79 % des gènes marqueurs d'iridophores soient dans un de ces deux clusters, soit 0.79 x 17 = 13,4 gènes. Dans une hypothèse restrictive comptant 6 gènes dans les clusters noir ou vert, 17 – 6 = 11 gènes sont également dans les clusters extrêmes. Dans une hypothèse permissive, qui est justifiable au vu des profils individuels, 17 – 11 = 6 gènes sont dans ce cas. On a alors une différence significative entre ces deux groupes de gènes, indiquant que les gènes marqueurs d'iridophores sont plus souvent représentés que la moyenne dans ces deux clusters noir et vert ($p = 0.02$, test de χ^2).

Que la différence soit significative ou pas, cette représentation permet de se rendre compte que de nombreux marqueurs d'iridophores ont un profil d'expression correspondant à une augmentation aux stades 5 et 6 par rapport au stade 4. De plus, aucun de ces gènes n'est représenté vers "le bas et le milieu de la branche gauche" du graphe (clusters gris, rose, jaune et cyan), correspondant aux gènes qui ont une expression croissante avant S4 puis décroissante après S4. Dans l'ensemble, l'expression des marqueurs d'iridophores est donc compatible avec le développement important des bandes blanches au cours du développement entre les stades 4 et 6.

On peut remarquer qu'un autre des gènes évoqués au-dessus a un profil d'expression faible jusqu'au stade 4 et en constante augmentation par la suite : TR-β. De même, les deux TRα ont une expression qui augmente entre les stades 4 et 6. Cela pourrait donc suggérer un contrôle des gènes de différentiation des iridophores par la voie des hormones thyroïdiennes. Cela reste évidemment à tester fonctionnellement, mais ce serait une première, à ma connaissance. En effet, jusqu'à présent, le rôle de ces hormones n'a été montré que dans la différentiation des mélanocytes et des xanthophores chez le Poisson-zèbre *D. rerio* (McMenamin *et al.*, 2014).

Autre point intéressant : sur la **figure 12** est également représenté le profil d'expression de *roshi-a*. Celui-ci est, nous l'avons montré, un probable nouveau marqueur d'iridophores. Son expression est compatible avec une augmentation entre S4 et S5/S6, et va de fait dans le sens de ce qui est avancé au-dessus (**Figure 12, cartouche en haut à droite**).

Les iridophores, chez le poisson-clown, se mettent en place massivement au cours de la métamorphose post-S4. On peut donc "s'étonner" de constater que de nombreux marqueurs d'iridophores sont exprimés de manière plus précoce, aux stades 1 à 3 (même si, ensuite, leur expression semble réaugmenter après le stade 4), et ont un profil d'expression en "creux" plutôt qu'en exponentielle : *gart*, *impdh1b, fhl2, roshi-a*... Cela pourrait correspondre, de manière un peu équivalente à ce qui est observé chez le Poisson-zèbre, à la présence de deux populations d'iridophores : des iridophores larvaires, qui seraient présents précocement au cours du développement, mais diffus et sans former de bandes, et des iridophores adultes, se mettant en place en même temps que les bandes (Singh and Nüsslein-Volhard, 2015).

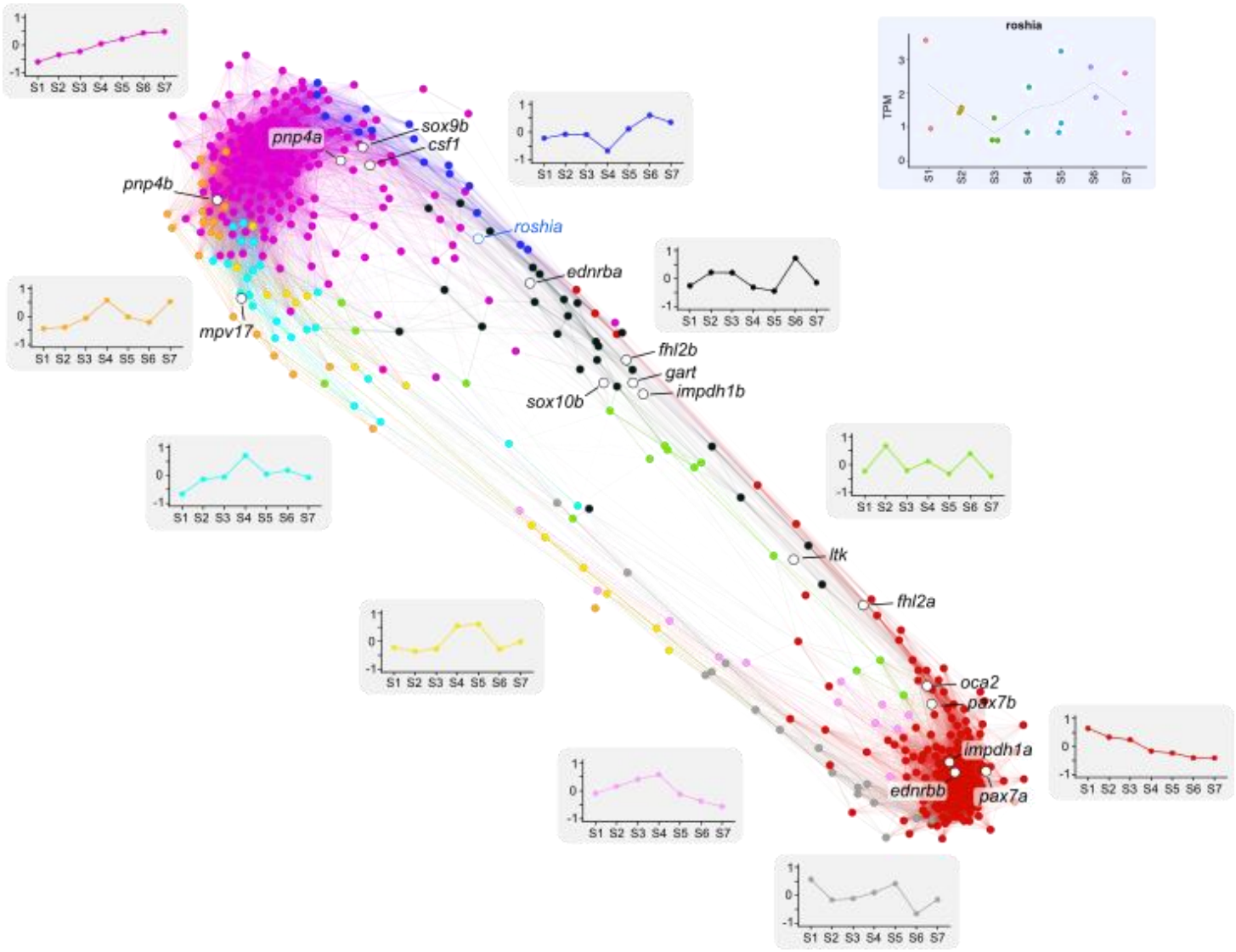


Figure 12. Réseau représentant les profils d'expression des marqueurs d'iridophores au cours du développement d'*A. ocellaris*. Les marqueurs reconnus dans la littérature sont les gènes en blanc cerclé de noir. Ils sont plus souvent qu'attendu par hasard dans la branche "droite" du graphe, c'est-à-dire avec un profil qui monte entre S4 et S5/S6. *Roshia* est également représenté et son profil individuel indiqué (cartouche en haut à droite).

CONCLUSION

Ce travail est toujours en cours et il faudra bien entendu des résultats fonctionnels complémentaires, ainsi que l'obtention d'une vue d'ensemble plus claire, pour parvenir à une ébauche d'article. On peut revenir, en conclusion, sur les questions posées initialement.

➔ *Question 1 : La définition des stades telle qu'effectuée sur la base de critères morphologiques et anatomiques se traduit-elle par des différences au niveau transcriptomique ? Est-il possible, en utilisant les résultats de RNA-Seq, d'inférer des marqueurs de stades ?*

Nous avons ici montré que le développement du poisson-clown *A. ocellaris* s'effectue selon des stades morphologiques dans l'ensemble compatibles avec les stades transcriptomiques. Toutefois, au niveau global, il semble que ce développement soit divisé en trois grandes étapes : avant le stade 4, le stade 4, et après le stade 4.

Nous avons cependant pu montrer que des gènes pouvaient individuellement servir de marqueurs de stades. Lorsqu'un seul gène ne suffit pas, une combinaison de trois gènes peut permettre de discriminer des individus de stades morphologiquement proches.

➔ *Question 2 : Dans quelle mesure le métabolisme thyroïdien est-il impliqué au cours du développement du poisson-clown* A. ocellaris *? De manière générale, comment les différents processus développementaux se mettent-ils en place au cours du développement ?*

Nous avons, au cours de ces travaux, montré que la voie des hormones thyroïdiennes semble être impliquée au cours du développement, selon des modalités exactes restant à préciser. Certains gènes cibles semblent avoir une expression corrélée ou anti-corrélée à celle de TR-β (*opn1lw*, *opn1sw*) alors que d'autres, habituellement régulés, ne semblent pas être influencés (*klf9*). Nous avons également mis en évidence de nombreux gènes qui pourraient être à l'origine du déterminisme endogène de l'entrée en métamorphose mais pour lesquels des analyses complémentaires seront nécessaires. Enfin, nous avons vu que c'est essentiellement entre les stades pré-S4 et S4 que de nombreux changements génétiques (ainsi que les processus biologiques qui leur sont associés) ont lieu.

➔ *Question 3 : Comment les voies génétiques de la pigmentation se déploient au cours du développement du poisson-clown* A. ocellaris *?*

Le développement de la bande blanche correspond à une prolifération importante des iridophores. Nous avons montré que les gènes marqueurs précédemment identifiés (chapitre 1), ainsi que nouvellement identifiés (*roshi-a*), ont des profils d'expression dans l'ensemble compatibles avec cette prolifération. Leur expression tôt au cours du développement embryonnaire pourrait indiquer qu'il existe deux types d'iridophores chez le poisson-clown, larvaires puis adultes, à l'instar de ce qui est observé chez le Poisson-zèbre *D. rerio*.

➔ *Question 4 : Comment les voies génétiques de la pigmentation et des hormones thyroïdiennes sont-elles reliées chez le poisson-clown* A. ocellaris *?*

Certains des gènes de la voie des hormones thyroïdiennes, en particulier TR-β, ont un profil d'expression indiquant une augmentation de l'expression après le stade développemental 4. Ce profil est corrélé à ceux de nombreux gènes de différentiation des iridophores. De plus, un lien causal entre développement des cellules pigmentaires et le métabolisme thyroïdien a déjà été montré chez le poisson-zèbre *D. rerio* (McMenamin *et al.*, 2014). Cependant, uniquement les xanthophores et mélanocytes étaient affectés. Cela suggèrerait donc ici, pour la première fois (à ma connaissance), un lien entre le développement des iridophores et le métabolisme thyroïdien. Il conviendrait maintenant de tester cela fonctionnellement.

RÉFÉRENCES

Anavy L, Levin M, Khair S, Nakanishi N, Fernandez-Valverde SL, Degnan BM, *et al.* (2014). BLIND ordering of large-scale transcriptomic developmental timecourses. *Development* **141**: 1161–1166.

Bedolla DE, Torre V (2011). A Component of Retinal Light Adaptation Mediated by the Thyroid Hormone Cascade (J-MA Lobaccaro, Ed.). *PLoS One* **6**: e26334.

Bray NL, Pimentel H, Melsted P, Pachter L (2016). Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**: 525–527.

Črnigoj Kristan K, Viero G, Dalla Serra M, Maček P, Anderluh G (2009). Molecular mechanism of pore formation by actinoporins. *Toxicon* **54**: 1125–1134.

Denver RJ, Williamson KE (2009). Identification of a thyroid hormone response element in the mouse *Krüppel-like factor 9* gene to explain its postnatal expression in the brain. *Endocrinology* **150**: 3935–3943.

Holzer G, Besson M, Lambert A, François L, Barth P, Gillet B, *et al.* (2017). Fish larval recruitment to reefs is a thyroid hormone-mediated metamorphosis sensitive to the pesticide chlorpyrifos. *Elife* **6**.

Kuo CT, Zhu S, Younger S, Jan LY, Jan YN (2006). Identification of E2/E3 Ubiquitinating Enzymes and Caspase Activity Regulating Drosophila Sensory Neuron Dendrite Pruning. *Neuron* **51**: 283–290.

Langfelder P, Horvath S (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**: 559.

Laudet V (2011). The Origins and Evolution of Vertebrate Metamorphosis. *Curr Biol* **21**: R726–R737.

Liu Y, Fu L, Chen D-G, Deeb SS (2007). Identification of novel retinal target genes of thyroid hormone in the human WERI cells by expression microarray analysis. *Vision Res* **47**: 2314–2326.

Love MI, Huber W, Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

McMenamin SK, Bain EJ, McCann AE, Patterson LB, Eom DS, Waller ZP, *et al.* (2014). Thyroid hormone-dependent adult pigment cell lineage and pattern in zebrafish. *Science (80-)* **345(602)**: 1358–1361.

Reddy PK, Lam TJ (1992). Effect of thyroid hormones on morphogenesis and growth of larvae and fry of telescopic-eye black goldfish, Carassius auratus. *Aquaculture* **107**: 383–394.

Reimand J, Arak T, Adler P, Kolberg L, Reisberg S, Peterson H, *et al.* (2016). g:Profiler — a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res* **44**: W83–W89.

Scarman AL, Hooper JD, Boucaut KJ, Sit ML, Webb GC, Normyle JF, *et al.* (2001). Organization and chromosomal localization of the murine Testisin gene encoding a serine protease temporally expressed during spermatogenesis. *Eur J Biochem* **268**: 1250–8.

Singh AP, Nüsslein-Volhard C (2015). Zebrafish Stripes as a Model for Vertebrate Colour Pattern Formation. *Curr Biol* **25**: R81–R92.

Stachorski L, Thangapandi V, Reinhardt D, Klusmann J-H (2014). Characterization of oncogenes on chromosome 21 identified by shRNA-based viability screening. *Mol Cell Pediatr* **1**: A20.

Suzuki R, Shimodaira H (2006). Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**: 1540–1542.

DISCUSSION GÉNÉRALE ET CONCLUSION

"Les hommes vont sur la Lune, ils envisagent de faire un voyage vers Mars dans le but d'en faire une planète habitée, mais le premier petit poisson venu reste une énigme."

Michel Onfray
Le désir ultramarin – Les Marquises après les Marquises

Résumé des travaux de thèse et annonce de la discussion

Au cours de ces travaux de thèse, j'ai (avec d'autres) tout d'abord caractérisé l'évolution des gènes de la pigmentation chez les Vertébrés, et en particulier chez les poissons Téléostéens, le groupe de Vertébrés probablement le plus divers en termes de pigmentation. J'ai notamment pu montrer que ces gènes appartiennent à des familles qui ne sont pas plus retenues que le reste du génome après la double duplication 1R/2R ayant eu lieu à la base des Vertébrés. Cependant, dans l'ensemble, ces gènes sont plus retenus que la moyenne lors des duplications de génome spécifiques des poissons Téléostéens. Cela est également le cas pour leurs paralogues, ce qui nous a conduit à proposer un modèle d'évolution pour ces familles de gènes.

Le catalogue de gènes pigmentaires ainsi établi a ensuite été utilisé pour étudier la nature cellulaire et moléculaire de la bande blanche chez un modèle émergent de poisson corallien, le poisson-clown *Amphiprion ocellaris*. Nous avons pu montrer que cette couleur blanche est due non pas à des leucophores mais à des iridophores. Au cours de cette étude, de potentiels nouveaux gènes marqueurs d'iridophores ont été caractérisés, et certains ont été fonctionnellement confirmés chez le Poisson-zèbre *Danio rerio*. L'un deux, en particulier, nous a paru intéressant. Baptisé *roshi-a*, ce gène a un patron évolutif hautement singulier qui suggère une acquisition par transfert horizontal à de multiples reprises au sein des animaux, depuis un hôte initial de nature potentiellement fongique.

Enfin, nous nous sommes intéressés à l'acquisition de la pigmentation du poisson-clown au cours de son développement, et le lien que ce phénomène clef de l'histoire de vie des poissons coralliens entretient avec la voie de signalisation des hormones thyroïdiennes. Cela a nécessité une caractérisation transcriptomique du développement, qui est sensiblement différente de sa définition morpho-anatomique

L'ensemble de ces travaux s'inscrit, de manière plus globale, dans une étude de l'évolution, ou plutôt *des évolutions*, de la pigmentation. (i) En premier lieu, une *évolution* au sens presque darwinien de "descendance avec modification", ici appliquée à l'étude des changements des génomes au cours du temps. (ii) Ensuite, une *évolution* au sens des modifications, des renouvellements des idées et des connaissances des gènes et des processus impliqués dans la pigmentation chez les Vertébrés. (iii) Enfin, une *évolution* au sens étymologique du terme (*evolutio*), c'est-à-dire un "déroulement", un changement au cours de l'ontogenèse d'un organisme.

Je discute donc les résultats selon ces trois angles du terme "évolution".

PLAN DÉTAILLÉ DE LA DISCUSSION

I. "Évolution" des gènes de la pigmentation et duplications de génomes

1. Importance des études évolutives dans l'étude de la pigmentation
- a. Importance des études évolutives afin d'obtenir une annotation adéquate des gènes de la pigmentation
- b. De l'importance des études "manuelles" lors de l'étude des gènes de la pigmentation
- c. Les études évolutives ne suffisent pas toujours à comprendre l'évolution d'un gène pigmentaire

2. WGD et rétention des gènes de la pigmentation
- a. L'étude d'autres WGD apporterait des comparaisons intéressantes avec celles étudiées ici
- b. Les gènes de la pigmentation pourraient, par leur forte rétention, influencer l'évolution d'autres gènes

3. Au-delà des taux de rétention : l'évolution des gènes de la pigmentation est due à d'autres facteurs

II. "Évolution" de l'étude du répertoire pigmentaire et de la compréhension de la pigmentation chez les Vertébrés

1. Pourquoi il est important d'en apprendre plus sur la pigmentation des Vertébrés (et autres)

2. Le répertoire pigmentaire s'enrichit constamment
- a. Études "omiques"
- b. Apport de nouvelles bases de données
- c. Les principales "lacunes" restant à combler dans le domaine

3. *Roshi-a*, un nouveau venu dans la liste des gènes de la pigmentation…
- a. Quelques réflexions de nomenclature
- b. La fonction de *roshi-a* n'est pas encore connue avec précision
- c. Un gène qui bénéficie d'un contexte favorable

4. … remettant partiellement en cause le modèle établi de l'état ancestral de la pigmentation des Vertébrés
- a. Types cellulaires de l'ancêtre des Vertébrés
- b. Patron de pigmentation de l'ancêtre des Vertébrés

III. "Évolution" de la pigmentation au cours du développement d'un poisson corallien, le poisson-clown *Amphiprion ocellaris*

1. La pigmentation se met en place pendant la métamorphose, probablement en lien avec l'action des TH

2. Importance de l'étude d'*A. ocellaris* pour l'étude du patterning de la pigmentation
- a. Hormones thyroïdiennes et patterning au cours du développement d'*Amphiprion ocellaris*
- b. Un contrôle par les TH pourrait expliquer la variation de la pigmentation au sein du genre *Amphiprion*

I. "Évolution" des gènes de la pigmentation et duplications de génomes

1. Importance des études évolutives dans l'étude de la pigmentation

a. Importance des études évolutives afin d'obtenir une annotation adéquate des gènes de la pigmentation

Lors de l'étude d'un gène, qu'elle soit au niveau évolutif ou fonctionnel, il est important de s'assurer qu'on parle du "bon" gène. Je ne cite pas ici tous ceux que j'ai rencontrés lors de ma thèse, mais les problèmes de nomenclature sont en particulier notables lors de l'étude des gènes chez les poissons Téléostéens, qui, en raison de la duplication de génome survenue à la base de cette lignée (Ts3R), ont souvent deux copies "a et b" d'un gène ancestral.

Par exemple, *gch1* est un gène impliqué dans la conversion du GTP en composés précurseurs de ptéridines des xanthophores (Pelletier *et al.*, 2001). Ce gène a un paralogue issu de la 1R/2R, *gch2* (Lorin *et al.*, 2018). Dans une étude chez la métamorphose du Cardeau hirame *Paralichthys olivaceus*, le paralogue étudié n'est pas clairement indiqué (Watanabe *et al.*, 2008), et, dans une étude de 2014 chez le médaka, les auteurs parlent de *gch* au lieu de *gch1*. De même, il n'y a pas "un" gène *tyrp1* impliqué dans la mélanogenèse chez le Turbot ou la Carpe, comme évoqué dans deux études récentes (Guillot *et al.*, 2012; Jiang *et al.*, 2014) ; les poissons Téléostéens ont en effet deux ohnologues *tyrp1a* et *tyrp1b* (Braasch *et al.*, 2009). Soit cela correspond à une perte secondaire dans ces lignées, et auquel cas cela devrait être mentionné ; soit cela est une non-prise en compte de la réalité de l'évolution de ces gènes. Dans ce dernier cas, cela peut fausser les résultats des auteurs car l'identité du gène dont l'expression est par exemple mesurée en qPCR ou RNA-Seq n'est alors pas vérifiable (paralogue "a" ou "b" ?). Qui plus est, la comparaison des résultats avec d'autres précédemment publiés est impossible.

Par ailleurs, les duplications de génomes peuvent parfois conduire à une divergence des fonctions entre les deux paralogues. C'est par exemple le cas pour le gène pigmentaire *mitf* : chez le Poisson-zèbre, l'ohnologue *mitfa* est impliqué dans la mélanogenèse des mélanocytes issus des crêtes neurales et l'ohnologue *mitfb* dans ceux de l'épithélium pigmenté de la rétine (Lister *et al.*, 2001). Ne pas considérer l'histoire évolutive d'un gène lors de son étude peut donc conduire à des incongruités lors des interprétations de résultats.

Lors d'études "omiques", la profusion de gènes impliqués est telle qu'une analyse fine du contexte évolutif, permettant une annotation adéquate, n'est pas faisable pour chaque gène. Toutefois, dans le cas d'études à plus petite échelle, et notamment fonctionnelles, je pense que cela devrait être plus systématiquement réalisé. Dans le cadre d'études à grande échelle, une telle approche devrait également être entreprise pour les principaux gènes d'intérêt. Par exemple, dans notre étude sur la bande blanche du poisson-clown, nous avons vérifié l'histoire évolutive des 10 gènes les plus différentiellement exprimés dans chaque condition. De plus, chez *Stegastes partitus*, les deux ohnologues de *trpm1*, un gène cible de *mitf* impliqué dans la mélanogenèse, sont tous deux annotés "*transient receptor potential cation channel subfamily M member 1-like*" (**Supp. Table 1 – Chapitre 2**). Seul leur identifiant change : LOC103374804 et LOC103353689, puisque ce sont deux gènes différents. Dans ce cas, l'annotation, au sens de "donner un nom", requiert une analyse phylogénétique.

b. De l'importance des études "manuelles" lors de l'étude des gènes de la pigmentation

Lors de ces travaux, nous avons opté pour une approche manuelle de la résolution des phylogénies. Ces approches sont plus longues et ne sont donc pas applicables à très grande échelle, mais, dans des cas particuliers, peuvent s'avérer plus adaptées, en particulier lorsqu'un gène est mal annoté ou que sa séquence est mal prédite chez une espèce.

Par exemple, afin de récupérer la séquence de *mc1r* chez une espèce qui n'est pas dans Ensembl, par exemple le Saumon, une approche possible est de récupérer l'ensemble des séquences de *mc1r* de Vertébrés disponibles dans Ensembl (mammifère y compris) et de les "blaster" sur le génome du Saumon *Salmo salar*. En tout logique, on obtient alors une ou plusieurs séquences de *mc1r* de saumon qui s'alignent relativement bien avec les autres séquences. Or, dans Ensembl (v92 et avant), suite à un problème d'annotation chez l'ornithorynque, *mc1r* est "artificiellement" fusionné au gène qui lui est adjacent sur le chromosome 11, *tubb3* (ENSOANG00000002346), produisant une séquence chimère. Le meilleur hit d'un BLAST sur le génome du Saumon ne correspond alors pas à *mc1r*, mais à *tubb3* – qui ne s'aligne évidemment pas bien avec la séquence de *mc1r* des autres espèces. Sans vérification manuelle de l'alignement, l'arbre ainsi obtenu n'aurait pas signifié grand-chose, du moins pour le saumon. Je note que ces cas sont toutefois très rares mais doivent parfois être envisagés.

c. Les études évolutives ne suffisent pas toujours à comprendre l'évolution d'un gène pigmentaire

Malgré les efforts appuyés parfois menés, la compréhension de l'histoire évolutive d'un gène de la pigmentation n'est pas toujours possible. C'est ainsi le cas d'une famille de gènes qui est pourtant clef dans la compréhension de la pigmentation chez les Vertébrés : la famille du gène *agouti*. Cette famille contient deux gènes chez les tétrapodes : *asip,* impliqué dans le patterning dorso-ventral de la pigmentation, et *agrp,* impliqué dans la régulation de la prise alimentaire. Elle en contient cependant quatre chez les Téléostéens : *asip1* et *agrp2,* impliqués dans la pigmentation, ainsi que *asip2* et *agrp1*, qui ne le sont pas. La détermination des relations évolutives entre ces gènes, et le rôle en particulier joué par les WGD, a fait l'objet d'une farouche correspondance entre deux équipes en 2011 (Braasch and Postlethwait, 2011; Schioth *et al.*, 2011). Arbitrée par une étude publiée l'année suivante (Västermark *et al.*, 2012), la conclusion de cette discussion est que… c'est compliqué, et qu'il faut probablement envisager des duplications à petite échelle.

Nous avons vu dans le cas de l'évolution de *roshi-a* que les deux paralogons chez le Médaka sont partiellement conservés par rapport au Lépisosté tacheté et, surtout, qu'au sein de chacun d'eux, des réarrangements ont eu lieu (**cf. Fig. 2 – Résultats / Chapitre 3**). Lorsque ces réarrangements sont couplés à des vitesses d'évolution variables entre paralogues et entre espèces, la reconstruction de l'histoire évolutive par des critères structuraux et phylogénétiques peut être alors délicate.

Dans le cas des gènes de la pigmentation, neuf familles de gènes n'ont pas pu être étudiées car leur histoire évolutive était trop complexe pour déterminer de manière non ambigüe s'il y avait eu 1R/2R ou non. Les travaux actuellement menés sur le génome de la lamproie pourraient permettre de mieux comprendre l'évolution des génomes de Vertébrés et en particulier résoudre l'épineuse question des mécanismes en jeu lors des paléo-polyploïdisations ayant eu lieu à la base de ce groupe (Smith *et al.*, 2013; Smith and Keinath, 2015).

2. WGD et rétention des gènes de la pigmentation

Nous l'avons vu en introduction, les WGD jouent un rôle majeur dans l'évolution des génomes. Les WGD génèrent en abondance un matériel génétique nouvellement disponible à la variation et à l'adaptation éventuelle. En cas de rétention des paralogues, de nouvelles fonctions peuvent être assurées tout en assurant l'ancienne fonction. D'un point de vue génomique, elles ne sont donc pas sans rappeler les "monstres prometteurs" défendus par Goldschmidt, même si leur effet phénotypique peut parfois se présenter après un long délai et dans des conditions environnementales particulières (Vanneste *et al.*, 2014).

a. L'étude d'autres WGD apporterait des comparaisons intéressantes avec celles étudiées ici

Au cours de cette étude nous avons étudié l'impact des WGD ayant eu lieu soit à la base des Vertébrés, soit dans quelques lignées spécifiques : les poissons Téléostéens, et les salmonidés. Il serait intéressant d'étudier les conséquences d'autres WGD sur les gènes de la pigmentation, que ce soit au sein des poissons Téléostéens (Carpe) ou chez les tétrapodes (Xénope) (Li, Hou, *et al.*, 2015; Session *et al.*, 2016). En particulier, la tendance observée d'un lien entre diversité pigmentaire et rétention des gènes de la pigmentation est-elle conservée, ou est-elle spécifique des "poissons" ? La rétention des gènes après la WGD spécifique du Xénope *X. laevis,* survenue il y a seulement 17-18 Ma, est élevée (autour de 56 %) par rapport à celle des Téléostéens, mais comparable avec celle des salmonidés (Session *et al.*, 2016). Elle serait un point de comparaison intéressant, car cette espèce n'est pas un poisson Téléostéen et possède moins de cellules pigmentaires.

Enfin, le groupe des insectes étant reconnu pour sa diversité exceptionnelle, notamment pigmentaire, il serait intéressant de comparer les conséquences des WGD chez les Vertébrés avec celles ayant eu lieu chez les Insectes et récemment identifiées (Li *et al.*, 2018).

b. Les gènes de la pigmentation pourraient, par leur forte rétention, influencer l'évolution d'autres gènes

Les conséquences d'une WGD sur l'évolution des gènes sont multiples. En particulier, les ohnologues issus d'une WGD sont souvent moins affectés par les duplications à petite échelle que les autres gènes du génome (Makino and McLysaght, 2010). L'effet des duplications ne s'arrête pas là, et une autre conséquence concerne, non pas les gènes issus d'une WGD, mais leurs *voisins*.

En effet, de manière très intéressante, autour d'un ohnologue, les autres gènes ont moins de chance d'être dupliqués (Makino *et al.*, 2013). Il y a ainsi, dans le génome des régions riches en ohnologues qui correspondent à des déserts de CNV (Copy Number Variation). Inversement, certaines familles très dupliquées, comme celle des récepteurs olfactifs, sont présentes dans un "désert d'ohnologues". Le mécanisme suggéré est uniquement putatif et basé sur l'observation que les CNV ne concernent en général pas un, mais plusieurs gènes. Or, il a été montré que la duplication d'un ohnologue était moins fréquente que pour les autres gènes à cause de leur contre-sélection par effet de dosage : une copie en plus a un effet délétère (Makino and McLysaght, 2010). L'idée serait donc que la duplication d'un gène voisin d'un ohnologue, impliquant également un ohnologue, soit, par effet de dosage, immédiatement contre-sélectionnée (Makino *et al.*, 2013).

Les gènes de la pigmentation sont, on l'a vu, souvent retenus. Cela signifie donc que les gènes *voisins*, sont, eux, moins souvent dupliqués que le reste du génome. Nous n'avons pas étudié ces gènes là mais il serait intéressant de vérifier cela et d'étudier dans quels processus ils sont impliqués.

On note toutefois qu'un gène constitue un contre-exemple à tout cela : *roshi-a.* (i) En premier lieu, ce gène montre une évolution convergente après une WGD et après une SSD, puisque chez le Lépisosté trois domaines, sauf le domaine 3, ont été perdus après la SSD de roshi. De même, l'un des deux ohnologues issus de la Ts3R, *roshi-ab*, a ancestralement perdu tous les domaines sauf le 3. (ii) De plus, chez l'épinoche à trois épines, dont je n'ai pas parlé dans le chapitre consacré à l'évolution de *roshi-a*, on trouve deux copies de *roshi-a* en quasi-tandem (un gène entre les deux) inversé, sur le même chromosome (chromosome 11).

3. Au-delà des taux de rétention : autres aspects susceptibles d'influer l'évolution des gènes de la pigmentation

Au cours de cette thèse, j'ai envisagé les conséquences des duplications de génomes essentiellement sous l'angle de la rétention. Cependant, même si c'est un aspect que je qualifierais de "préalable" à ceux évoqués plus loin (sans gène, rien – ou presque !), il ne suffit pas à lui seul pour comprendre les bases génétiques de l'adaptation.

En premier lieu, pour qu'un gène remplisse sa fonction, il faut qu'il soit exprimé. Or, après une duplication de génome, il a été montré chez l'Arabette que 70% des paires avaient des niveaux d'expression dissymétriques ; chez le Coton, c'est jusqu'à 99.4% (Panchy *et al.*, 2016). Il y a donc rapidement une divergence des patrons d'expression entre les paralogues.

Parmi une multitude de grandeurs, le niveau d'expression d'un gène est celle qui est la plus intimement corrélée à son taux d'évolution (Cherry, 2010). Ainsi, un gène fortement exprimé a une évolution plus contrainte qu'un gène faiblement exprimé, car l'impact phénotypique d'une modification du premier risque d'être plus important. Au cours de cette thèse, nous n'avons pas étudié de manière systématique ni les différences d'expression, ni de taux d'évolution, entre paralogues. Ces mesures pourraient être intéressantes à mener sur les familles de gènes de la pigmentation, afin de déterminer si, entre les gènes de la pigmentation et leurs paralogues, des différences notables sont observables.

II. "Évolution" de l'étude du répertoire pigmentaire et de la compréhension de la pigmentation chez les Vertébrés

1. Pourquoi il est important d'en apprendre plus sur la pigmentation des Vertébrés (et autres organismes)

Nous l'avons vu en introduction : l'étude de la pigmentation, ou, plus généralement, de la couleur, est, historiquement fondamentale à l'essor de la biologie moderne. Ce rôle "patrimonial" pourrait en lui seul suffire à justifier de continuer à étudier les processus impliqués, mais il y a plus. En effet, encore aujourd'hui, ce champ disciplinaire est à la croisée de nombreuses disciplines (Cuthill *et al.*, 2017). Depuis la cancérologie et la compréhension des mélanomes, à la génétique fondamentale et l'étude de croisements, en passant par la biologie du développement, la pigmentation, et plus encore la couleur, est au cœur de nombreuses études. En évolution, elle permet en particulier de répondre à des questions fondamentales sur le processus adaptatif. Enfin, l'étude de la couleur de différentes populations d'espèces de mammifères polyphéniques (pelage brun en été et blanc en hiver), qui ont du mal à s'adapter au rythme du changement climatique actuel, a même été la base de la proposition récente d'un programme de conservation (Mills *et al.*, 2018).

2. Le répertoire pigmentaire s'enrichit constamment

a. Études "omiques"

L'étude des bases génétiques de la pigmentation s'enrichit constamment, notamment suite à l'essor d'études à grandes échelles (San-Jose and Roulin, 2017). Les Genome-Wide Association Studies permettent d'identifier de nouveaux candidats (Hysi *et al.*, 2018), de même que les études transcriptomiques, telle que celle ayant conduit à la mise en évidence du rôle de *alx3* dans le patterning pigmentaire chez les rongeurs (Mallarino *et al.*, 2016).

Chez le poisson-clown *Amphiprion ocellaris*, une approche transcriptomique nous a également permis de mettre en évidence des nouveaux marqueurs d'iridophores. La profusion des gènes identifiés en RNA-Seq est intéressante, mais elle est telle qu'essayer de tous les caractériser fonctionnellement est impossible. Afin de mieux cibler les candidats potentiels, et de réduire le champ des possibles, une méthode que nous avons partiellement appliquée pourrait consister à systématiquement croiser les données obtenues par plusieurs études chez diverses espèces. Cela réduirait le spectre de gènes candidats, et on s'attend à obtenir des marqueurs généralement conservés. Ainsi, *roshi-a* avait été mis en évidence chez le Poisson-zèbre parmi une liste de quasiment 300 marqueurs *potentiels* d'iridophores (Higdon *et al.*, 2013). C'est en réalisant une autre étude, sur une autre espèce, que nous avons de nouveau identifié ce gène, qui s'est en effet avéré fonctionnellement intéressant.

De même, nous avons mis en évidence le rôle potentiel d'*alx4b* dans le développement des iridophores. Le choix de ce gène parmi la liste de candidats s'est en partie appuyé sur le fait qu'un gène proche, *alx3*, est un marqueur de mélanocytes chez les rongeurs, et que les gènes de la famille *alx* sont des gènes exprimés dans la différentiation de cellules de crêtes neurales crâniennes (Lumb *et al.*, 2017). Il était donc vraisemblable que *alx4b* ait aussi un rôle dans le développement des iridophores, des cellules dérivées des mêmes précurseurs des crêtes neurales et ontogénétiquement proches des mélanocytes chez les Téléostéens (Kimura *et al.*, 2014).

"Croiser" les études ne suffit pas toujours : c'est ainsi le cas de *apod1a*, que nous avons identifié et potentiellement également impliqué chez les cichlidés (Gu *et al.*, 2017). Toutefois, fonctionnellement, il n'apparaît pas être un gène majeur dans la différentiation des iridophores (**Fig. 5 – Résultats / Chapitre 1**).

b. Apport de nouvelles bases de données

De nouvelles bases de données émergent et viennent compléter, de manière plus spécifique, celles déjà connues sur la pigmentation (notamment Gene Ontology). Par exemple, Gephebase, une base de données qui recense les relations génotype-phénotype, contient aujourd'hui de nombreux gènes associés au trait "pigmentation" (Martin and Orgogozo, 2013). Quatre d'entre eux n'ont pas été inclus dans notre analyse (cf. chapitre 1) suite à une découverte trop tardive de cet outil : *tbx3*, *herc2*, *ckn2a, zeb1* et *psmb7*. Au-delà du nombre et du nom des gènes, un des intérêts d'une telle base de données est que pour aucun de ces gènes, le terme de Gene Ontology "pigmentation" n'est associé, malgré leur implication avérée dans la pigmentation. Un recensement des gènes de pigmentation en utilisant uniquement l'outil de Gene Ontology ne les inclurait donc pas. L'émergence de bases de données plus adaptées représente donc un atout considérable dans ce domaine d'étude.

c. Les principales "lacunes" restant à combler dans le domaine

D'un point de vue génétique, beaucoup est aujourd'hui connu sur quelques organismes clefs de laboratoire : souris, poisson-zèbre. Cependant, au sein même des laboratoires, un autre organisme est considérablement utilisé : le Xénope *Xenopus laevis*. Cet organisme représente un rare modèle d'Amphibien qui pourrait donc être plus largement étudié du point de vue de sa pigmentation. De trop rares études, à mon sens, se sont intéressées aux bases cellulaires et génétiques de la pigmentation du Xénope, qui pourrait éclairer sous un nouvel angle la compréhension de l'évolution des Vertébrés (Williams *et al.*, 2017). Cette espèce présente en effet des xanthophores, des iridophores et des mélanocytes et pourrait constituer un point de comparaison entre tétrapodes et Téléostéens qui serait, à mon avis, intéressant (Fukuzawa, 2010). Cela permettrait notamment d'étudier si les bases génétiques et ontogéniques sont communes, et confirmer ou infirmer les hypothèses généralement avancées sur l'état ancestral de la pigmentation des Vertébrés (voir ci-dessous). Par ailleurs, la lignée du Xénope a récemment connu une WGD, et, dans le cas particulier des relations entre WGD et pigmentation, ce modèle serait d'un intérêt certain (Session *et al.*, 2016).

Hors des laboratoires, beaucoup reste encore à découvrir chez les "reptiles", pour lesquels les bases cellulaires sont relativement bien connues (Kuriyama *et al.*, 2006; Teyssier *et al.*, 2015) mais ou par contre, génétiquement, "tout reste à faire".

Chez les Téléostéens, hormis le Poisson-zèbre et le Médaka, rien, ou presque, n'est connu au niveau génétique. En particulier, chez les poissons coralliens, où la diversité des chromatophores est majeure, on ne connait rien des relations ontogénétiques entre les différents types de chromatophores. Ces relations pourraient, là encore, éclairer d'un nouveau jour la compréhension actuelle du développement des cellules pigmentaires, cantonnée aux deux espèces de poissons modèles citées plus haut. Une étude récente chez le Médaka a d'ailleurs montré qu'un même gène, *sox5,* induit la prolifération des xanthophores de Médaka et inhibe les "mêmes" xanthophores de Poisson-zèbre. Il est donc important, à mon sens, d'étudier de nouveaux modèles pour avoir une compréhension intégrée des processus.

L'outil CRISPR/Cas9, s'il est encore loin d'une application à grande échelle sur n'importe quelle espèce non-modèle, est toutefois prometteur (Doudna and Charpentier, 2014). Il serait en particulier intéressant de tester, sur des espèces de poisson corallien non-modèles possédant des types cellulaires additionnels, l'effet de la délétion de certains marqueurs de lignée cellulaire pigmentaire. Cela pourrait par exemple permettre de savoir si les érythro-iridophores sont plus proches des iridophores ou des érythophores/xanthophores. Cela pourrait donc permettre d'enlever les "?" de la figure 2 de l'article publié dans G3 (**Résultats / Chapitre 1**)**,** c'est-à-dire de mieux comprendre les interrelations entre chromatophores.

L'article publié dans G3 constitue, dans cette optique, un cadre évolutif intéressant pour les futures études du domaine.

3. *Roshi-a*, un nouveau venu dans la liste des gènes de la pigmentation

a. Quelques réflexions de nomenclature

Au cours de cette thèse, le gène *roshi-a* a été "découvert". En réalité, ce gène était déjà annoté dans le génome sous l'identifiant "si:ch211-256m1.8". Dans la littérature, à ma connaissance, il est fait mention de ce gène dans une seule étude, qui ne s'intéresse pas à *roshi-a*, mais à son voisin, *mief1*, dupliqué chez une espèce invasive hybride de Chabot (Dennenmoser *et al.*, 2017). Ce nom est un nom automatiquement donné lors des processus d'annotation mais, pour des raisons évidentes de praticité, ne peut pas être utilisé tel quel.

Nommer un gène est un exercice délicat et le nom choisi peut s'avérer finalement partiellement inadapté ou véhiculer des représentations trop biaisées de sa fonction (Seringhaus *et al.*, 2008). Par exemple, le facteur de transcription découvert en tant que marqueur de myoblastes, *myoD*, a un nom qui suggère un rôle majeur dans le développement du muscle. Alors que les mutants de Poisson-zèbre ne sont en effet pas viables (Hinits *et al.*, 2011), les mutants KO pour *myoD* ont paradoxalement un phénotype peu marqué chez la Souris (Rudnicki *et al.*, 1992).

De plus, donner à "si:ch211-256m1.8" un nom qui évoque de trop près les iridophores serait peu adapté à une étude de ce gène chez la Lingule ou l'Hydre, où le gène est également trouvé mais chez lesquelles on ne trouve pas d'iridophores. En général, donner un nom trop amusant, ou qui n'a rien à voir avec sa fonction est également peu recommandé, au cas où le gène joue finalement un rôle dans une maladie grave – même si cela est peu probable pour *roshi-a*, qui n'est pas présent chez l'Homme (Schwartz, 2006; Seringhaus *et al.*, 2008). Finalement, donner un nom qui n'évoque rien de prime abord est également déconseillé ; les noms de *p53* ou des clusters de différentiation CD1, CD2, CD3, CD4… ont été très humoristiquement critiqués par (Petsko, 2002). Il faut également vérifier que le nom ne soit pas

déjà utilisé, sous peine de générer de nouveaux problèmes d'annotation. Enfin, un clin d'œil à une histoire personnelle est parfois plébiscité (Petsko, 2002).

Le nom de *roshi-a* est un peu à la croisée de ces différentes opinions. C'est le nom d'un personnage de Dragon Ball Z qui porte en permanence des lunettes de soleil qui le protègent de cet astre lumineux. C'est donc un nom qui n'est pas franchement humoristique, qui n'évoque pas de manière trop ostentatoire sa fonction (mais qui y est tout de même relié, puisque les iridophores réfléchissent la lumière), qui n'est pas déjà utilisé et qui est facile à prononcer. Enfin, ce choix est issu d'une promesse faite à un ami il y a quelques années, avant même de commencer à travailler sur la pigmentation, laquelle était : "Si je trouve un gène, promis, je lui donnerai le nom d'un personnage de DBZ." Ayman : c'est chose faite !

Pour l'instant, la fonction précise, au niveau moléculaire, de *roshi-a*, est inconnue. Comme le suggère (Petsko, 2002), les gènes les mieux nommés sont souvent ceux dont le nom évoque directement la fonction telles que la "tyrosinase" ou "l'hémoglobine". Roshi est probablement une enzyme avec de multiples activités, et il sera délicat de lui trouver un nom court décrivant l'ensemble de ses activités. De plus, souvent, lorsqu'un nom est utilisé, l'usage a tendance à le consacrer, au détriment du bon sens (*p53*…) ?

En 2006, un médecin interviewé sur les problèmes que les noms "amusants" peuvent provoquer lors de l'annonce du gène impliqué dans la maladie du patient, soulignait : "*People get to be very, very fond of the names of things they've discovered. They don't like somebody telling them what to call it.*" (Schwartz, 2006). Avec *roshi-a*, je ne veux pas tomber dans un narcissime abusif : les arguments ci-dessus sont des arguments (qui se veulent) objectifs, en étant certes empreints d'histoire personnelle, mais si l'étude de la fonction du gène consacre un nouveau nom plus adapté, je ne verrai aucun problème à cela !

b. La fonction de *roshi-a* n'est pas encore connue avec précision

Au cours de ces travaux, nous avons montré que *roshi-a* est impliqué dans la différentiation des iridophores – au moins. Cependant, plus d'études fonctionnelles sont requises. En particulier, il n'existe pas de lignée mutante stable pour ce gène, et les mutants F0 injectés atteignent rarement la maturité sexuelle (V. Lewis, comm. pers.). Ce gène pourrait donc être impliqué dans d'autres processus que le seul développement des iridophores.

Il serait en outre intéressant de déterminer si les mutants *roshi-a* F0 ont réellement moins d'iridophores ou si les iridophores sont présents mais avec un défaut d'accumulation de plaquettes de purines. Il faudrait pour cela réaliser cette injection de sgRNA non pas chez des individus sauvages mais chez des individus mutants GFP pour un marqueur d'iridophores. Si les iridophores se développent normalement, hormis leurs plaquettes de purines, les cellules devraient rester fluorescentes.

On note que ce n'est pas la première fois qu'un gène de la pigmentation est découvert avant que sa fonction soit mieux comprise. C'est le cas par exemple de *tpcn2*, marqueur mélanogénique découvert en 2008 et seulement caractérisé fonctionnellement en 2016 (Sulem *et al.*, 2008; Ambrosio *et al.*, 2016). C'est également le cas pour au moins deux autres marqueurs mélanogéniques, *trpm1* (Oancea *et al.*, 2009) et *slc24a5* (Lamason *et al.*, 2005; Ginger *et al.*, 2008). Dans les années qui viennent, nul doute que de nouvelles études viendront éclairer la fonction de ce gène.

c. Un gène qui bénéficie d'un contexte scientifique favorable

L'étude de *roshi-a* bénéficie d'un contexte favorable. Outre la double mise en évidence de ce gène, par deux études indépendantes sur deux organismes différents (*D. rerio* et *A. ocellaris*), de nouvelles études permettent de mieux cerner

l'histoire évolutive du gène et qui permettent de répondre à des questions auxquelles il aurait été impossible de répondre il y a encore 10 ans.

En premier lieu, au sein des animaux, de nombreux génomes dans des lignées différentes des Vertébrés ont été publiés ces dernières années (Cao *et al.*, 2013; Schwager *et al.*, 2017). De plus, diverses données transcriptomiques sont également disponibles (Zapata *et al.*, 2015). Des phylogénies actualisées sont également parues récemment, notamment chez les poissons ou chez les arthropodes, et permettent de mieux cerner les temps de divergence entre principales lignées (Regier *et al.*, 2010; Betancur-R *et al.*, 2017; Sharma *et al.*, 2018).

Même si cela nécessitera des travaux complémentaires, l'évolution de *roshi-a* semble être liée au groupe des champignons. "Par chance", ce groupe est aujourd'hui celui où l'effort de séquençage a été le plus intense au sein des eucaryotes (Fitzpatrick, 2012). De plus, des phylogénies détaillées sont apparues récemment et viennent éclairer les relations entre les différents groupes de champignons, précisant la compréhension de l'évolution de *roshi-a*.

4. *Roshi-a* remet partiellement en cause le modèle traditionnel de l'état ancestral de la pigmentation chez les Vertébrés

J'envisage ici l'état ancestral sous deux angles : les types cellulaires, ainsi que les patrons.

a. Types cellulaires de l'ancêtre des Vertébrés

Outre sa fonction dans le développement d'un type cellulaire majeur des Vertébrés, *roshi-a* est intéressant car sa distribution évolutive est atypique. C'est en effet un gène qui a probablement été transmis à de multiples reprises par transfert horizontal, selon des mécanismes discutés dans le chapitre 3. De plus, et peut-être *surtout*, c'est un gène dont la distribution au sein des Vertébrés remet partiellement en question l'état ancestral de la pigmentation dans ce groupe.

Au sein des Vertébrés, les iridophores ont déjà été décrits chez les Amphibiens, les Lépidosauriens, et les poissons Téléostéens. Ils ont en outre été décrits chez des Actinoptérygiens non-Téléostéens, tels que les Acipenseriformes (esturgeons) (Zarnescu, 2007) et les Lépisostéiformes (I. Braasch, comm. pers.). La pensée dominante actuellement est donc de considérer que tous les Ostéichthyens, ou "poissons osseux", partageaient ancestralement mélanocytes, xanthophores et iridophores (Mellgren and Johnson, 2002) (**Hypothèse A – Fig. 59**). Cette hypothèse conservatrice peut même être étendue à "l'ancêtre des Vertébrés" puisque les lamproies, des Cyclostomes, ont également des iridophores (Wright and McBurney, 1992). Par ailleurs, et cela va dans le sens de cette hypothèse, l'étude que nous avons menée sur la présence/absence des gènes de pigmentation au sein des Vertébrés montre que la vingtaine de marqueurs d'iridophores aujourd'hui connus sont dans l'ensemble conservés dans l'ensemble des Vertébrés (Lorin *et al.*, 2018).

Néanmoins, quelques exceptions existent. En effet, *roshi-a,* qui semble être indispensable à la différentiation des iridophores, est spécifique des actinoptérygiens. De plus, comme discuté dans le chapitre 3 des résultats, au moins trois autres marqueurs d'iridophores identifiés chez les poissons Téléostéens ne sont pas conservés chez tous les Vertébrés : *pnp4a*, *mpv17* et *csf1*. Si les iridophores des Vertébrés étaient homologues, on s'attendrait au contraire à observer une conservation des voies moléculaires, et donc de l'ensemble des gènes. Il semblerait donc que les poissons actinoptérygiens aient des voies de différentiation partiellement différentes des autres Vertébrés.

La méconnaissance actuelle de la fonction de ce gène, au niveau moléculaire, ne permet pas d'affirmer avec certitude que les iridophores des Téléostéens ne sont pas homologues de ceux des Vertébrés. Cette distribution évolutive "singulière" est toutefois une observation "déclic" qui nous a conduit à imaginer deux autres scénarios pour l'état ancestral

de la pigmentation (Fig. 59). Nous avons également rassemblé des observations qui viennent étayer ces scénarios alternatifs. Pour résumer :

- *roshi-a*¸ ainsi que d'autres gènes, ne sont pas conservés dans l'ensemble des Vertébrés, semblant indiquer que les voies de différentiation diffèrent sensiblement d'un groupe de Vertébrés à l'autre
- il existe, chez certains poissons Téléostéens, des types pigmentaires dichromatiques : des cyano-érythrophores, qui contiennent des pigments bleus (de nature inconnue) et des ptéridines/caroténoïdes ; des érythro-iridophores, qui contiennent des caroténoïdes et des purines (Goda and Fujii, 1995; Goda *et al.*, 2011, 2013). Les leucophores eux-mêmes, chez le Médaka, contiennent à la fois des purines et des ptéridines à l'état embryonnaire (Kimura *et al.*, 2014), et pourraient même être considérés comme partiellement "dichromatiques". Il existe enfin dans ce clade des cellules qualifiées de "iridophore-like" qui sont des iridophores à l'agencement différent des plaquettes de purines (Goda, 2017). Il semblerait donc qu'au sein des Téléostéens, il y ait une grande "plasticité évolutive" des chromatophores avec une fréquente association entre plusieurs pigments (mosaïcisme) au sein d'une même cellule
- cette plasticité est également présente chez les Amphibiens, où des cellules mosaïques possédant des caractéristiques de mélanocytes, de xanthophores et d'iridophores ont déjà été documentées (Bagnara, 1983). Ces cellules pigmentaires mosaïques (MPC = Mosaic Pigment Cells) présentent en effet des mélanosomes, des plaquettes de purines, des xanthosomes ainsi que des caroténoïdes.

Deux scénarios sont alors possibles :

1. Si on considère l'état ancestral comme supposé jusqu'ici, avec le trio xantho-/mélano-/iridophores, il faudrait alors imaginer que de multiples pertes aient eu lieu dans chaque lignée. Il faudrait également imaginer que les voies de différentiation des iridophores aient également évolué de manière différente, et que *roshi-a* (par exemple) se soit rendu indispensable chez les actinoptérygiens ; puis que, enfin, au moins chez les Amphibiens et les Téléostéens, des types cellulaires mosaïques soient apparus à partir de ces trois types ancestraux.

2. Une hypothèse alternative (**Hypothèse C – Fig. 59**) est de considérer que l'ancêtre des Vertébrés possédait des mélanocytes et un type cellulaire mosaïque. Chez les lamproies, les iridophores seraient apparus à partir d'une MPC contenant des plaquettes de purines – MPC qui aurait ensuite été perdue, puisqu'aucune MPC n'a été décrite dans cette lignée. Les xanthophores auraient quant à eux pu apparaître à la base des ostéichthyens à partir de ces MPC, qui contiennent des caroténoïdes et des ptéridines. Ensuite, chacune des lignées de tétrapodes aurait alors subi des pertes d'un ou de plusieurs de ces types cellulaires, et les iridophores seraient apparus indépendamment à partir des MPC chez les Lépidosauriens et les Amphibiens ; chez les Lépidosauriens, les MPC auraient été perdues et elles auraient été gardées chez les Amphibiens. Enfin, chez les poissons actinoptérygiens, les MPC auraient pu être à l'origine de l'apparition des iridophores, des leucophore et des érythro-iridophores (avant de disparaître), pendant que cyanophores et cyano-érythrophores apparaissaient indépendamment dans certaines lignées.

Ces scénarios évolutifs sont dans l'ensemble hautement discutables, et sont en partie probablement erronés. Toutefois, l'accumulation de données à la fois génétiques et cellulaires sur diverses lignées au sein des Vertébrés porte à croire que le scénario de l'évolution des cellules pigmentaires imaginé dans les années 2000 doit être reconsidéré. On conçoit donc à ce titre l'importance que revêt une meilleure connaissance des processus de différentiation des iridophores et des autres types cellulaires chez l'ensemble des Vertébrés, et pas seulement chez les Téléostéens.

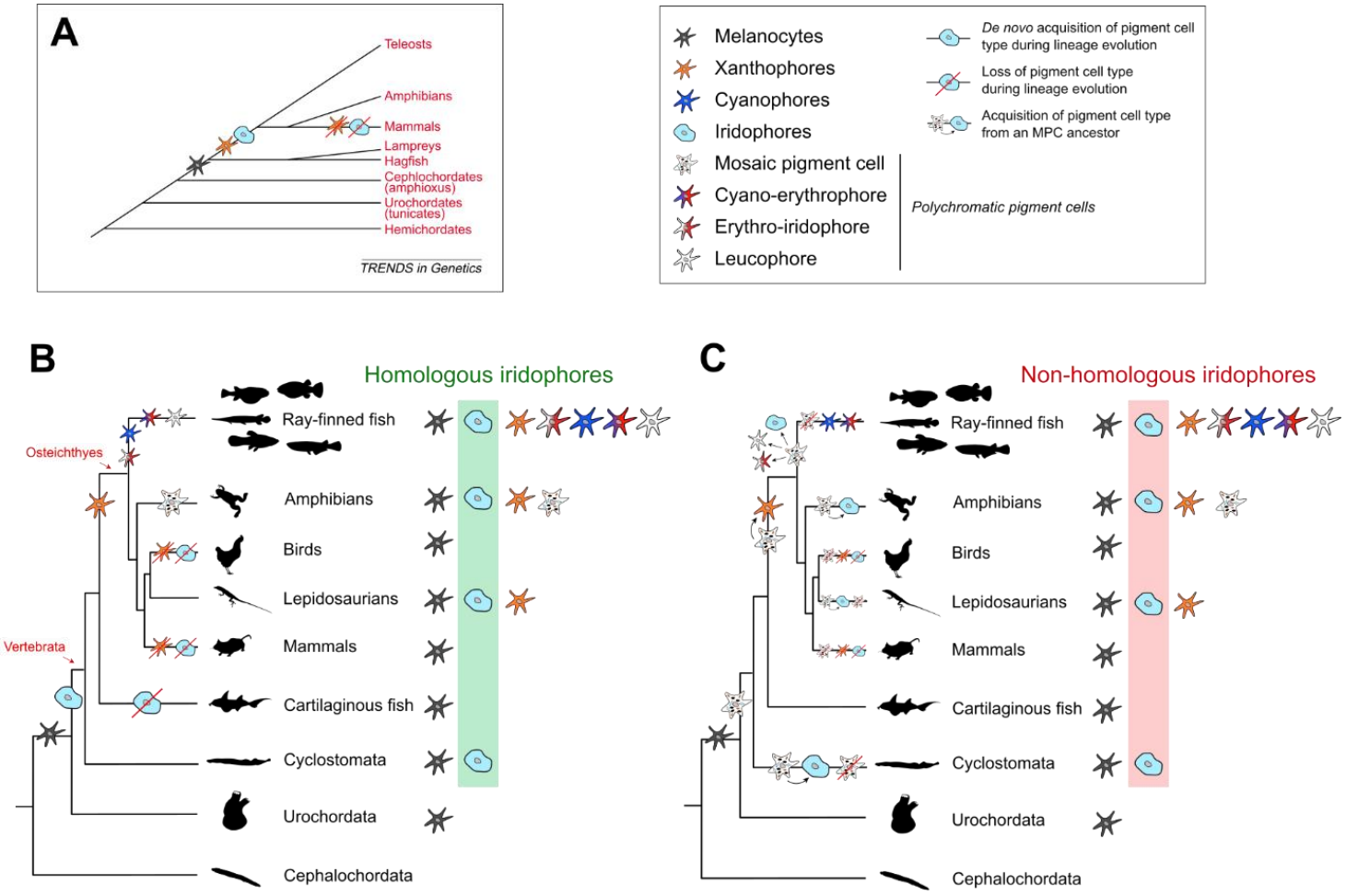


Figure 59 - Hypothèses pour l'évolution des types cellulaires pigmentaires chez les Vertébrés
L'hypothèse A est issue d'un article de 2002 (Mellgren and Johnson 2002). L'hypothèse B suppose une origine unique aux iridophores. L'hypothèse C suppose quant à elle des apparitions multiples, indépendantes, à partir d'un type cellulaire mosaïque ancestral. Ces hypothèses sont décrites plus en longueur dans le texte.

b. Patron de pigmentation de l'ancêtre des Vertébrés

D'un point de vue du patron de pigmentation dorso-ventral, il est vraisemblable que l'antagonisme *mc1r/asip* (ou *asip1*) ait déjà été à l'œuvre chez le paléo-Vertébré. Chez les Mammifères, l'antagonisme *mc1r/asip* est à l'origine de la balance entre phéo- et eumélanine (Manceau *et al.*, 2010). Chez les Actinoptérygiens, *mc1r/asip1* contrôle le ratio entre mélanocytes du côté dorsal et iridophores du côté ventral (Ceinos *et al.*, 2015; Cal *et al.*, 2017). Si les bases cellulaires et moléculaires de ce patron dorso-ventral diffèrent, le contrôle génétique semble lui conservé.

Les autres patrons de pigmentation (en points, en bandes, ...) semblent tellement variables chez les Vertébrés qu'il paraît difficile de se hasarder à proposer un état ancestral.

III. "Évolution" de la pigmentation au cours du développement d'un poisson corallien, le poisson-clown *Amphiprion ocellaris*

La dernière acceptation du terme "évolution" que j'envisage ici est celle du "déroulement". Elle n'est pas sans rappeler l'étymologie de "développement", qui consiste à "dé-velopper", à "déplier" un embryon en un organisme adulte (Pradeu *et al.*, 2016).

Je discute donc ici de quelques aspects du développement post-embryonnaire du poisson-clown *Amphiprion ocellaris*. Ce poisson corallien est étudié au sein de l'équipe de Vincent Laudet pour répondre à des questions fondamentales sur les bases génétiques et cellulaires de la pigmentation. Nous avons en effet souligné la nécessité de nouveaux organismes expérimentaux afin de mieux comprendre la diversité de la pigmentation chez les Vertébrés.

Par ailleurs, cette espèce est également une espèce de poisson corallien. Il n'existe actuellement pas d'espèce modèle de poisson corallien – à ma connaissance. Or, les enjeux qui concernent ces espèces sont colossaux, et il y a urgence à mieux comprendre les processus intervenant lors de leur développement. Pour l'essentiel des poissons coralliens, le cycle de développement est caractérisé par une métamorphose, qui est entre autres souvent intimement liée à des changements majeurs de la pigmentation (Leis and McCormick, 2002).

1. La pigmentation se met en place pendant la métamorphose, probablement en lien avec l'action des TH

Chez de nombreux Vertébrés, le développement, et en particulier la métamorphose, est un processus finement régulé par l'axe hypothalamo-hypophysaire et en particulier la voie des hormones thyroïdiennes (Laudet, 2011). Une perturbation de cette voie (notamment par des pesticides) sur des larves de poisson corallien en milieu naturel a des conséquences sur leur développement, et en particulier sur la mise en place de leur pigmentation (Holzer *et al.*, 2017).

L'étude de cette voie passe en général en premier lieu par (i) une mesure des taux d'hormones chez les individus et (ii) une mesure des niveaux d'expression des récepteurs aux hormones. Au cours de cette thèse, pour des raisons d'ordre pratique, et notamment l'obtention d'assez de larves pour avoir une représentation statistique, la mesure des taux d'hormones n'a pas été réalisée ; elle est en cours et menée dans le cadre du doctorat de Natacha Roux, à l'OOB.

Les niveaux d'expression des TR ont pu être déterminés par une analyse RNA-Seq (**Fig. 7 – Résultats / Chapitre 4**). On n'observe pas, comme c'est le cas pour certaines espèces de Vertébrés (y compris de poissons), de pic d'expression bien net au moment du climax de la métamorphose, qui a lieu au niveau des stades 4 à 6. On observe toutefois, pour TR-β, une augmentation de l'expression à partir du stade 4. Pour les deux TR-α, on peut également observer une légère augmentation entre les stades 4 et 6, qui reste à confirmer.

Au sein des Téléostéens, le lien entre la métamorphose et le niveau d'expression des TR est très variable, mais, en général, au moins un des TR a un profil qui forme un pic au moment du climax (**Fig. 60**). De ce point de vue, le poisson-clown ne serait pas si différent, avec les deux TR-α qui augmentent à partir du S4 jusqu'au S6 et diminuent au S7, et TR-β qui lui aussi augmente, sans diminuer au S7. On note que tout cela n'a été mesuré que pour n = 2 voire n = 3 points par stade, et il faudrait maintenant reproduire ces mesures sur un plus grand nombre de larves/juvéniles. De même, les profils d'expression étant parfois hautement variables selon les tissus (Harada *et al.*, 2008), il serait intéressant de procéder à des mesures sur des échantillons tissulaires pour mieux cerner les différentes potentialités de réponse de l'organisme.

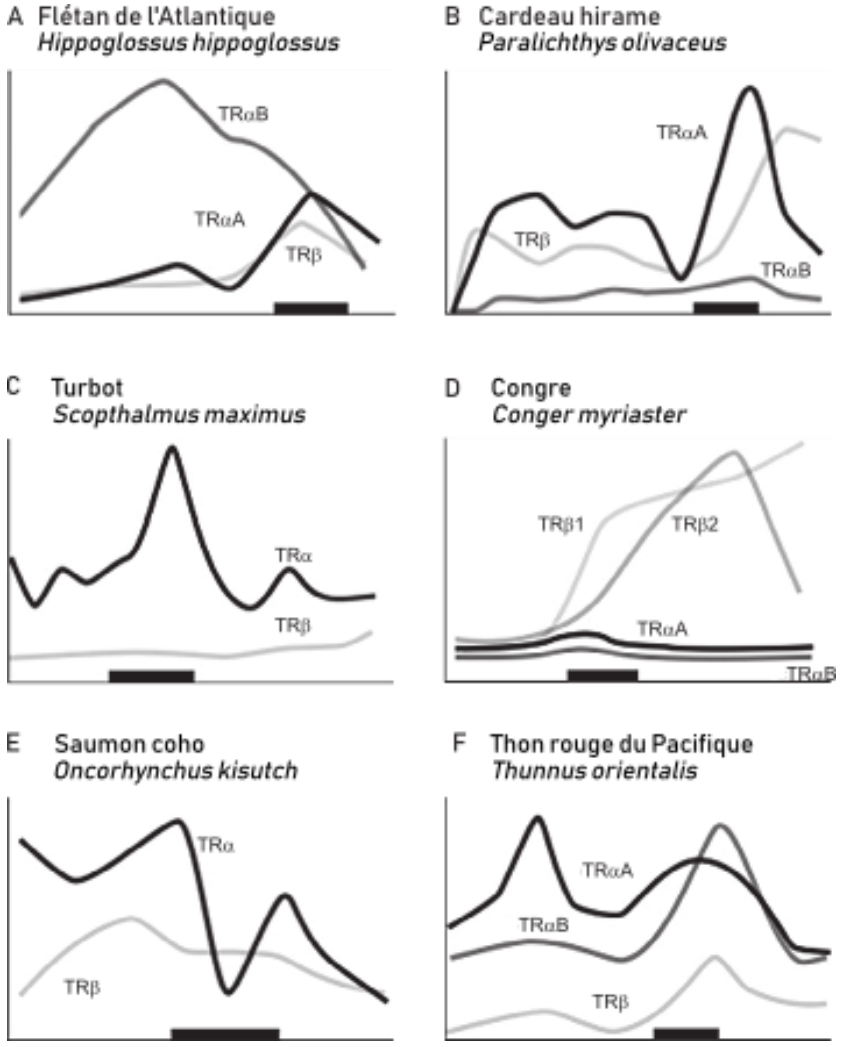


Figure 60 - Niveaux d'expression relatifs des TR au cours du développement de différents poissons Téléostéens.

L'expression a été mesurée sur larve ou juvénile total(e) pour toutes les figures, sauf pour le Saumon et le Cardeau (cerveau et tête, respectivement). La barre noire correspond au climax de la métamorphose. *Adapté de McMenamin and Parichy (2013).*

À ce titre, les récepteurs des hormones thyroïdiennes ont des profils d'expression comparables à ceux des gènes marqueurs d'iridophores. Quel que soit leur patron d'expression avant le stade 4 (peut-être lié à une population d'iridophores des stades précoces), nous avons en effet montré que les marqueurs d'iridophores ont dans l'ensemble une expression qui augmente à partir du stade 4 et jusqu'au stade 6. Cela suggère un contrôle de l'expression de ces gènes, qui vont participer à la mise en place de la bande blanche, par les hormones thyroïdiennes. Ce serait une première, à ma connaissance, car, jusqu'à présent, les études sur les traitements aux hormones thyroïdiennes des larves de poissons n'ont pas identifié d'effets sur les iridophores mais uniquement sur les xanthophores et leucophores (McMenamin *et al.*, 2014; Guillot *et al.*, 2016).

Là encore, ce travail n'est qu'exploratoire et il faudrait confirmer cela par une multiplication de l'échantillonnage. De plus, il faudrait non pas montrer une corrélation mais une relation causale entre ces voies biologiques. Cette relation est déjà morphologiquement visible, puisque des traitements aux hormones accélèrent l'apparition des bandes (**Introduction – Fig. 54**). Des traitements pharmacologiques menés par N. Roux et P. Salis à l'OOB, bloquant notamment l'action des TH, devraient permettre d'éclairer ces points dans les mois à venir.

2. Importance de l'étude du poisson-clown pour l'étude du patterning de la pigmentation

a. Hormones thyroïdiennes et patterning au cours du développement d'*Amphiprion ocellaris*

Un traitement aux hormones thyroïdiennes, s'il accélère l'apparition de la bande blanche médiane, semble également perturber la mise en place du patron global de la pigmentation puisque, cinq jours après le début du traitement, la bande blanche médiane est moins développée que chez le contrôle, et s'interrompt parfois. Elle se forme donc "plus vite, mais moins bien".

De manière générale, au cours du développement du poisson-clown *Amphiprion ocellaris*, il y a une forte variabilité intraspécifique qui se traduit par des individus aux patrons de pigmentation variables. Certains individus sont ainsi "misbar" (bande blanche médiane interrompue), ont une bande blanche médiane réduite à un point ou n'ont parfois même pas de bande blanche médiane (T. Lorin, obs. pers.). D'après MM. O. Soulat et P. Masanet, aquariophiles à Canet-en-Roussillon, il est possible d'augmenter la fréquence de ces phénotypes en modifiant les conditions d'alimentation des larves. Changement de régime alimentaire et traitement aux hormones thyroïdiennes ayant le même effet, il est donc probable que ce signal environnemental (changement du régime alimentaire) agisse sur le patron pigmentaire par l'intermédiaire des hormones thyroïdiennes.

Chez le Poisson-zèbre *D. rerio*, les modifications du métabolisme thyroïdien influent directement sur le développement des cellules pigmentaires (McMenamin *et al.*, 2014). Les xanthophores prolifèrent ainsi dans le cas d'une suractivation de la voie, soit par ajout de T4, soit par une mutation rendant les poissons hyperthyroïdiens. Si la voie est bloquée (hypothyroïdisme), les xanthophores sont moins nombreux que chez le sauvage. Il est donc possible, chez le poisson-clown, que la réduction de la bande blanche suite à un traitement à la T3 soit due, en réalité, à une sur-prolifération des xanthophores qui envahissent l'espace normalement occupé par la bande blanche. Cela irait partiellement contre l'hypothèse évoquée plus haut d'une activation de la différentiation des iridophores par ces mêmes hormones. On pourrait donc imaginer un processus en deux étapes : (i) au début de la métamorphose, les TH influeraient sur le développement des iridophores, qui initieraient le patterning des bandes ; (ii) une fois le patterning établi, les TH activeraient la prolifération des xanthophores des interbandes.

Enfin, au cours de cette étude, nous avons montré que *gja5* (*cx41.8*), un gène impliqué dans le patterning des bandes chez le poisson-zèbre, était également surexprimé dans le tissu blanc du poisson-clown. Chez *D. rerio*, le patterning dépend largement d'un système de réaction-diffusion, notamment lié à l'expression de *gja5*. Ce mécanisme de réaction-diffusion n'est pas impliqué chez le poisson-clown (Salis et al., subm.), et on peut donc s'étonner de la surexpression d'un tel gène dans le tissu blanc. Cela pourrait donc indiquer une nouvelle forme de régulation du patterning par *gja5* chez le poisson-clown, qui resterait toutefois à étudier.

b. Un contrôle par les TH pourrait expliquer la variation de la pigmentation au sein du genre *Amphiprion*

De manière intéressante, chez *D. albolineatus*, une espèce proche de *D. rerio*, les individus ont naturellement une surprolifération des xanthophores et sont considérés comme semblables à des mutants hyperthyroïdiens de *D. rerio* (McMenamin *et al.*, 2014). D'un point de vue pigmentaire, des individus de *D. albolineatus* sans thyroïde ont toujours des xanthophores alors que les TH sont nécessaires au développement des xanthophores chez *D. rerio*. Tout porte donc à croire que des changements de dosage et de sensibilité au sein de la voie des hormones thyroïdiennes aient joué un rôle dans l'évolution des patrons de pigmentation du genre *Danio*.

Dans le genre *Amphiprion,* comme évoqué en introduction **(Fig. 46)**, il existe également de multiples patrons de pigmentation, en particulier du point de vue du nombre de bandes. Il serait donc possible que des sensibilités différentes aux hormones thyroïdiennes influent sur le patterning, selon des mécanismes encore largement à étudier. Chez l'Épinoche à trois épines *Gasterosteus aculeatus*, il a été montré que des modifications en *cis* de TSH-β2 dans les écotypes d'eau douce ont eu lieu par rapport aux écotypes marins et seraient impliqués dans l'adaptation à la vie en eau douce (Kitano *et al.*, 2010). Dans le cadre du genre *Amphiprion*, les changements de patterning entre espèces proches sont parfois rapides (quelques centaines de milliers d'années à quelques millions d'années) ; on pourrait donc imaginer qu'une modification en *cis* de certains gènes de la voie des hormones thyroïdiennes intervienne également. Parmi les modifications possibles des séquences régulatrices, il en est qui ont déjà été mises en évidence dans le cas précis de la pigmentation : les éléments transposables, impliqués dans l'apparition des egg-spots chez les cichlidés (Santos *et al.*, 2014). Il serait donc intéressant d'étudier les polymorphismes d'insertion d'éléments autour des gènes de la voie des hormones thyroïdiennes comme des gènes de la pigmentation. Cela nécessitera toutefois l'obtention de nouveaux génomes au sein du genre *Amphiprion*. Avant 2018, il n'y en avait pas, et trois sont parus depuis le début de 2018 ; gageons que dans deux ou trois ans, de l'eau aura coulé sous les ponts !

Fin

ANNEXE 1

Article de vulgarisation sur les cellules des crêtes neurales

Au cours d'un de mes comités de suivi de thèse, après avoir appris que je voulais devenir enseignant, Frédéric Flamant et Marc Robinson-Rechavi m'ont tous deux suggéré d'écrire un article de vulgarisation en rapport avec ma thèse. Intéressé par la formation continue des enseignants du secondaire, j'ai souhaité écrire un article ciblant en particulier les professeurs de Sciences de la Vie et de la Terre (SVT). Intéressé par la biologie du développement et l'évolution des Vertébrés, je me suis naturellement orienté vers la rédaction d'un article traitant des cellules des crêtes neurales.

Après des retours de Jean-Nicolas Volff sur une première version de l'article, j'ai proposé ce texte à Pascal Combemorel, éditeur du site institutionnel Planet-Vie ayant pour vocation de publier des articles pour les enseignants de SVT (https://planet-vie.ens.fr/). P. Combemorel m'a grandement aidé à le rendre plus accessible pour des enseignants du secondaire, et l'article a finalement été publié au tout début du mois de mai 2018 (https://planet-vie.ens.fr/article/2528/cellules-cretes-neurales-seule-chose-interessante-vertebres).

Je joins ici l'équivalent en format pdf du texte publié en ligne. Je tiens à remercier chaleureusement Pascal Combemorel pour son travail de relecture ainsi que de mise en page *ad hoc* du pdf.

ARTICLE

Les cellules des crêtes neurales : « la seule chose intéressante » chez les Vertébrés ?

Thibault Lorin[1]

[1] École normale supérieure de Lyon

Les Vertébrés sont un groupe d'animaux comptant plus de 60 000 espèces auquel nous, humains, appartenons. Ce groupe très diversifié, apparu il y a environ 500 millions d'années, est constitué d'espèces ayant colonisé de nombreux milieux de vie sur Terre. Ils diffèrent en cela de leurs proches parents à l'écologie moins diversifiée. D'après certains auteurs, le succès évolutif des Vertébrés pourrait s'expliquer par la présence de cellules particulières qui constituent une innovation évolutive de ce groupe : les cellules des crêtes neurales. Dans cet article, nous envisageons quelques aspects du développement et du devenir de ces cellules, qui peuvent se différencier en des types cellulaires très variés. Nous discutons également de l'apport d'études récentes à la compréhension des mécanismes ayant permis l'apparition de ce type cellulaire unique, « seule chose intéressante » chez les Vertébrés.

Les Vertébrés constituent un taxon d'animaux Bilatériens de plus de 60 000 espèces qui possèdent, à au moins un moment de leur développement, des vertèbres ou des rudiments de vertèbres [1]. Parmi celles-ci, environ 30 000 sont des poissons téléostéens quand les Mammifères ne représentent qu'entre 5 000 et 6 000 espèces (Fig. 1).

Les Vertébrés sont apparus il y a 500 millions d'années [2] et ont colonisé aujourd'hui tous les milieux de vie sur Terre, depuis les grands fonds marins (la donzelle *Typhlonus nasus*) jusqu'aux sommets des montagnes himalayennes (l'oie *Anser indicus*), depuis les déserts les plus arides (le lézard *Apaurosaura anchieta*) jusque dans les contrées les plus glaciales (le manchot empereur *Aptenodytes forsteri*).

D'après certains auteurs l'innovation évolutive ayant permis la radiation des Vertébrés et l'exploitation d'une aussi grande diversité de niches écologiques est un territoire cellulaire assez restreint, ne comptant que peu de cellules pendant l'essentiel du développement embryonnaire (comparé aux autres territoires) et se mettant en place assez tard pendant le développement : les *cellules des crêtes neurales* (CCN).

Par la suite nous présenterons en premier lieu ce que sont ces cellules, et ce qu'elles deviennent au cours du développement. Nous verrons ensuite quels sont les signaux qui les induisent et comment elles migrent et se différencient. Nous discuterons enfin de l'apport de travaux récents sur l'origine évolutive de ces cellules, intimement liées à l'histoire des Vertébrés – notre histoire.

1 Les crêtes neurales : des territoires cellulaires restreints aux dérivés multiples

1.1 Les crêtes neurales se forment à la bordure de la plaque neurale

Durant le développement embryonnaire, le système nerveux central se met en place lors d'une étape appelée neurulation. Suite à l'invagination de la plaque neurale en gouttière neurale, les territoires adjacents à celle-ci se retrouvent en position haute et ont donc été appelés *crêtes neurales*. Après la fermeture du tube neural, les CCN se séparent de la partie dorsale de celui-ci – on parle de « délamination » – et commencent à migrer (Fig. 2).

Les crêtes neurales se forment tout le long du tube neural, de la partie antérieure de l'embryon (future tête) jusqu'à la partie postérieure de celui-ci (future queue). Plusieurs domaines sont distingués en fonction de leur devenir : les crêtes neurales du crâne, du cœur (parfois rattachées aux premières), du tronc et le regroupement des crêtes neurales vagales (du cou) et sacrées (Fig. 2). Ces domaines se chevauchent partiellement et des cellules provenant de la même région peuvent avoir des devenirs distincts (exemple des crêtes neurales cardiaques et vagales).

1.2 Les crêtes neurales forment des types cellulaires et des structures diverses

Les CCN ne sont présentes qu'à l'état embryonnaire, et se différencient (après migration) en de multiples types cellulaires au cours du développement (ces cellules sont dites « multipotentes »). Les dérivés de ces cellules diffèrent quelque peu selon la position de celles-ci dans l'em-

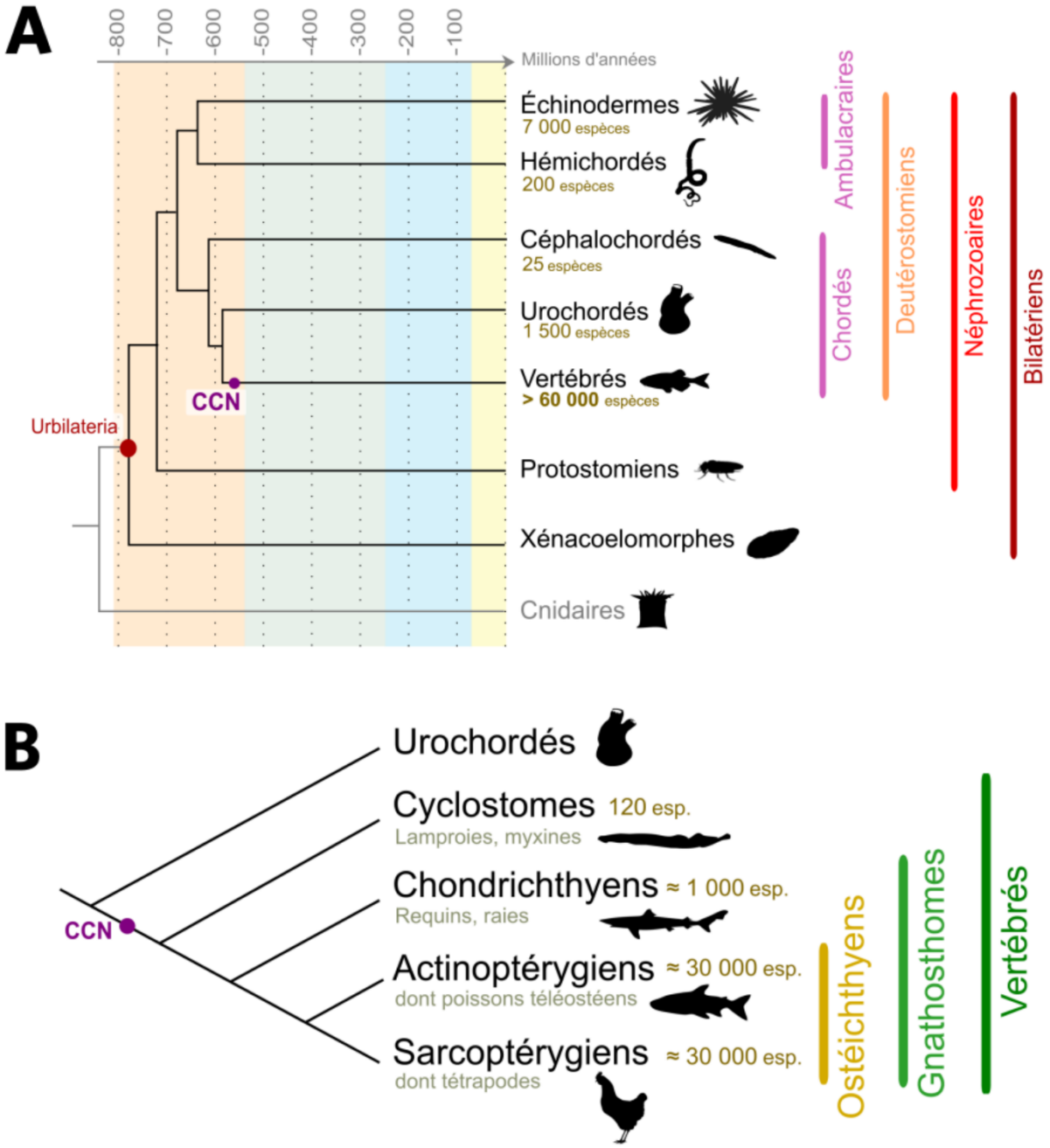


Figure 1 | Phylogénies simplifiées des Bilatériens et des Vertébrés. (A) Phylogénie simplifiée des organismes Bilatériens. L'ancêtre hypothétique des Bilatériens est appelé *Urbilateria*. Celui-ci a donné naissance aux Xénacoelomorphes, un groupe aujourd'hui réduit et dont la position systématique est discutée, et aux Néphrozoaires, eux-mêmes divisés en Protostomiens et en Deutérostomiens. Nous-mêmes, Vertébrés, appartenons aux Deutérostomiens Chordés, et les Urochordés sont, chez les Chordés, nos plus proches cousins. Un des premiers Vertébrés connu est *Myllokunmigia* (faune de Chengjang, environ 520 millions d'années). Le nombre d'espèces de Deutérostomiens ainsi que l'apparition des cellules des crêtes neurales (CCN) sont indiqués. Les âges sont en millions d'années. D'après Cannon *et al.*, 2016. (B) Phylogénie simplifiée des Vertébrés. Les Vertébrés sont divisés en deux grands taxons : les Cyclostomes (aussi appelés Agnathes, ou Vertébrés sans mâchoire) et les Gnathostomes (Vertébrés avec mâchoire). Au sein des Gnathostomes, les Chondrichthyens (poissons cartilagineux) sont le groupe frère des Ostéichthyens (poissons osseux). Les Tétrapodes constituent la majorité des Sarcoptérygiens (Vertébrés à membres charnus), le groupe frère des Actinoptérygiens (poissons à nageoires rayonnées). Le nombre d'espèces ainsi que l'apparition des cellules des crêtes neurales (CCN) sont indiqués. Auteur : Thibault Lorin

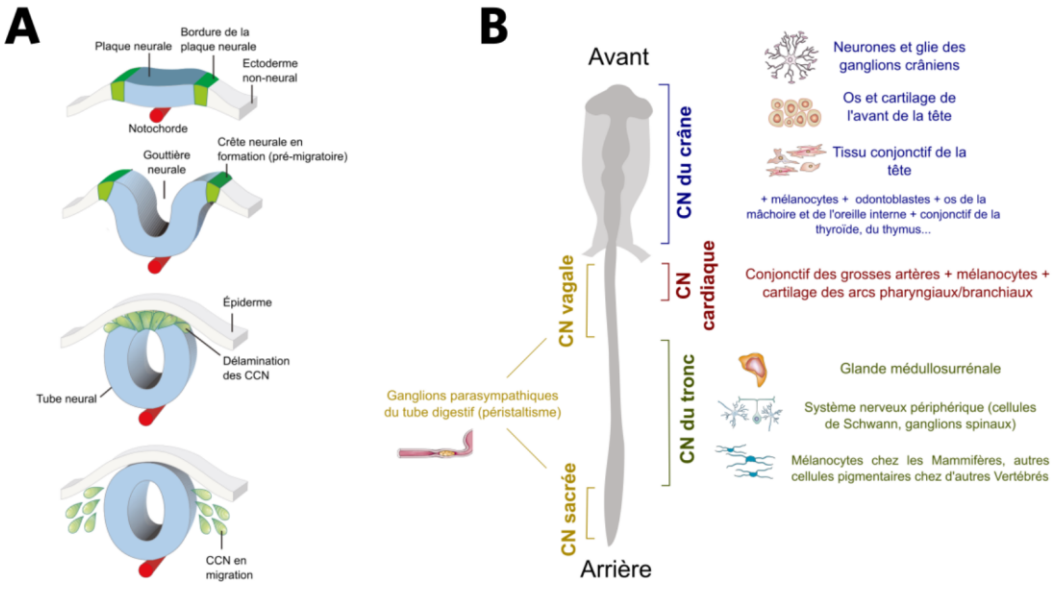


Figure 2 | Les crêtes neurales. **(A) Embryologie des cellules des crêtes neurales (CCN). Les CCN (en vert) sont originaires des bordures de la plaque neurale (un épaississement du neurectoderme, en bleu), qui, lorsqu'elles se replient au cours de la neurulation, forment des « crêtes » entourant la gouttière neurale. Après l'invagination du neurectoderme, les cellules de la partie dorsale du tube neural se détachent (délamination) et commencent à migrer : ce sont les CCN. (B) Les différents domaines de la crête neurale (CN) ainsi que leurs principaux dérivés. Adapté de Simoes-Costa *et al.*, 2015 et de Knecht *et al.*, 2002. Auteur : Thibault Lorin**

bryon (Fig. 2). Par exemple, les cellules de la région vagale et sacrée ont des potentialités assez restreintes et deviennent principalement les ganglions parasympathiques du tube digestif (importants pour le péristaltisme, c'est-à-dire les contractions musculaires synchronisées des viscères). Les cellules des crêtes neurales du tronc ont des devenirs plus variés, tandis que la palme est remportée par les CCN du crâne qui peuvent tout à la fois permettre la formation de différents tissus conjonctifs – tissus adipeux (adipocytes), cartilage (chondroblastes) et tissu osseux (ostéoblastes) – mais aussi de tissu musculaire (myocytes), de glandes sécrétrices (thyroïde, parathyroïde), de cellules nerveuses (neurones et cellules gliales) et de cellules pigmentaires (mélanocytes). On peut par ailleurs noter que ces dernières sont produites par toutes les régions de la crête neurale (crête neurale sacrée et vagale exceptées)[3].

Les cellules des crêtes neurales partagent donc une même origine développementale, des capacités migratoires et une multipotence élevée[4]. Elles sont donc caractérisées par une certaine « cohérence embryologique » qui les distinguent en cela des trois autres feuillets embryonnaires des Bilatériens : endoderme, mésoderme et ectoderme. Cela a conduit certains auteurs à les qualifier de « quatrième feuillet embryonnaire ».

Feuillets embryonnaires

Les feuillets embryonnaires sont les territoires cellulaires mis en place au cours de la gastrulation, au nombre de trois (ectoderme, mésoderme et endoderme) chez les animaux triploblastiques, de deux (ectoderme et endoderme) chez les animaux diploblastiques. Ils donneront naissance à l'ensemble des tissus et organes. Les cellules des crêtes neurales sont parfois considérées comme un quatrième feuillet.

Qui plus est, une théorie postule que l'acquisition des cellules des crêtes neurales et de leurs dérivés *aurait permis* le passage d'une physiologie de « filtreur benthique », caractéristique de l'ancêtre des Chordés (et que l'on retrouve encore actuellement chez les Céphalochordés et les Urochordés), à une physiologie « active », éventuellement prédatrice, chez l'ancêtre des Vertébrés :

— acquisition d'une vraie tête portant de nombreux organes sensoriels et, souvent, des mâchoires ;
— de cellules pigmentaires impliquées dans le camouflage et les relations inter et intra-spécifiques ;
— d'une glande médullosurrénale libérant des hormones de stress permettant l'ajustement du métabolisme à un milieu changeant rapidement…

À la cohérence embryologique évoquée plus haut se doublerait donc une cohérence « fonctionnelle » des crêtes neurales, justifiant encore davantage de les reconnaître en tant que quatrième feuillet embryonnaire[5].

Certains auteurs refusent cependant cette dénomination dans la mesure où la mise en place des crêtes neurales se fait bien plus tardivement que celle des trois autres feuillets. « Quatrième feuillet » ou non, c'est une question de nomenclature qui ne doit pas éclipser le rôle primordial de ces cellules dans la physiologie et l'écologie des Vertébrés.

2 Induction, migration et différenciation des cellules des crêtes neurales

2.1 La spécification des CCN met en jeu différents signaux

Les crêtes neurales ne sont distinguables morphologiquement qu'à partir de la formation de la gouttière neurale. Auparavant, elles ne constituent qu'un territoire cellulaire qui ne peut être reconnu anatomiquement, à l'interface entre la plaque neurale et le futur épiderme. Quand la spécification de ce territoire cellulaire a-t-telle lieu ? Au moment de la neurulation et de la formation des crêtes neurales ou avant, dès le stade gastrula ?

Des expériences menées chez le poulet ont montré que des cellules de gastrula, explantées hors de l'embryon et cultivées seules, pouvait déjà se différencier en cellules pigmentaires produisant de la mélanine (mélanocytes)[6]. Cela signifie donc que la spécification des cellules des crêtes neurales a lieu avant même que la crête neurale ne soit distinguable dans l'embryon.

Spécification et détermination

Un territoire est dit *spécifié* si, lorsqu'on l'explante hors de l'embryon, il devient ce qu'il serait devenu en restant sur l'embryon : il conserve son devenir présomptif. Par contre, greffé ailleurs (hétérotopiquement) sur l'embryon, son devenir change et devient celui de sa nouvelle localisation.
Un territoire est dit *déterminé* s'il est spécifié et si, greffé sur un autre embryon (ou ailleurs sur le même), il conserve son devenir présomptif.
Ainsi, un territoire déterminé est spécifié, mais un territoire spécifié n'est pas pour autant déterminé.

Les signaux moléculaires intervenant dans la spécification de la future crête neurale sont sécrétés par l'ectoderme ventral ainsi que par certaines régions du mésoderme. Ce sont des protéines telles que les BMP (*bone morphogenetic proteins*), Wnt, et les FGF (*fibroblast growth factors*). Elles activent l'expression de gènes d'identité des crêtes neurales, tels que les facteurs de transcription *Pax7* et *Sox10* (Fig. 3)

2.2 La migration des cellules des crêtes neurales le long de l'axe dorso-ventral

Toutes les cellules des crêtes neurales effectuent des migrations. Celles-ci peuvent être suivies en fusionnant un gène marqueur de ces cellules (par exemple, *Pax7*) avec un gène codant la GFP (*green fluorescent protein*, ou protéine fluorescente verte). Ainsi, les cellules exprimant *Pax7* seront fluorescentes. Les résultats montrent que les CCN migrent depuis le côté dorsal (position du tube neural chez les Vertébrés) vers le côté ventral de l'organisme (Fig. 4).

Comment des cellules qui étaient situées dans la crête neurale, et donc intégrées au sein d'un tissu embryonnaire (le neurectoderme), acquièrent-elles des propriétés migratrices et se déplacent-elles jusqu'à leur position finale dans l'organisme ? Ce processus migratoire est constitué de plusieurs étapes : (1) le détachement du tissu initial (2) l'acquisition de capacités migratoires par la cellule (3) la réception de signaux indiquant vers où migrer (4) la migration *per se* et (5) la réception de signaux indiquant que la destination est atteinte et que la migration doit s'arrêter.

La transition épithélio-mésenchymateuse

La perte d'adhérence entre les cellules d'un épithélium et l'acquisition par celles-ci de propriétés migratoires est un événement qualifié de transition épithélio-mésenchymateuse (TEM). Ce processus a lieu de manière normale au cours du développement, par exemple dans le cas des cellules des crêtes neurales ou lors de la gastrulation, mais peut également avoir lieu de manière anormale dans certains cancers dont les cellules prolifèrent et se mettent à migrer (métastases)[7]. L'étude de la TEM est donc un champ d'investigation actuellement très actif.

Toute perturbation d'une seule de ces étapes peut causer des défauts de migration et des maladies. Par exemple, le défaut de migration des cellules des crêtes neurales à l'origine des futurs ganglions nerveux du tube digestif peut provoquer un dysfonctionnement du péristaltisme du côlon et causer la maladie de Hirschprung, dont les principaux symptômes sont une mauvaise évacuation des selles et une constipation. Cette maladie est, chez certaines espèces, parfois associée à des défauts de pigmentation du pelage, indiquant que les mélanocytes (d'autres dérivés de CCN) sont également affectés.

2.3 Détachement du neurectoderme et acquisition des propriétés migratoires

Alors qu'elles sont encore intégrées dans la partie dorsale du tube neural, les cellules des crêtes neurales reçoivent des signaux inducteurs de ce dernier – signaux déjà à l'origine de la spécification de ces cellules – les BMP et Wnt. Ceux-ci activent des gènes cibles qui induisent une diminution de l'expression des cadhérines (des molécules d'adhérence en surface de la cellule) et la dissociation des cellules du tissu.

Par ailleurs, Wnt active l'expression de petites protéines GTPases (qui hydrolysent le GTP) de la famille des Rho GTPases. Ces protéines permettent de réorganiser le cytosquelette d'actine, nécessaire à la future migration cellulaire (Fig. 5). Le signal qui déclenche la migration est quant à lui original : les CCN se repoussent en effet les unes les autres par **inhibition de contact** (« je t'aime » / « moi non plus »). Dès qu'elles se touchent, les deux partenaires repartent à l'opposé l'une de l'autre. Chaque cellule voulant éviter les autres, cela provoque une « fuite en avant » et un mouvement en masse de ce type cellulaire, en direction ventrale, depuis le tube neural.

2.4 Les cellules des crêtes neurales empruntent différentes voies de migration

Les tissus que les cellules des crêtes neurales traversent sont formés de cellules possédant dans leur membrane des signaux moléculaires spécifiques (éphrines, sémaphorines) auxquels chaque CCN répond différemment. Selon son état de différenciation et les molécules de surface qu'elle-même porte, elle continue son chemin dans le tissu ou elle l'évite. La combinaison des facteurs tissulaires et propres aux CCN est donc à l'origine des voies de migration empruntées.

Les modalités de migration des cellules des crêtes neurales sont variables et diffèrent selon les espèces de Vertébrés, selon la temporalité de leur migration (plus ou moins précoce) et selon leur position le long de la crête neurale (plus ou moins antérieure). Toutefois, deux voies principales de migration sont classiquement distinguées : une

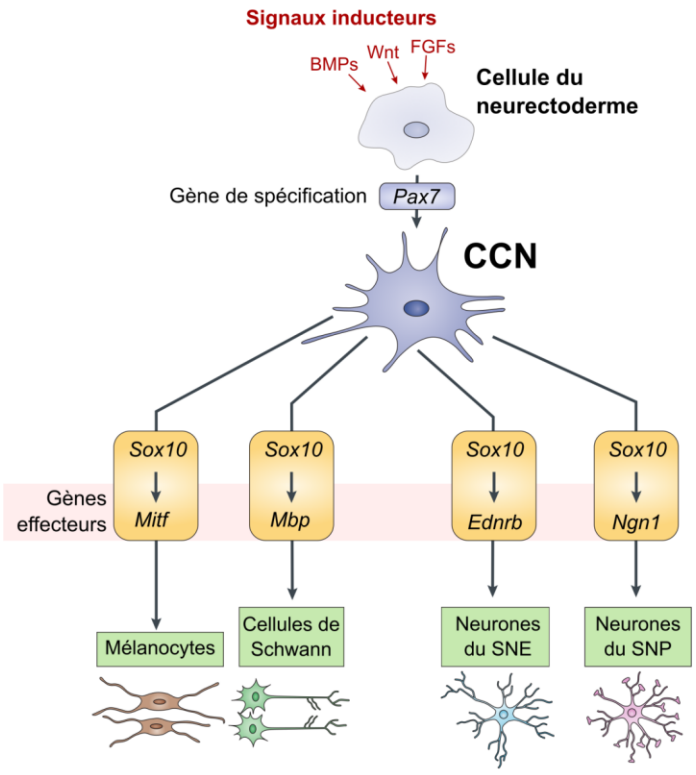


Figure 3 | Spécification des cellules des crêtes neurales (CCN) au cours du développement. D'après Basch et *al.*, 2006, Sauka-Spengler *et al.*, 2008 et Gilbert *et al.*, 2014. Auteur : Thibault Lorin

voie **ventrale** (les CCN se dirigent d'abord ventralement, à l'opposé du tube neural), et une voie **dorso-latérale** (les CCN se dirigent le long du futur dos).

Il n'est pas possible de détailler ici toutes les voies empruntées par chacun des dérivés de CCN (elles sont trop nombreuses). Nous présentons donc ici les principales voies migratoires empruntées par les mélanoblastes (précurseurs des mélanocytes) au cours du développement chez la souris *Mus musculus* (Fig. 6). Une première voie dorso-latérale est empruntée, entre les somites et l'épiderme en développement. Plus tard au cours du développement, les mélanoblastes rejoignent l'épiderme chez les Mammifères. Les mélanocytes adultes ont toutefois chez la souris une deuxième origine : ils peuvent en effet être issus de la différenciation de cellules précurseures des cellules de Schwann, qui constituent chez les Vertébrés la gaine de myéline des prolongements des neurones du système nerveux périphérique[8]. Ces cellules précurseures empruntent quant à elles une voie de migration ventrale.

2.5 La différenciation des cellules des crêtes neurales en de multiples types cellulaires

Après leur migration, des combinaisons spécifiques de facteurs de transcription restreignent peu à peu le destin des cellules des crêtes neurales en activant l'expression de **gènes effecteurs**. Par exemple, les cellules exprimant le facteur de transcription Mitf (produit du gène *microphtalmia*), en plus de Sox10, deviendront des mélanocytes. Les cellules exprimant la neurogénine 1, un autre facteur de transcription, deviendront quant à elles des neurones sensoriels du système nerveux périphérique[9] (Fig. 3). C'est une combinaison de facteurs de transcription qui peu à peu permet une restriction des destins cellulaires jusqu'à un type cellulaire différencié.

Cependant, les potentialités d'un précurseur sont-elles intrinsèques à celui-ci (détermination autonome), déterminées par l'environnement cellulaire ou influencées par ces deux facteurs ?

Des expériences de transplantation ont montré que quand des cellules des crêtes neurales du tronc (ne donnant normalement pas de cartilage) sont transplantées au niveau de la crête neurale du crâne, elles migrent et atteignent les sites de formation de cartilage mais ne se différencient jamais en cartilage. Cependant, elles peuvent former des mélanocytes, des neurones ou de la glie. Ainsi, les cellules des crêtes neurales du tronc semblent être privées d'un potentiel de formation de cartilage (le tissu osseux est également concerné). Cela est lié à leur position initiale dans l'embryon : plus postérieures, l'expression des gènes *Hox* y est différente, et leurs capacités de formation de tissu cartilagineux et osseux est perdue. D'ailleurs, si des explants de CCN du tronc sont cultivés en présence de facteurs qui leur font perdre l'expression de Hoxd10 (qu'elles expriment normalement), ces cellules se différencient sur le long terme, en deux semaines, en chondroblastes – les cellules à l'origine du cartilage[10].

Les cellules de la crête neurale du crâne peuvent quant à elles donner du cartilage en plus de tous les autres dérivés issus de la crête neurale du tronc. Toutefois, transplantées au niveau du tronc, elles se différencient dans tous les types cellulaires de cette région de l'embryon mais *pas* en cellules osseuses ni cartilagineuses : elles manquent donc des signaux qui induiraient leur différenciation en ces types cellulaires, et leur détermination est donc partiellement environnementale.

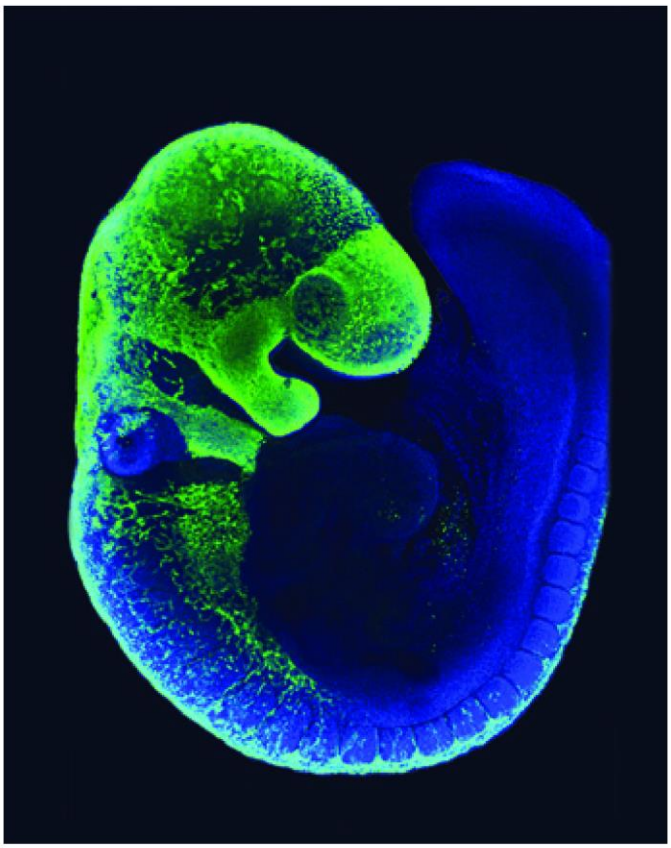


Figure 4 | Cellules des crêtes neurales (CCN) en migration. Les CCN ont été marquées par fluorescence verte (protéine fluorescente verte, ou GFP). On peut observer qu'elles forment une bande continue du côté dorsal, depuis laquelle émergent des files de cellules qui se déplacent ventralement. La tête est en particulier recouverte de CCN, soulignant le rôle essentiel que ces cellules jouent dans la formation de la tête des Vertébrés (os et cartilage, tissu conjonctif, ganglions crâniens…). Cette image a été fournie par Amanda Barlow et Paul Trainor, Stowers Institute for Medical Research www.stowers.org

3 L'origine évolutive des cellules des crêtes neurales éclaire l'évolution des Vertébrés

Les dérivés de CCN sont, on l'a vu, partout ! Il semble qu'il n'y ait pas un endroit du corps où, en cherchant à quelques cellules alentour, on ne tombe sur un dérivé de CCN… Mais comment les organismes faisaient-ils « avant » la crête neurale ? En d'autres termes, si ce type cellulaire semble être une innovation majeure apparue à la base des Vertébrés, existe-t-il des cellules leur ressemblant ailleurs chez les Chordés ? Chez les Bilatériens (Fig. 1) ? Quels sont les événements ayant pu provoquer l'apparition d'un nouveau type cellulaire ? Et enfin, existe-t-il des variations de ce « quatrième feuillet embryonnaire » caractéristique des Vertébrés ?

3.1 Des cellules ayant des caractères de CCN chez des Chordés autres que les Vertébrés

Les CCN sont une innovation évolutive des Vertébrés et la question de leur origine évolutive interpelle. Sont-elles un nouveau type cellulaire apparu *ex nihilo* ? À l'inverse, trouve-t-on des cellules qui partagent quelques caractéristiques de CCN chez les autres Chordés ? Enfin, si les dérivés de CCN sont parfois des types cellulaires nouveaux (par exemple, les odontoblastes produisant la dentine), ils ressemblent souvent à des types cellulaires déjà existants (adipocytes, chondroblastes, fibroblastes, neurones, myocytes…). Il serait donc possible d'imaginer que ces types cellulaires aient pu se retrouver à l'emplacement de la crête neurale et aient acquis des propriétés migratrices. Auquel cas la crête neurale ne serait pas vraiment une « innovation évolutive » mais plutôt le rassemblement d'une mosaïque d'éléments déjà présents. L'étude des cellules des autres Chordés, dont les Urochordés, permet d'éclairer ces questions.

Les autres Chordés partagent un plan d'organisation similaire : un système nerveux centrale en position dorsale ; une notochorde ; un tronc segmenté. Cependant, jusqu'à ces dernières années, aucune lignée cellulaire équivalente aux CCN n'avait été trouvée chez ces organismes.

En 2012, une étude a montré que chez l'Urochordé *Ciona intestinalis* certaines cellules pigmentées mélaniques trouvent leur origine au bord de la plaque neurale et migrent ensuite… mais très peu, sur une distance de l'ordre de quelques cellules seulement ! L'expression artificielle d'un seul gène, *twist*, peut cependant induire une migration plus importante de ces cellules [11].

En 2015, une deuxième étude a montré, toujours chez *C. intestinalis*, qu'il existait un type de neurones sensoriels, appelés neurones bipolaires à queue (*bipolar tail neuron*, ou BTN, en anglais), qui (1) proviennent de cellules de la bordure de la plaque neurale, (2) migrent, (3) se différencient en neurones sensoriels entre des cellules sensorielles épidermiques et des moto-neurones [12]. On s'approche ici de la définition de cellule de crête neurale !

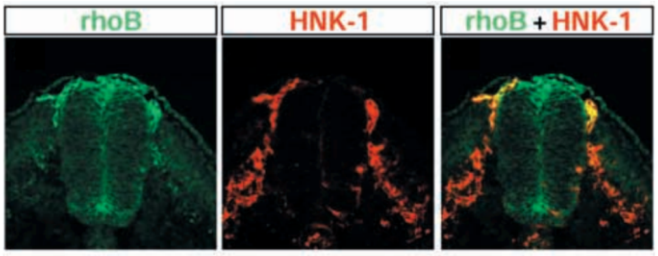


Figure 5 | Coupe transversale d'un embryon de poulet révélant par immunofluorescence la localisation des protéines rhoB et HNK-1. Les cellules des crêtes neurales en migration expriment la protéine de surface HNK-1 (au milieu, en rouge), alors que les CCN en préparation pour la migration, dans la partie dorsale du tube neural, expriment fortement rhoB (à gauche, en vert). Les cellules du tube neural, plus ventrales, ainsi que d'autres types cellulaires expriment également rhoB, mais faiblement. La fluorescence jaune (à droite) traduit la colocalisation des deux protéines. On constate que les cellules qui migrent vers le côté ventral de l'embryon perdent l'expression de rhoB. Cette protéine est nécessaire à la ré-organisation du cytosquelette en préparation à la migration des CCN. Auteur : J.P. Liu, T.M. Jessell. Source : Development
http://dev.biologists.org/content/125/24/5055

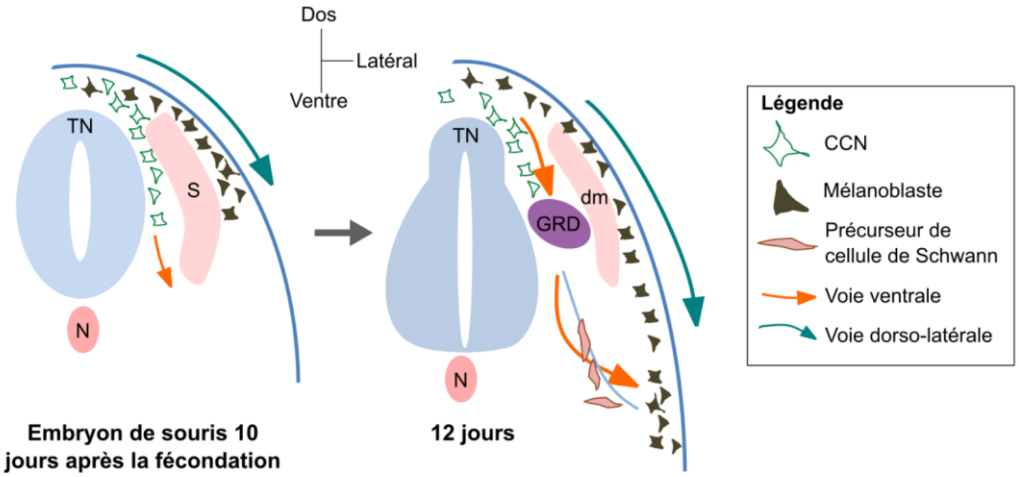


Figure 6 | Les mélanocytes adultes chez la souris proviennent de deux types principaux de voies de migration des cellules des crêtes neurales (CCN) : la voie dorsolatérale et la voie ventrale. Les CCN précurseurs des mélanocytes empruntent soit une voie dorsolatérale directe, soit une voie ventrale indirecte. Dans ce second cas, les mélanocytes proviennent alors de la différenciation de précurseurs des cellules de Schwann qui migrent le long des nerfs (représentés par un trait bleu clair). S = somite; dm = dermomyotome (partie des somites qui forme le derme et les muscles du dos et des membres); N = notochorde; TN = tube neural; GRD = ganglion de la racine dorsale. Adapté de Mort *et al.*, 2015. Auteur : Thibault Lorin

Les auteurs proposent donc que l'ancêtre commun aux Urochordés et aux Vertébrés avait déjà une plaque neurale dont les bords produisaient des cellules pigmentaires (cf. étude de 2012) ainsi que des cellules ayant déjà des potentialités migratrices (cf. étude de 2015). Les auteurs concluent que la crête neurale a évolué par l'acquisition d'un progéniteur multipotent « en amont » de progéniteurs multiples qui étaient déjà présents dans la plaque neurale mais qui avaient des devenirs plus restreints (Fig. 7).

Cooptation

La cooptation correspond à l'acquisition par un gène ou un réseau de gènes de nouvelles fonctions, différentes de leurs fonctions ancestrales. Par exemple, si les gènes Hox sont utilisés dans la spécification de l'axe antéro-postérieur chez les Bilatériens, ils ont été « réutilisés » pour la détermination de la polarité proximo-distale des membres des Tétrapodes.

Les recherches n'ont pour l'instant pas montré l'existence de cellules des crêtes neurales chez les autres Deutérostomiens (Céphalochordés y compris) : de nombreux types cellulaires ayant des similitudes partielles ont été discutés au sein de la communauté scientifique, et peut-être que les années à venir éclaireront ce point sous un jour nouveau.

L'étude de l'origine des CCN est un champ de recherche actuellement très actif; ainsi, aujourd'hui, on ne cherche plus uniquement les *cellules* qui, chez les « invertébrés » (Bilatériens non-Vertébrés), pourraient ressembler aux CCN, mais des *gènes*, ou plutôt des réseaux de régulation génique (*gene regulatory network* = GRN) impliqués dans la différenciation de ce type cellulaire (un réseau de régulation génique est un ensemble de régulateurs moléculaires qui interagissent entre eux et avec d'autres molécules pour contrôler l'expression de gènes). Il a ainsi été récemment démontré que les cellules de la bordure de la plaque neurale expriment presque les mêmes gènes chez tous les Bilatériens, qu'ils soient Protostomiens ou Deutérostomiens. Les mécanismes moléculaires nécessaires à la spécification de la bordure de la plaque neurale en neurones périphériques étaient donc déjà proba-

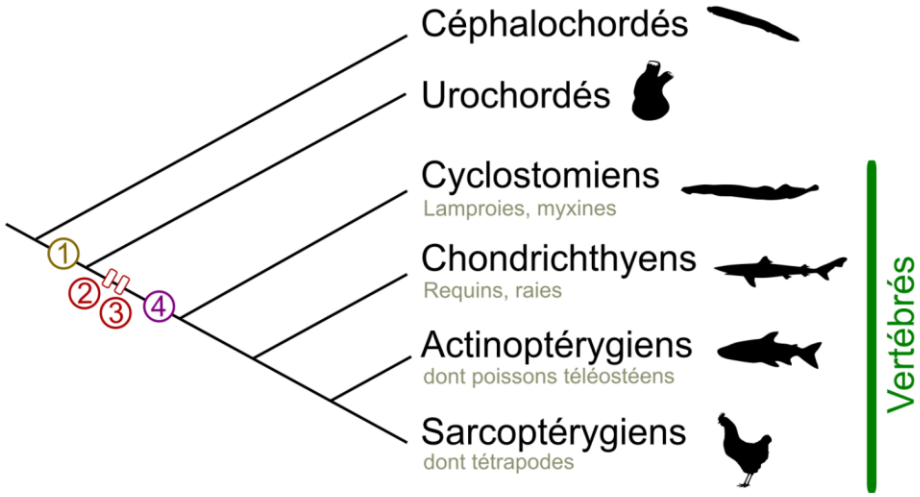


Figure 7 | Phylogénie simplifiée des Chordés et principaux événements concernant les cellules des crêtes neurales (CCN). 1. Acquisition de cellules possédant quelques caractères de cellules des crêtes neurales (ces cellules proviennent de la bordure de la plaque neurale, migrent mais ne se différencient qu'en quelques types cellulaires). 2. 1R : première duplication du génome. 3. 2R : seconde duplication du génome. 4. Acquisition de cellules des crêtes neurales au sens strict : ces cellules donnent des dérivés beaucoup plus nombreux que ceux retrouvés chez les Urochordés. Auteur : Thibault Lorin

blement présents chez l'ancêtre des Bilatériens (« Urbilateria »)[13]. De plus, dans l'ensemble, les gènes de spécification des CCN en d'autres types cellulaires sont présents chez tous les Bilatériens.

En conclusion, il est vraisemblable que les CCN soient issues de la co-optation de divers programmes génétiques de différenciation (en neurones, mélanocytes, cellules gliales…) par des cellules aux potentialités migratrices. Ces programmes de différenciation auraient été rajoutés, tels des « plug-ins », à un programme de spécification de progéniteur de crête neurale[14].

3.2 Les événements évolutifs ayant permis la formation de CCN

La formation d'un nouveau type cellulaire nécessite des modifications majeures des réseaux de régulation génique : des groupes entiers de gènes (modules), encodant une fonction donnée (par exemple, la différenciation en neurone), doivent être mis en lien avec des modules de migration cellulaire et des modules de multipotence, pour qu'une seule cellule puisse assurer toutes ces fonctions. Si on assimile un module fonctionnel à une « boîte », cela nécessite donc de ré-agencer et connecter des boîtes originellement indépendantes (et mises en œuvre dans des types cellulaires différents), en un ensemble de boîtes cohérent dans une seule cellule… Pas si simple !

L'origine de tels remaniements des réseaux de régulation génique fait toujours débat, mais la double duplication de génome à la base des Vertébrés est un événement majeur ayant pu conduire à de tels changements.

En effet, après la séparation des lignées Urochordés / Vertébrés, une double duplication de génome a eu lieu au sein de la lignée des Vertébrés. Cette double duplication de génome (appelée 1R/2R, pour « premier round puis second round de duplication »), encore imparfaitement comprise, a eu pour conséquence de multiplier par deux, puis encore par deux (donc par quatre) l'ensemble des gènes de notre ancêtre. Aujourd'hui, 500 millions d'années plus tard, nombre de ces gènes ont été perdus, mais il reste dans notre génome des gènes paralogues (deux gènes sont dits paralogues s'ils sont issus de la duplication d'un gène ancestral) issus de ces deux duplications. Par exemple, la portion de chromosome contenant les gènes homéotiques est présente en une seule copie chez les Ciones (*Ciona sp.*, Urochordés), contre 4 chez les Vertébrés.

Vertébrés et duplications de génome

Au cours de l'évolution des Vertébrés, il n'y a pas eu que deux événements de duplication de génome. Par exemple, au sein des tétrapodes, chez les Amphibiens, la lignée des xénopes a elle aussi été marquée par une duplication de génome (voir ou revoir à ce propos L'histoire évolutive du xénope révélée par son génome). De même, à la base des poissons Téléostéens, une duplication de génome a eu lieu il y a environ 350 Ma, et au sein même de ce groupe, il y a encore eu des événements de duplication de génome, par exemple chez les Salmonidés (saumons, truites…) il y a 100 Ma[1]. Ainsi, les duplications de génomes n'ont pas eu lieu uniquement à la base des Vertébrés mais ont concerné de nombreux taxons au cours de l'évolution.

Le lien entre ces duplications et les réseaux de régulation génique n'est pas clair, mais un mécanisme proposé serait que le quadruplement de l'ensemble des gènes d'un organisme ait pu permettre la modification en profondeur des interactions entre les gènes des différents modules, voire même mettre des modules sous contrôle d'autres modules. En somme, l'ensemble des gènes issus d'une duplication de génome serait

un « substrat » à partir duquel la variation serait possible : l'une des copies d'un gène serait libre de varier, tandis que l'autre copie pourrait garder la fonction ancestrale du gène (ou également varier)[15]. Cela est non seulement vrai pour les séquences codantes mais également pour les séquences régulatrices.

Un scénario possible à l'origine des cellules des crêtes neurales serait donc que des copies de gènes originellement indépendants (dans des modules différents) passent sous le contrôle de séquences régulatrices similaires et deviennent ainsi co-régulés (Fig . 8).

3.3 L'évolution des CCN au sein des Vertébrés

La crête neurale des Vertébrés, si elle est une de leur spécificité, n'est pas pour autant une entité immuable au sein de ce groupe. Les types cellulaires produits par les crêtes neurales des différents taxons de Vertébrés ne sont pas toujours les mêmes. Par exemple, chez les poissons téléostéens, il existe de multiples types de cellules pigmentaires, en plus des mélanocytes : des cellules iridescentes (iridophores), rouge-orange (érythrophores et xanthophores) et des cellules blanches (leucophores). Ces cellules dérivent toutes d'un même précurseur mais accumulent des pigments différents[16] (Fig. 9).

En particulier, on l'a vu, la crête neurale du tronc a un devenir plus restreint que la crête neurale du crâne et n'est pas capable de produire de l'os et du cartilage… ou plutôt, elle ne le *serait plus*. Certains auteurs (et c'est là encore discuté) affirment ainsi que les CCN du tronc ont perdu leur capacité à produire des tissus durs chez la majorité des Vertébrés, et que les plaques dermiques de certains Vertébrés fossiles, comme les « ostracodermes » (groupe non-monophylétique) et les placodermes, auraient pu provenir des crêtes neurales thoraciques[17]. L'avantage évolutif d'une telle perte, s'il y en a un, n'est pas documenté. Des travaux récents sur les Chondrichthyens (« poissons cartilagineux ») montrent que la dentine des denticules dermiques (petites dents sur le corps des requins et des raies) provient d'odontoblastes issus de cellules des crêtes neurales du tronc[17]. Ainsi, ces cellules paraissent avoir conservé une fonction ancestrale de production de dentine uniquement chez les requins et raies.

Enfin, chez les lamproies (des Cyclostomes), les ganglions digestifs semblent être issus des cellules des crêtes neurales du tronc et non des vagales et sacrées comme chez les Gnathostomes[18]. Les auteurs suggèrent que les CCN vagales ont acquis cette fonction à la base des Gnathostomes, montrant là encore une différence entre CCN des différentes lignées de Vertébrés.

Conclusion

L'étude des cellules des crêtes neurales et de leurs dérivés est donc à l'interface entre la biologie cellulaire et du développement, l'évolution et l'écologie, et s'avère assez fascinante. Ces cellules, par leurs potentialités migratrices et les multiples dérivés qu'elles génèrent (multipotence), ont pu contribuer à l'acquisition de structures développementales propres aux Vertébrés : crâne, mâchoire, glandes hormonales (surrénales), structures sensorielles (oreille interne)… Celles-ci ont alors permis à ces organismes d'exploiter de nouvelles niches écologiques. Des variations de ces structures ont permis l'adaptation des populations aux caractéristiques (physico-chimiques et biologiques) de leur milieu, fournissant en cela le substrat permettant la diversification et, à terme, la spéciation au sein du clade des Vertébrés. C'est à ce titre que certains auteurs ont qualifié les cellules des crêtes neurales de « seule chose intéressante chez les Vertébrés » (*the only interesting thing about vertebrates*) (Thorogood, 1989), et qu'on parle parfois aujourd'hui pour désigner ce territoire cellulaire de « quatrième feuillet embryonnaire ».

L'évolution des CCN est complexe et semble être due à des changements de réseaux de régulation génique provoqués par les événements de duplication de génome survenus à la base des Vertébrés. Ces événements ont pu induire l'acquisition, par un seul type cellulaire, de modules fonctionnels qui étaient ancestralement présents dans des types cellulaires différents. L'évolution des CCN est toutefois toujours en cours, et les grandes lignées de Vertébrés ont des CCN qui ne possèdent pas les mêmes spécificités.

On ne saurait toutefois résumer le succès évolutif des Vertébrés à ce seul groupe de cellules[19]. En particulier, les placodes ectodermiques crâniennes constituent un autre type cellulaire dont dérivent certaines structures sensorielles de la tête (nerfs olfactifs, oreille interne, bourgeons du goût, cristallin). Ces structures, embryologiquement proches des cellules des crêtes neurales (également issues de la bordure de la plaque neurale, mais, elles, ectodermiques, et non pas neurectodermiques), ont probablement elles aussi joué un rôle dans la diversification de notre lignée et dans l'adaptation aux multiples niches écologiques.

Lexique

Facteur de transcription : une protéine possédant un domaine de fixation à l'ADN et influant sur la transcription d'au moins un gène cible.

Gènes homéotiques : les gènes homéotiques (ou gènes *Hox*) sont des gènes impliqués dans le développement et l'agencement relatif des grandes structures de l'organisme.

Poisson : le terme « poisson » regroupe tous les animaux à nageoires capables de se déplacer activement dans l'eau. Ce terme ne désigne pas un groupe monophylétique mais plutôt une réalité écologique. Les Téléostéens, en revanche, constituent un taxon exclusivement représenté par des poissons et représentent la moitié des espèces de Vertébrés.

Taxon : groupe d'espèces qui partage un ancêtre commun exclusif (monophylétique). Un taxon peut contenir de nombreuses espèces (le taxon des Bilatériens contient plusieurs millions d'espèces) ou restreint (le taxon des Hominidés ne contient que six espèces).

Références

1. Berthelot, C, Brunet, F, Chalopin, D, Juanchich, A, Bernard, M, Noël, B, Bento, P, Da Silva, C, Labadie, K, Alberti, A, Aury, J.-M, Louis, A, Dehais, P, Bardou, P, Montfort, J, Klopp, C, Cabau, C, Gaspin, C, Thorgaard, G. H, Boussaha, M, Quillet, E, Guyomard, R, Galiana, D, Bobe, J, Volff, J.-N, Genêt, C, Wincker, P, Jaillon, O, Crollius, H. R, & Guiguen, Y. (2014) The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nature Communications* **5**.
2. Janvier, P. (2011) Comparative Anatomy : All Vertebrates Do Have Vertebrae. *Current Biology* **21**, R661--R663.
3. Cichorek, M, Wachulska, M, Stasiewicz, A, & Tyminska, A. (2013) Skin melanocytes : biology and development. *Advances in Dermatology and Allergology* **1**, 30--41.
4. Hall, B. K & Gillis, J. A. (2013) Incremental evolution of the neural crest, neural crest cells and neural crest-derived skeletal tissues. *Journal of Anatomy* **222**, 19--31.
5. Gans, C & Northcutt, R. G. (1983) Neural Crest and the Origin of Vertebrates : A New Head. *Science* **220**, 268--273.
6. Basch, M. L, Bronner-Fraser, M, & García-Castro, M. I. (2006) Specification of the neural crest occurs during gastrulation and requires Pax7. *Nature* **441**, 218--222.
7. Brabletz, T, Kalluri, R, Nieto, M. A, & Weinberg, R. A. (2018) EMT in cancer. *Nature Reviews Cancer* **18**, 128--134.

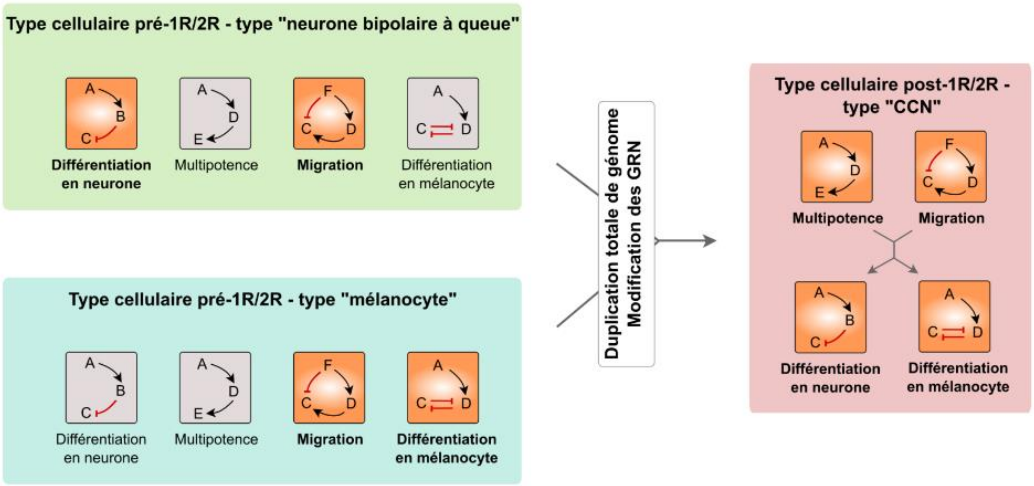


Figure 8 | Exemple de modification possible de réseau de régulation génique (GRN) après la double duplication de génome à la base des Vertébrés. Des types cellulaires ancestraux (deux sont représentés, en vert et en bleu) avaient chacun des potentialités différentes en fonction des modules de gènes exprimés (représentés en orange; les modules inactifs sont en gris). Chaque module est constitué de plusieurs gènes, représentés par des lettres, qui sont impliqués dans la même fonction cellulaire, avec des interactions propres entre les gènes d'un même module : activation (flèche noire) ou inhibition (flèche rouge). On note qu'un gène peut être impliqué dans plusieurs voies et donc être présent dans plusieurs modules (par exemple, le gène A). Les modules activés diffèrent selon le type cellulaire. Par exemple, les cellules précurseures des neurones bipolaires à queue ont deux modules activés : « migration » et « différenciation en neurone ». Lors du double événement de duplication de génome à la base des Vertébrés (1R/2R), les relations entre les modules ont pu être modifiées et certains modules ont été activés ou se sont retrouvés sous le contrôle d'autres modules. Par exemple, le module « multipotence » a été activé et les modules de différenciation ont été mis sous le contrôle des modules « multipotence » et « migration ». C'est une explication possible (certes théorique et difficile à tester) pour l'apparition d'un nouveau type cellulaire. CCN : cellules des crêtes neurales. Auteur : Thibault Lorin

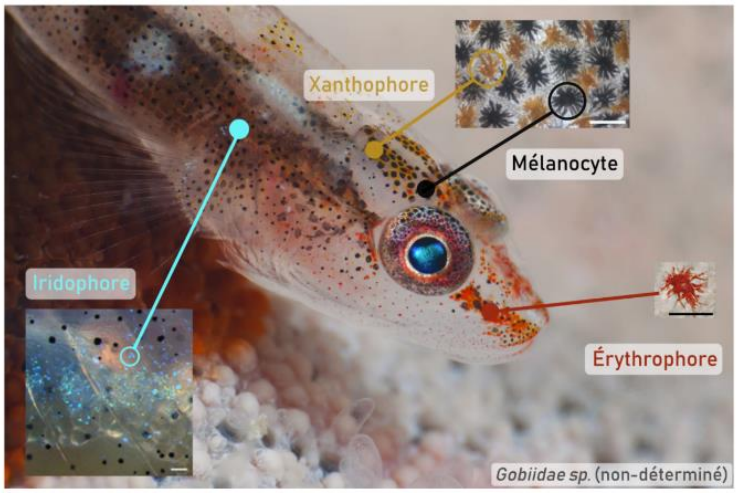


Figure 9 | Les cellules des crêtes neurales ont des devenirs variables au sein des Vertébrés : exemple des cellules pigmentaires. Les cellules des crêtes neurales sont à l'origine des cellules pigmentaires. Alors que les Mammifères n'en ont qu'un seul type (les mélanocytes), les poissons téléostéens peuvent en compter plusieurs. Par exemple, chez certaines larves de poissons coralliens (ici une espèce indéterminée de Gobiidae), on peut observer quatre types de cellules pigmentaires : des mélanocytes noirs, à l'instar des Mammifères, mais également des xanthophores (jaune-orange), des érythrophores (rouges) et des iridophores (argentés). La larve mesure environ 2 cm de longueur totale (1 cm visible sur le cliché) ; barres d'échelles des encarts : 50 µm. Droits réservés pour la photo de larve : Germain Boussarie. Référence de l'encadré « iridophores » : http://dev.biologists.org/content/140/14/2997. Référence de l'encadré « M/X » : http://jeb.biologists.org/content/early/2016/06/15/jeb.139501. Référence de l'encadré « érythrophore » : https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0749.2005.00267.x. Compilation : T. Lorin

8. Jessen, K. R, Mirsky, R, & Lloyd, A. C. (2015) Schwann Cells : Development and Role in Nerve Repair. *Cold Spring Harbor Perspectives in Biology* **7**, a020487.
9. Sauka-Spengler, T & Bronner-Fraser, M. (2008) A gene regulatory network orchestrates neural crest formation. *Nature Reviews Molecular Cell Biology* **9**, 557--568.
10. Abzhanov, A, Tzahor, E, Lassar, A. B, & Tabin, C. J. (2003) Dissimilar regulation of cell differentiation in mesencephalic (cranial) and sacral (trunk) neural crest cells in vitro. *Development* **130**, 4567--4579.
11. Abitua, P. B, Wagner, E, Navarrete, I. A, & Levine, M. (2012) Identification of a rudimentary neural crest in a non-vertebrate chordate. *Nature* **492**, 104--107.
12. Stolfi, A, Ryan, K, Meinertzhagen, I. A, & Christiaen, L. (2015) Migratory neuronal progenitors arise from the neural plate borders in tunicates. *Nature* **527**, 371--374.
13. Li, Y, Zhao, D, Horie, T, Chen, G, Bao, H, Chen, S, Liu, W, Horie, R, Liang, T, Dong, B, Feng, Q, Tao, Q, & Liu, X. (2017) Conserved gene regulatory module specifies lateral neural borders across bilaterians. *Proceedings of the National Academy of Sciences* **114**, E6352--E6360.
14. Green, S. A, Simoes-Costa, M, & Bronner, M. E. (2015) Evolution of vertebrates as viewed from the crest. *Nature* **520**, 474--482.
15. Kassahn, K. S, Dang, V. T, Wilkins, S. J, Perkins, A. C, & Ragan, M. A. (2009) Evolution of gene function and regulatory control after whole-genome duplication : Comparative analyses in vertebrates. *Genome Research* **19**, 1404--1418.
16. Kimura, T, Nagao, Y, Hashimoto, H, Yamamoto-Shiraishi, Y.-i, Yamamoto, S, Yabe, T, Takada, S, Kinoshita, M, Kuroiwa, A, & Naruse, K. (2014) Leucophores are similar to xanthophores in their specification and differentiation processes in medaka. *Proceedings of the National Academy of Sciences* **111**, 7343--7348.
17. Gillis, J. A, Alsema, E. C, & Criswell, K. E. (2017) Trunk neural crest origin of dermal denticles in a cartilaginous fish. *Proceedings of the National Academy of Sciences* **114**, 13200--13205.
18. Green, S. A, Uy, B. R, & Bronner, M. E. (2017) Ancient evolutionary origin of vertebrate enteric neurons from trunk-derived neural crest. *Nature* **544**, 88--91.
19. Shimeld, S. M & Holland, P. W. H. (2000) Vertebrate innovations. *Proceedings of the National Academy of Sciences* **97**, 4449--4452.

ANNEXE 2

Poster présenté au congrès Marine Evolution 2018

(15-17 Mai 2018, Strömstad, Suède)

Specific retention of pigmentation genes and their families in teleost fish after whole genome duplications in vertebrates

Thibault Lorin[1]*, Frédéric G. Brunet[1], Vincent Laudet[2], Jean-Nicolas Volff[1]

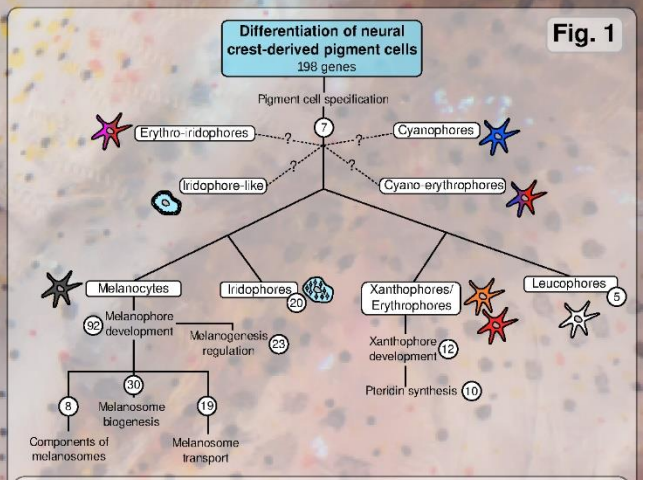
[1] Institut de Génomique Fonctionnelle de Lyon, École Normale Supérieure de Lyon, UMR 5242 CNRS, UCB Lyon 1, Université de Lyon
[2] Observatoire Océanologique de Banyuls-sur-Mer, UMR CNRS 7232 BIOM; Sorbonne Université
* thibault.lorin@ens-lyon.org

Background. Vertebrate pigmentation is a highly diverse trait mainly determined by neural crest cell derivatives **(Fig. 1)**. It has been suggested that two rounds (1R/2R) of whole-genome duplications (WGDs) at the basis of vertebrates allowed changes in gene regulation associated with neural crest evolution [1]. Subsequently, the teleost fish lineage experienced other WGDs, including the teleost-specific Ts3R before teleost radiation and the more recent Ss4R at the basis of salmonids [2,3]. As the teleost lineage harbors the highest number of pigment cell types and pigmentation diversity in vertebrates **(Figs. 1 and 2)**, WGDs might have contributed to the evo-lution and diversification of the pigmentation gene (PG) repertoire in teleosts.

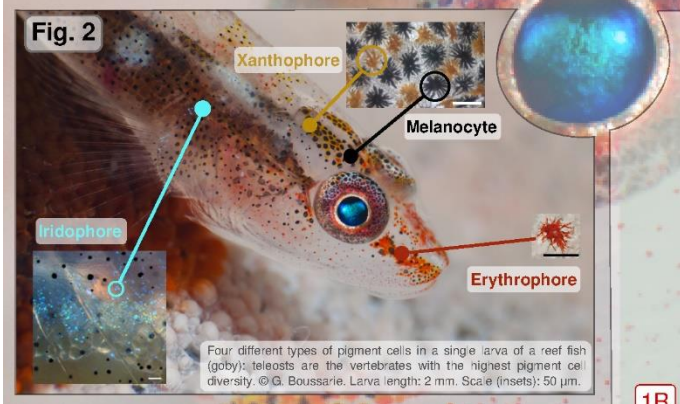
Methods. We first reviewed the existing literature to obtain the list of described vertebrate PGs to date (Fig. 1). We then compared the impact of the basal vertebrate 1R/2R duplications with that of the teleost-specific Ts3R and salmonid-specific Ss4R WGDs on 181 gene families containing PGs.

Results. We show that PGs have been globally more frequently retained as duplicates than other genes after Ts3R and Ss4R but not after the early 1R/2R **(Fig. 3)**. This is only true after Ts3R for PGs in multigenic 1R/2R families **(Fig. 4C)** but not in monogenic families (Fig. 4B), possibly linked to the molecular nature of their proteins (e.g. DNA binding transcriptional regulators). However, in salmonids, after Ss4R, there is no observed difference between PGs in monogenic and multigenic families (Fig. 4D-G), probably due to the more recent occurrence of this event (Fig. 3). Finally, interestingly, the high retention is also observed after both Ts3R and Ss4R for non-pigmentary paralogs of PGs (Fig. 4C', F', G').

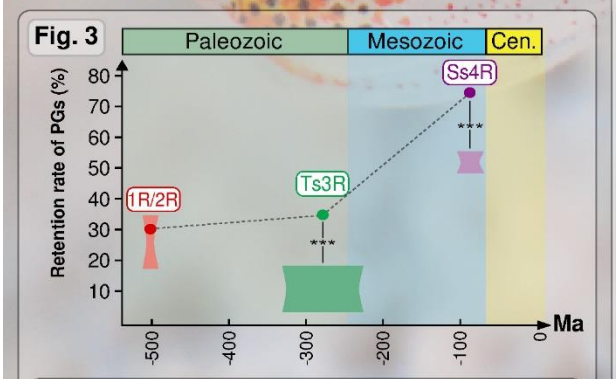
Significance. Taken together, our results support a major role of WGDs in the diversification of the PG repertoire in the teleost lineage, with a possible link with the diversity of pigment cell lineages observed in these animals compared to other vertebrates. The absence of difference between PGs and their paralogs in multigenic families suggests that the function in pigmentation is not the sole key driver of gene retention after WGDs. Additionnally, after a duplication event, gene conversion and ectopic recombination can homogenize the sequences of both duplicates [4]. Thus, even if they are not directly involved in pigmentation, the presence of non-pigmentary paralogous copies might favor the maintenance and evolution of PGs.



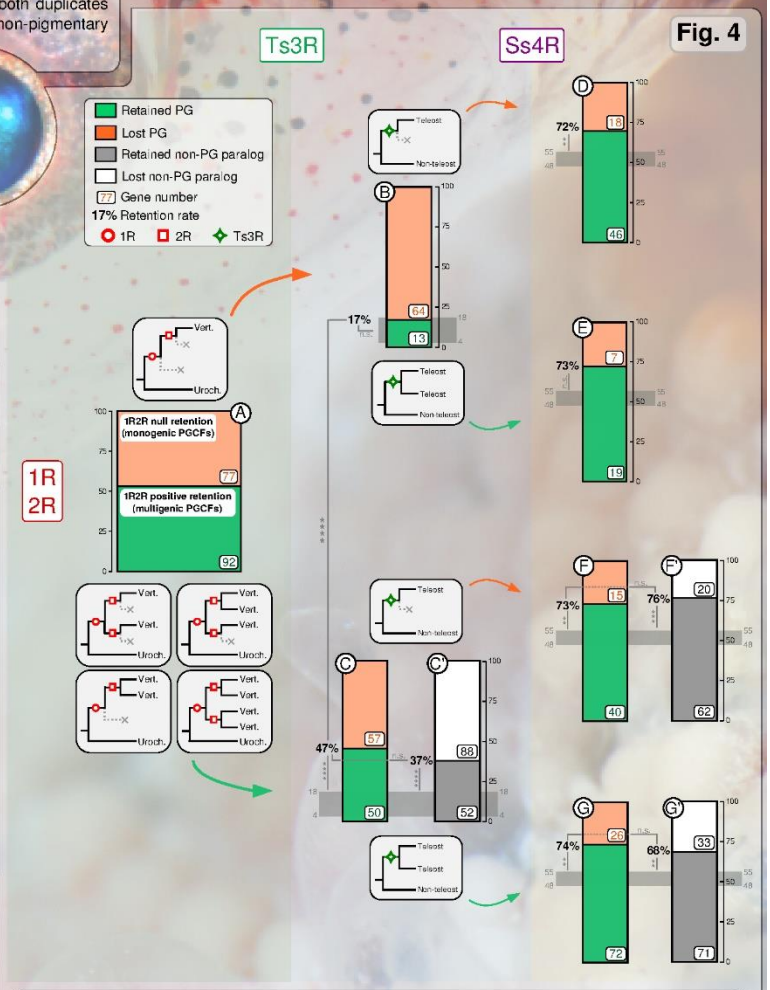
Classification of 198 vertebrate PGs according to functions and cell types. Circles indicate numbers of PGs in (sub)categories. Ontogenetic relationships between pigment cells are as proposed by [5] and [6]. Question marks indicate that these newly described pigment cell that may originate from NC, but that there is no experimental evidence for that yet. While Mammals only have melanocytes, teleost fish harbour all these types of pigment cells.



Four different types of pigment cells in a single larva of a reef fish (goby): teleosts are the vertebrates with the highest pigment cell diversity. © G. Boussarie. Larva length: 2 mm. Scale (insets): 50 µm.



PGs are more retained than genome average after Ts3R and Ss4R, but not 1R/2R. At each WGD event, heights of shaded areas represent the genome-wide retention rates estimates based on literature estimates. Widths represent the temporal estimates. Dots represent the retention rate of PGs only. Pigmentation repertoire retention rate is higher than genomic average after Ts3R- and Ss4R-WGDs. ***: $p < 10^{-7}$ (χ^2-test).



Retention rates for PGs from monogenic and multigenic families after successive WGDs in vertebrates. **Green** and **orange** boxes (respectively **black** and **white**) represent number of genes that are whether retained or not after the considered WGD for PGs (respectively non-PGs paralogs in multigenic PGCFs). Total gene number per dataset is indicated. Retention rates are indicated (between 0 and 100%) for Ts3R and Ss4R. Horizontal grey bands represent the genome average retention rate for Ts3R (4-18%) and Ss4R (48-55%). *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ****: $p < 0.0001$; n.s.: not significant at a 0.05 level (χ^2-test).

References.
[1] Green et al. (2015) *Evolution of vertebrates as viewed from the crest* (Nature)
[2] Jaillon et al. (2004) *Genome duplication in the teleost fish* T. nigroviridis (Nature)
[3] Berthelot et al. (2014) *The rainbow trout genome* (Nat. Commun.)
[4] McGrath et al. (2014) *Differential retention and divergent resolution of duplicate genes following WGD* (Genome Res.)
[5] Braasch et al. (2009) *Pigmentation pathway evolution after WGD in fish* (GBE)
[6] Kimura et al. (2014) *Leucophores are similar to xanthophores* (PNAS)

Publication associated to this poster: Lorin et al. (2018) G3 Journal

ANNEXE 3

Revue sur la pigmentation des poissons coralliens

Au cours de ma thèse, j'ai eu l'opportunité de participer à la rédaction d'une revue sur la pigmentation des poissons coralliens. Cette revue doit être soumise vers la mi-mai dans le journal Trends in Ecology and Evolution. Ma contribution a été de participer à la discussion des principaux axes en amont de la rédaction de la revue et d'écrire un paragraphe sur la diversité des types cellulaires chez les poissons coralliens. J'ai également réalisé la figure 2 et rédigé la Box 1. J'ai, enfin, participé aux relectures de la version "sub-finale", que je joins ci-après, et qui ne sera plus beaucoup modifiée avant soumission.

A Magic Trait in Magic Fish: Understanding Pigmentation Patterns using Coral Reef Fish

Pauline Salis[1], Thibault Lorin[2], Vincent Laudet[1*] and Bruno Frédérich[3]

1: Observatoire Océanologique de Banyuls-sur-Mer, UMR CNRS 7232 BIOM; Sorbonne Université ;1, avenue Pierre Fabre, 66650 Banyuls-sur-Mer, France

2: Institut de Génomique Fonctionnelle de Lyon, École Normale Supérieure de Lyon, UMR 5242 CNRS, Université Claude Bernard Lyon I, Université de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France

3: Laboratory of Functional and Evolutionary Morphology, UR FOCUS, University of Liège, Liège, Belgium

* Corresponding author: vincent.laudet@obs-banyuls.fr

Key words: Pigmentation pattern, Ecology, Evolution, Development, Coral reef fish

ABSTRACT

Pigmentation patterns provide an easy access to phenotypic diversity. Their study has allowed to tackle a wide array of questions such as the diversification of complex patterns, the adaptive value of camouflage/mimicry or the constraints acting on phenotypic evolution. Teleost fish living in coral reefs offer a unique opportunity to address such questions because they are ecologically and phylogenetically diverse and have the largest variety of pigmentation cell types known in vertebrates. They also offer relevant experimental models that allow the discrimination of ecological, developmental and evolutionary processes at work. Here, we lay out how the study of pigmentation patterns in coral reef fish, integrated in an Eco/Evo/Devo perspective, actually brings new data and insights on century-old questions.

Highlights

- *Pigmentation patterns have a wide range of functions in teleost ranging from protection against predators (camouflage, mimicry), social interactions (mate choice, dominance) to physical protection (thermoregulation, photoprotection).*
- *In coral reef fish, there is a tremendous diversity of pigmentation patterns such as stripes, eyes spot, etc. Interestingly, there is a convergence on pigmentation patterns across various taxonomic levels. Thus, the diversity of observed pigmentation patterns allows the study of the ultimate selective forces shaping the evolution of such patterns and offers a better understanding of the adaptations occurring in this complex ecosystem. However, such studies cannot provide information on proximate mechanisms allowing their emergence during development.*
- *It is important now to re-visit pigmentation pattern in an Eco/Evo/Devo perspective by combining ecological analysis with genomic and/or developmental analysis. Such studies will now allow to provide an integrated view of evolution in action.*

Questions regarding the diversity, evolution, and ecological significance of pigmentation patterns have caught scientist's attention for centuries [1]. Indeed, these patterns have a wide range of functions in animals from protection against predators (camouflage, mimicry), social interactions (mate choice, dominance) to physical protection (thermoregulation, photoprotection). Colors are often the results of multiple components either pigment-based and/or structurally-based [1].

The study of pigmentation patterns in animals has allowed to tackle many biological questions. For example, they were used to understand how developmental processes generate complex phenotypes as well as their variation [2]. Colors are highly valuable functional traits for the study of natural as well as social and sexual selection [3–5]. Finally, the study of pigmentation loss in extreme environments (*e.g.* caves) has launched the debate about the cost of maintaining colors, providing arguments for the selective value of complex patterns [6].

Pigmentation has been studied using a wide variety of animal models from hexapoda to vertebrates [7–10]. Drosophila and mouse are still important models to study pigmentation genes [11,12]. However, over the last years, teleost fish also became efficient systems for addressing questions related to pigmentation patterns [13]. Zebrafish and, in a lesser extent, medaka are helpful models for combining genetic manipulations with live imaging, and their study provided new insights in the development of pigmentation patterns [14,15].

While mammals only possess melanocytes, teleost fish have the highest number of pigment cell (also called chromatophores) types [16,17]. This chromatophore diversity is directly related to the diversity of color and the diversity of genes underlying pigmentation. The list of those genes increased in recent years (Box 1) [18] and it was recently shown that the whole genome duplication of teleost fish played a major role in the diversification of their pigment gene repertoire [19].

The richness of pigmentation patterns in teleost fish is another major advantage to test eco-evolutionary hypotheses. Experimental models, such as zebrafish, do not capture alone the remarkable diversity of phenotype and behavior observed across teleost. In contrast, coral reef fish offer an amazing diversity of original pigmentation patterns. These include dark or flamboyant colors, and can be made of a diverse combination of spots, stripes, lines, eyespots, etc. In addition, these fish live in a complex environment with extremely rich intra and inter-specific communication, and their pigmentation patterns may vary according to developmental stage, sex, social status, ecological niche, etc. Therefore, coral reef fish offer plethora of situations in which variation of pigment patterns can be studied.

Finally, extensive phylogenetic studies now provide a good comparative framework to investigate the evolution of color diversity in coral reef fish. Well resolved time-tree with large taxon sampling of various coral reef fish families are flourishing (*e.g.* wrasses: [20], butterflyfishes: [21], damselfishes: [22], angelfishes: [23], boxfishes: [24], triggerfishes: [25], etc.) and comprehensive phylogenies of bony fishes are now available [26]). In parallel, numerous phylogeographic studies illustrated color divergence in coral reef fish associated or not to genetic divergence [27,28]. All these studies provide the framework for testing hypotheses at both shallow and deep evolutionary time.

In this review, we will show how the analyses of functions from pigmentation patterns in coral reef fish combined with developmental knowledge and phylogenetic information will give new insights into processes generating complex phenotypes. For this, we will first focus on the diversity of coloration in coral reef fish, emphasizing their function and evolution. Then, we will summarize the knowledge generated in zebrafish on the development of pigmentation. From these two views, we will finally discuss how coral reef fishes constitute excellent models to integrate the eco/evo/devo perspectives toward a new synthesis that will provide a better understanding of the emergence of pigmentation patterns.

Diversity and Function of the Coral Reef Fish Flamboyant Patterns

Pigmentation patterns of coral reef fish are extremely diverse, from uniform body color such as the blue/green damselfish *Chromis viridis* to the complex pattern seen in the clown triggerfish *Balistapus conspicillum*. The latter combines a series of large ventral white spots, with a dorsal yellow shield punctuated with small brown spots. Between these two extremes, there are bicolor fish (*e.g. Pseudochromis pacagnellae*), striped fish (*e.g. Acanthurus triostegus*) and dotted fish (*e.g. Cephalopholis miniata*). This high diversity of patterns is related to important functions for camouflage and communication [29]. Here, we synthesize functions of widespread markings such as stripes, eyespots and spots, and we discuss their associated diversification.

Functions of stripes, spots and eyespots

The "prey-predator" relationship certainly led to a large variety of pigmentation patterns [30]. The shoaling behavior of *Lutjanus* species inhabiting open water is often linked to color stripes oriented rostro-caudally (Fig. 1A and there are several hypotheses about how these stripes function. Stripes disrupt the outline of the fish (disruptive coloration), increase the confusion effect [31], and serve as cues for intra-school orientation [32]. A phylogenetic comparative study in Chaetodontidae provided evidence that the evolution of body stripes is associated with social behavior and diet [33]. More social Chaetodontidae, *i.e.* living in groups, have fewer diagonal stripes, while species with greater dietary diversity have more of these markings.

Another conspicuous pattern is the eyespots or ocelli, *i.e.* circular markings that contrast with the surrounding area. The roles of eyespots in predator-prey interactions were extensively studied in moths and butterflies and it would be expected to act in similar ways in fish. Eyespots have been linked to various antipredatory functions, such as deterring hunting predators to initiate an attack (intimidation hypothesis) or diverting their attacks toward less vital body parts (deflective hypothesis) [34]. For example, it was largely assumed that the large eyespot of the comet fish *Calloplesiops altivelis* has such an antipredatory function (Fig. 1B. However, the eyespots function for predatory defense is not so obvious in all coral reef fish. No support for the adaptive value of eyespots on the posterior dorsal fin as an effective antipredator mechanism was detected in the juveniles of the ambon damselfish *Pomacentrus amboinensis*. These markings serve as a signal of subordinance from juveniles to reduce aggression by mature males living in coral heads [35]. Moreover, the function of eyespots in *P. amboinensis* changes over ontogeny. Indeed, some mature males of *P. amboinensis* retain eyespots, when others do not (*i.e.* the mature dominant males), and adopt a deceptive appearance [36]. These studies from *P. amboinensis* illustrate that markings may have multiple roles and their functional significance may change during ontogeny.

Eye stripes, *i.e.* a bar that runs through the eye and matching the eye color, are another usual component of pigmentation pattern in coral reef fish. Eye stripes may serve to camouflage the eyes from predators, hence hiding a primary target [37]. Interestingly, the comparative study from Kelley et al. [33] revealed that eye stripes were always associated with eyespots during the evolution of butterflyfishes. It is generally assumed that the combination of both marks strengthen the disruptive effect [33,38].

Convergence and modularity of pigmentation patterns

Coral reef fish are characterized by a great taxonomic diversity, encompassing a minimum of 60 families of acanthomorph fishes [39]. This diversity can be exploited to identify convergence on pigmentation patterns across various taxonomic levels. Figure 1 provides some examples of convergent evolution: the pattern made of vertical black stripes which evolved in Pomacentridae and Acanthuridae, or the pattern of horizontal white bars that is observed in Plotosidae and Apogonidae. The black stripes pattern evolved at least three times independently within Pomacentridae (genera *Dascyllus*, *Abudefduf* and *Chrysiptera*; Fig. X) illustrating that convergence may also occurs even at lower phylogenetic level.

Beyond the repeated appearance of some markings across coral reef fish evolution mimicry cases are striking examples of convergence. However, it is still challenging to question the main drivers producing the mimic's mode-like appearance. The *mimicry hypothesis* proposes that the mimic resembles its model and the mimicry target is the agent of natural selection that produced that resemblance in the mimic. An alternative scenario, [40] is based on the assumption that the "model" and "mimic" have independently evolved a common appearance, and this coincidental resemblance attracts the "mimic" to the "model" by social trapping. A large proportion of "mimic-model" pairs would then be the result of such coincidental resemblances but this certainly asks for additional comparative studies to test both mimicry and social-trap hypotheses in this diversity of coral reef fish [41,42].

Very often, even the most complex pigmentation patterns can be interpreted by the combination of several simpler elements/markings. This property offers a unique opportunity to explore the evolution of pigmentation patterns through the biological concepts of integration and modularity [43]. The above-mentioned comparative study of Chaetodontidae provided a first demonstration that some markings evolved differently: eyespots are evolutionary labile whereas eye stripes are more phylogenetically conserved [33]. Correlated evolution of some specific markings, such as "spots and eye

stripes" or "eyespot and adjacent eye stripe" in Chaetodontidae [33], allows the suggestion of ultimate and proximate mechanisms driving the pigmentation patterns. This modular characteristic of pigmentation patterns can be very helpful to better understand the mechanisms and the constraints that control their elaboration.

Understanding the ontology of pigmentation patterns using fish models

The diversity of pigmentation patterns observed in coral reef fish allows the study of the ultimate selective forces shaping the evolution of such patterns and offers a better understanding of the adaptations occurring in this complex ecosystem. However, such studies cannot provide information on proximate mechanisms allowing their emergence during development. Up to now, cellular and molecular studies have mainly been carried out using zebrafish (*Danio rerio)*, a widely used model. Thanks to the genetic and live imaging tools developed in this species, it has been possible to investigate the mechanisms underlying pigmentation pattern formation.

The cellular context

In zebrafish, three distinct types of chromatophores are present: black melanophores, yellow xanthophores, and iridescent iridophores [44]. As in most teleost fish, the zebrafish shows two very different pigmentation patterns during ontogeny: a larval pattern and an adult one. The larval pattern is formed by a transparent with subtle yellowish colored body (due to scattered xanthophores) covered by loose longitudinal bands of melanophores, in the dorsal and ventral apex as well as laterally at the level of the myoseptum (Fig. 3(A)). At the onset of metamorphosis, the adult pattern starts developing. It is composed of longitudinal black stripes with melanophores and iridophores contrasting with inter-stripe regions containing xanthophores and iridophores (Fig. 3(A)).

The generation of the adult pigmentation pattern is complex because the origin of pigment cells found at the adult stage varies. Experimental genetic analyses revealed that the largest number of melanophores and iridophores found in adult (often called metamorphic chromatophores) differentiate during metamorphosis and later [45,15,46], whereas almost all adult xanthophores differentiate during the larval stage [47]. Additionally, the melanophores found in adults have a dual origin: the largest number of melanophores differentiate only at the adult stage whereas a minority corresponds to persisting embryonic melanophores. [48]. These results demonstrated that the underlying genetic architectures of the larval and adult patterns are only partially overlapping.

An important feature of this two steps-process that corresponds to a metamorphosis, is the role of thyroid hormones (TH). As in other teleosts, these hormones trigger and coordinate this elaborate transformation. Interestingly, the different types of chromatophores are differentially sensitive to alterations of TH levels. For example, treatment with TH leads to marked xanthophores excess and deficiency in melanophores in adults [49]. The role of TH is thus central for controlling the differentiation and the ultimate presence of the three types of chromatophores, generating the observed pattern.

Cell-cell interactions are instrumental for patterning

Genetic studies in zebrafish have revealed that the interactions among the three types of chromatophores play a key role to create the pigmentation pattern. For example, in mutants deficient for xanthophores (*pfeffer* mutants), the melanophores stripes are re-organized into spots [50]. Mutants in which two chromatophores types have been deleted (e.g. *shady::pfeffer* mutants) reveal that the single remaining chromatophore type is not able to form the precise pattern [50].

Moreover, such interdependency is also important for sustaining formation and/or survival of chromatophores. For example, it was shown that iridophores promote and sustain melanophores differentiation [45,50], whereas depletion of xanthophores leads to a reduction in melanophores number [51].

The Turing model can predict the formation of zebrafish adult pigmentation pattern

These dynamics of cells interactions were interestingly predicted by Turing models (Box 2), which became a standard for the modelling of complex pattern formation. Interestingly, the Turing model was validated in zebrafish. In fact, artificial disturbance of the striped pattern induces changes in the pattern predicted by the Turing model. It has been observed that when the pigment cells of the tail are killed, regeneration of stripes occurs but without the original directionality [52]. Moreover, ablation experiments of chromatophore types in different regions leads to the disruption of various short-range and long-range interactions, characteristics essential for the Turing model. For example, when part of a xanthophores stripe is ablated, only xanthophores will arise in the cleared area, but only so long as an adjacent melanophores stripe is present (Fig. 3(B)- upper panel). If adjacent melanophores are absent, melanophores will readily emerge in the former xanthophores domain (Fig. 3(B)- middle panel) [51]. Conversely, the inhibition of xanthophores causes a decrease of melanophores in distant regions, creating a long-range inhibition loop (Fig. 3(B)- Bottom panel) [51]. This suggests that long-range interactions, *e.g.* xanthophores promoting melanophores emergence and melanophores inhibiting other melanophores, play some part in setting up final width of the stripes and showed that this network of interactions possesses the properties necessary to form the Turing pattern (Fig. 3(C) [51].
Another fundamental characteristic of Turing model is that the number of repeated stripes or spots is intimately connected to the size, and therefore the growth of the organism. Such characteristics were observed in a reef fish [53] and in long-fin zebrafish mutant (for which the fins never stop growing) which continues to form perfectly new stripes as the fins grow [46].

If cell interdependency shapes the width of the stripes, the global directionality of the pattern has to be established. Stripe direction is specified by some biological indicators. Accordingly, the pigmentation pattern of zebrafish body trunk needs some initial conditions: the horizontal myoseptum in which the iridophores precursors migrate to form the first horizontal stripe provides the initial mediator for stripe patterning. Melanophores and xanthophores, which will develop later, are influenced by this stripe of iridophores. The crucial influence of the horizontal myoseptum was illustrated by *choker* mutants (fish for which the myoseptum is lost), for which the adult fish develops into a labyrinth-like pattern [50,54]. The identification of the molecular factors involved in these interactions is still ongoing.

Evolution of pigmentation patterns

The study of *D. rerio* and their closely related species showing different pigmentation patterns allowed deciphering some evolutionary mechanisms controlling the evolution of pigment patterns. For example, an interesting case is provided by *D. albolineatus,* a non-striped *Danio* species characterized by the presence of intermingled populations of the three pigment cells. In this species, differentiation of xanthophores occurs earlier than in *D. rerio* because of an increased expression of *csf1*, a growth factor supplied by iridophores and other cells in the skin [2]. The induced increased and earlier population of xanthophores modifies the positioning information provided to melanophores compared to *D. rerio.* Consequently, *D. albolineatus* individuals do not form stripes. [2]. These data illustrate how changes of expression of key molecular factors coupled with changes in cell-cell interactions can lead to the evolution of a new pigmentation pattern.

Coral Reef Fish Harbor the Largest Diversity of Pigment Cells

If zebrafish with its unique toolkit is an excellent model to understand the development of repeated striped pattern, it must be emphasized that the situation in coral reef fish is much more complex since some fish pigmentation pattern does not involve a repeated motif suggesting a pre-pattern mechanism. In addition to the three chromatophore types observed in zebrafish (i.e. melanophores, xanthophores, and iridophores), coral reef fish exhibit many other chromatophores. They are thus of particular importance to fully grasp the range of possible pigmentation systems in vertebrates. Coral reef fish are also providing the only known case of dichromatic pigment cells. The erythro-iridophores, found in the diadem dottyback *Pseudochromis diadema*, contain both a reddish carotenoid pigment and reflecting platelets similar to those found in iridophores [55]. The mandarin fish *S. splendidus* also possesses dichromatic cells, the cyano-erythrophores [56]. Lastly, it is interesting to note that the mechanism allowing color change of some species have started to be analyzed. The

chameleon sand tilefish *Hoplolatilus chlupatyi* can exhibit color change from blue to red in a matter of few seconds and this very fast color change is linked to a novel type of iridophores in which the reflecting platelets are concentrated in the periphery of the cell. Adrenergic stimulation leads to changes in the reflecting platelet organization and therefore in fish color [16].

All these data clearly indicate that, if zebrafish will remain an unsurpassed model to understand in cellular and molecular terms how complex pigmentation patterns are generated, we must enlarge our perspective and start to use coral reef fish as model systems if we want to fully understand the diversity of pigmentation patterns and their associated ecological or behavioral specialization.

Integrating Ecology with Evo/Devo in Coral Reef Fishes

A full understanding of complex phenotypic patterns needs the integration of the ultimate causes of these patterns (*i.e.* their ecological function) with the proximate factors that govern their appearance (*i.e.* cellular and molecular mechanisms).

The Turing mechanisms that allow the understanding of repeated stripe formation in zebrafish can be applied to other species. This is the case of some marine angelfishes (*Pomacanthus* spp) that have been instrumental in testing the validity of these models [53]. Other authors have developed more complex models with two reaction-diffusion systems coupled together in such a way that one of them produces a simple Turing pattern of spots or stripes, and the other produces wave fronts. These models produce series of patterns that are seen in many species of coral reef fish such as the mandarin fish or the pufferfish [57,58].

To our best knowledge, very few study globally explored the ecological function of colors and associated cellular and molecular basis for pigmentation in coral reef fish (*e.g.* [59]). A transcriptomic analysis of divergent color morphs (bright red versus dark red; both with small blue spots) present in the leopard coral grouper (*Plectropomus leopardus*) detected 34 genes differentially expressed between the two morphs but no functional studies have yet been performed [60]. However, coral reef fish offer excellent materials to reconcile ecological and developmental analysis needed for our understanding of complex patterns. Two cases beautifully illustrate the potential of these types of studies.

The first case concerns phenotypic plasticity. This common phenomenon in coral reef fish is well exemplified by the dusky dottyback *Pseudochromis fuscus*, a small predatory fish [59,61]. This species can exhibit numerous uniform color morphs from orange to brown to yellow, pink and gray. Yellow and brown morphs are present at the Great Barrier Reef (Lizard Islands, Australia), The yellow morph inhabits living coral heads with yellow damselfishes (*e.g. Pomacentrus amboiensis*) whereas the brown one is associated with brown damselfish species (*e.g. P. chrysurus*) on coral rubbles. Experimentations illustrated that yellow morphs can transformed in brown morphs within 2 weeks if translocated from living corals to coral rubbles [62]. Strikingly, however, the dottyback does not change color because of the environment but because of the presence of colored damselfishes. The advantages of this strategy are double for *P. fuscus*. First, by mimicking adults of a damselfish species, it increases its predation success on the juveniles of this species. Second, the color change helps the dottyback to escape its own predator by providing a habitat-associated crypsis. The study of associated cellular mechanisms revealed that this change in color is explained by effects on two specialized cell populations as it is the proportion of xanthophores compared to melanophores that is modified [61]. In a fascinating follow-up study, this color change has been placed upon an ontogenetic trajectory and it has been shown that dottybacks change color twice during development: once during metamorphosis, when a pelagic translucid larvae is transformed into a grey juvenile, and when the large-enough juvenile starts its mimicry strategy and select either yellow or brown victims [59]. These changes are also accompanied by changes of the visual system illustrating that the mimicry strategy is deeply integrated at various level of the fish morphology and physiology [59].

The second case study, the radiation of the caribbean hamlets (*Hypoplectrus* spp), shows how color polymorphisms will allow understanding the ecological and developmental basis of phenotypic adaptation. Detailed analysis of their radiation revealed that a single trait, pigmentation pattern, has driven incipient speciation in this fish [63]. It is often considered that, as a predatory fish, *Hypoplectrus* mimics harmless fish in order to increase their predation success on their preys ([63] but see [40]). Genetic analyses allowed to identify divergent loci among color morphs [64] [65] [66]. Among them, an analysis using SNPs identify the HoxC cluster as being associated with color variation [65].

However, Hox genes have never been associated with pigmentation defect in teleost fish but they have been linked to body pigmentation and eyespot formation in insects [67]. Developmental studies are much needed to better understand the role, if any, that these genes could play in the divergence of pigmentation patterns.

The cases presented above show that pigmentation can indeed be considered as a "magic trait" that may easily lead to speciation [68]. Coral reef fish provide a large diversity of patterns and processes acting to generate divergent pigmentation patterns and promote diversification. Combining ecological analysis with genomic and/or developmental analysis will now allow to provide an integrated view of evolution in action using these "magic fish" as model systems.

ACKNOWLEDGEMENTS

We would like to thank Germain Boussarie for providing larval goby and diamant madarin pictures.

REFERENCES

1 Price, A.C. *et al.* (2008) Pigments , Patterns , and Fish Behavior. *Zebrafish* 5, 297–307

2 Patterson, L.B. *et al.* (2014) Pigment cell interactions and differential xanthophore recruitment underlying zebrafish stripe reiteration and Danio pattern evolution. *Nat. Commun.* 5, 5299

3 Kodric-Brown, A. (1998) Sexual Dichromatism and Temporary Color Changes in the Reproduction of Fishes. *Am. Zool.* 38, 70–81

4 Dunn, P.O. *et al.* (2015) Natural and sexual selection act on different axes of variation in avian plumage color. *Sci. Adv.* 1,

5 Tobias, J.A. *et al.* (2012) The evolution of female ornaments and weaponry: social selection, sexual selection and ecological competition. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 2274–2293

6 Protas, M.E. and Patel, N.H. (2008) Evolution of coloration patterns. *Annu. Rev. Cell Dev. Biol.* 24, 425–446

7 Mallarino, R. *et al.* (2016) Developmental mechanisms of stripe patterns in rodents. *Nature* 539, 518–523

8 Gompel, N. *et al.* (2005) Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in Drosophila. *Nature* 433, 481–487

9 Hof, A.E.V. t. *et al.* (2016) The industrial melanism mutation in British peppered moths is a transposable element. *Nature* 534, 102–105

10 Santos, M.E. *et al.* (2014) The evolution of cichlid fish egg-spots is linked with a cis-regulatory change. *Nat. Commun.* 5,

11 Hoekstra, H.E. (2006) Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity (Edinb).* 97, 222–234

12 Arnoult, L. *et al.* (2013) Emergence and diversification of fly pigmentation through evolution of a gene regulatory module. *Science* 339, 1423–6

13 Streisinger, G. *et al.* (1981) Production of clones of homozygous diploid zebra fish (Brachydanio rerio). *Nature* 291, 293–296

14 Lamoreux, M.L. *et al.* (2005) Pigment pattern formation in the medaka embryo. *Pigment Cell Res.* 18, 64–73

15 Singh, A.P. *et al.* (2014) Proliferation, dispersal and patterned aggregation of iridophores in the skin prefigure striped colouration of zebrafish. *Nat. Cell Biol.* 16, 607–14

16 Goda, M. (2017) Rapid integumental color changes due to novel iridophores in the chameleon sand tilefish Hoplolatilus chlupatyi. *Pigment Cell Melanoma Res.* 30, 368–371

17 Wucherer, M.F. and Michiels, N.K. (2012) A fluorescent chromatophore changes the level of fluorescence in a reef fish. *PLoS One* 7,

18 Lamason, R.L. (2005) SLC24A5, a Putative Cation Exchanger, Affects Pigmentation in Zebrafish and Humans. *Science (80-.).* 310, 1782–1786

19 Lorin, T. *et al.* (2018) Teleost Fish-Specific Preferential Retention of Pigmentation Gene-Containing Families After Whole Genome Duplications in Vertebrates. *G3 Genes|Genomes|Genetics* DOI: 10.1534/g3.118.200201

20 Baliga, V.B. and Law, C.J. (2016) Cleaners among wrasses: Phylogenetics and evolutionary patterns of cleaning behavior within Labridae. *Mol. Phylogenet. Evol.* 94, 424–435

21 Cowman, P.F. and Bellwood, D.R. (2011) Coral reefs as drivers of cladogenesis: Expanding coral reefs, cryptic extinction events, and the development of biodiversity hotspots. *J. Evol. Biol.* 24, 2543–2562

22 Frédérich, B. *et al.* (2013) Iterative ecological radiation and convergence during the evolutionary history of damselfishes (Pomacentridae). *Am. Nat.* 181, 94–113

23 Frédérich, B. *et al.* (2017) Body shape convergence driven by small size optimum in marine angelfishes. *Biol. Lett.* 13, 20170154

24 Santini, F. *et al.* (2013) A multilocus molecular phylogeny of boxfishes (Aracanidae, Ostraciidae; Tetraodontiformes). *Mol. Phylogenet. Evol.* 66, 153–160

25 Santini, F. *et al.* (2013) A new multi-locus timescale reveals the evolutionary basis of diversity patterns in triggerfishes and filefishes (Balistidae, Monacanthidae; Tetraodontiformes). *Mol. Phylogenet. Evol.* 69, 165–176

26 Betancur-R, R. *et al.* (2017) Phylogenetic classification of bony fishes. *BMC Evol. Biol.* 17, 162

27 Bernardi, G. *et al.* (2002) Species boundaries, populations and colour morphs in the coral reef three-spot damselfish (Dascyllus trimaculatus) species complex. *Proc. R. Soc. B Biol. Sci.* 269, 599–605

28 Bowen, B.W. *et al.* Phylogeography of two Atlantic squirrelfishes (family Holocentridae): Exploring links between pelagic larval duration and population connectivity. , *Marine Biology*, 149. (2006) , 899–913

29 Marshall, N.J. (2000) Communication and camouflage with the same “bright” colours in reef fishes. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*

355, 1243–1248
30 Endler, J.A. (1980) Natural Selection on Color Patterns in *Poecilia reticulata*. *Evolution (N. Y).* 34, 76–91
31 McRobert, S.P. and Bradner, J. (1998) The influence of body coloration on shoaling preferences in fish. *Anim. Behav.* 56, 611–615
32 McFarland, W.N. *et al.* (1979) The influence of light on the twilight migrations of grunts. *Environ. Biol. Fishes* 4, 9–22
33 Kelley, J.L. *et al.* (2013) Spots and stripes: ecology and colour pattern evolution in butterflyfishes. *Proc. Biol. Sci.* 280, 20122730
34 Stevens, M. The role of eyespots as anti-predator mechanisms, principally demonstrated in the Lepidoptera. , *Biological Reviews of the Cambridge Philosophical Society*, 80. (2005) , 573–588
35 Gagliano, M. (2008) On the spot: The absence of predators reveals eyespot plasticity in a marine fish. *Behav. Ecol.* 19, 733–739
36 Gagliano, M. and Depczynski, M. (2013) Spot the Difference: Mimicry in a Coral Reef Fish. *PLoS One* 8,
37 Barlow, G.W. (1972) The Attitude of Fish Eye-Lines in Relation to Body Shape and to Stripes and Bars. *Copeia* 1972, 4
38 Kjernsmo, K. *et al.* (2016) Adaptive constellations of protective marks: Eyespots, eye stripes and diversion of attacks by fish. *Anim. Behav.* 111, 189–195
39 Price, S.A. *et al.* (2015) Phylogenetic insights into the history and diversification of fishes on reefs. *Coral Reefs* 34, 997–1009
40 Robertson, D.R. (2013) Who Resembles Whom? Mimetic and Coincidental Look-Alikes among Tropical Reef Fishes. *PLoS One* 8,
41 Moland, E. *et al.* (2005) Ecology and evolution of mimicry in coral reef fishes. *Oceanogr. Mar. Biol. an Annu. Rev.* at <http://apps.webofknowledge.com/full_record.do?product=UA&search_mode=GeneralSearch&qid=3&SID=W1Im5CDA943dGglEnMC&page=1&doc=1>
42 Randall, J.E. A review of mimicry in marine fishes. , *Zoological Studies*, 44. (2005) , 299–328
43 Klingenberg, C.P. (2008) Morphological Integration and Developmental Modularity. *Annu. Rev. Ecol. Evol. Syst.* 39, 115–132
44 Hirata, M. *et al.* (2003) Pigment cell organization in the hypodermis of zebrafish. *Dev. Dyn.* 227, 497–503
45 Patterson, L.B. and Parichy, D.M. (2013) Interactions with Iridophores and the Tissue Environment Required for Patterning Melanophores and Xanthophores during Zebrafish Adult Pigment Stripe Formation. *PLoS Genet.* 9, 1–14
46 van Eeden, F.J. *et al.* (1996) Genetic analysis of fin formation in the zebrafish, Danio rerio. *Development* 123, 255–262
47 Mahalwar, P. *et al.* (2014) Local reorganization of xanthophores fine-tunes and colors the striped pattern of zebrafish. *Science* 345, 1362–4
48 Johnson, S.L. *et al.* (2011) Mitfa is required at multiple stages of melanocyte differentiation but not to establish the melanocyte stem cell. *Dev. Biol.* 350, 405–413
49 S. K. McMenamin *et al.* (2014) Thyroid hormone-dependent adult pigment cell lineage and pattern in zebrafish. *Science* 345, 1358–61
50 Frohnhöfer, H.G. *et al.* (2013) Iridophores and their interactions with other chromatophores are required for stripe formation in zebrafish. *Development* 140, 2997–3007
51 Nakamasu, A. *et al.* (2009) Interactions between zebrafish pigment cells responsible for the generation of Turing patterns. *Pnas* 106, 8429–8434
52 Parichy, D.M. and Turner, J.M. (2003) Zebrafish puma mutant decouples pigment pattern and somatic metamorphosis. *Dev. Biol.* 256, 242–257
53 Kondo, S. and Asai, R. (1995) A reaction diffusion wave on the skin of the marine angelfish Pomacanthus. *Nature* 376, 765–768
54 Svetic, V. *et al.* (2007) Sdf1a patterns zebrafish melanophores and links the somite and melanophore pattern defects in choker mutants. *Development* 134, 1011–1022
55 Goda, M. *et al.* (2011) Integumental reddish-violet coloration owing to novel dichromatic chromatophores in the teleost fish, Pseudochromis diadema. *Pigment Cell Melanoma Res.* 24, 614–617
56 Goda, M. *et al.* (2013) Novel dichromatic chromatophores in the integument of the mandarin fish Synchiropus splendidus. *Biol. Bull.* 224, 14–17
57 Barrio, R.A. *et al.* (2009) Modeling the skin pattern of fishes. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* 79,
58 Kondo, S. and Miura, T. (2010) Reaction-Diffusion Model as a Framework for Understanding Biological Pattern Formation. *Science* 329, 1616–1620
59 Cortesi, F. *et al.* (2016) From crypsis to mimicry: changes in colour and the configuration of the visual system during ontogenetic habitat transitions in a coral reef fish. *J. Exp. Biol.* 219, 2545–2558
60 Wang, L. *et al.* (2015) In silico comparative transcriptome analysis of two color morphs of the common coral trout (Plectropomus leopardus). *PLoS One* 10,
61 Cortesi, F. *et al.* (2015) Phenotypic plasticity confers multiple fitness benefits to a mimic. *Curr. Biol.* 25, 949–954
62 Messmer, V. *et al.* (2005) Genetic and ecological characterisation of colour dimorphism in a coral reef fish. *Environ. Biol. Fishes* 74, 175–183
63 Puebla, O. *et al.* (2007) Colour pattern as a single trait driving speciation in Hypoplectrus coral reef fishes? *Proc. R. Soc. B Biol. Sci.* 274, 1265–1271
64 Holt, B.G. *et al.* (2011) Searching for speciation genes: Molecular evidence for selection associated with colour morphotypes in the caribbean reef fish genus hypoplectrus. *PLoS One* 6,
65 Puebla, O. *et al.* (2014) Genomic atolls of differentiation in coral reef fishes (Hypoplectrus spp., Serranidae). *Mol. Ecol.* 23, 5291–5303
66 Picq, S. *et al.* (2016) Population genomics of local adaptation versus speciation in coral reef fishes (Hypoplectrus spp, Serranidae). *Ecol. Evol.* 6, 2109–2124
67 Jeong, S. *et al.* (2006) Regulation of Body Pigmentation by the Abdominal-B Hox Protein and Its Gain and Loss in Drosophila Evolution. *Cell* 125, 1387–1399

68 Servedio, M.R. *et al.* Magic traits in speciation: “magic” but not rare? , *Trends in Ecology and Evolution*, 26. (2011) , 389–397
69 Wakamatsu, Y. *et al.* (2001) The see-through medaka: A fish model that is transparent throughout life. *Proc. Natl. Acad. Sci.* 98, 10046–10050
70 Michiels, N.K. *et al.* (2008) Red fluorescence in reef fish: A novel signalling mechanism? *BMC Ecol.* 8,
71 Goda, M. and Fujii, R. (1995) Blue Chromatophores in Two Species of Callionymid Fish. *Zoolog. Sci.* 12, 811–813
72 Ban, E. *et al.* (2005) The signaling pathway in photoresponses that may be mediated by visual pigments in erythrophores of Nile tilapia. *Pigment Cell Res.* 18, 360–369
73 Bennett, D.C. and Lamoreux, M.L. (2003) The Color Loci of Mice - A Genetic Century. *Pigment Cell Res.* 16, 333–344
74 Schartl, M. *et al.* (2016) What is a vertebrate pigment cell? *Pigment Cell Melanoma Res.* 29, 8–14
75 San-Jose, L.M. and Roulin, A. (2017) Genomics of coloration in natural animal populations. *Philos. Trans. R. Soc. B Biol. Sci.* 372, 20160337
76 Lister, J.A. *et al.* (1999) nacre encodes a zebrafish microphthalmia-related protein that regulates neural-crest-derived pigment cell fate. *Development* 126, 3757–67
77 Cal, L. *et al.* (2017) BAC Recombineering of the Agouti Loci from Spotted Gar and Zebrafish Reveals the Evolutionary Ancestry of Dorsal-Ventral Pigment Asymmetry in Fish. *J. Exp. Zool. Part B Mol. Dev. Evol.* DOI: 10.1002/jez.b.22748
78 Fadeev, A. *et al.* (2016) Zebrafish Leucocyte tyrosine kinase controls iridophore establishment, proliferation and survival. *Pigment Cell Melanoma Res.* 29, 284–296
79 Nagao, Y. *et al.* (2014) Sox5 Functions as a Fate Switch in Medaka Pigment Cell Development. *PLoS Genet.* 10,
80 Nagao, Y. *et al.* (2018) Distinct interactions of Sox5 and Sox10 in fate specification of pigment cells in medaka and zebrafish. *PLOS Genet.* 14, e1007260
81 McGrath, C.L. *et al.* (2014) Differential retention and divergent resolution of duplicate genes following whole-genome duplication. *Genome Res.* 24, 1665–1675
82 Ohno, S. (1970) *Evolution by gene duplication.*,
83 Braasch, I. *et al.* (2009) Pigmentation pathway evolution after whole-genome duplication in fish. *Genome Biol. Evol.* 1, 479–493
84 Turing, A.M. (1952) The Chemical Basis of Morphogenesis THE CHEMICAL BASIS OF MORPHOGENESIS. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 237, 37–72
85 Hiscock, T.W. and Megason, S.G. (2015) Mathematically guided approaches to distinguish models of periodic patterning. *Development* 142, 409–419
86 Painter, K.J. *et al.* (1999) Stripe formation in juvenile Pomacanthus explained by a generalized Turing mechanism with chemotaxis. *Proc. Natl. Acad. Sci.* 96, 5549–5554
87 Yamaguchi, M. *et al.* (2007) Pattern regulation in the stripe of zebrafish suggests an underlying dynamic and autonomous mechanism. *Proc. Natl. Acad. Sci. U. S. A.* 104, 4790–4793
88 Sick, S. *et al.* (2006) WNT and DKK determine hair follicle spacing through a reaction-diffusion mechanism. *Science (80-.).* 314, 1447–1450
89 Harris, M.P. *et al.* (2005) Molecular evidence for an activator-inhibitor mechanism in development of embryonic feather branching. *Proc. Natl. Acad. Sci. U. S. A.* 102, 11734–11739
90 Miura, T. and Shiota, K. (2002) Depletion of FGF acts as a lateral inhibitory factor in lung branching morphogenesis in vitro. *Mech. Dev.* 116, 29–38
91 Salazar-Ciudad, I. Tooth patterning and evolution. , *Current Opinion in Genetics and Development*, 22. (2012) , 585–592

Figure 1 Illustrations of some pigmentation patterns in reef fishes. The snapper, *Lutjanus kasmira* (photo credit – Sutton); (B) the comet fish, *Calloplesiops altivelis* (photo credit – Guido Poppe); The vertical black stripped pattern is observed in (C) the surgeonfish, *Acanthurus triostegus* (photo credit – Merlier) and the three damselfishes (D) *Abudefduf sexfasciatus* (photo credit - Merlier); (E) *Dascyllus aruanus* (photo credit – Merlier); (F) *Chrysiptera annulata* (photo credit – De Vroe). The horizontal white bars evolved in (G) the eel catfish, *Plotosus lineatus* (photo credit – Bourgeon) and (H) the cardinalfish, *Ostorhinchus nigrofasciatus*.

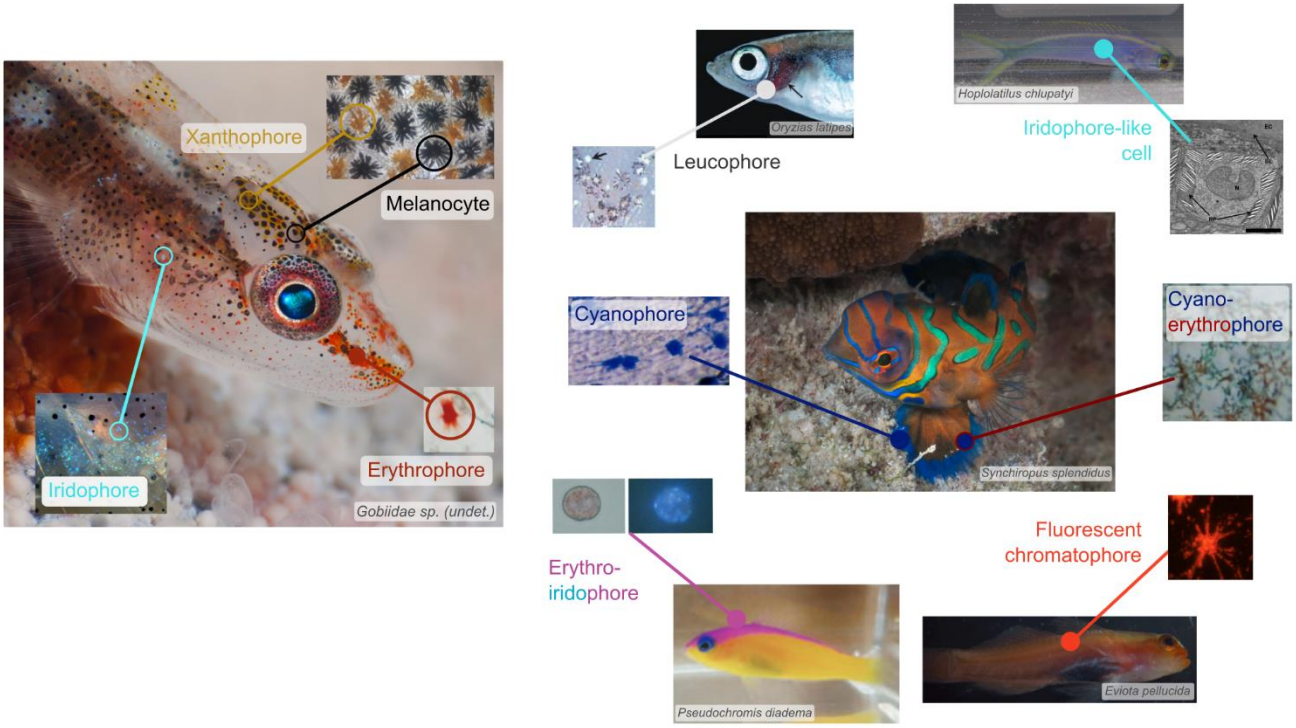


Figure 2 Coral reef fish harbor a high diversity of pigment cells. Of all pigment cells present in teleost fish, all but leucophores have been described in coral reef fish.
O. latipes: [69]; *H. chlupatyi*: [16]; *P. diadema*: [55]; *E. pellucida*:[70].
Pictures of chromatophores come from the following articles : [16,17,50,55,56,59,69,71,72]
Photo credit for individuals: Germain Boussarie (goby larva and *S. splendidus*).

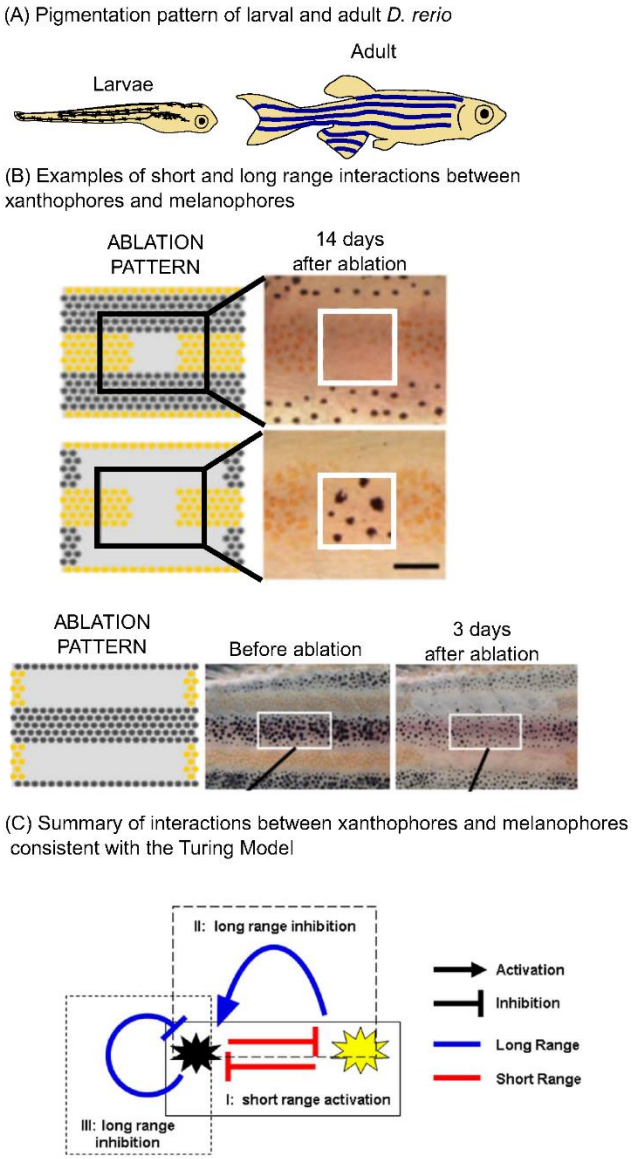


Figure 3. Understanding the ontology of pigmentation patterns using *Danio rerio*
(A) Pigmentation pattern of larval (left) and adult *D. rerio* (right). (B) Ablation experiments showing long and short range interactions between xanthophores and melanophores. (C) Cartoon summarizing interactions between xanthophores and melanophores interactions consistent with Turing Model: I. a short-range activation resulting on a negative feedback loop between xanthophores and melanophores; II. an overall long-range inhibition resulting on a long-range positive effect of xanthophores on melanophores and III. a long range auto-inhibition of melanophores [51].

BOXES

Box 1: Pigmentation Genes

Pigmentation patterns are controlled by genes deployed during the development of pigment cells [11]. In vertebrates, pigment cells are neural crest cell (NCC) derivatives, and the acquisition of a functional, pigmentary NCC-derived cell, is a multiple step process that requires a fine orchestration of the expression of specific set of genes.

Pigmentation genes have mainly been studied in Mammals, for which only melanocytes have been described. Genes involved in (i) melanocyte differentiation, (ii) melanosome biogenesis (occurring once the melanocyte is formed), (iii) melanogenesis regulation and (iv) melanosome transport are often distinguished "chronologically" [73].

The situation is more complex in other vertebrates, and in particular teleost, that have more pigment cell types displaying very diverse in terms of shape, ontology, and pigment biochemistry [74]. Studies in zebrafish and medaka have identified genes involved in specific teleost chromatophore differentiation as well as genes involved in processes conserved in all vertebrates [75]. Altogether, this makes a total of ca. 200 genes known to be involved in pigmentation to date. Some genes, such as *mitf* (important for melanocyte development) and *agouti* (that controls dorso-ventral patterning), have conserved mechanisms of action throughout vertebrates [76,77]. Others are specifically involved in teleost fish; *e. g* both *ltk* and *sox5* are known to be required for iridophores and xanthophores development in teleost [78,79]. Of note, recent work has shown that a very same gene can be involved in the development of a same pigment cell type in conflicting ways in different fish species. For instance, xanthophore differentiation requires the expression of *sox5* in medaka and its repression in zebrafish [80].

During vertebrates' evolution, the pigmentation gene repertoire has been shaped, as all other genes, by several whole-genome duplications (WGDs). After a WGD event, genes are either retained or lost. The retention pattern greatly varies with the function of the encoded protein, and genes that are retained in two copies often provide the raw material for the acquisition of new functions [81,82]. Interestingly, pigmentation genes were not more retained than the genome average after the WGDs that occurred in the ancestor of all vertebrates [19,83]. On the other hand, pigmentation genes were more retained than genome average after the teleost fish specific WGDs that occurred in the ancestor of all teleost. This pigmentation gene repertoire is thus expected to be linked to the highest pigment cell diversity and the great diversity of pigmentation patterns observed in teleosts.

Box 2: Turing Models

Originally introduced by the mathematician Alan Turing in 1952, the Turing or reaction-diffusion model (RD), explains the spontaneous formation of periodic biological patterns [84,85]. This model involves two diffusing molecules that are interacting: a slowly-diffusing activator and a rapidly-diffusing inhibitor. As the inhibitor molecule diffuses more rapidly than the activator, it impairs activation at long range (Fig. Box2(A)). If the activator is sufficiently strong it is in sufficient amount to preclude its inhibition at short range. It is the balance between the reaction of the two molecules and their diffusion that explains how various periodic patterns can spontaneously emerge from an initially homogeneous pattern. The parameters that can vary in models (relative strengths of the activator and inhibitor and their diffusion abilities), explain the wide variety of patterns (stripes, spots, etc.).

An illustration of the RD has been provided in the juvenile angelfish *Pomacanthus imperator* and *Pomacanthus semicirculatus* [53]. Young juveniles of *Pomacanthus semicirculatus* display three vertical white stripes on a dark background. During growth, new stripes insert between the preexisting ones, and this process is repeated several time to give rise to the final pattern. This dynamic change was predicted by the RD. The same authors also show how rearrangement of the parallel striped pattern of the adult angelfish can be predicted. When the fish grows the number of lines increase proportionally to body size and the space between them remain constant (Fig. Box2(B)). By incorporating in the model key aspects such as cell growth and movement, authors explain in a detailed manner the dynamic of stripes formation [86]. More recently, similar arrangements the zebrafish stripes were also shown to be consistent with a RD [58,87].

RD has also been applied in a variety of other biological systems. As it is particularly easy to implement in a simple two dimensional space they have been used to better understand the formation of several ectodermal appendages such as

hair follicle spacing in mouse [88], or feathers patterning in birds [89]. More complex systems such as branching morphogenesis in the lung, or teeth patterning have also been explored [90,91]. By changing the parameters and initial conditions of the systems, RD can generate a virtually unlimited variety of spatial pattern [58]. We thus expect that a large proportion of pigmentation patterns observed in coral reef fish could be explained through RD.

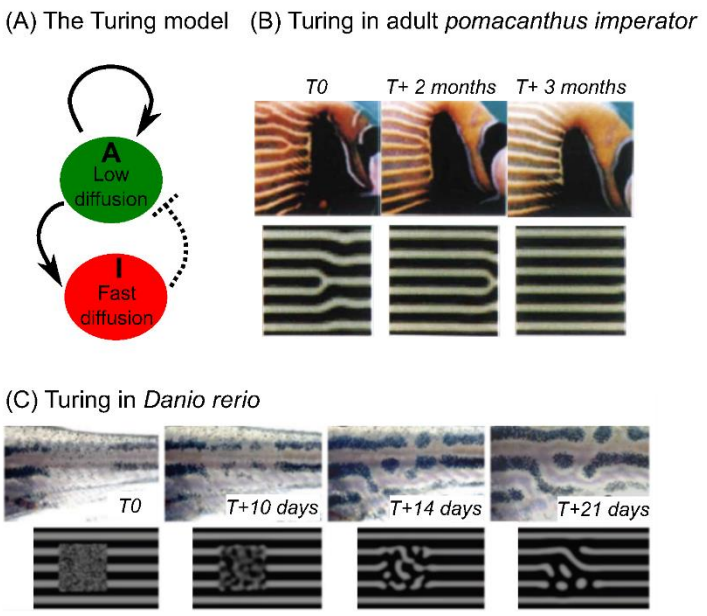


Figure Box 2. Stripe formation is predicted by Turing model in fishes
(A) The activator stimulates the production of both itself and its inhibitor (arrows). The inhibitor turns off the production of the activator (dashed line). As the inhibitor molecule diffuses more rapidly than the activator, it impairs activation at long range. If the activator is sufficiently strong it is in sufficient amount to preclude its inhibition at short range. (B) Rearrangement of the stripe of the same adult *Pomacanthus imperator* (up panel) and its computer simulation (down panel): as they grow, the number of lines increases proportionally to body size whereas the width remain constant. At to, *P. imperator* contains a branching point, during growth, the branching point move horizontally (to the anterior) like a zip resulting in its fusion and thus in the addition of a new line[53]. (C) Regeneration of labyrinthine pattern of the same adult zebrafish after laser ablation (ablation of the three pigment cells) (up panel) and its computer simulation (down panel): the stripes developed but the directionality is lost [87].

RÉFÉRENCES BIBLIOGRAPHIQUES

Abol-Munafi A, Norazmi-Lokman N, Asma N, Sarmiza S, Abduh M (2011). Histological study on the gonad of the protandrous anemonefish Amphiprion ocellaris. *J Anim Vet Adv* **10**: 3031–3036.

Adams K (2013). Genomic Clues to the Ancestral Flowering Plant. *Science (80-)* **342**: 1456–1457.

Ahn S, Tanksley SD (1993). Comparative linkage maps of the rice and maize genomes. *Proc Natl Acad Sci U S A* **90**: 7980–4.

Ainsworth SJ, Stanley RL, Evans DJR (2010). Developmental stages of the Japanese quail. *J Anat* **216**: 3–15.

Albalat R, Cañestro C (2016). Evolution by gene loss. *Nat Rev Genet* **17**: 379–391.

Albertin W, Marullo P (2012). Polyploidy in fungi: evolution after whole-genome duplication. *Proc R Soc B Biol Sci* **279**: 2497–2509.

Alexandrou MA, Swartz BA, Matzke NJ, Oakley TH (2013). Genome duplication and multiple evolutionary origins of complex migratory behavior in Salmonidae. *Mol Phylogenet Evol* **69**: 514–523.

Alfaro ME, Santini F, Brock CD (2007). Do reefs drive diversification in marine teleosts? Evidence from the pufferfish and their allies (Order Tetraodontiformes). *Evolution (N Y)* **61**: 2104–2126.

Allen BM (1916). The Results of Extirpation of the Anterior Lobe of the Hypophysis and of the Thyroid of Rana Pipiens Larvæ. *Science (80-)* **44**: 755–758.

Allen GR (1972). *The Anemonefishes: Their Classification and Biology*. T. F. H. Publication Inc., Neptune City, New Jersey.

Allen GR (1991). *Damselfishes of the world*. Mergus Publishes, Melle, Germany.

Allender CJ, Seehausen O, Knight ME, Turner GF, Maclean N (2003). Divergent selection during speciation of Lake Malawi cichlid fishes inferred from parallel radiations in nuptial coloration. *Proc Natl Acad Sci* **100**: 14074–14079.

Allendorf FW, Danzmann RG (1997). Secondary tetrasomic segregation of MDH-B and preferential pairing of homeologues in rainbow trout. *Genetics* **145**: 1083–92.

Almany GR, Berumen ML, Thorrold SR, Planes S, Jones GP (2007). Local Replenishment of Coral Reef Fish Populations in a Marine Reserve. *Science (80-)* **316**: 742–744.

Almany GR, Planes S, Thorrold SR, Berumen ML, Bode M, Saenz-Agudelo P, *et al.* (2017). Larval fish dispersal in a coral-reef seascape. *Nat Ecol Evol* **1**: 148.

Alonso WJ (2016). Evolution of bright colours in animals: worlds of prohibition and oblivion. *F1000Research* **4**: 115.

Altincicek B, Kovacs JL, Gerardo NM (2012). Horizontally transferred fungal carotenoid genes in the two-spotted spider mite Tetranychus urticae. *Biol Lett* **8**: 253–257.

Ambrosio AL, Boyle JA, Aradi AE, Christian KA, Di Pietro SM (2016). TPC2 controls pigmentation by regulating melanosome pH and size. *Proc Natl Acad Sci* **113**: 5622–5627.

Amores A, Force A, Yan YL, Joly L, Amemiya C, Fritz A, *et al.* (1998). Zebrafish hox clusters and vertebrate genome evolution. *Science* **282**: 1711–4.

Amoutzias GD, He Y, Gordon J, Mossialos D, Oliver SG, Van de Peer Y (2010). Posttranslational regulation impacts the fate of duplicated genes. *Proc Natl Acad Sci* **107**: 2967–2971.

Amundsen (2000). Why are female birds ornamented? *Trends Ecol Evol* **15**: 149–155.

Amundsen T, Forsgren E (2001). Male mate choice selects for female coloration in a fish. *Proc Natl Acad Sci* **98**: 13155–13160.

Anavy L, Levin M, Khair S, Nakanishi N, Fernandez-Valverde SL, Degnan BM, *et al.* (2014). BLIND ordering of large-scale transcriptomic developmental timecourses. *Development* **141**: 1161–1166.

Andersson M, Simmons LW (2006). Sexual selection and mate choice. *Trends Ecol Evol* **21**: 296–302.

Aparicio S, Chapman J, Stupka E, Putnam N, Chia J, Dehal P, *et al.* (2002). Whole-Genome Shotgun Assembly and Analysis of the Genome of Fugu rubripes. *Science (80-)* **297**: 1301–1310.

Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* **408**: 796–815.

Arnoult L, Su KFY, Manoel D, Minervino C, Magrina J, Gompel N, *et al.* (2013). Emergence and Diversification of Fly Pigmentation Through Evolution of a Gene Regulatory Module. *Science (80-)* **339**: 1423–1426.

Arvedlund M, Brolund TM, Nielsen LE (2003). Morphology and cytology of the olfactory organs in small juvenile Dascyllus aruanus and Amphiprion ocellaris (Pisces: Pomacentridae). *J Mar Biol Assoc UK* **83**: 1321–1326.

Arvedlund M, Bundgaard I, Nielsen LE (2000). Host imprinting in anemonefishes (Pisces: Pomacentridae): does it dictate spawning site preferences? *Environ Biol fishes*: 203–213.

Arvedlund M, McCormick M, Ainsworth T (2000). Effects of photoperiod on growth of larvae and juveniles of the anemonefish Amphiprion melanopus. *Naga, WorldFish Cent* **23**: 18–23.

Arvedlund M, McCormick M, Fautin D, Bildsøe M (1999). Host recognition and possible imprinting in the anemonefish Amphiprion melanopus (Pisces:Pomacentridae). *Mar Ecol Prog Ser* **188**: 207–218.

Aury J-M, Jaillon O, Duret L, Noel B, Jubin C, Porcel BM, *et al.* (2006). Global trends of whole-genome duplications revealed by the ciliate Paramecium tetraurelia. *Nature* **444**: 171–178.

Avella MA, Olivotto I, Gioacchini G, Maradonna F, Carnevali O (2007). The role of fatty acids enrichments in the larviculture of false percula clownfish Amphiprion ocellaris. *Aquaculture* **273**: 87–95.

Avella MA, Olivotto I, Silvi S, Place AR, Carnevali O (2010). Effect of dietary probiotics on clownfish: a molecular approach to define how lactic acid bacteria modulate development in a marine fish. *Am J Physiol Integr Comp Physiol* **298**: R359–R371.

Badouin H, Gouzy J, Grassa CJ, Murat F, Staton SE, Cottret L, *et al.* (2017). The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature* **546**: 148–152.

Bagnara JT (1983). Developmental Aspects of Vertebrate Chromatophores. *Am Zool* **23**: 465–478.

Bagnara JT, Fernandez PJ, Fujii R (2007). On the blue coloration of vertebrates. *Pigment Cell Res* **20**: 14–26.

Bagnara JT, Hadley EM (1973). *Chromatophores and Color Change. The Comparative Physiology of Animal Pigmentation*. Englewood Cliffs: Prentice-Hall.

Bagnara JT, Matsumoto J (2006). Comparative Anatomy and Physiology of Pigment Cells in Nonmammalian Tissues, 2nd ed. In: *The Pigmentary System*, Blackwell Publishing Ltd: Oxford, UK, pp 11–59.

Bagnara JT, Taylor JD, Hadley ME (1968). The dermal chromatophore unit. *J Cell Biol* **38**: 67–79.

Baldwin CC (2013). The phylogenetic significance of colour patterns in marine teleost larvae. *Zool J Linn Soc* **168**: 496–563.

Balment RJ, Lu W, Weybourne E, Warne JM (2006). Arginine vasotocin a key hormone in fish physiology and behaviour: A review with insights from mammalian models. *Gen Comp Endocrinol* **147**: 9–16.

Ban E, Kasai A, Sato M, Yokozeki A, Hisatomi O, Oshima N (2005). The signaling pathway in photoresponses that may be mediated by visual pigments in erythrophores of Nile tilapia. *Pigment Cell Res* **18**: 360–369.

Bar-Joseph Z, Gitter A, Simon I (2012). Studying and modelling dynamic biological processes using time-series gene expression data. *Nat Rev Genet* **13**: 552–564.

Baron M, Yanai I (2017). It's about time: studying gene regulatory programs across serial organs. *Genome Biol* **18**: 30.

Barreira AS, Lagorio MG, Lijtmaer DA, Lougheed SC, Tubaro PL (2012). Fluorescent and ultraviolet sexual dichromatism in the blue-winged parrotlet (A Kitchener, Ed.). *J Zool* **288**: 135–142.

Barsh GS (2003). What Controls Variation in Human Skin Color? *PLoS Biol* **1**: e27.

Bates HB (1862). Contributions to an insect fauna of the Amazon valley

(Lepidoptera: Heliconidae). *Trans Linn Soc London* **23**: 495–556.

Bateson W (1902). *Mendel's principles of heredity*. Cambridge University Press.

Baumel A, Ainouche ML, Levasseur JE (2001). Molecular investigations in populations of Spartina anglica C.E. Hubbard (Poaceae) invading coastal Brittany (France). *Mol Ecol* **10**: 1689–1701.

Beck CW (1997). Effect of changes in resource level on age and size at metamorphosis in Hyla squirella. *Oecologia* **112**: 187–192.

Beldade R, Holbrook SJ, Schmitt RJ, Planes S, Malone D, Bernardi G (2012). Larger female fish contribute disproportionately more to self-replenishment. *Proc R Soc B Biol Sci* **279**: 2116–2121.

Bellwood DR (1998). What are reef fishes? - Comment on the report by D. R. Robertson: Do coral-reef fish faunas have a distinctive taxonomic structure? (Coral Reefs 17: 179-186). *Coral Reefs* **17**: 187–189.

Bellwood DR, Hughes TP, Connolly SR, Tanner J (2005). Environmental and geometric constraints on Indo-Pacific coral reef biodiversity. *Ecol Lett* **8**: 643–651.

Bellwood DR, Hughes TP, Folke C, Nyström M (2004). Confronting the coral reef crisis. *Nature* **429**: 827–33.

Benjamini Y, Speed TP (2012). Summarizing and correcting the GC content bias in high-throughput sequencing. *Nucleic Acids Res* **40**: e72–e72.

Bennett DC, Lamoreux ML (2003). The Color Loci of Mice - A Genetic Century. *Pigment Cell Res* **16**: 333–344.

Bergé-Lefranc JL, Cartouzou G, Malthiéry Y, Perrin F, Jarry B, Lissitzky S (1981). Cloning of four DNA fragments complementary to human thyroglobulin messenger RNA. *Eur J Biochem* **120**: 1–7.

Berger MF, Levin JZ, Vijayendran K, Sivachenko A, Adiconis X, Maguire J, *et al.* (2010). Integrative analysis of the melanoma transcriptome. *Genome Res* **20**: 413–427.

Bernardi G (2011). Monophyletic origin of brood care in damselfishes. *Mol Phylogenet Evol* **59**: 245–248.

Bernardi G, Crane NL, Longo GC, Quiros AL (2017). The ecology of Altrichthys azurelineatus and A. curatus, two damselfishes that lack a pelagic larval phase. *Environ Biol Fishes* **100**: 111–120.

Berthelot C, Brunet F, Chalopin D, Juanchich A, Bernard M, Noël B, *et al.* (2014). The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nat Commun* **5**.

Besson M, Salis P, Laudet V, Lecchini D (2018). Complete and rapid reversal of the body color pattern in juveniles of the convict surgeonfish Acanthurus triostegus at Moorea Island (French Polynesia). *Coral Reefs* **37**: 31–35.

Betancur-R R, Wiley EO, Arratia G, Acero A, Bailly N, Miya M, *et al.* (2017). Phylogenetic classification of bony fishes. *BMC Evol Biol* **17**: 162.

Béziers P, Ducrest A-L, Simon C, Roulin A (2017). Circulating testosterone and feather-gene expression of receptors and metabolic enzymes in relation to melanin-based colouration in the barn owl. *Gen Comp Endocrinol* **250**: 36–45.

Bianco A, Kim B (2006). Deiodinases: implications of the local control of thyroid hormone action. *J Clin Invest* **116**: 2571–2579.

Birchler JA, Riddle NC, Auger DL, Veitia RA (2005). Dosage balance in gene regulation: biological implications. *Trends Genet* **21**: 219–226.

Blaxter JHS (1988). Pattern and Variety in Development. In: *The Physiology of Developing Fish*, Academic Press, San Diego, pp 1–58.

Bolker JA (1995). Model systems in developmental biology. *BioEssays* **17**: 451–455.

Bolker J (2012). Model organisms: There's more to life than rats and flies. *Nature* **491**: 31–33.

Boorse GC, Denver RJ (2004). Expression and hypophysiotropic actions of corticotropin-releasing factor in Xenopus laevis. *Gen Comp Endocrinol* **137**: 272–282.

Botía JA, Vandrovcova J, Forabosco P, Guelfi S, D'Sa K, Hardy J, *et al.* (2017). An additional k-means clustering step improves the biological features of WGCNA gene co-expression networks. *BMC Syst Biol* **11**: 47.

Boto L (2014). Horizontal gene transfer in the acquisition of novel traits by metazoans. *Proc R Soc B Biol Sci* **281**: 20132450.

Bower C, Turner D, Biever R (1887). A standardized method of propagating the marine fish parasite, Amyloodinium ocellatum. *J Parasitol* **73**: 85–88.

Braasch I, Brunet F, Volff J-N, Schartl M (2009). Pigmentation pathway evolution after whole-genome duplication in fish. *Genome Biol Evol* **1**: 479–493.

Braasch I, Gehrke AR, Smith JJ, Kawasaki K, Manousaki T, Pasquier J, *et al.* (2016). The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. *Nat Genet* **48**: 427–437.

Braasch I, Peterson SM, Desvignes T, McCluskey BM, Batzel P, Postlethwait JH (2015). A new model army: Emerging fish models to study the genomics of vertebrate Evo-Devo. *J Exp Zool Part B Mol Dev Evol* **324**: 316–341.

Braasch I, Postlethwait JH (2011). The teleost agouti-related protein 2 gene is an ohnolog gone missing from the tetrapod genome. *Proc Natl Acad Sci* **108**: E47–E48.

Braasch I, Schartl M, Volff J-N (2007). Evolution of pigment synthesis pathways by gene and genome duplication in fish. *BMC Evol Biol* **7**: 74.

Braasch I, Volff JN, Schartl M (2008). The evolution of teleost pigmentation and the fish-specific genome duplication. *J Fish Biol* **73**: 1891–1918.

Bradley D, Xu P, Mohorianu I-I, Whibley A, Field D, Tavares H, *et al.* (2017). Evolution of flower color pattern through selection on regulatory small RNAs. *Science (80-)* **358**: 925–928.

Bray NL, Pimentel H, Melsted P, Pachter L (2016). Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**: 525–527.

Brenner M, Hearing VJ (2008). The Protective Role of Melanin Against UV Damage in Human Skin†. *Photochem Photobiol* **84**: 539–549.

Bright LJ, Gout J-F, Lynch M (2017). Early stages of functional diversification in the Rab GTPase gene family revealed by genomic and localization studies in Paramecium species (PJ Brennwald, Ed.). *Mol Biol Cell* **28**: 1101–1110.

Brolund TM, Nielsen LE, Arvedlund M (2003). Do juvenile Amphiprion ocellaris (Pisces: Pomacentridae) recognize conspecifics by chemical or visual cues? *J Mar Biol Assoc UK* **83**: S0025315403008385.

Brooker RM, Dixson DL (2016). Assessing the role of olfactory cues in the early life history of coral reef fish: current methods and future directions. In: *Chemical Signals in Vertebrates*,, pp 229–244.

Brown DD (1997). The role of thyroid hormone in zebrafish and axolotl development. *Proc Natl Acad Sci U S A* **94**: 13011–6.

Brown DD, Cai L (2007). Amphibian metamorphosis. *Dev Biol* **306**: 20–33.

Brumbaugh R (2017). Protecting Million Dollar Reefs is Key to Sustaining Global Tourism.

Brunet FG, Roest Crollius H, Paris M, Aury J, Gibert P, Laudet OJ V., *et al.* (2006). Gene Loss and Evolutionary Rates Following Whole-Genome Duplication in Teleost Fishes. *Mol Biol Evol* **23**: 1808–1816.

Brunet FG, Volff J-N, Schartl M (2016). Whole Genome Duplications Shaped the Receptor Tyrosine Kinase Repertoire of Jawed Vertebrates. *Genome Biol Evol* **8**: 1600–1613.

Burgus R, Dunn TF, Desiderio D, Guillemin R (1969). Molecular structure of the hypothalamic hypophysiotropic TRF factor of ovine origin: mass spectrometry demonstration of the PCA-His-Pro-NH2 sequence. *C R Acad Sci Hebd Seances Acad Sci D* **269**: 1870–3.

Buston P (2003). Size and growth modification in clownfish. *Nature* **424**: 145–146.

Buston PM, García MB (2007). An extraordinary life span estimate for the clown anemonefish Amphiprion percula. *J Fish Biol* **70**: 1710–1719.

Cai L, Brown DD (2004). Expression of type II iodothyronine deiodinase marks the time that a tissue responds to thyroid hormone-induced metamorphosis in Xenopus laevis. *Dev Biol* **266**: 87–95.

Cal L, Megías M, Cerdá-reverter J miguel, Postlethwait J h., Braasch I, Rotllant J (2017). BAC Recombineering of the Agouti Loci from Spotted Gar and Zebrafish Reveals the Evolutionary Ancestry of Dorsal-Ventral Pigment Asymmetry in Fish. *J Exp Zool Part B Mol Dev Evol*.

Campinho MA, Galay-Burgos M, Sweeney GE, Power DM (2010). Coordination of deiodinase and thyroid hormone receptor expression during the larval to juvenile transition in sea bream (Sparus aurata, Linnaeus). *Gen Comp Endocrinol* **165**: 181–194.

Cao Z, Yu Y, Wu Y, Hao P, Di Z, He Y, *et al.* (2013). The genome of Mesobuthus martensii reveals a unique adaptation model of arthropods. *Nat Commun* **4**.

Caro T (2005). The Adaptive Significance of Coloration in Mammals. *Bioscience* **55**: 125–136.

Caro T (2017). Wallace on Coloration: Contemporary Perspective and Unresolved Insights. *Trends Ecol Evol* **32**: 23–30.

Casane D, Rétaux S (2016). Evolutionary Genetics of the Cavefish Astyanax mexicanus. In:, pp 117–159.

Casas L, Saborido-Rey F, Ryu T, Michell C, Ravasi T, Irigoien X (2016). Sex Change in Clownfish: Molecular Insights from Transcriptome Analysis. *Sci Rep* **6**: 35461.

Casas L, Saenz-Agudelo P, Irigoien X (2018). High-Throughput Sequencing and Linkage Mapping of a Clownfish Genome Provide Insights on the Distribution of Molecular Players Involved in Sex Change. *Sci Rep* **8**: 4073.

Ceinos RM, Guillot R, Kelsh RN, Cerdá-Reverter JM, Rotllant J (2015). Pigment patterns in adult fish result from superimposition of two largely independent pigmentation mechanisms. *Pigment Cell Melanoma Res* **28**: 196–209.

Chalopin D, Volff J-N (2017). Analysis of the spotted gar genome suggests absence of causative link between ancestral genome duplication and transposable element diversification in teleost fish. *J Exp Zool Part B Mol Dev Evol* **328**: 629–637.

Chamala S, Chanderbali AS, Der JP, Lan T, Walts B, Albert VA, *et al.* (2013). Assembly and Validation of the Genome of the Nonmodel Basal Angiosperm Amborella. *Science (80-)* **342**: 1516–1517.

Chapman BA, Bowers JE, Feltus FA, Paterson AH (2006). Buffering of crucial functions by paleologous duplicated genes may contribute cyclicality to angiosperm genome duplication. *Proc Natl Acad Sci* **103**: 2730–2735.

Chaudhari N, Hahn W (1983). Genetic expression in the developing brain. *Science (80-)* **220**: 924–928.

Chen ZJ (2010). Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci* **15**: 57–71.

Chen T-H, Hsieh C-Y (2017). Fighting Nemo: Effect of 17α-ethinylestradiol (EE2) on aggressive behavior and social hierarchy of the false clown anemonefish Amphiprion ocellaris. *Mar Pollut Bull* **124**: 760–766.

Chen S-C, Robertson RM, Hawryshyn CW (2013). Possible Involvement of Cone Opsins in Distinct Photoresponses of Intrinsically Photosensitive Dermal Chromatophores in Tilapia Oreochromis niloticus (K-W Koch, Ed.). *PLoS One* **8**: e70342.

Chen S-C, Xiao C, Troje NF, Robertson RM, Hawryshyn CW (2015). Functional characterisation of the chromatically antagonistic photosensitive mechanism of erythrophores in the tilapia Oreochromis niloticus. *J Exp Biol* **218**: 748–756.

Cheney KL, Newport C, McClure EC, Marshall NJ (2013). Colour vision and response bias in a coral reef fish. *J Exp Biol* **216**: 2967–2973.

Cherry JL (2010). Expression Level, Evolutionary Rate, and the Cost of Expression. *Genome Biol Evol* **2**: 757–769.

Choat JH, Bellwood DR (1991). Reef Fishes: Their History and Evolution. In: *The Ecology of Fishes on Coral Reefs*, Elsevier, pp 39–66.

Clarke JT, Lloyd GT, Friedman M (2016). Little evidence for enhanced phenotypic evolution in early teleosts relative to their living fossil sister group. *Proc Natl Acad Sci* **113**: 11531–11536.

Clement S, Lichtenbert J, Kohler C (2001). Stripping clowns: induced meristic changes in common clownfish (Amphiprion ocellaris). *Bull l'Institut océanographique Monaco*.

Clobert J, Baguette M, Benton TG, Bullock JM (2012). *Dispersal ecology and evolution*.

Cloonan N, Forrest ARR, Kolle G, Gardiner BBA, Faulkner GJ, Brown MK, *et al.* (2008). Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat Methods* **5**: 613–619.

Clutton-Brock T (2010). We do not need a Sexual Selection 2.0—nor a theory of Genial Selection. *Anim Behav* **79**: e7–e10.

Cobbs C, Heath J, Stireman JO, Abbot P (2013). Carotenoids in unexpected places: Gall midges, lateral gene transfer, and carotenoid biosynthesis in animals. *Mol Phylogenet Evol* **68**: 221–228.

Cold Spring Harbor Laboratory Press (2010). *Emerging Model Organisms: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Ed.).

Colleye O, Frederich B, Vandewalle P, Casadevall M, Parmentier E (2009). Agonistic sounds in the skunk clownfish Amphiprion akallopisos : size-related variation in acoustic features. *J Fish Biol* **75**: 908–916.

Colleye O, Iwata E, Parmentier E (2016). Clownfishes. In: Frédérich B, Parmentier E (eds) *Biology of Damselfishes*, Taylor & Francis, pp 246–266.

Colleye O, Parmentier E (2012). Overview on the Diversity of Sounds Produced by Clownfishes (Pomacentridae): Importance of Acoustic Signals in Their Peculiar Way of Life (SCF Neuhauss, Ed.). *PLoS One* **7**: e49179.

Colleye O, Vandewalle P, Lanterbecq D, Lecchini D, Parmentier E (2011). Interspecific variation of calls in clownfishes: degree of similarity in closely related species. *BMC Evol Biol* **11**: 365.

Conant GC, Birchler JA, Pires JC (2014). Dosage, duplication, and diploidization: clarifying the interplay of multiple models for duplicate gene evolution over time. *Curr Opin Plant Biol* **19**: 91–98.

Conant GC, Wolfe KH (2008). Turning a hobby into a job: How duplicated genes find new functions. *Nat Rev Genet* **9**: 938–950.

Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, *et al.* (2016). A survey of best practices for RNA-seq data analysis. *Genome Biol* **17**: 13.

Cooper WJ, Westneat MW (2009). Form and function of damselfish skulls: rapid and repeated evolution into a limited number of trophic niches. *BMC Evol Biol* **9**: 24.

Cortesi F, Feeney WE, Ferrari MCO, Waldie PA, Phillips GAC, McClure EC, *et al.* (2015). Phenotypic Plasticity Confers Multiple Fitness Benefits to a Mimic. *Curr Biol* **25**: 949–954.

Cortesi F, Musilová Z, Stieb SM, Hart NS, Siebeck UE, Cheney KL, *et al.* (2016). From crypsis to mimicry: changes in colour and the configuration of the visual system during ontogenetic habitat transitions in a coral reef fish. *J Exp Biol* **219**: 2545–2558.

Costanza R, de Groot R, Sutton P, van der Ploeg S, Anderson SJ, Kubiszewski I, *et al.* (2014). Changes in the global value of ecosystem services. *Glob Environ Chang* **26**: 152–158.

Cott H (1940). *Adaptive Coloration in Animals*. Oxford University Press.

Cowen RK, Paris CB, Srinivasan A (2005). Scaling of Connectivity in Marine Populations. *Science (80-)* **311**: 522–527.

Cowen RK, Sponaugle S (2009). Larval Dispersal and Marine Population Connectivity. *Annu Rev Mar Sci* **1**: 443–66.

Cowman PF, Bellwood DR (2011). Coral reefs as drivers of cladogenesis: expanding coral reefs, cryptic extinction events, and the development of biodiversity hotspots. *J Evol Biol* **24**: 2543–2562.

Crawford NG, Kelly DE, Hansen MEB, Beltrame MH, Fan S, Bowman SL, *et al.* (2017). Loci associated with skin pigmentation identified in African populations. *Science (80-)* **358**: eaan8433.

Curran K, Lister JA, Kunkel GR, Prendergast A, Parichy DM, Raible DW (2010). Interplay between Foxd3 and Mitf regulates cell fate plasticity in the zebrafish neural crest. *Dev Biol* **344**: 107–118.

Cuthill IC, Allen WL, Arbuckle K, Caspers B, Chaplin G, Hauber ME, *et al.* (2017). The biology of color. *Science (80-)* **357**: eaan0221.

Cuthill IC, Stevens M, Sheppard J, Maddocks T, Párraga CA, Troscianko TS (2005). Disruptive coloration and background pattern matching. *Nature* **434**: 72–74.

D'Alba L, Kieffer L, Shawkey MD (2012). Relative contributions of pigments and biophotonic nanostructures to natural color production: a case study in

budgerigar (Melopsittacus undulatus) feathers. *J Exp Biol* **215**: 1272–1277.

D'Mello S, Finlay G, Baguley B, Askarian-Amiri M (2016). Signaling Pathways in Melanogenesis. *Int J Mol Sci* **17**: 1144.

Darwin C (1868). *The Variation of Animals and Plants Under Domestication*. John Murray, London.

Darwin C (1871). *The Descent of Man, and Selection in Relation to Sex*. John Murray, London.

David L, Blum S, Feldman M, Lavi U, Hillel J (2003). Recent Duplication of the Common Carp (Cyprinus carpio L.) Genome as Revealed by Analyses of Microsatellite Loci. *Mol Biol Evol* **20**: 1425–1434.

Davis J, Petrov D (2005). Do disparate mechanisms of duplication add similar genes to the genome? *Trends Genet* **21**: 548–551.

De'ath G, Fabricius KE, Sweatman H, Puotinen M (2012). The 27-year decline of coral cover on the Great Barrier Reef and its causes. *Proc Natl Acad Sci* **109**: 17995–17999.

DeAngelis R, Gogola J, Dodd L, Rhodes JS (2017). Opposite effects of nonapeptide antagonists on paternal behavior in the teleost fish Amphiprion ocellaris. *Horm Behav* **90**: 113–119.

DeAngelis RS, Rhodes JS (2016). Sex Differences in Steroid Hormones and Parental Effort across the Breeding Cycle in Amphiprion ocellaris. *Copeia* **104**: 586–593.

Dehal P, Boore JL (2005). Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate (P Holland, Ed.). *PLoS Biol* **3**: e314.

Demeneix B (2017). *Cocktail toxique - Comment les perturbateurs endocriniens empoisonnent notre cerveau*. Odile Jacob.

Dennenmoser S, Sedlazeck FJ, Iwaszkiewicz E, Li X-Y, Altmüller J, Nolte AW (2017). Copy number increases of transposable elements and protein-coding genes in an invasive fish of hybrid origin. *Mol Ecol* **26**: 4712–4724.

Denver RJ (1988). Several hypothalamic peptides stimulate in vitro thyrotropin secretion by pituitaries of anuran amphibians. *Gen Comp Endocrinol* **72**: 383–93.

Dhayanithi N, Kumar T, Kathiresan K (2010). Effect of neem extract against the bacteria isolated from marine fish. *J Env Biol* **31**: 409–412.

Dickson D (1999). Gene estimate rises as US and UK discuss freedom of access. *Nature* **401**: 311–311.

Djurdjevič I, Kreft ME, Sušnik Bajec S (2015). Comparison of pigment cell ultrastructure and organisation in the dermis of marble trout and brown trout, and first description of erythrophore ultrastructure in salmonids. *J Anat* **227**: 583–595.

Dodsworth S, Chase MW, Leitch AR (2016). Is post-polyploidization diploidization the key to the evolutionary success of angiosperms? *Bot J Linn Soc* **180**: 1–5.

Doherty PJ (2002). Variable replenishment and the dynamics of reef fish populations. In: *Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem*,, pp 327–355.

Domyan ET, Shapiro MD (2017). Pigeonetics takes flight: Evolution, development, and genetics of intraspecific variation. *Dev Biol* **427**: 241–250.

Donoghue PCJ, Purnell M (2005). Genome duplication, extinction and vertebrate evolution. *Trends Ecol Evol* **20**: 312–319.

Dooley CM, Mongera A, Walderich B, Nusslein-Volhard C (2013). On the embryonic origin of adult melanophores: the role of ErbB and Kit signalling in establishing melanophore stem cells in zebrafish. *Development* **140**: 1003–1013.

Le Douarin NM, Dupin E (2003). Multipotentiality of the neural crest. *Curr Opin Genet Dev* **13**: 529–536.

Doudna JA, Charpentier E (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science (80-)* **346**: 1258096–1258096.

Douglas GM, Gos G, Steige KA, Salcedo A, Holm K, Josephs EB, *et al.* (2015). Hybrid origins and the earliest stages of diploidization in the highly successful recent polyploid Capsella bursa-pastoris. *Proc Natl Acad Sci* **112**: 2806–2811.

Dreiss AN, Séchaud R, Béziers P, Villain N, Genoud M, Almasi B, *et al.* (2016). Social huddling and physiological thermoregulation are related to melanism in the nocturnal barn owl. *Oecologia* **180**: 371–381.

Duboule D (1994). Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Dev Suppl*: 135–42.

Ducrest A, Keller L, Roulin A (2008). Pleiotropy in the melanocortin system, coloration and behavioural syndromes. *Trends Ecol Evol* **23**: 502–510.

Dujon B (1996). The yeast genome project: what did we learn? *Trends Genet* **12**: 263–270.

Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, *et al.* (2004). Genome evolution in yeasts. *Nature* **430**: 35–44.

Elliott JK, Lougheed SC, Bateman B, McPhee LK, Boag PT (1999). Molecular phylogenetic evidence for the evolution of specialization in anemonefishes. *Proc R Soc B Biol Sci* **266**: 677–685.

Emerson SB, Cooper TA, Ehleringer JR (1990). Convergence in Reflectance Spectra Among Treefrogs. *Funct Ecol* **4**: 47.

Ensminger P (2001). *Life Under the Sun*. Yale University Press.

Eom DS, Bain EJ, Patterson LB, Grout ME, Parichy DM (2015). Long-distance communication by specialized cellular projections during pigment pattern development and evolution. *Elife* **4**.

Eom DS, Parichy DM (2017). A macrophage relay for long-distance signaling during postembryonic tissue remodeling. *Science (80-)* **355**: 1317–1320.

Evans RM, Mangelsdorf DJ (2014). Nuclear Receptors, RXR, and the Big Bang. *Cell* **157**: 255–266.

Eyre BD, Cyronak T, Drupp P, De Carlo EH, Sachs JP, Andersson AJ (2018). Coral reefs will transition to net dissolving before end of century. *Science (80-)* **359**: 908–911.

Eytan RI, Evans BR, Dornburg A, Lemmon AR, Lemmon EM, Wainwright PC, *et al.* (2015). Are 100 enough? Inferring acanthomorph teleost phylogeny using Anchored Hybrid Enrichment. *BMC Evol Biol* **15**: 113.

Fautin D, Allen GR (1992). *Field Guide to Anemone Fishes and Their Host Sea Anemones*. Western Australian Museum.

Fautin D, Allen GR (1997). *Anemone fishes and their host sea anemones*. Western Australian Museum.

Feitosa JLL, Concentino AM, Teixeira SF, Ferreira BP (2012). Food resource use by two territorial damselfish (Pomacentridae: Stegastes) on South-Western Atlantic algal-dominated reefs. *J Sea Res* **70**: 42–49.

Ferraresso S, Bonaldo A, Parma L, Cinotti S, Massi P, Bargelloni L, *et al.* (2013). Exploring the larval transcriptome of the common sole (Solea solea L.). *BMC Genomics* **14**: 315.

Ferris SD, Whitt GS (1977). Duplicate gene expression in diploid and tetraploid loaches (Cypriniformes, Cobitidae). *Biochem Genet* **15**: 1097–112.

Fisher R (1958). *The Genetical Theory of Natural Selection, 2nd ed.* Dover, N. Y.

Fisher R (2005). Swimming speeds of larval coral reef fishes: Impacts on self-recruitment and dispersal. *Mar Ecol Prog Ser* **285**: 223–232.

Fisher R, Bellwood D, Job S (2000). Development of swimming abilities in reef fish larvae. *Mar Ecol Prog Ser* **202**: 163–173.

Fisher R, Leis JM, Clark DL, Wilson SK (2005). Critical swimming speeds of late-stage coral reef fish larvae: variation within species, among species and between locations. *Mar Biol* **147**: 1201–1212.

Fitzpatrick DA (2012). Horizontal gene transfer in fungi. *FEMS Microbiol Lett* **329**: 1–8.

Fleisch VC, Schonthaler HB, von Lintig J, Neuhauss SCF (2008). Subfunctionalization of a Retinoid-Binding Protein Provides Evidence for Two Parallel Visual Cycles in the Cone-Dominant Zebrafish Retina. *J Neurosci* **28**: 8208–8216.

Flot J-F, Hespeels B, Li X, Noel B, Arkhipova I, Danchin EGJ, *et al.* (2013). Genomic evidence for ameiotic evolution in the bdelloid rotifer Adineta vaga. *Nature* **500**: 453–457.

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J (1999).

Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**: 1531–45.

Fraser RH, Currie DJ (1996). The Species Richness-Energy Hypothesis in a System Where Historical Factors Are Thought to Prevail: Coral Reefs. *Am Nat* **148**: 138–159.

Frédérich B, Fabri G, Lepoint G, Vandewalle P, Parmentier E (2009). Trophic niches of thirteen damselfishes (Pomacentridae) at the Grand Récif of Toliara, Madagascar. *Ichthyol Res* **56**: 10–17.

Frédérich B, Parmentier E (2016). *Biology of Damselfishes* (B Frédérich and E Parmentier, Eds.). CRC Press.

Frédérich B, Sorenson L, Santini F, Slater GJ, Alfaro ME (2013). Iterative Ecological Radiation and Convergence during the Evolutionary History of Damselfishes (Pomacentridae). *Am Nat* **181**: 94–113.

Freeling M, Thomas BC (2006). Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. *Genome Res* **16**: 805–814.

Fricke HW (1973). Individual partner recognition in fish: Field studies onAmphiprion bicinctus. *Naturwissenschaften* **60**: 204–204.

Fricke HW (1979). Mating System, Resource Defence and Sex Change in the Anemonefish Amphiprion akallopisos. *Z Tierpsychol* **50**: 313–326.

Fricke HW, Fricke S (1977). Monogamy and sex change by aggressive dominance in coral reef fish. *Nature* **266**: 830–832.

Fricke H-W, Holzberg S (1974). Social units and hermaphroditism in a pomacentrid fish. *Naturwissenschaften* **61**: 367–368.

Friedman R, Hughes AL (2003). The Temporal Distribution of Gene Duplication Events in a Set of Highly Conserved Human Gene Families. *Mol Biol Evol* **20**: 154–161.

Frith C (2016). *Charles Darwin's Life With Birds: His Complete Ornithology*. Oxford University Press.

Froese R, Pauly D (2018). FishBase.

Fuentes I, Stegemann S, Golczyk H, Karcher D, Bock R (2014). Horizontal genome transfer as an asexual path to the formation of new species. *Nature* **511**: 232–235.

Fujii R (1993). Cytophysiology of Fish Chromatophores. In: International Review of Cytology, pp 191–255.

Fukuzawa T (2004). Unusual leucophore-like cells specifically appear in the lineage of melanophores in the periodic albino mutant of Xenopus laevis. *Pigment Cell Res* **17**: 252–261.

Fukuzawa T (2010). Unusual development of light-reflecting pigment cells in intact and regenerating tail in the periodic albino mutant of Xenopus laevis. *Cell Tissue Res* **342**: 53–66.

Fumey J, Hinaux H, Noirot C, Thermes C, Rétaux S, Casane D (2018). Evidence for late Pleistocene origin of Astyanax mexicanus cavefish. *BMC Evol Biol* **18**: 43.

Furlow JD, Neff ES (2006). A developmental switch induced by thyroid hormone: Xenopus laevis metamorphosis. *Trends Endocrinol Metab* **17**: 40–47.

Gachelin G (2008). De l'animal expérimental au modèle en recherche biomédicale. *Rev Philos France Let* **133**: 319.

Gallardo MH, Bickham JW, Honeycutt RL, Ojeda RA, Köhler N (1999). Discovery of tetraploidy in a mammal. *Nature* **401**: 341–341.

Galván I, Solano F (2016). Bird Integumentary Melanins: Biosynthesis, Forms, Function and Evolution. *Int J Mol Sci* **17**: 520.

Garber M, Grabherr MG, Guttman M, Trapnell C (2011). Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat Methods* **8**: 469–477.

Gardner T a, Côté IM, Gill J a, Grant A, Watkinson AR (2003). Long-term region-wide declines in Caribbean corals. *Science* **301**: 958–960.

Gaut BS, Doebley JF (1997). DNA sequence evidence for the segmental allotetraploid origin of maize. *Proc Natl Acad Sci U S A* **94**: 6809–14.

Geissler EN, Ryan MA, Housman DE (1988). The dominant-white spotting (W) locus of the mouse encodes the c-kit proto-oncogene. *Cell* **55**: 185–192.

Genin A, Yahel G, Reidenbach M, Monismith S, Koseff J (2002). Intense Benthic Grazing on Phytoplankton in Coral Reefs Revealed Using the Control Volume Approach. *Oceanography* **15**: 90–96.

Gerlach G, Atema J, Kingsford MJ, Black KP, Miller-Sims V (2007). Smelling home can prevent dispersal of reef fish larvae. *Proc Natl Acad Sci* **104**: 858–863.

Ginger RS, Askew SE, Ogborne RM, Wilson S, Ferdinando D, Dadd T, *et al.* (2008). SLC24A5 Encodes a trans -Golgi Network Protein with Potassium-dependent Sodium-Calcium Exchange Activity That Regulates Human Epidermal Melanogenesis. *J Biol Chem* **283**: 5486–5495.

Glasauer SMK, Neuhauss SCF (2014). Whole-genome duplication in teleost fishes and its evolutionary consequences. *Mol Genet Genomics* **289**: 1045–1060.

Goda M (2017). Rapid integumental color changes due to novel iridophores in the chameleon sand tilefish Hoplolatilus chlupatyi. *Pigment Cell Melanoma Res* **30**: 368–371.

Goda M, Fujii R (1995). Blue Chromatophores in Two Species of Callionymid Fish. *Zoolog Sci* **12**: 811–813.

Goda M, Fujiyoshi Y, Sugimoto M, Fujii R (2013). Novel Dichromatic Chromatophores in the Integument of the Mandarin Fish Synchiropus splendidus. *Biol Bull* **224**: 14–17.

Goda M, Ohata M, Ikoma H, Fujiyoshi Y, Sugimoto M, Fujii R (2011). Integumental reddish-violet coloration owing to novel dichromatic chromatophores in the teleost fish, Pseudochromis diadema. *Pigment Cell Melanoma Res* **24**: 614–617.

Godin J-GJ, McDonough HE (2003). Predator preference for brightly colored males in the guppy: a viability cost for a sexually selected trait. *Behav Ecol* **14**: 194–200.

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, *et al.* (1996). Life with 6000 Genes. *Science (80-)* **274**: 546–567.

Gompel N, Prud'homme B, Wittkopp PJ, Kassner VA, Carroll SB (2005). Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in Drosophila. *Nature* **433**: 481–487.

Gopurappilly R, Ogawa S, Parhar IS (2013). Functional Significance of GnRH and Kisspeptin, and Their Cognate Receptors in Teleost Reproduction. *Front Endocrinol (Lausanne)* **4**.

Gout J-F, Kahn D, Duret L (2010). The Relationship among Gene Expression, the Evolution of Gene Dosage, and the Rate of Protein Evolution (JK Pritchard, Ed.). *PLoS Genet* **6**: e1000944.

Gout J-F, Lynch M (2015). Maintenance and Loss of Duplicated Genes by Dosage Subfunctionalization. *Mol Biol Evol* **32**: 2141–2148.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, *et al.* (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **29**: 644–652.

Gracheva EO, Ingolia NT, Kelly YM, Cordero-Morales JF, Hollopeter G, Chesler AT, *et al.* (2010). Molecular basis of infrared detection by snakes. *Nature* **464**: 1006–1011.

Graham LA, Li J, Davidson WS, Davies PL (2012). Smelt was the likely beneficiary of an antifreeze gene laterally transferred between fishes. *BMC Evol Biol* **12**: 190.

Graham LA, Lougheed SC, Ewart KV, Davies PL (2008). Lateral Transfer of a Lectin-Like Antifreeze Protein Gene in Fishes (M Isalan, Ed.). *PLoS One* **3**: e2616.

Green SA, Bronner ME (2013). Gene duplications and the early evolution of neural crest development. *Semin Cell Dev Biol* **24**: 95–100.

Green AL, Maypa AP, Almany GR, Rhodes KL, Weeks R, Abesamis RA, *et al.* (2015). Larval dispersal and movement patterns of coral reef fishes, and implications for marine reserve network design. *Biol Rev* **90**: 1215–1247.

Green SA, Simoes-Costa M, Bronner ME (2015). Evolution of vertebrates as viewed from the crest. *Nature* **520**: 474–482.

De Groef B, Van der Geyten S, Darras VM, Kühn ER (2006). Role of corticotropin-releasing hormone as a thyrotropin-releasing factor in non-mammalian vertebrates. *Gen Comp Endocrinol* **146**: 62–68.

Grunwald DJ, Eisen JS (2002). Headwaters of the zebrafish — emergence of a new model vertebrate. *Nat Rev Genet* **3**: 717–724.

Gu L, Zhang C, Xia C (2017). Revelation of the Genetic Basis for Convergent Innovative Anal Fin Pigmentation Patterns in Cichlid Fishes. *bioRxiv*.

Gudernatsch JF (1912). Feeding Experiments on tadpoles. *Arch für Entwicklungsmechanik der Org* **35**: 457–483.

Guillot R, Ceinos RM, Cal R, Rotllant J, Cerdá-Reverter JM (2012). Transient Ectopic Overexpression of Agouti-Signalling Protein 1 (Asip1) Induces Pigment Anomalies in Flatfish (J V. Planas, Ed.). *PLoS One* **7**: e48526.

Guillot R, Muriach B, Rocha A, Rotllant J, Kelsh RN, Cerdá-Reverter JM (2016). Thyroid Hormones Regulate Zebrafish Melanogenesis in a Gender-Specific Manner (Z Yin, Ed.). *PLoS One* **11**: e0166152.

Guo B (2017). Complex Genes Are Preferentially Retained After Whole-Genome Duplication in Teleost Fish. *J Mol Evol* **84**: 253–258.

Haffter P, Odenthal J, Mullins MC, Lin S, Farrell MJ, Vogelsang E, *et al.* (1996). Mutations affecting pigmentation and shape of the adult zebrafish. *Dev Genes Evol* **206**: 260–276.

Hahn MW (2009). Distinguishing Among Evolutionary Models for the Maintenance of Gene Duplicates. *J Hered* **100**: 605–617.

Hakes L, Pinney JW, Lovell SC, Oliver SG, Robertson DL (2007). All duplicates are not equal: the difference between small-scale and genome duplication. *Genome Biol* **8**: R209.

Hall BK (2000). The neural crest as a fourth germ layer and vertebrates as quadroblastic not triploblastic. *Evol Dev* **2**: 3–5.

Hamada H, Watanabe M, Lau HE, Nishida T, Hasegawa T, Parichy DM, *et al.* (2014). Involvement of Delta/Notch signaling in zebrafish adult pigment stripe patterning. *Development* **141**: 318–324.

Hamburger V, Hamilton HL (1951). A series of normal stages in the development of the chick embryo. *Dev Dyn* **195**: 231–272.

Hamilton WJ, Peterman RM (1971). Countershading in the colourful reef fish Chaetodon lunula: Concealment, communication or both? *Anim Behav* **19**: 357–364.

Harada M, Yoshinaga T, Ojima D, Iwata M (2008). cDNA cloning and expression analysis of thyroid hormone receptor in the coho salmon Oncorhynchus kisutch during smoltification. *Gen Comp Endocrinol* **155**: 658–667.

Harmelin-Vivien ML (2002). Energetics and Fish Diversity on Coral Reefs. In: *Coral Reef Fishes*, Elsevier, pp 265–274.

Hashimshony T, Feder M, Levin M, Hall BK, Yanai I (2015). Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. *Nature* **519**: 219–222.

Havelka M, Bytyutskyy D, Symonová R, Ráb P, Flajšhans M (2016). The second highest chromosome count among vertebrates is observed in cultured sturgeon and is associated with genome plasticity. *Genet Sel Evol* **48**: 12.

Heard SB, Hauser DL (1995). Key evolutionary innovations and their ecological mechanisms. *Hist Biol* **10**: 151–173.

He X, Zhang J (2005). Rapid Subfunctionalization Accompanied by Prolonged and Substantial Neofunctionalization in Duplicate Gene Evolution. *Genetics* **169**: 1157–1164.

Hegarty MJ, Hiscock SJ (2008). Genomic Clues to the Evolutionary Success of Polyploid Plants. *Curr Biol* **18**: R435–R444.

Heimeier RA, Das B, Buchholz DR, Fiorentino M, Shi Y-B (2010). Studies on Xenopus laevis intestine reveal biological pathways underlying vertebrate gut adaptation from embryo to adult. *Genome Biol* **11**: R55.

Herranz R, Larkin OJ, Hill RJ, Lopez-Vidriero I, van Loon JJ, Medina F (2013). Suboptimal evolutionary novel environments promote singular altered gravity responses of transcriptome during Drosophila metamorphosis. *BMC Evol Biol* **13**: 133.

Hetem RS, de Witt BA, Fick LG, Fuller A, Kerley GIH, Meyer LCR, *et al.* (2009). Body temperature, thermoregulatory behaviour and pelt characteristics of three colour morphs of springbok (Antidorcas marsupialis). *Comp Biochem Physiol Part A Mol Integr Physiol* **152**: 379–388.

Higdon CW, Mitra RD, Johnson SL (2013). Gene Expression Analysis of Zebrafish Melanocytes, Iridophores, and Retinal Pigmented Epithelium Reveals Indicators of Biological Function and Developmental Origin (YF Leung, Ed.). *PLoS One* **8**: e67801.

Higuera C, Gardiner KJ, Cios KJ (2015). Self-Organizing Feature Maps Identify Proteins Critical to Learning in a Mouse Model of Down Syndrome (Y Herault, Ed.). *PLoS One* **10**: e0129126.

Hill G (2006). Female choice for ornamental coloration. In: Hill GE, McGraw KJ (eds) *Bird Coloration: Function and evolution2*, Harvard University Press, pp 137–200.

Hill G, McGraw K (2006). *Bird Coloration: Mechanisms and measurements*. Harvard University Press.

Hinits Y, Williams VC, Sweetman D, Donn TM, Ma TP, Moens CB, *et al.* (2011). Defective cranial skeletal development, larval lethality and haploinsufficiency in Myod mutant zebrafish. *Dev Biol* **358**: 102–112.

Hittinger CT, Carroll SB (2007). Gene duplication and the adaptive evolution of a classic genetic switch. *Nature* **449**: 677–681.

Ho ALFC, O'Shea SK, Pomeroy HF (2013). Dietary esterified astaxanthin effects on color, carotenoid concentrations, and compositions of clown anemonefish, Amphiprion ocellaris, skin. *Aquac Int* **21**: 361–374.

Hoegh-Guldberg O, Mumby PJ, Hooten A, Steneck RS, Greenfield P, Gomez E, *et al.* (2007). Coral reefs under rapid climate change and ocean acidification. *Science (80-)* **318**: 1737–1742.

van Hoek MJA, Hogeweg P (2009). Metabolic Adaptation after Whole Genome Duplication. *Mol Biol Evol* **26**: 2441–2453.

Hoekstra HE (2006). Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity (Edinb)* **97**: 222–234.

Hof AE van't, Campagne P, Rigden DJ, Yung CJ, Lingley J, Quail MA, *et al.* (2016). The industrial melanism mutation in British peppered moths is a transposable element. *Nature* **534**: 102–105.

Hoffman EA, Schueler FW, Jones AG, Blouin MS (2006). An analysis of selection on a colour polymorphism in the northern leopard frog. *Mol Ecol* **15**: 2627–2641.

Holzer G (2015). *Origine de la signalisation thyroïdienne chez les métazoaires et implication dans leur métamorphose.*

Holzer G, Besson M, Lambert A, François L, Barth P, Gillet B, *et al.* (2017). Fish larval recruitment to reefs is a thyroid hormone-mediated metamorphosis sensitive to the pesticide chlorpyrifos. *Elife* **6**.

Holzer G, Laudet V (2015). Thyroid Hormones: A Triple-Edged Sword for Life History Transitions. *Curr Biol* **25**: R344–R347.

Hopwood N (2007). A history of normal plates, tables and stages in vertebrate embryology. *Int J Dev Biol* **51**: 1–26.

Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, *et al.* (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**: 498–503.

Huang H, Cai L, Remo BF, Brown DD (2001). Timing of metamorphosis and the onset of the negative feedback loop between the thyroid gland and the pituitary is controlled by type II iodothyronine deiodinase in Xenopus laevis. *Proc Natl Acad Sci* **98**: 7348–7353.

Huang R, Hippauf F, Rohrbeck D, Haustein M, Wenke K, Feike J, *et al.* (2012). Enzyme functional evolution through improved catalysis of ancestrally nonpreferred substrates. *Proc Natl Acad Sci* **109**: 2966–2971.

Hubbard JK, Uy JAC, Hauber ME, Hoekstra HE, Safran RJ (2010). Vertebrate pigmentation: from underlying genes to adaptive function. *Trends Genet* **26**: 231–239.

Hughes AL (1994). The Evolution of Functionally Novel Proteins after Gene Duplication. *Proc R Soc B Biol Sci* **256**: 119–124.

Hughes TP, Baird AH, Bellwood DR, Card M, Connolly SR, Folke C, *et al.* (2003). Climate change, human impacts, and the resilience of coral reefs. *Science* **301**: 929–33.

Hughes TP, Kerry J, Álvarez-Noriega M, Álvarez-Romero J, Anderson K, Baird A, *et al.* (2017). Global warming and recurrent mass bleaching of corals. *Nature* **543**: 373–386.

Hurley IA, Mueller RL, Dunn KA, Schmidt EJ, Friedman M, Ho RK, *et al.* (2007).

A new time-scale for ray-finned fish evolution. *Proc R Soc B Biol Sci* **274**: 489–498.

Hysi PG, Valdes AM, Liu F, Furlotte NA, Evans DM, Bataille V, *et al.* (2018). Genome-wide association meta-analysis of individuals of European ancestry identifies new loci explaining a substantial fraction of hair color variation and heritability. *Nat Genet*.

Ihmels J, Bergmann S, Gerami-Nejad M, Yanai I, McClellan M, Berman J, *et al.* (2005). Rewiring of the Yeast Transcriptional Network Through the Evolution of Motif Usage. *Science (80-)* **309**: 938–940.

Innan H, Kondrashov F (2010). The evolution of gene duplications: classifying and distinguishing between models. *Nat Rev Genet* **11**: 97–108.

Inoue J, Sato Y, Sinclair R, Tsukamoto K, Nishida M (2015). Rapid genome reshaping by multiple-gene loss after whole-genome duplication in teleost fish suggested by mathematical modeling. *Proc Natl Acad Sci* **112**: 14918–14923.

Irion U, Singh AP, Nüsslein-Volhard C (2016). The Developmental Genetics of Vertebrate Color Pattern Formation. In: Wassarman PM (ed) *Essays on Developmental Biology, Part B*, Academic Press, pp 141–169.

Irisson J-O, Lecchini D (2008). In situ observation of settlement behaviour in larvae of coral reef fishes at night. *J Fish Biol* **72**: 2707–2713.

Ito S (2003). A Chemist's View of Melanogenesis. *Pigment Cell Res* **16**: 230–236.

Iwata E, Mikami K, Manbo J, Moriya-Ito K, Sasaki H (2012). Social Interaction Influences Blood Cortisol Values and Brain Aromatase Genes in the Protandrous False Clown Anemonefish, Amphiprion ocellaris. *Zoolog Sci* **29**: 849–855.

Iwata E, Nagai Y, Hyoudou M, Sasaki H (2008). Social Environment and Sex Differentiation in the False Clown Anemonefish, Amphiprion ocellaris. *Zoolog Sci* **25**: 123–128.

Iwata E, Nagai Y, Sasaki H (2010). Social rank modulates brain arginine vasotocin immunoreactivity in false clown anemonefish (Amphiprion ocellaris). *Fish Physiol Biochem* **36**: 337–345.

Jackson IJ (1988). A cDNA encoding tyrosinase-related protein maps to the brown locus in mouse. *Proc Natl Acad Sci U S A* **85**: 4392–6.

Jackson GA, Strathmann RR (1981). Larval Mortality from Offshore Mixing as a Link between Precompetent and Competent Periods of Development. *Am Nat* **118**: 16–26.

Jacob F (1977). Evolution and tinkering. *Science (80-)* **196**: 1161–1166.

Jacob H, Pouil S, Lecchini D, Oberhänsli F, Swarzenski P, Metian M (2017). Trophic transfer of essential elements in the clownfish Amphiprion ocellaris in the context of ocean acidification (P Bustamante, Ed.). *PLoS One* **12**: e0174344.

Jacquier A (2009). The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. *Nat Rev Genet* **10**: 833–844.

Jaillon O, Aury J-M, Brunet F, Petit J-L, Stange-Thomann N, Mauceli E, *et al.* (2004). Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature* **431**: 946–957.

Janvier P (2011). Comparative Anatomy: All Vertebrates Do Have Vertebrae. *Curr Biol* **21**: R661–R663.

Jiang Y, Zhang S, Xu J, Feng J, Mahboob S, Al-Ghanim KA, *et al.* (2014). Comparative Transcriptome Analysis Reveals the Genetic Basis of Skin Color Variation in Common Carp (AT Slominski, Ed.). *PLoS One* **9**: e108200.

Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, *et al.* (2011). Ancestral polyploidy in seed plants and angiosperms. *Nature* **473**: 97–100.

Johnsen S (2012). *The Optics of Life A Biologist's Guide to Light in Nature*. Princeton University Press.

Joly S, Heenan PB, Lockhart PJ (2009). A Pleistocene inter-tribal allopolyploidization event precedes the species radiation of Pachycladon (Brassicaceae) in New Zealand. *Mol Phylogenet Evol* **51**: 365–372.

Jones GP, McCormick MI, Srinivasan M, Eagle J V (2004). Coral decline threatens fish biodiversity in marine reserves. *Proc Natl Acad Sci U S A* **101**: 8251–3.

Jones GP, Planes S, Thorrold SR (2005). Coral Reef Fish Larvae Settle Close to Home. *Curr Biol* **15**: 1314–1318.

Justin Marshall N (2000). Communication and camouflage with the same 'bright' colours in reef fishes. *Philos Trans R Soc B Biol Sci* **355**: 1243–1248.

Kaessmann H (2010). Origins, evolution, and phenotypic impact of new genes. *Genome Res* **20**: 1313–1326.

Kaessmann H, Vinckenbosch N, Long M (2009). RNA-based gene duplication: mechanistic and evolutionary insights. *Nat Rev Genet* **10**: 19–31.

Kalinka AT, Varga KM, Gerrard DT, Preibisch S, Corcoran DL, Jarrells J, *et al.* (2010). Gene expression divergence recapitulates the developmental hourglass model. *Nature* **468**: 811–814.

Kawakami Y, Nozaki J, Seoka M, Kumai H, Ohta H (2008). Characterization of thyroid hormones and thyroid hormone receptors during the early development of Pacific bluefin tuna (*Thunnus orientalis*). *Gen Comp Endocrinol* **155**: 597–606.

Kelsh RN (2004). Genetics and Evolution of Pigment Patterns in Fish. *Pigment Cell Res* **17**: 326–336.

Kelsh RN, Harris ML, Colanesi S, Erickson CA (2009). Stripes and belly-spots—A review of pigment cell morphogenesis in vertebrates. *Semin Cell Dev Biol* **20**: 90–104.

Kenny NJ, Chan KW, Nong W, Qu Z, Maeso I, Yip HY, *et al.* (2016). Ancestral whole-genome duplication in the marine chelicerate horseshoe crabs. *Heredity (Edinb)* **116**: 190–199.

Kiessling W, Simpson C, Foote M (2010). Reefs as Cradles of Evolution and Sources of Biodiversity in the Phanerozoic. *Science (80-)* **327**: 196–198.

Kimmel CB (1989). Genetics and early development of zebrafish. *Trends Genet* **5**: 283–288.

Kimura T, Nagao Y, Hashimoto H, Yamamoto-Shiraishi Y -i., Yamamoto S, Yabe T, *et al.* (2014). Leucophores are similar to xanthophores in their specification and differentiation processes in medaka. *Proc Natl Acad Sci* **111**: 7343–7348.

Kirkpatrick M (1982). Sexual selection and the evolution of female choice. *Evolution (N Y)* **36**: 1–12.

Kitano J, Lema SC, Luckenbach JA, Mori S, Kawagishi Y, Kusakabe M, *et al.* (2010). Adaptive Divergence in the Thyroid Hormone Signaling Pathway in the Stickleback Radiation. *Curr Biol* **20**: 2124–2130.

Kleppe L, Edvardsen RB, Furmanek T, Taranger GL, Wargelius A (2014). Global transcriptome analysis identifies regulated transcripts and pathways activated during oogenesis and early embryogenesis in atlantic cod. *Mol Reprod Dev* **81**: 619–635.

Knecht AK, Bronner-Fraser M (2002). Induction of the neural crest: a multigene process. *Nat Rev Genet* **3**: 453–461.

Knowles DG, McLysaght A (2009). Recent de novo origin of human protein-coding genes. *Genome Res* **19**: 1752–1759.

Kobel HR, Du Pasquier L (1986). Genetics of polyploid Xenopus. *Trends Genet* **2**: 310–315.

Kocher TD (2004). Adaptive evolution and explosive speciation: the cichlid fish model. *Nat Rev Genet* **5**: 288–298.

Kodric-Brown A (1989). Dietary carotenoids and male mating success in the guppy: an environmental component to female choice. *Behav Ecol Sociobiol* **25**: 393–401.

Kodric-Brown A (1998). Sexual Dichromatism and Temporary Color Changes in the Reproduction of Fishes. *Am Zool* **38**: 70–81.

Kollias N (1994). The spectroscopy of human melanin pigmentation. In: Zeise L, Chedeke MR, Fitzpatrick TB (eds) *Melanin: Its Role in Human Photoprotection*, Valdenmar Publishing, pp 31–38.

Kondo S, Asai R (1995). A reaction–diffusion wave on the skin of the marine angelfish Pomacanthus. *Nature* **376**: 765–768.

Kondo S, Miura T (2010). Reaction-Diffusion Model as a Framework for Understanding Biological Pattern Formation. *Science (80-)* **329**: 1616–1620.

Kornberg TB (2017). Macrophages Help Cells Connect to Pattern Zebrafish Stripes. *Dev Cell* **40**: 520–521.

Kottler VA, Fadeev A, Weigel D, Dreyer C (2013). Pigment Pattern Formation in the Guppy, Poecilia reticulata, Involves the Kita and Csf1ra Receptor Tyrosine

Kinases. *Genetics* **194**: 631–646.

Kottler VA, Künstner A, Schartl M (2015). Pheomelanin in fish? *Pigment Cell Melanoma Res* **28**: 355–356.

Kuriyama T, Miyaji K, Sugimoto M, Hasegawa M (2006). Ultrastructure of the Dermal Chromatophores in a Lizard (Scincidae: Plestiodon latiscutatus) with Conspicuous Body and Tail Coloration. *Zoolog Sci* **23**: 793–799.

DE L. Brooke M (1998). Ecological factors influencing the occurrence of 'flash marks' in wading birds. *Funct Ecol* **12**: 339–346.

Lamason RL, Mohideen M-APK, Mest JR, Wong AC, Norton HL, Aros MC, *et al.* (2005). SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science* **310**: 1782–6.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature* **409**: 860–921.

Larhammar D, Risinger C (1994). Molecular Genetic Aspects of Tetraploidy in the Common Carp Cyprinus carpio. *Mol Phylogenet Evol* **3**: 59–68.

Laudet V (2011). The Origins and Evolution of Vertebrate Metamorphosis. *Curr Biol* **21**: R726–R737.

Laurent S, Salamin N, Robinson-Rechavi M (2017). No evidence for the radiation time lag model after whole genome duplications in Teleostei (A von Haeseler, Ed.). *PLoS One* **12**: e0176384.

Leclercq E, Taylor JF, Migaud H (2009). Morphological skin colour changes in teleosts. *Fish Fish* **11**: 159–193.

Lee AP, Kerk SY, Tan YY, Brenner S, Venkatesh B (2011). Ancient Vertebrate Conserved Noncoding Elements Have Been Evolving Rapidly in Teleost Fishes. *Mol Biol Evol* **28**: 1205–1215.

Leis JM, McCormick MI (2002). The Biology, Behavior, and Ecology of the Pelagic Larval Stage of Coral Reef Fishes. In: *Coral Reef Fishes*, Elsevier, pp 171–199.

Leis JM, Siebeck U, Dixson DL (2011). How Nemo Finds Home: The Neuroecology of Dispersal and of Population Connectivity in Larvae of Marine Fishes. *Integr Comp Biol* **51**: 826–843.

Leitch IJ, Leitch AR (2013). Genome Size Diversity and Evolution in Land Plants. In: *Plant Genome Diversity Volume 2*, Springer Vienna: Vienna, pp 307–322.

Leonelli S, Ankeny RA (2013). What makes a model organism? *Endeavour* **37**: 209–212.

Levin M, Anavy L, Cole AG, Winter E, Mostov N, Khair S, *et al.* (2016). The mid-developmental transition and the evolution of animal body plans. *Nature* **531**: 637–641.

Levin M, Hashimshony T, Wagner F, Yanai I (2012). Developmental Milestones Punctuate Gene Expression in the Caenorhabditis Embryo. *Dev Cell* **22**: 1101–1108.

Li Z, Baniaga AE, Sessa EB, Scascitelli M, Graham SW, Rieseberg LH, *et al.* (2015). Early genome duplications in conifers and other seed plants. *Sci Adv* **1**: e1501084–e1501084.

Li J-T, Hou G-Y, Kong X-F, Li C-Y, Zeng J-M, Li H-D, *et al.* (2015). The fate of recent duplicated genes following a fourth-round whole genome duplication in a tetraploid fish, common carp (Cyprinus carpio). *Sci Rep* **5**: 8199.

Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, *et al.* (2010). The developmental dynamics of the maize leaf transcriptome. *Nat Genet* **42**: 1060–1067.

Li Z, Tiley GP, Galuska SR, Reardon CR, Kidder TI, Rundell RJ, *et al.* (2018). Multiple large-scale gene and genome duplications during the evolution of hexapods. *Proc Natl Acad Sci*: 201710791.

Lien S, Gidskehaug L, Moen T, Hayes BJ, Berg PR, Davidson WS, *et al.* (2011). A dense SNP-based linkage map for Atlantic salmon (Salmo salar) reveals extended chromosome homeologies and striking differences in sex-specific recombination patterns. *BMC Genomics* **12**: 615.

Lien S, Koop BF, Sandve SR, Miller JR, Kent MP, Nome T, *et al.* (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature* **533**: 200–205.

Lieske E, Myers R (2002). *Coral Reef Fishes - Caribbean, Indian Ocean and Pacific Ocean, including Red Sea* (E Lieske and R Myers, Eds.).

Liew HJ, Ambak M, Abol-Munafi A, Chuah T (2006). Embryonic development of clownfish Amphiprion ocellaris under laboratory condition. *J Sustain Sci Manag* **1**: 64–73.

Lin SJ, Foley J, Jiang TX, Yeh CY, Wu P, Foley A, *et al.* (2013). Topology of Feather Melanocyte Progenitor Niche Allows Complex Pigment Patterns to Emerge. *Science (80-)* **340**: 1442–1445.

Lister JA, Close J, Raible DW (2001). Duplicate mitf Genes in Zebrafish: Complementary Expression and Conservation of Melanogenic Potential. *Dev Biol* **237**: 333–344.

Litsios G, Kostikova A, Salamin N (2014). Host specialist clownfishes are environmental niche generalists. *Proc R Soc B Biol Sci* **281**: 20133220–20133220.

Litsios G, Pearman P, Lanterbecq D, Tolou N, Salamin N (2014). The radiation of the clownfishes has two geographical replicates. *J Biogeogr* **41**: 2140–2149.

Litsios G, Salamin N (2014). Hybridisation and diversification in the adaptive radiation of clownfishes. *BMC Evol Biol* **14**: 245.

Litsios G, Sims CA, Wüest RO, Pearman PB, Zimmermann NE, Salamin N (2012). Mutualism with sea anemones triggered the adaptive radiation of clownfishes. *BMC Evol Biol* **12**: 212.

Lokki J, Saura A (1979). Polyploidy in insect evolution. *Basic Life Sci* **13**: 277–312.

Lopes RJ, Johnson JD, Toomey MB, Ferreira MS, Araujo PM, Melo-Ferreira J, *et al.* (2016). Genetic Basis for Red Coloration in Birds. *Curr Biol* **26**: 1427–1434.

Lopes SS, Yang X, Müller J, Carney TJ, McAdow AR, Rauch G-J, *et al.* (2008). Leukocyte Tyrosine Kinase Functions in Pigment Cell Development (G Barsh, Ed.). *PLoS Genet* **4**: e1000026.

Lorgen M, Casadei E, Król E, Douglas A, Birnie MJ, Ebbesson LOE, *et al.* (2015). Functional Divergence of Type 2 Deiodinase Paralogs in the Atlantic Salmon. *Curr Biol* **25**: 936–941.

Lorin T, Brunet FG, Laudet V, Volff J-N (2018). Teleost Fish-Specific Preferential Retention of Pigmentation Gene-Containing Families After Whole Genome Duplications in Vertebrates. *G3 Genes|Genomes|Genetics*: g3.200201.2018.

Losey GS, Cronin TW, Goldsmith TH, Hyde D, Marshall NJ, McFarland WN (1999). The UV visual world of fishes: a review. *J Fish Biol* **54**: 921–943.

Losey GS, McFarland WN, Loew ER, Zamzow JP, Nelson PA, Marshall NJ (2003). Visual Biology of Hawaiian Coral Reef Fishes. I. Ocular Transmission and Visual Pigments (WL Montgomery, Ed.). *Copeia* **2003**: 433–454.

Losos JB (2011). Convergence, adaptation and constraint. *Evolution (N Y)* **65**: 1827–1840.

Lumb R, Buckberry S, Secker G, Lawrence D, Schwarz Q (2017). Transcriptome profiling reveals expression signatures of cranial neural crest cells arising from different axial levels. *BMC Dev Biol* **17**: 5.

Lunghi E, Monti A, Binda A, Piazzi I, Salvadori M, Cogoni R, *et al.* (2017). Cases of albinism and leucism in amphibians in Italy: new reports. *Nat Hist Sci* **4**: 73.

Lynch M, Conery JS (2000). The evolutionary fate and consequences of duplicate genes. *Science* **290**: 1151–5.

Lynch M, Force AG (2000). The Origin of Interspecific Genomic Incompatibility via Gene Duplication. *Am Nat* **156**: 590–605.

MacArthur R, Levins R (1964). Competition, habitat selection and character displacement in a patchy environment. *Proc Natl Acad Sci U S A* **6**: 1207–1210.

Macaulay IC, Svensson V, Labalette C, Ferreira L, Hamey F, Voet T, *et al.* (2016). Single-Cell RNA-Sequencing Reveals a Continuous Spectrum of Differentiation in Hematopoietic Cells. *Cell Rep* **14**: 966–977.

Machuca I, Esslemont G, Fairclough L, Tata JR (1995). Analysis of structure and expression of the Xenopus thyroid hormone receptor-beta gene to explain its autoinduction. *Mol Endocrinol* **9**: 96–107.

Macqueen DJ, Johnston IA (2014). A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proc R Soc B Biol Sci* **281**: 20132881–20132881.

Madduppa H, Timm J, Kochzius M (2018). Reduced genetic diversity in clown anemonefish (Amphiprion ocellaris) in exploited reefs of the Spermonde Archipelago, Indonesia. *Front Mar Sci*.

Madeira C, Madeira D, Diniz MS, Cabral HN, Vinagre C (2017). Comparing biomarker responses during thermal acclimation: A lethal vs non-lethal approach in a tropical reef clownfish. *Comp Biochem Physiol Part A Mol Integr Physiol* **204**: 104–112.

Maderspacher F (2012). Colour Patterns: Channelling Turing. *Curr Biol* **22**: R266–R268.

Maere S, De Bodt S, Raes J, Casneuf T, Van Montagu M, Kuiper M, *et al.* (2005). Modeling gene and genome duplications in eukaryotes. *Proc Natl Acad Sci* **102**: 5454–5459.

Mahalwar P, Walderich B, Singh AP, Nusslein-Volhard C (2014). Local reorganization of xanthophores fine-tunes and colors the striped pattern of zebrafish. *Science (80-)* **345**: 1362–1364.

Maini PK, Baker RE, Chuong C-M (2006). Developmental Biology: The Turing Model Comes of Molecular Age. *Science (80-)* **314**: 1397–1398.

Makino T, Hokamp K, McLysaght A (2009). The complex relationship of gene duplication and essentiality. *Trends Genet* **25**: 152–155.

Makino T, McLysaght A (2010). Ohnologs in the human genome are dosage balanced and frequently associated with disease. *Proc Natl Acad Sci* **107**: 9270–9274.

Makino T, McLysaght A, Kawata M (2013). Genome-wide deserts for copy number variation in vertebrates. *Nat Commun* **4**.

Mallarino R, Henegar C, Mirasierra M, Manceau M, Schradin C, Vallejo M, *et al.* (2016). Developmental mechanisms of stripe patterns in rodents. *Nature* **539**: 518–523.

Mallarino R, Linden TA, Linnen CR, Hoekstra HE (2017). The role of isoforms in the evolution of cryptic coloration in Peromyscus mice. *Mol Ecol* **26**: 245–258.

Manceau M, Domingues VS, Linnen CR, Rosenblum EB, Hoekstra HE (2010). Convergence in pigmentation at multiple levels: mutations, genes and function. *Philos Trans R Soc B Biol Sci* **365**: 2439–2450.

Manceau M, Domingues VS, Mallarino R, Hoekstra HE (2011). The Developmental Role of Agouti in Color Pattern Evolution. *Science (80-)* **331**: 1062–1065.

Des Marais DL, Rausher MD (2008). Escape from adaptive conflict after duplication in an anthocyanin pathway gene. *Nature* **454**: 762–765.

Marburger S, Alexandrou MA, Taggart JB, Creer S, Carvalho G, Oliveira C, *et al.* (2018). Whole genome duplication and transposable element proliferation drive genome expansion in Corydoradinae catfishes. *Proc R Soc B Biol Sci* **285**: 20172732.

Marcet-Houben M, Gabaldón T (2015). Beyond the Whole-Genome Duplication: Phylogenetic Evidence for an Ancient Interspecies Hybridization in the Baker's Yeast Lineage (LD Hurst, Ed.). *PLOS Biol* **13**: e1002220.

Marcionetti A, Rossier V, Bertrand JAM, Litsios G, Salamin N (2018). First draft genome of an iconic clownfish species (Amphiprion frenatus). *Mol Ecol Resour*.

Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y (2008). RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res* **18**: 1509–1517.

Mariscal RN (1966). The symbiosis between tropical sea anemones and fishes: a review. In: *The Galápagos: Proceedings of the Symposia of the Galápagos International Scientific Project*, University of California Press, p 318.

del Marmol V, Beermann F (1996). Tyrosinase and related proteins in mammalian pigmentation. *FEBS Lett* **381**: 165–168.

Marshall NJ (2000). The Visual Ecology of Reef Fish Colours. In: Espmark Y, Amundsen T, Rosenqvist G (eds) *Animal Signals: Signalling and Signal Design in Animal Communication*, Tapir, Norway, p 83-.

Marshall NJ, Jennings K, McFarland WN, Loew ER, Losey GS (2003). Visual Biology of Hawaiian Coral Reef Fishes. II. Colors of Hawaiian Coral Reef Fish (WL Montgomery, Ed.). *Copeia* **2003**: 455–466.

Martin A, Courtier-Orgogozo V (2017). Morphological Evolution Repeatedly Caused by Mutations in Signaling Ligand Genes. In: *Diversity and Evolution of Butterfly Wing Patterns*, Springer Singapore: Singapore, pp 59–87.

Martin A, Orgogozo V (2013). The loci of repeated evolution: A catalog of genetic hotspots of phenotypic variation. *Evolution (N Y)* **67**: 1235–1250.

Masly JP, Jones CD, Noor MAF, Locke J, Orr HA (2006). Gene Transposition as a Cause of Hybrid Sterility in Drosophila. *Science (80-)* **313**: 1448–1450.

Mathavan S, Lee SGP, Mak A, Miller LD, Murthy KRK, Govindarajan KR, *et al.* (2005). Transcriptome Analysis of Zebrafish Embryogenesis Using Microarrays. *PLoS Genet* **1**: e29.

Mäthger LM, Land MF, Siebeck UE, Marshall NJ (2003). Rapid colour changes in multilayer reflecting stripes in the paradise whiptail, Pentapodus paradiseus. *J Exp Biol* **206**: 3607–13.

Mazurais D, Darias M, Zambonino-Infante JL, Cahu CL (2011). Transcriptomics for understanding marine fish larval development 1 This review is part of a virtual symposium on current topics in aquaculture of marine fish and shellfish. *Can J Zool* **89**: 599–611.

McClintock B (1950). The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci* **36**: 344–355.

McCormick M (1999). Delayed metamorphosis of a tropical reef fish (Acanthurus triostegus):a field experiment. *Mar Ecol Prog Ser* **176**: 25–38.

McCormick MI, Makey LJ, Dufour V (2002a). Comparative study of metamorphosis in tropical reef fishes. *Mar Biol* **141**: 841–853.

McCormick MI, Makey L, Dufour V (2002b). Comparative study of metamorphosis in tropical reef fishes. *Mar Biol* **141**: 841–853.

McGrath CL, Gout J-F, Doak TG, Yanagi A, Lynch M (2014). Insights into Three Whole-Genome Duplications Gleaned from the Paramecium caudatum Genome Sequence. *Genetics* **197**: 1417–1428.

McGrath CL, Gout J-F, Johri P, Doak TG, Lynch M (2014). Differential retention and divergent resolution of duplicate genes following whole-genome duplication. *Genome Res* **24**: 1665–1675.

McGraw KJ (2006). Mechanics of Uncommon colors: Pterins, Porphyrins and Psittacofulvins. In: Hill GE, McGraw KJ (eds) *Bird Coloration: Function and evolution*, Harvard University Press, pp 354–398.

McGraw KJ, Nogare MC (2004). Carotenoid pigments and the selectivity of psittacofulvin-based coloration systems in parrots. *Comp Biochem Physiol - B Biochem Mol Biol* **138**: 229–233.

McGraw KJ, Toomey MB, Nolan PM, Morehouse NI, Massaro M, Jouventin P (2007). A description of unique fluorescent yellow pigments in penguin feathers. *Pigment Cell Res* **20**: 301–304.

McLysaght A, Hokamp K, Wolfe KH (2002). Extensive genomic duplication during early chordate evolution. *Nat Genet* **31**: 200–204.

McMenamin SK, Parichy DM (2013). Metamorphosis in Teleosts. *Curr Top Dev Biol* **103**: 127-165.

McMenamin SK, Bain EJ, McCann AE, Patterson LB, Eom DS, Waller ZP, *et al.* (2014). Thyroid hormone-dependent adult pigment cell lineage and pattern in zebrafish. *Science (80-)* **345(602)**: 1358–1361.

Meadows MG, Anthes N, Dangelmayer S, Alwany MA, Gerlach T, Schulte G, *et al.* (2014). Red fluorescence increases with depth in reef fishes, supporting a visual function, not UV protection. *Proc R Soc B Biol Sci* **281**: 20141211–20141211.

Mebs D (2009). Chemical biology of the mutualistic relationships of sea anemones with fish and crustaceans. *Toxicon* **54**: 1071–1074.

Mehta TK, Ravi V, Yamasaki S, Lee AP, Lian MM, Tay B-H, *et al.* (2013). Evidence for at least six Hox clusters in the Japanese lamprey (Lethenteron japonicum). *Proc Natl Acad Sci* **110**: 16044–16049.

Meinhardt H, Gierer A (2000). Pattern formation by local self-activation and lateral inhibition. *BioEssays* **22**: 753–760.

Mellgren EM, Johnson SL (2002). The evolution of morphological complexity in zebrafish stripes. *Trends Genet* **18**: 128–34.

Menter DG, Obika M, Tchen TT, Taylor JD (1979). Leucophores and iridophores ofFundulus heteroclitus: Biophysical and ultrastructural properties. *J Morphol* **160**: 103–119.

Meyer A, Van de Peer Y (2005). From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *BioEssays* **27**: 937–945.

Meyer A, Schartl M (1999). Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol* **11**: 699–704.

Michiels NK, Anthes N, Hart NS, Herler J, Meixner AJ, Schleifenbaum F, *et al.* (2008). Red fluorescence in reef fish: A novel signalling mechanism? *BMC Ecol* **8**: 16.

Milam EL (2010). *Looking for a Few Good Males*. John Hopkins University Press.

Milam EL, Millstein RL, Potochnik A, Roughgarden JE (2011). Sex and sensibility: The role of social selection. *Metascience* **20**: 253–277.

Militz TA, Foale S (2017). The 'Nemo Effect': Perception and reality of Finding Nemo 's impact on marine aquarium fisheries. *Fish Fish* **18**: 596–606.

Mills LS, Bragina E V., Kumar A V., Zimova M, Lafferty DJR, Feltner J, *et al.* (2018). Winter color polymorphisms identify global hot spots for evolutionary rescue from climate change. *Science (80-)* **359**: 1033–1036.

Mills MG, Patterson LB (2009). Not just black and white: Pigment pattern development and evolution in vertebrates. *Semin Cell Dev Biol* **20**: 72–81.

Mitchell JS (2003). Mobility of Stichodactyla Gigantea Sea Anemones and Implications for Resident False Clown Anemonefish, Amphiprion Ocellaris. *Environ Biol Fishes* **66**: 85–90.

Mitchell J (2005). Queue selection and switching by false clown anemonefish, Amphiprion ocellaris. *Anim Behav* **69**: 643–652.

Miyagawa-Kohshima K, Miyahara H, Uchida S, Odoriba S, Okabe D, Baba Y, *et al.* (2014). Embryonic learning of chemical cues via the parents' host in anemonefish (Amphiprion ocellaris). *J Exp Mar Bio Ecol* **457**: 160–172.

Moberg FF, Folke C (1999). Ecological goods and services of coral reef ecosystems. *Ecol Econ* **29**: 215–233.

Moran NA, Jarvik T (2010). Lateral Transfer of Genes from Fungi Underlies Carotenoid Production in Aphids. *Science (80-)* **328**: 624–627.

Morgan TH (1910). Sex limited inheritance in Drosophila. *Science (80-)* **32**: 120–122.

Morgan TH (1911). Notes on two different crosses between races of pigeons. *Biol Bull* **21**: 215–221.

Morrison RL, Rand MS, Frost-Mason SK (1995). Cellular Basis of Color Differences in Three Morphs of the Lizard Sceloporus undulatus erythrocheilus. *Copeia* **1995**: 397.

Morrison RL, Sherbrooke WC, Frost-Mason SK (1996). Temperature-Sensitive, Physiologically Active Iridophores in the Lizard Urosaurus ornatus: An Ultrastructural Analysis of Color Change. *Copeia* **1996**: 804.

Mort RL, Jackson IJ, Patton EE (2015). The melanocyte lineage in development and disease. *Development* **142**: 620–632.

Mortazavi A, Schwarz EM, Williams B, Schaeffer L, Antoshechkin I, Wold BJ, *et al.* (2010). Scaffolding a Caenorhabditis nematode genome with RNA-seq. *Genome Res* **20**: 1740–1747.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**: 621–628.

Mountjoy KG, Robbins LS, Mortrud MT, Cone RD (1992). The cloning of a family of genes that encode the melanocortin receptors. *Science* **257**: 1248–51.

Moyer J, Nakazano A (1978). Protandrous hermaphroditism in six species of the anemonefish genus Amphiprion in Japan. *Japanese J Ichthyol* **25**: 101–106.

Müller F (1879). Ituna and Thyridia: a remarkable case of mimicry in butterflies. *Trans Entomol Soc London*: 20–29.

Mumby PJ (2006). The impact of exploiting grazers (Scaridae) on the dynamics of Caribbean coral reefs. *Ecol Appl* **16**: 747–69.

Mundy NI, Stapley J, Bennison C, Tucker R, Twyman H, Kim K-W, *et al.* (2016). Red Carotenoid Coloration in the Zebra Finch Is Controlled by a Cytochrome P450 Gene Cluster. *Curr Biol* **26**: 1435–1440.

Muñoz MM, Crawford NG, McGreevy TJ, Messana NJ, Tarvin RD, Revell LJ, *et al.* (2013). Divergence in coloration and ecological speciation in the Anolis marmoratus species complex. *Mol Ecol* **22**: 2668–2682.

Munro JL, Gaut VC, Thompson R, Reeson PH (1973). The spawning seasons of Caribbean reef fishes. *J Fish Biol* **5**: 69–84.

Murata M, Miyagawa-Kohshima K, Nakanishi K, Naya Y (1986). Characterization of Compounds That Induce Symbiosis Between Sea Anemone and Anemone Fish. *Science (80-)* **234**: 585–587.

Myrberg AA, Fuiman LA (2002). The Sensory World of Coral Reef Fishes. In: *Coral Reef Fishes*, Elsevier, pp 123–148.

Nadeau NJ, Pardo-Diaz C, Whibley A, Supple MA, Saenko S V., Wallbank RWR, *et al.* (2016). The gene cortex controls mimicry and crypsis in butterflies and moths. *Nature* **534**: 106–110.

Nagao Y, Suzuki T, Shimizu A, Kimura T, Seki R, Adachi T, *et al.* (2014). Sox5 Functions as a Fate Switch in Medaka Pigment Cell Development. *PLoS Genet* **10**.

Nagao Y, Takada H, Miyadai M, Adachi T, Seki R, Kamei Y, *et al.* (2018). Distinct interactions of Sox5 and Sox10 in fate specification of pigment cells in medaka and zebrafish. *PLoS Genet* **14**: e1007260

Nakamasu A, Takahashi G, Kanbe A, Kondo S (2009). Interactions between zebrafish pigment cells responsible for the generation of Turing patterns. *Proc Natl Acad Sci* **106**: 8429–8434.

Nakashima K, Yamada L, Satou Y, Azuma J, Satoh N (2004). The evolutionary origin of animal cellulose synthase. *Dev Genes Evol* **214**: 81–88.

Nakatani Y, Takeda H, Kohara Y, Morishita S (2011). Reconstruction of the Vertebrate Ancestral Genome Reveals Dynamic Genome Reorganization in Early Vertebrates. In: *Medaka*, Springer Japan: Tokyo, pp 307–322.

Naville M, Chalopin D, Volff J-N (2014). Interspecies Insertion Polymorphism Analysis Reveals Recent Activity of Transposable Elements in Extant Coelacanths (H Escriva, Ed.). *PLoS One* **9**: e114382.

Naville M, Volff J-N (2016). Endogenous Retroviruses in Fish Genomes: From Relics of Past Infections to Evolutionary Innovations? *Front Microbiol* **7**.

Near TJ, Dornburg A, Eytan RI, Keck BP, Smith WL, Kuhn KL, *et al.* (2013). Phylogeny and tempo of diversification in the superradiation of spiny-rayed fishes. *Proc Natl Acad Sci* **110**: 12738–12743.

Near TJ, Eytan RI, Dornburg A, Kuhn KL, Moore JA, Davis MP, *et al.* (2012). Resolution of ray-finned fish phylogeny and timing of diversification. *Proc Natl Acad Sci* **109**: 13698–13703.

Negri AP, Uthicke S, F. Mueller J (2011). Chemical Pollution on Coral Reefs: Exposure and Ecological Effects. In: W. van Dam J (ed) *Ecological Impacts of Toxic Chemicals (Open Access)*, BENTHAM SCIENCE PUBLISHERS, pp 187–211.

Nelson JS, Grande TC, Wilson MVH (2016). *Fishes of the World*, 5th edn. John Wiley & Sons, Inc: Hoboken, NJ, USA.

Neme R, Amador C, Yildirim B, McConnell E, Tautz D (2017). Random sequences are an abundant source of bioactive RNAs or peptides. *Nat Ecol Evol* **1**: 127.

Nilsson Sköld H, Aspengren S, Wallin M (2013). Rapid color change in fish and amphibians - function, regulation, and emerging applications. *Pigment Cell Melanoma Res* **26**: 29–38.

Nishioka M, Ueda H (1985). Electron-microscopic observation on the dermal chromatophores of normal frogs and three kinds of color variants in Rhacophorus schlegelii. *Sci Rep Lab Amphib Biol ,Hiroshima Univ* **7**: 123–155.

Niu D, Wang F, Xie S, Sun F, Wang Z, Peng M, *et al.* (2016). Developmental Transcriptome Analysis and Identification of Genes Involved in Larval Metamorphosis of the Razor Clam, Sinonovacula constricta. *Mar Biotechnol* **18**: 168–175.

Norman CS, O'Gorman L, Gibson J, Pengelly RJ, Baralle D, Ratnayaka JA, *et al.* (2017). Identification of a functionally significant tri-allelic genotype in the Tyrosinase gene (TYR) causing hypomorphic oculocutaneous albinism (OCA1B). *Sci Rep* **7**: 4415.

Nossa CW, Havlak P, Yue J-X, Lv J, Vincent KY, Brockmann HJ, *et al.* (2014). Joint assembly and genetic mapping of the Atlantic horseshoe crab genome reveals ancient whole genome duplication. *Gigascience* **3**: 9.

Nunes VL, Miraldo A, Beaumont MA, Butlin RK, Paulo OS (2011). Association of Mc1r variants with ecologically relevant phenotypes in the European ocellated lizard, Lacerta lepida. *J Evol Biol* **24**: 2289–2298.

Oancea E, Vriens J, Brauchi S, Jun J, Splawski I, Clapham DE (2009). TRPM1

Forms Ion Channels Associated with Melanin Content in Melanocytes. *Sci Signal* **2**: ra21-ra21.

Odenthal J, Rossnagel K, Haffter P, Kelsh RN, Vogelsang E, Brand M, *et al.* (1996). Mutations affecting xanthophore pigmentation in the zebrafish, Danio rerio. *Development* **123**: 391–8.

Odiorne JM (1933). The Occurrence of Guanophores in Fundulus. *Proc Natl Acad Sci U S A* **19**: 750–4.

Ogden JC, Quinn TP (1984). Migration in coral reef fishes: ecological significance and orientation mechanisms. In: *Mechanisms of migration in fishes,*, pp 293–308.

Ohno S (1970). *Evolution by gene duplication.* Springer.

Ohno S, Muramoto J, Christian L, Atkin NB (1967). Diploid-tetraploid relationship among old-world members of the fish family Cyprinidae. *Chromosoma* **23**: 1–9.

Oliphant LW (1987). Pteridines and purines as major pigments of the avian iris. *Pigment cell Res* **1**: 129–31.

Oliphant LW, Hudon J (1993). Pteridines as reflecting pigments and components of reflecting organelles in vertebrates. *Pigment cell Res* **6**: 205–8.

Oliver SG, van der Aart QJM, Agostoni-Carbone ML, Aigle M, Alberghina L, Alexandraki D, *et al.* (1992). The complete DNA sequence of yeast chromosome III. *Nature* **357**: 38–46.

Oliver KR, McComb JA, Greene WK (2013). Transposable Elements: Powerful Contributors to Angiosperm Evolution and Diversity. *Genome Biol Evol* **5**: 1886–1901.

Olivotto I, Di Stefano M, Rosetti S, Cossignani L, Pugnaloni A, Giantomassi F, *et al.* (2011). Live prey enrichment, with particular emphasis on HUFAs, as limiting factor in false percula clownfish (Amphiprion ocellaris, Pomacentridae) larval development and metamorphosis: Molecular and biochemical implications. *Comp Biochem Physiol Part A Mol Integr Physiol* **159**: 207–218.

Ollerton J, McCollin D, Fautin DG, Allen GR (2007). Finding NEMO: nestedness engendered by mutualistic organization in anemonefish and their hosts. *Proc R Soc B Biol Sci* **274**: 591–598.

Olsson M, Stuart-Fox D, Ballen C (2013). Genetics and evolution of colour patterns in reptiles. *Semin Cell Dev Biol* **24**: 529–541.

Orr HA (1990). 'Why Polyploidy is Rarer in Animals Than in Plants' Revisited. *Am Nat* **136**: 759–770.

Osborn TC, Chris Pires J, Birchler JA, Auger DL, Jeffery Chen Z, Lee H-S, *et al.* (2003). Understanding mechanisms of novel gene expression in polyploids. *Trends Genet* **19**: 141–147.

Otto SP, Whitton J (2000). Polyploid Incidence and Evolution. *Annu Rev Genet* **34**: 401–437.

Ou J, Deng H-M, Zheng S-C, Huang L-H, Feng Q-L, Liu L (2014). Transcriptomic analysis of developmental features of Bombyx mori wing disc during metamorphosis. *BMC Genomics* **15**: 820.

Painter KJ, Maini PK, Othmer HG (1999). Stripe formation in juvenile Pomacanthus explained by a generalized turing mechanism with chemotaxis. *Proc Natl Acad Sci U S A* **96**: 5549–5554.

Panchy N, Lehti-Shiu MD, Shiu S-H (2016). Evolution of gene duplication in plants. *Plant Physiol*: pp.00523.2016.

Pandolfi JM, Connolly SR, Marshall DJ, Cohen AL (2011). Projecting Coral Reef Futures Under Global Warming and Ocean Acidification. *Science (80-)* **333**: 418–422.

Pantalacci S, Guéguen L, Petit C, Lambert A, Peterkovà R, Sémon M (2017). Transcriptomic signatures shaped by cell proportions shed light on comparative developmental biology. *Genome Biol* **18**: 29.

Pantalacci S, Sémon M (2015). Transcriptomics of developing embryos and organs: A raising tool for evo-devo. *J Exp Zool Part B Mol Dev Evol* **324**: 363–371.

Papatsenko D, Levine M, Goltsev Y (2011). Clusters of Temporal Discordances Reveal Distinct Embryonic Patterning Mechanisms in Drosophila and Anopheles (MW Kirschner, Ed.). *PLoS Biol* **9**: e1000584.

Papp B, Pál C, Hurst LD (2003). Dosage sensitivity and the evolution of gene families in yeast. *Nature* **424**: 194–197.

Parichy DM (2006). Evolution of danio pigment pattern development. *Heredity (Edinb)* **97**: 200–210.

Parichy DM (2007). Homology and the evolution of novelty duringDanio adult pigment pattern development. *J Exp Zool Part B Mol Dev Evol* **308B**: 578–590.

Parichy DM, Elizondo MR, Mills MG, Gordon TN, Engeszer RE (2009). Normal table of postembryonic zebrafish development: Staging by externally visible anatomy of the living fish. *Dev Dyn* **238**: 2975–3015.

Parichy DM, Ransom DG, Paw B, Zon LI, Johnson SL (2000). An orthologue of the kit-related gene fms is required for development of neural crest-derived xanthophores and a subpopulation of adult melanocytes in the zebrafish, Danio rerio. *Development* **127**: 3031–44.

Parichy DM, Turner JM (2003). Zebrafish puma mutant decouples pigment pattern and somatic metamorphosis. *Dev Biol* **256**: 242–257.

Parmentier E, Colleye O, Mann D (2009). Hearing ability in three clownfish species. *J Exp Biol* **212**: 2023–2026.

Parmentier E, Kéver L, Casadevall M, Lecchini D (2010). Diversity and complexity in the acoustic behaviour of Dacyllus flavicaudus (Pomacentridae). *Mar Biol* **157**: 2317–2327.

Pasquier J, Braasch I, Batzel P, Cabau C, Montfort J, Nguyen T, *et al.* (2017). Evolution of gene expression after whole-genome duplication: New insights from the spotted gar genome. *J Exp Zool Part B Mol Dev Evol* **328**: 709–721.

Pebusque MJ, Coulier F, Birnbaum D, Pontarotti P (1998). Ancient large-scale genome duplications: phylogenetic and linkage analyses shed light on chordate genome evolution. *Mol Biol Evol* **15**: 1145–1159.

Van de Peer Y, Maere S, Meyer A (2009). The evolutionary significance of ancient genome duplications. *Nat Rev Genet* **10**: 725–732.

Pelletier I, Bally-Cuif L, Ziegler I (2001). Cloning and developmental expression of zebrafish GTP cyclohydrolase I. *Mech Dev* **109**: 99–103.

Petsko GA (2002). What's in a name? *Genome Biol* **3**: comment1005.1.

Phillips GAC, How MJ, Lange JE, Marshall NJ, Cheney KL (2017). Disruptive colouration in reef fish: does matching the background reduce predation risk? *J Exp Biol* **220**: 1962–1974.

Piault R, Gasparini J, Bize P, Jenni-Eiermann S, Roulin A (2009). Pheomelanin-Based Coloration and the Ability to Cope with Variation in Food Supply and Parasitism. *Am Nat* **174**: 548–556.

De Pinna MCC (1996). Teleostean Monophyly. In: *Interrelationships of Fishes*, Elsevier, pp 147–162.

Pirarat N, Pratakpiriya W, Jongnimitpaiboon K, Sajjawiriyakul K, Rodkhum C, Chansue N (2011). Lymphocystis disease in cultured false clown anemonefish (Amphiprion ocellaris). *Aquaculture* **315**: 414–416.

Planes S, Jones GP, Thorrold SR (2009). Larval dispersal connects fish populations in a network of marine protected areas. *Proc Natl Acad Sci* **106**: 5693–5697.

Postlethwait JH, Yan Y-L, Gates MA, Horne S, Amores A, Brownlie A, *et al.* (1998). Vertebrate genome evolution and the zebrafish gene map. *Nat Genet* **18**: 345–349.

Poulton E (1890). *The Colours of Animals: Their Meaning and Use, Especially Considered in the Case of Insects*. D. Appleton, 1890.

Powell RA (1982). Evolution of Black-Tipped Tails in Weasels: Predator Confusion. *Am Nat* **119**: 126–131.

Power D., Llewellyn L, Faustino M, Nowell M., Björnsson BT, Einarsdottir I., *et al.* (2001). Thyroid hormones in growth and development of fish. *Comp Biochem Physiol Part C Toxicol Pharmacol* **130**: 447–459.

Pradeu T, Laplane L, Prévot K, Hoquet T, Reynaud V, Fusco G, *et al.* (2016). Defining 'Development'. In: Wassarman PM (ed) *Essays on Developmental Biology, Part B*, Current Topics in Developmental Biology, pp 171–183.

Price SA, Claverie T, Near TJ, Wainwright PC (2015). Phylogenetic insights into the history and diversification of fishes on reefs. *Coral Reefs* **34**: 997–1009.

Price SA, Holzman R, Near TJ, Wainwright PC (2011). Coral reefs promote the evolution of morphological diversity and ecological novelty in labrid fishes. *Ecol Lett* **14**: 462–469.

Protas ME, Hersey C, Kochanek D, Zhou Y, Wilkens H, Jeffery WR, *et al.* (2006). Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nat Genet* **38**: 107–111.

Prum RO (2012). Aesthetic evolution by mate choice: Darwin's really dangerous idea. *Philos Trans R Soc B Biol Sci* **367**: 2253–2265.

Prum RO, Torres RH, Williamson S, Dyck J (1998). Coherent light scattering by blue feather barbs. *Nature* **396**: 28–29.

Prum RO, Torres R, Williamson S, Dyck J (1999). Two-dimensional Fourier analysis of the spongy medullary keratin of structurally coloured feather barbs. *Proc R Soc B Biol Sci* **266**: 13–22.

Putnam NH, Butts T, Ferrier DEK, Furlong RF, Hellsten U, Kawashima T, *et al.* (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature* **453**: 1064–1071.

Randall J. (2005). A Review of Mimicry in Marine Fishes. *Zool Stud* **44**: 299–328.

Randall J., Fautin DG (2002). Fishes other than anemonefishes that associate with sea anemones. *Coral Reefs* **21**: 188–190.

Ravi V, Venkatesh B (2018). The Divergent Genomes of Teleosts. *Annu Rev Anim Biosci* **6**: 47–68.

Reed RD, Papa R, Martin A, Hines HM, Counterman BA, Pardo-Diaz C, *et al.* (2011). optix Drives the Repeated Convergent Evolution of Butterfly Wing Pattern Mimicry. *Science (80-)* **333**: 1137–1141.

Regier JC, Shultz JW, Zwick A, Hussey A, Ball B, Wetzer R, *et al.* (2010). Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* **463**: 1079–1083.

Relyea RA (2007). Getting out alive: how predators affect the decision to metamorphose. *Oecologia* **152**: 389–400.

Richards WJ, Lindeman KC (1987). Recruitment dynamics of reef fishes: planktonic processes, settlement and demersal ecologies, and fishery analysis. *Bull Mar Sci* **41**: 392–410.

Riley DR, Sieber KB, Robinson KM, White JR, Ganesan A, Nourbakhsh S, *et al.* (2013). Bacteria-Human Somatic Cell Lateral Gene Transfer Is Enriched in Cancer Samples (JA Eisen, Ed.). *PLoS Comput Biol* **9**: e1003107.

Robertson DR (1998). Do coral-reef fish faunas have a distinctive taxonomic structure? *Coral Reefs* **17**: 179–186.

Robertson FM, Gundappa MK, Grammes F, Hvidsten TR, Redmond AK, Lien S, *et al.* (2017). Lineage-specific rediploidization is a mechanism to explain time-lags between genome duplication and evolutionary diversification. *Genome Biol* **18**: 111.

Robinson-Rechavi M, Marchand O, Escriva H, Bardet P-L, Zelus D, Hughes S, *et al.* (2001). Euteleost Fish Genomes are Characterized by Expansion of Gene Families. *Genome Res* **11**: 781–788.

Rohrlich ST, Porter KR (1972). Fine structural observations relating to the production of color by the iridophores of a lizard, Anolis carolinensis. *J Cell Biol* **53**: 38–52.

Rohwer S (1978). Passerine Subadult Plumages and the Deceptive Acquisition of Resources: Test of a Critical Assumption. *Condor* **80**: 173.

Rosenblum EB, Rompler H, Schoneberg T, Hoekstra HE (2010). Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proc Natl Acad Sci* **107**: 2113–2117.

Roulin A (2009). Covariation between eumelanic pigmentation and body mass only under specific conditions. *Naturwissenschaften* **96**: 375–382.

Roulin A (2015). Condition-dependence, pleiotropy and the handicap principle of sexual selection in melanin-based colouration. *Biol Rev* **91**: 328–348.

Roulin A, Ducrest A-L (2013). Genetics of colouration in birds. *Semin Cell Dev Biol* **24**: 594–608.

Rousseau K, Dufour S (2012). Introduction to Fish First and Secondary Metamorphoses. In: *Metamorphosis in Fish*, Science Publishers, p .

Roux J, Rosikiewicz M, Robinson-Rechavi M (2015). What to compare and how: Comparative transcriptomics for Evo-Devo. *J Exp Zool Part B Mol Dev Evol* **324**: 372–382.

Rowen L, Mahairas G, Hood L (1997). Sequencing the human genome. *Science (80-)* **278**: 605–7.

Rowland HM (2011). The history, theory and evidence for a cryptic function of countershading. In: Stevens M, Merilaita S (eds) *Animal Camouflage: from mechanisms to function.*, Cambridge University Press: Cambridge, pp 53–72.

Rowland HM, Mappes J, Ruxton GD, Speed MP (2010). Mimicry between unequally defended prey can be parasitic: evidence for quasi-Batesian mimicry. *Ecol Lett* **13**: 1494–1502.

Rudnicki MA, Braun T, Hinuma S, Jaenisch R (1992). Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development. *Cell* **71**: 383–390.

Sachs LM, Jones PL, Havis E, Rouse N, Demeneix BA, Shi Y-B (2002). Nuclear receptor corepressor recruitment by unliganded thyroid hormone receptor in gene repression during Xenopus laevis development. *Mol Cell Biol* **22**: 8527–38.

Saenko S V, Teyssier J, van der Marel D, Milinkovitch MC (2013). Precise colocalization of interacting structural and pigmentary elements generates extensive color pattern variation in Phelsuma lizards. *BMC Biol* **11**: 105.

Safi R, Bertrand S, Marchand O, Duffraisse M, de Luze A, Vanacker J-M, *et al.* (2004). The Axolotl (Ambystoma mexicanum), a Neotenic Amphibian, Expresses Functional Thyroid Hormone Receptors. *Endocrinology* **145**: 760–772.

Salse J (2016). Ancestors of modern plant crops. *Curr Opin Plant Biol* **30**: 134–142.

Salvat B (1992). Coral reefs - a challenging ecosystem for human societies. *Glob Environ Chang* **2**: 12–18.

San-Jose LM, Roulin A (2017). Genomics of coloration in natural animal populations. *Philos Trans R Soc B Biol Sci* **372**: 20160337.

Sanger F, Nicklen S, Coulson AR (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**: 5463–7.

Santana SE, Lynch Alfaro J, Alfaro ME (2012). Adaptive evolution of facial colour patterns in Neotropical primates. *Proc R Soc B Biol Sci* **279**: 2204–2211.

Santini F, Harmon LJ, Carnevale G, Alfaro ME (2009). Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. *BMC Evol Biol* **9**: 194.

Santos E, Braasch I, Boileau N, Meyer BS, Bo A, Belting H, *et al.* (2014). The evolution of cichlid fish egg-spots is linked with a cis-regulatory change. *Nat Commun.*

Scannell DR, Frank AC, Conant GC, Byrne KP, Woolfit M, Wolfe KH (2007). Independent sorting-out of thousands of duplicated gene pairs in two yeast species descended from a whole-genome duplication. *Proc Natl Acad Sci* **104**: 8397–8402.

Schally A V, Redding TW, Bowers CY, Barrett JF (1969). Isolation and properties of porcine thyrotropin-releasing hormone. *J Biol Chem* **244**: 4077–88.

Schartl M, Larue L, Goda M, Bosenberg MW, Hashimoto H, Kelsh RN (2016). What is a vertebrate pigment cell? *Pigment Cell Melanoma Res* **29**: 8–14.

Schioth HB, Vastermark A, Cone RD (2011). Reply to Braasch and Postlethwait: Evolutionary origin of the teleost A2 agouti genes (agouti signaling protein 2 and agouti-related protein 2) remains unclear. *Proc Natl Acad Sci* **108**: E49–E50.

Schluter D (2000). *The ecology of adaptive radiation.* Oxford University Press.

Schoener TW (1974). Resource Partitioning in Ecological Communities. *Science (80-)* **185**: 27–39.

Schranz M, Mohammadin S, Edger PP (2012). Ancient whole genome duplications, novelty and diversification: the WGD Radiation Lag-Time Model. *Curr Opin Plant Biol* **15**: 147–153.

Schreiber AM, Cai L, Brown DD (2005). Remodeling of the intestine during metamorphosis of Xenopus laevis. *Proc Natl Acad Sci* **102**: 3720–3725.

Schreiber AM, Specker JL (1998). Metamorphosis in the summer flounder (*Paralichthys dentatus*): stage-specific developmental response to altered thyroid status. *Gen Comp Endocrinol* **111**: 156–66.

Schunter C, Welch MJ, Nilsson GE, Rummer JL, Munday PL, Ravasi T (2018). An interplay between plasticity and parental phenotype determines impacts of ocean acidification on a reef fish. *Nat Ecol Evol* **2**: 334–342.

Schwager EE, Sharma PP, Clarke T, Leite DJ, Wierschin T, Pechmann M, *et al.* (2017). The house spider genome reveals an ancient whole-genome duplication during arachnid evolution. *BMC Biol* **15**: 62.

Schwartz J (2006). Sonic Hedgehog' sounded funny, at first. *New York Times*.

Scott A, Dixson DL (2016). Reef fishes can recognize bleached habitat during settlement: sea anemone bleaching alters anemonefish host selection. *Proc R Soc B Biol Sci* **283**: 20152694.

Scott A, Francisco B (2006). Observations on the feeding behaviour of resident anemonefish during host sea anemone spawning. *Coral Reefs* **25**: 451–451.

Seehausen O, Terai Y, Magalhaes IS, Carleton KL, Mrosso HDJ, Miyagi R, *et al.* (2008). Speciation through sensory drive in cichlid fish. *Nature* **455**: 620–626.

Semon M, Wolfe KH (2008). Preferential subfunctionalization of slow-evolving genes after allopolyploidization in Xenopus laevis. *Proc Natl Acad Sci* **105**: 8333–8338.

Senar J (2006). Color displays as intrasexual signals of aggression and dominance. In: Hill G., McGraw K (eds) *Bird Coloration: Function and evolution*, Harvard University Press, pp 87–136.

Seringhaus MR, Cayting PD, Gerstein MB (2008). Uncovering trends in gene naming. *Genome Biol* **9**: 401.

Session AM, Uno Y, Kwon T, Chapman JA, Toyoda A, Takahashi S, *et al.* (2016). Genome evolution in the allotetraploid frog Xenopus laevis. *Nature* **538**: 336–343.

Shamim G, Ranjan SK, Pandey DM, Ramani R (2014). Biochemistry and biosynthesis of insect pigments. *Eur J Entomol* **111**: 149–164.

Sharma PP, Baker CM, Cosgrove JG, Johnson JE, Oberski JT, Raven RJ, *et al.* (2018). A revised dated phylogeny of scorpions: Phylogenomic support for ancient divergence of the temperate Gondwanan family Bothriuridae. *Mol Phylogenet Evol* **122**: 37–45.

Sharon G, Benharroch D, Kachko L, Reis-Hevlin N, Zilberg D (2015). Liposarcoma in clownfish, Amphiprion ocellaris Cuvier, produced in indoor aquaculture. *J Fish Dis* **38**: 575–580.

Shawkey MD, Hill GE (2005). Carotenoids need structural colours to shine. *Biol Lett* **1**: 121–124.

Shendure J (2008). The beginning of the end for microarrays? *Nat Methods* **5**: 585–587.

Shendure J, Mitra RD, Varma C, Church GM (2004). Advanced sequencing technologies: methods and goals. *Nat Rev Genet* **5**: 335–344.

Shibahara S, Tomita Y, Sakakura T, Nager C, Chaudhuri B, Müller R (1986). Cloning and expression of cDNA encoding mouse tyrosinase. *Nucleic Acids Res* **14**: 2413–27.

Sick S, Reinker S, Timmer J, Schlake T (2006). WNT and DKK Determine Hair Follicle Spacing Through a Reaction-Diffusion Mechanism. *Science (80-)* **314**: 1447–1450.

Siebeck UE (2014). Communication in the Ultraviolet: Unravelling the Secret Language of Fish. In: *Biocommunication of Animals*, Springer Netherlands: Dordrecht, pp 299–320.

Siebeck UE, Parker AN, Sprenger D, Mäthger LM, Wallis G (2010). A Species of Reef Fish that Uses Ultraviolet Patterns for Covert Face Recognition. *Curr Biol* **20**: 407–410.

Simoes-Costa M, Bronner ME (2015). Establishing neural crest identity: a gene regulatory recipe. *Development* **142**: 242–257.

Simons MJP, Cohen AA, Verhulst S (2012). What Does Carotenoid-Dependent Coloration Tell? Plasma Carotenoid Level Signals Immunocompetence and Oxidative Stress State in Birds–A Meta-Analysis (JJ Bolhuis, Ed.). *PLoS One* **7**: e43088.

Simons MJP, Verhulst S (2011). Zebra finch females prefer males with redder bills independent of song rate—a meta-analysis. *Behav Ecol* **22**: 755–762.

Simpson SD (2005). Homeward Sound. *Science (80-)* **308**: 221–221.

Simpson SD, Meekan MG, Larsen NJ, McCauley RD, Jeffs A (2010). Behavioral plasticity in larval reef fish: orientation is influenced by recent acoustic experiences. *Behav Ecol* **21**: 1098–1105.

Singh AP, Dinwiddie A, Mahalwar P, Schach U, Linker C, Irion U, *et al.* (2016). Pigment Cell Progenitors in Zebrafish Remain Multipotent through Metamorphosis. *Dev Cell* **38**: 316–330.

Singh AP, Nüsslein-Volhard C (2015). Zebrafish Stripes as a Model for Vertebrate Colour Pattern Formation. *Curr Biol* **25**: R81–R92.

Singh AP, Schach U, Nüsslein-Volhard C (2014). Proliferation, dispersal and patterned aggregation of iridophores in the skin prefigure striped colouration of zebrafish. *Nat Cell Biol* **16**: 604–611.

Siroki J, Hodurkova J, Negrutiu I, Vyskot B (1999). Functional and Structural Chromosome Analyses in Autotetraploid Silene latifolia. *Ann Bot* **84**: 633–638.

Skelhorn J, Rowland HM, Speed MP, Ruxton GD (2010). Masquerade: Camouflage Without Crypsis. *Science (80-)* **327**: 51–51.

Skrabanek L, Wolfe KH (1998). Eukaryote genome duplication - where's the evidence? *Curr Opin Genet Dev* **8**: 694–700.

Slominski A, Tobin DJ, Shibahara S, Wortsman J (2004). Melanin Pigmentation in Mammalian Skin and Its Hormonal Regulation. *Physiol Rev* **84**: 1155–1228.

Smith MM (1987). Molecular evolution of the Saccharomyces cerevisiae histone gene loci. *J Mol Evol* **24**: 252–9.

Smith JJ, Antonacci F, Eichler EE, Amemiya CT (2009). Programmed loss of millions of base pairs from a vertebrate genome. *Proc Natl Acad Sci* **106**: 11212–11217.

Smith JJ, Keinath MC (2015). The sea lamprey meiotic map improves resolution of ancient vertebrate genome duplications. *Genome Res* **25**: 1081–1090.

Smith JJ, Kuraku S, Holt C, Sauka-Spengler T, Jiang N, Campbell MS, *et al.* (2013). Sequencing of the sea lamprey (Petromyzon marinus) genome provides insights into vertebrate evolution. *Nat Genet* **45**: 415–421.

Smith JJ, Timoshevskaya N, Ye C, Holt C, Keinath MC, Parker HJ, *et al.* (2018). The sea lamprey germline genome provides insights into programmed genome rearrangement and vertebrate evolution. *Nat Genet* **50**: 270–277.

Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, *et al.* (2009). Polyploidy and angiosperm diversification. *Am J Bot* **96**: 336–348.

Soltis PS, Marchant DB, Van de Peer Y, Soltis DE (2015). Polyploidy and genome evolution in plants. *Curr Opin Genet Dev* **35**: 119–125.

Songwut P, Sataporn D, Opas T (2014). A simple method for cell culture of 'Nemo' ocellaris clownfish (Amphiprion ocellaris, Cuvier 1830). *Cell Biol Int Rep* **21**: 39–45.

Sonneborn TM (1975). The Paramecium aurelia Complex of Fourteen Sibling Species. *Trans Am Microsc Soc* **94**: 155.

Soucy SM, Huang J, Gogarten JP (2015). Horizontal gene transfer: building the web of life. *Nat Rev Genet* **16**: 472–482.

de Souza FSJ, Bumaschny VF, Low MJ, Rubinstein M (2005). Subfunctionalization of Expression and Peptide Domains Following the Ancient Duplication of the Proopiomelanocortin Gene in Teleost Fishes. *Mol Biol Evol* **22**: 2417–2427.

Spalding M, Green E, Ravilious C (2007). *World Atlas of Coral Reefs*. University of California, Berkeley (California).

Spalding MD, Grenfell AM (1997). New estimates of global and regional coral reef areas. *Coral Reefs* **16**: 225–230.

Speed MP (2000). Warning signals, receiver psychology and predator memory. *Anim Behav* **60**: 269–278.

Spring J (1997). Vertebrate evolution by interspecific hybridisation--are we polyploid? *FEBS Lett* **400**: 2–8.

Stanganello E, Hagemann AIH, Mattes B, Sinner C, Meyen D, Weber S, *et al.* (2015). Filopodia-based Wnt transport during vertebrate tissue patterning. *Nat Commun* **6**: 5846.

Stavenga DG, Leertouwer HL, Marshall NJ, Osorio D (2011). Dramatic colour changes in a bird of paradise caused by uniquely structured breast feather

barbules. *Proc R Soc B Biol Sci* **278**: 2098–2104.

Steiner CC, Rompler H, Boettger LM, Schoneberg T, Hoekstra HE (2008). The Genetic Basis of Phenotypic Convergence in Beach Mice: Similar Pigment Patterns but Different Genes. *Mol Biol Evol* **26**: 35–45.

Steiner CC, Weber JN, Hoekstra HE (2007). Adaptive Variation in Beach Mice Produced by Two Interacting Pigmentation Genes (MA. Noor, Ed.). *PLoS Biol* **5**: e219.

Stern DL, Orgogozo V (2009). Is Genetic Evolution Predictable? *Science (80-)* **323**: 746–751.

Stevens M, Merilaita S (2009). Animal camouflage: current issues and new perspectives. *Philos Trans R Soc B Biol Sci* **364**: 423–427.

Stieb SM, Cortesi F, Sueess L, Carleton KL, Salzburger W, Marshall NJ (2017). Why UV vision and red vision are important for damselfish (Pomacentridae): structural and expression variation in opsin genes. *Mol Ecol* **26**: 1323–1342.

Stobutzki I, Bellwood D (1997). Sustained swimming abilities of the late pelagic stages of coral reef fishes. *Mar Ecol Prog Ser* **149**: 35–41.

Streelman JT, Alfaro M, Westneat MW, Bellwood DR, Karl SA (2002). Evolutionary history of the parrotfishes: biogeography, ecomorphology, and comparative diversity. *Evolution* **56**: 961–71.

Streisinger G, Walker C, Dower N, Knauber D, Singer F (1981). Production of clones of homozygous diploid zebra fish (Brachydanio rerio). *Nature* **291**: 293–6.

Sturm RA, Teasdale RD, Box NF (2001). Human pigmentation genes: identification, structure and consequences of polymorphic variation. *Gene* **277**: 49–62.

Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Jakobsdottir M, *et al.* (2008). Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet* **40**: 835–837.

Svartman M, Stone G, Stanyon R (2005). Molecular cytogenetics discards polyploidy in mammals. *Genomics* **85**: 425–430.

Szczebak JT, Henry RP, Al-Horani FA, Chadwick NE (2013). Anemonefish oxygenate their anemone hosts at night. *J Exp Biol* **216**: 970–976.

Tamayo P, Slonim D, Mesirov J, Zhu Q, Kitareewan S, Dmitrovsky E, *et al.* (1999). Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc Natl Acad Sci U S A* **96**: 2907–12.

Tan MH, Au KF, Yablonovitch AL, Wills AE, Chuang J, Baker JC, *et al.* (2013). RNA sequencing reveals a diverse and dynamic repertoire of the Xenopus tropicalis transcriptome over development. *Genome Res* **23**: 201–216.

Tan MH, Austin CM, Hammer MP, Lee YP, Croft LJ, Gan HM (2018). Finding Nemo: Hybrid assembly with Oxford Nanopore and Illumina reads greatly improves the Clownfish (Amphiprion ocellaris) genome assembly. *Gigascience*.

Tanaka Y (1996). Social selection and the evolution of animal signals. *Evolution (N Y)* **50**: 512–523.

Taniguchi M, Miura K, Iwao H, Yamanaka S (2001). Quantitative Assessment of DNA Microarrays—Comparison with Northern Blot Analyses. *Genomics* **71**: 34–39.

Tank DC, Eastman JM, Pennell MW, Soltis PS, Soltis DE, Hinchliff CE, *et al.* (2015). Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. *New Phytol* **207**: 454–467.

Tarafder AK, Bolasco G, Correia MS, Pereira FJC, Iannone L, Hume AN, *et al.* (2014). Rab11b Mediates Melanin Transfer between Donor Melanocytes and Acceptor Keratinocytes via Coupled Exo/Endocytosis. *J Invest Dermatol* **134**: 1056–1066.

Tassabehji M, Newton VE, Read AP (1994). Waardenburg syndrome type 2 caused by mutations in the human microphthalmia (MITF) gene. *Nat Genet* **8**: 251–255.

Tata JR (2008). Getting hooked on thyroid hormone action: A semi-autobiographical account. *J Biosci* **33**: 653–667.

Taylor JS, Raes J (2004). Duplication and Divergence: The Evolution of New Genes and Old Ideas. *Annu Rev Genet* **38**: 615–643.

Teichmann SA, Babu MM (2004). Gene regulatory network growth by duplication. *Nat Genet* **36**: 492–496.

Teyssier J, Saenko S V., van der Marel D, Milinkovitch MC (2015). Photonic crystals cause active colour change in chameleons. *Nat Commun* **6**: 6368.

Thomas DB, McGoverin CM, McGraw KJ, James HF, Madden O (2013). Vibrational spectroscopic analyses of unique yellow feather pigments (spheniscins) in penguins. *J R Soc Interface* **10**: 20121065–20121065.

Thomas DB, McGraw KJ, Butler MW, Carrano MT, Madden O, James HF (2014). Ancient origins and multiple appearances of carotenoid-pigmented feathers in birds. *Proc R Soc B Biol Sci* **281**: 20140806–20140806.

Thompson DA, Roy S, Chan M, Styczynsky MP, Pfiffner J, French C, *et al.* (2013). Evolutionary principles of modular gene regulation in yeasts. *Elife* **2**.

Thompson A, Vo D, Comfort C, Zakon HH (2014). Expression Evolution Facilitated the Convergent Neofunctionalization of a Sodium Channel Gene. *Mol Biol Evol* **31**: 1941–1955.

Thorgaard GH, Bailey GS, Williams D, Buhler DR, Kaattari SL, Ristow SS, *et al.* (2002). Status and opportunities for genomics research with rainbow trout. *Comp Biochem Physiol B Biochem Mol Biol* **133**: 609–46.

Thorisson K (1994). Is metamorphosis a critical interval in the early life of marine fishes? *Environ Biol Fishes* **40**: 23–36.

Thresher RE (1984). *Reproduction in reef fishes*. THF Publications.

Thresher RE, Colin PL, Bell LJ (1989). Planktonic Duration, Distribution and Population Structure of Western and Central Pacific Damselfishes (Pomacentridae). *Copeia* **1989**: 420.

Tibbetts EA, Dale J (2004). A socially enforced signal of quality in a paper wasp. *Nature* **432**: 218–222.

Tinbergen J, Wilts BD, Stavenga DG (2013). Spectral tuning of Amazon parrot feather coloration by psittacofulvin pigments and spongy structures. *J Exp Biol* **216**: 4358–4364.

Tisch D, Schmoll M (2010). Light regulation of metabolic pathways in fungi. *Appl Microbiol Biotechnol* **85**: 1259–1277.

Toomey MB, Lopes RJ, Araújo PM, Johnson JD, Gazda MA, Afonso S, *et al.* (2017). High-density lipoprotein receptor SCARB1 is required for carotenoid coloration in birds. *Proc Natl Acad Sci* **114**: 5219–5224.

Toral GM, Figuerola J, Negro JJ (2008). Multiple ways to become red: Pigment identification in red feathers using spectrometry. *Comp Biochem Physiol Part B Biochem Mol Biol* **150**: 147–152.

Trifonov VA, Romanenko SS, Beklemisheva VR, Biltueva LS, Makunin AI, Lemskaya NA, *et al.* (2016). Evolutionary plasticity of acipenseriform genomes. *Chromosoma* **125**: 661–668.

Turing AM (1952). The Chemical Basis of Morphogenesis. *Philos Trans R Soc B Biol Sci* **237**: 37–72.

Uyeno T, Smith GR (1972). Tetraploid origin of the karyotype of catostomid fishes. *Science* **175**: 644–6.

Vanneste K, Maere S, Van de Peer Y (2014). Tangled up in two: a burst of genome duplications at the end of the Cretaceous and the consequences for plant evolution. *Philos Trans R Soc B Biol Sci* **369**: 20130353.

Västermark Å, Krishnan A, Houle ME, Fredriksson R, Cerdá-Reverter JM, Schiöth HB (2012). Identification of Distant Agouti-Like Sequences and Re-Evaluation of the Evolutionary History of the Agouti-Related Peptide (AgRP) (V Laudet, Ed.). *PLoS One* **7**: e40982.

Venkatesh B, Lee AP, Ravi V, Maurya AK, Lian MM, Swann JB, *et al.* (2014). Elephant shark genome provides unique insights into gnathostome evolution. *Nature* **505**: 174–179.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, *et al.* (2001). The Sequence of the Human Genome. *Science (80-)* **291**: 1304–1351.

Victor BC (1986). Delayed Metamorphosis with Reduced Larval Growth in a Coral Reef Fish (Thalassoma bifasciatum). *Can J Fish Aquat Sci* **43**: 1208–1213.

Vinckenbosch N, Dupanloup I, Kaessmann H (2006). Evolutionary fate of

retroposed gene copies in the human genome. *Proc Natl Acad Sci* **103**: 3220–3225.

Vitt JL, Calwell PJ (2013). *Herpetology, Fourth Edition: An Introductory Biology of Amphibians and Reptiles*.

Voldoire E, Brunet F, Naville M, Volff J-N, Galiana D (2017). Expansion by whole genome duplication and evolution of the sox gene family in teleost fish (H Vaudry, Ed.). *PLoS One* **12**: e0180936.

Volff J-N (2005). Genome evolution and biodiversity in teleost fish. *Heredity (Edinb)* **94**: 280–294.

Voordeckers K, Pougach K, Verstrepen KJ (2015). How do regulatory networks evolve and expand throughout evolution? *Curr Opin Biotechnol* **34**: 180–188.

Vorbach BS, Wolf JC, Yanong RP (2018). Odontomas in two long-finned ocellaris clownfish (Amphiprion ocellaris). *J Vet Diagnostic Investig* **30**: 136–139.

De Vries GJ, Panzica GC (2006). Sexual differentiation of central vasopressin and vasotocin systems in vertebrates: Different mechanisms, similar endpoints. *Neuroscience* **138**: 947–955.

Wagner CE, Harmon LJ, Seehausen O (2012). Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* **487**: 366–369.

Wagner GP, Pavlicev M, Cheverud JM (2007). The road to modularity. *Nat Rev Genet* **8**: 921–931.

Wagner JT, Singh PP, Romney AL, Riggs CL, Minx P, Woll SC, *et al.* (2018). The genome of Austrofundulus limnaeus offers insights into extreme vertebrate stress tolerance and embryonic development. *BMC Genomics* **19**: 155.

Wainwright PC, Bellwood DR (2002). Ecomorphology of Feeding in Coral Reef Fishes. In: *Coral Reef Fishes*, Elsevier, pp 33–55.

Wakamatsu Y, Pristyazhnyuk S, Kinoshita M, Tanaka M, Ozato K (2001). The see-through medaka: A fish model that is transparent throughout life. *Proc Natl Acad Sci* **98**: 10046–10050.

Walderich B, Singh AP, Mahalwar P, Nüsslein-Volhard C (2016). Homotypic cell competition regulates proliferation and tiling of zebrafish pigment cells during colour pattern formation. *Nat Commun* **7**: 11462.

Wallace AR (1877). The colours of animals and plants. *Am Nat* **11**: 713–728.

Wallace AR (1879). The Protective Colours of Animals. In: Smith CH (ed) *Science for All*,

Wallace AR (1895). *Natural selection and tropical nature, 2nd edn.* New York, NY: Macmillan and Co.

Walpita CN, Crawford AD, Janssens EDR, Van der Geyten S, Darras VM (2009). Type 2 Iodothyronine Deiodinase Is Essential for Thyroid Hormone-Dependent Embryonic Development and Pigmentation in Zebrafish. *Endocrinology* **150**: 530–539.

Wang Z, Gerstein M, Snyder M (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* **10**: 57–63.

Wang J-T, Li J-T, Zhang X-F, Sun X-W (2012). Transcriptome analysis reveals the time of the fourth round of genome duplication in common carp (Cyprinus carpio). *BMC Genomics* **13**: 96.

Wang Y, Lu Y, Zhang Y, Ning Z, Li Y, Zhao Q, *et al.* (2015). The draft genome of the grass carp (Ctenopharyngodon idellus) provides insights into its evolution and vegetarian adaptation. *Nat Genet* **47**: 625–631.

Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, *et al.* (2011). The genome of the mesopolyploid crop species Brassica rapa. *Nat Genet* **43**: 1035–1039.

Wapinski I, Pfeffer A, Friedman N, Regev A (2007). Natural history and evolutionary principles of gene duplication in fungi. *Nature* **449**: 54–61.

Warren IA, Naville M, Chalopin D, Levin P, Berger CS, Galiana D, *et al.* (2015). Evolutionary impact of transposable elements on genomic diversity and lineage-specific innovation in vertebrates. *Chromosom Res* **23**: 505–531.

Wasik BR, Liew SF, Lilien DA, Dinwiddie AJ, Noh H, Cao H, *et al.* (2014). Artificial selection for structural color on butterfly wings and comparison with natural evolution. *Proc Natl Acad Sci* **111**: 12109–12114.

Watanabe M, Iwashita M, Ishii M, Kurachi Y, Kawakami A, Kondo S, *et al.* (2006). Spot pattern of leopard Danio is caused by mutation in the zebrafish connexin41.8 gene. *EMBO Rep* **7**: 893–897.

Watanabe M, Kondo S (2012). Changing clothes easily: connexin41.8 regulates skin pattern variation. *Pigment Cell Melanoma Res* **25**: 326–330.

Watanabe M, Kondo S (2015). Is pigment patterning in fish skin determined by the Turing mechanism? *Trends Genet* **31**: 88–96.

Watanabe K, Washio Y, Fujinami Y, Aritaki M, Uji S, Suzuki T (2008). Adult-type pigment cells, which color the ocular sides of flounders at metamorphosis, localize as precursor cells at the proximal parts of the dorsal and anal fins in early larvae. *Dev Growth Differ* **50**: 731–741.

Watson AC, Siemann LA, Hanlon RT (2014). Dynamic camouflage by Nassau groupers Epinephelus striatus on a Caribbean coral reef. *J Fish Biol* **85**: 1634–1649.

Wellington GM, Victor BC (1989). Planktonic larval duration of one hundred species of Pacific and Atlantic damselfishes (Pomacentridae). *Mar Biol* **101**: 557–567.

Werner T, Koshikawa S, Williams TM, Carroll SB (2010). Generation of a novel wing colour pattern by the Wingless morphogen. *Nature* **464**: 1143–1148.

Wertheim B, Beukeboom LW, van de Zande L (2013). Polyploidy in Animals: Effects of Gene Expression on Sex Determination, Evolution and Ecology. *Cytogenet Genome Res* **140**: 256–269.

West-Eberhard M (2003). *Developmental Plasticity and Evolution*. Oxford University Press.

Wilbur HM, Collins JP (1973). Ecological Aspects of Amphibian Metamorphosis: Nonnormal distributions of competitive ability reflect selection for facultative metamorphosis. *Science (80-)* **182**: 1305–1314.

Williams RM, Winkfein RJ, Ginger RS, Green MR, Schnetkamp PP, Wheeler GN (2017). A functional approach to understanding the role of NCKX5 in Xenopus pigmentation (EM Verheyen, Ed.). *PLoS One* **12**: e0180465.

Wolf JB, Brodie III ED, Moore AJ (1999). Interacting Phenotypes and the Evolutionary Process. II. Selection Resulting from Social Interactions. *Am Nat* **153**: 254–266.

Wolf Horrell EM, Boulanger MC, D'Orazio JA (2016). Melanocortin 1 Receptor: Structure, Function, and Regulation. *Front Genet* **7**.

Wolfe KH (2001). Yesterday's polyploids and the mystery of diploidization. *Nat Rev Genet* **2**: 333–341.

Wolfe KH (2015). Origin of the Yeast Whole-Genome Duplication. *PLOS Biol* **13**: e1002221.

Wolfe KH, Shields DC (1997). Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**: 708–713.

Wright GM, McBurney KM (1992). Changes in the ventral dermis and development of iridophores in the anadromous sea lamprey, Petromyzon marinus, during metamorphosis: an ultrastructural study. *Histol Histopathol* **7**: 237–50.

Wu X, Qi X (2010). Genes encoding hub and bottleneck enzymes of the Arabidopsis metabolic network preferentially retain homeologs through whole genome duplication. *BMC Evol Biol* **10**: 145.

Wucherer MF, Michiels NK (2012). A Fluorescent Chromatophore Changes the Level of Fluorescence in a Reef Fish (J Zeil, Ed.). *PLoS One* **7**: e37913.

Wucherer MF, Michiels NK (2014). Regulation of red fluorescent light emission in a cryptic marine fish. *Front Zool* **11**: 1.

Xu P, Zhang X, Wang X, Li J, Liu G, Kuang Y, *et al.* (2014a). Genome sequence and genetic diversity of the common carp, Cyprinus carpio. *Nat Genet* **46**: 1212–1219.

Xu P, Zhang X, Wang X, Li J, Liu G, Kuang Y, *et al.* (2014b). Genome sequence and genetic diversity of the common carp, Cyprinus carpio. *Nat Genet* **46**: 1212–1219.

Y. Sadovy de Mitcheson and P.L. Colin (2012). Reef Fish Spawning Aggregations: Biology, Research and Management. In: *Reef Fish Spawning Aggregations: Biology, Research and Management*, Vol 565, pp 405–565.

Yaeger C, Ros AM, Cross V, DeAngelis RS, Stobaugh DJ, Rhodes JS (2014). Blockade of arginine vasotocin signaling reduces aggressive behavior and c-Fos

expression in the preoptic area and periventricular nucleus of the posterior tuberculum in male Amphiprion ocellaris. *Neuroscience* **267**: 205–218.

Yamagata Y, Yamamoto E, Aya K, Win KT, Doi K, Sobrizal, *et al.* (2010). Mitochondrial gene in the nuclear genome induces reproductive barrier in rice. *Proc Natl Acad Sci* **107**: 1494–1499.

Yamano K, Miwa S (1998). Differential gene expression of thyroid hormone receptor alpha and beta in fish development. *Gen Comp Endocrinol* **109**: 75–85.

Yanai I, Peshkin L, Jorgensen P, Kirschner MW (2011). Mapping Gene Expression in Two Xenopus Species: Evolutionary Constraints and Developmental Flexibility. *Dev Cell* **20**: 483–496.

Yang H, Zhou Y, Gu J, Xie S, Xu Y, Zhu G, *et al.* (2013). Deep mRNA Sequencing Analysis to Capture the Transcriptome Landscape of Zebrafish Embryos and Larvae (Y Herault, Ed.). *PLoS One* **8**: e64058.

Yaoita Y, Shi Y, Brown D (1990). Xenopus laevis alpha and beta thyroid hormone receptors. *Proc Natl Acad Sci U S A* **87**: 8684.

Yasir I, Qin JG (2007). Embryology and early ontogeny of an anemonefish Amphiprion ocellaris. *J Mar Biol Assoc UK* **87**: 1025.

Yasir I, Qin JG (2009a). Effect of Light Intensity on Color Performance of False Clownfish, Amphiprion ocellaris Cuvier. *J World Aquac Soc* **40**: 337–350.

Yasir I, Qin JG (2009b). Impact of Background on Color Performance of False Clownfish, Amphiprion ocellaris, Cuvier. *J World Aquac Soc* **40**: 724–734.

Yasir I, Qin JG (2010). Effect of dietary carotenoids on skin color and pigments of false clownfish, Amphiprion ocellaris, Cuvier. *J World Aquac Soc* **41**: 308–318.

Yúfera M, Halm S, Beltran S, Fusté B, Planas J V., Martínez-Rodríguez G (2012). Transcriptomic Characterization of the Larval Stage in Gilthead Seabream (Sparus aurata) by 454 Pyrosequencing. *Mar Biotechnol* **14**: 423–435.

Zahavi A (1975). Mate selection-a selection for a handicap. *J Theor Biol* **53**: 205–14.

Zahavi A (1977). The cost of honesty (further remarks on the handicap principle). *J Theor Biol* **67**: 603–5.

Zapata F, Goetz FE, Smith SA, Howison M, Siebert S, Church SH, *et al.* (2015). Phylogenomic Analyses Support Traditional Relationships within Cnidaria (RE Steele, Ed.). *PLoS One* **10**: e0139068.

Zarnescu O (2007). Ultrastructure of the skin melanophores and iridophores in paddlefish, Polyodon spathula. *Micron* **38**: 81–84.

Zhang B, Horvath S (2005). A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* **4**: Article17.

Zhang Y, Liu J, Fu W, Xu W, Zhang H, Chen S, *et al.* (2017). Comparative Transcriptome and DNA methylation analyses of the molecular mechanisms underlying skin color variations in Crucian carp (Carassius carassius L.). *BMC Genet* **18**: 95.

Zhang H, Okamoto N, Ikeda Y (1995). Two c-myc genes from a tetraploid fish, the common carp (Cyprinus carpio). *Gene* **153**: 231–236.

Zhang H, Ravi V, Tay B-H, Tohari S, Pillai NE, Prasad A, *et al.* (2017). Lampreys, the jawless vertebrates, contain only two ParaHox gene clusters. *Proc Natl Acad Sci* **114**: 9146–9151.

Zhou X, Ren L, Meng Q, Li Y, Yu Y, Yu J (2010). The next-generation sequencing technology and application. *Protein Cell* **1**: 520–536.

LISTE DES ABRÉVIATIONS

1R/2R : 1er et 2ème round de WGD à la base des Vertébrés

ACTH : Adreno-Cortico-Trophic Hormone

ACP : Analyse en Composantes Principales

AMPc : Adénosine Monophosphate cyclique

CCaGD : Common carp genome duplication

DHI : Dihydroxyindole

DHICA : acide dihydroxyindole carboxylique

DOPA : dihydroxyphénylalanine

GTP : Guanosine Triphosphate

HGT : Transfert Horizontal de Gènes

Ma : Million d'années

Mpb : millions de paires de bases

MPC : Cellule pigmentaire mosaïque

MSH : melanocyte-stimulating hormone (mélanocortine)

qPCR : PCR quantitative

RGL : Perte de gènes réciproque

Ss4R : 4ème round de WGD à la base des Salmonidés

SSD : Small-Scale duplication

TH : Thyroid hormones

TR : Thyroid hormones receptors

Ts3R : 3ème round de WGD à la base des Téléostéens

TSH : thyroid-stimulating hormone (thyréotropine)

WGD : Whole-Genome Duplication