



HAL
open science

**De la recherche fondamentale sur les bases génomiques
de la diversification du vivant à des initiatives
appliquées de gestion de la biodiversité: contribution au
développement d'approches plus intégratives**

Joris Bertrand

► **To cite this version:**

Joris Bertrand. De la recherche fondamentale sur les bases génomiques de la diversification du vivant à des initiatives appliquées de gestion de la biodiversité: contribution au développement d'approches plus intégratives. Sciences du Vivant [q-bio]. Université de Perpignan Via Domitia, 2022. tel-03853817

HAL Id: tel-03853817

<https://theses.hal.science/tel-03853817v1>

Submitted on 15 Nov 2022

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.



Distributed under a Creative Commons Attribution 4.0 International License



Université de Perpignan Via Domitia

2022

Mémoire présenté en vue de l'obtention de
l'Habilitation à Diriger des Recherches

Par **Joris BERTRAND**

Maître de Conférences au
Laboratoire Génome & Développement des Plantes



**De la recherche fondamentale sur les bases génomiques de la
diversification du vivant à des initiatives appliquées de gestion de la
biodiversité : contribution au développement d'approches plus intégratives**

Soutenu publiquement le 30 juin 2022

Devant le jury composé de :

Dr. Ivan SCOTTI
Dr. Chloé DELMAS
Dr. Maxime BONHOMME
Prof. Stéphanie MANEL
Prof. Jérôme BOISSIER

Rapporteur externe
Rapporteur externe
Rapporteur externe
Examineur externe
Examineur interne

Résumé

La compréhension de l'évolution des populations et des espèces passe par l'étude des liens entre les génotypes, les phénotypes qu'ils déterminent et la *fitness* que ces derniers confèrent à un individu donné. Historiquement, l'écologie et l'évolution se sont principalement attelées à décrypter les liens entre variation phénotypique et *fitness*, souvent en considérant les bases génétiques de la variation phénotypique comme une « boîte noire » dont l'accès était hors d'atteinte. Au mieux, le génotype était considéré comme un marqueur pour expliquer les patrons de distribution géographique de la variation phénotypique et éventuellement pour retracer l'histoire des populations et des espèces. La biologie fonctionnelle s'est quant à elle plutôt focalisée sur l'étude des mécanismes par lesquels le génotype façonnait le phénotype. Or faire le lien entre phénotype et *fitness* dans un contexte réaliste peut nécessiter une étude de la performance de ces phénotypes (et des génotypes qui les déterminent) en populations naturelles. L'avènement des approches dites « -omiques » a, en partie, permis de lever les verrous technologiques qui ont longtemps freiné la mise en place d'une recherche résolument intégrative. Or, en dépit de cette révolution dans le domaine des sciences du vivant, trouver des modèles biologiques pertinents pour mettre en place ce type de recherche demeure un défi majeur. Dans certains cas, un modèle peut paraître tout à fait approprié pour travailler en populations naturelles mais ses caractéristiques intrinsèques limitent la possibilité d'accéder à l'intégralité de l'information génomique et/ou de son utilisation en conditions contrôlées. Dans d'autres cas, le modèle a tout du candidat idéal en conditions expérimentales mais présente des limites concernant la validation de la pertinence des phénotypes en populations naturelles. Le modèle biologique parfait n'existant tout simplement pas, je propose dans ce manuscrit à la fois un retour d'expérience personnel et une feuille de route à propos de pistes possibles pour entreprendre et mener une recherche intégrative et potentiellement fructueuse pour inférer ces liens entre génotypes, phénotypes et *fitness* et mieux appréhender les processus qui sous-tendent l'émergence et le maintien de la diversité biologique.

Remerciements & avant-propos

Avant toute chose, je voudrais remercier les membres du jury pour avoir accepté d'être rapporteurs ou examinateurs de ce travail. **MERCI** donc à Chloé DELMAS, Stéphanie MANEL, Ivan SCOTTI, Maxime BONHOMME et Jérôme BOISSIER. Ce mémoire présenté en vue de l'obtention de l'Habilitation à Diriger des Recherches ne ressemble certainement pas tout à fait à celui que j'aurais imaginé. J'ai l'impression qu'il a été rédigé trop tôt, trop vite et avec une prise de recul insuffisante. Il y a toute cette frustration de ne pas y avoir inclus des résultats encore trop préliminaires ou ceux qui vont arriver au cours des mois prochains. J'essaye de me rassurer en me persuadant que cela prouve que les choses bougent ou en écoutant les retours d'expérience de mes collègues qui pour la plupart, me disent avoir eu le même sentiment au moment de la soumission de leur propre manuscrit. Je vous livre donc ici une feuille de route plus qu'un *corpus* de résultats en ayant hâte de pouvoir prendre en compte votre expertise et échanger avec vous tous pour avancer dans la bonne direction.

Je voudrais témoigner de toute ma reconnaissance à Jean-Marc DERAGON (Directeur d'Unité du LGDP lors de mon recrutement) et Olivier PANAUD (Directeur de l'équipe Analyse du Génome & Évolution m'ayant accueilli à cette période) de ne pas s'être opposé à mon envie d'essayer de lancer mon propre projet de recherche sur *Ophrys*. Je n'oublierai jamais la moue d'Olivier lorsque je lui ai annoncé la taille de génome que mentionnait alors la littérature concernant *Ophrys*... Mais après m'avoir brièvement suggéré d'oublier mon idée, il m'a finalement hébergé dans son « hôtel à projets » et soutenu par la suite pour faire en sorte que j'obtienne les financements nécessaires pour lui faire prendre son envol. **MERCI** aussi à Marie MIROUZE qui a pris la direction de l'équipe MANGO (*Mechanisms of Adaptation and GenOmic*s) depuis, et se doit de « guider ce troupeau de chats » (*herding cats* selon les termes d'Olivier) que nous formons. Je reste persuadé que le jour viendra où nous aurons l'occasion de faire des choses intéressantes sur les variants structuraux. C'est juste que je pars de très (très) loin... Du coup, **MERCI** à mes collègues génomiciens et biologistes fonctionnels du LGDP bien que nous ne parlions pas toujours la même langue. Les sciences du vivant, c'est tout bonnement vertigineux... **MERCI** également à mes collègues enseignant-chercheurs du département de biologie et plus généralement de l'UFR des Sciences Exactes et Expérimentales. Et tant qu'on y est, **MERCI** également à nos personnels administratifs en premier lieu desquels Myriam FACI puis Joëlle VIGIER et Élisabeth GOETSCHY. Encore une fois, pardon pour mon côté (quasi-)maladif à avoir bien du mal à suivre des règles pour lesquelles je n'arrive pas à me persuader du bien fondé.

MERCI à Christelle LLAURO, notre bonne fée du laboratoire pour son aide précieuse dès qu'il faut enfile la blouse blanche et **MERCI** à tous les autres personnels techniques du LGDP en premier lieu desquels Laëticia BARRIAT et Michèle LAUDIÉ qui ont directement contribué au projet *Ophrys*. **MERCI** aussi au personnel des plateformes BioEnvironnement : Margot DOBERVA, Bio2mar via son plateau MSXR : Delphine RAVIGLIONE et Jules KOTARBAT et BioPIC : David PECQUEUR. **MERCI** aussi à Marie-Christine CARPENTIER et à tous ces collègues qui savent coder autrement qu'avec les pieds et essayent de me convaincre jour après jour qu'il ne faut pas juste se satisfaire d'un code qui tourne (Moïne EL BAIDOURI, Martin ROSALIE et Anaïs GIBERT).

MERCI à tous les étudiants que j'ai eu l'opportunité d'encadrer depuis mon recrutement. Dans l'ordre chronologique: Jean-Loup, Florian, Pascaline, Hugo, Dominique, Pauline, Manon, Céleste et Morgane, je voudrais vous remercier pour votre patience et votre bienveillance à toute épreuve face à mon manque chronique de disponibilité... Je promets d'essayer de ne jamais devenir l'encadrant que j'aurais détesté avoir et je dois dire que souvent, j'ai déjà le sentiment que c'est plutôt mal parti... **MERCI** aussi à toutes ces cohortes d'étudiants perpignanais face auxquels je me retrouve trop souvent avec des enseignements pas assez bien préparés mais qui me redonnent (généralement) le goût d'enseigner au fur et à mesure que l'heure tourne.

MERCI à tous mes collaborateurs sans qui mon projet aurait sans doute bien battu de l'aile avant de prendre son envol. **MERCI** à Rosy BUSCAIL pour avoir partagé sans réserve toutes ces connaissances sur les orchidées et les endroits où on les trouvait. **MERCI** à Bertrand SCHATZ, Michel BAGUETTE et Cédric BERTRAND pour cette mutualisation de connaissances et de compétences qui inclu désormais les *Ophrys* et leurs pollinisateurs en passant par leurs gènes et leurs odeurs.

MERCI à Perpignan et sa région. A la mer, à la montagne, à tout ce qui est entre les deux et à tous ces trésors de biodiversité, source intarissable d'inspiration pour le biologiste, et l'être humain.

MERCI à ma famille et à mes amis pour m'avoir soutenu (et supporté) durant toutes ces années. J'ai cette chance inouïe d'exercer le métier dont j'avais toujours rêvé. Je pense avant tout à Anaïs avec qui je passe désormais plus de 95% de mon temps, à la maison ou au labo, en sachant que plus de la moitié de ce temps est sans nul doute consacré à notre travail. Notre petit Élio a déjà bien compris que pour se faire entendre, le mieux était de s'imposer dans la conversation en parlant plus fort que tout le monde pour prendre la parole et formuler des hypothèses scientifiques à tester pour alimenter le débat.

Je ne remercie pas la touche «E» du clavier de mon ordinateur qui après 4 ans d'usage intensif se fait régulièrement la malle et aura sérieusement mis à l'épreuve ma foi pendant la rédaction de ce manuscrit.

Je ne remercie pas non plus celles et ceux qui établissent les règles concernant les missions dans l'administration française mais derrière l'ineptie que constitue la myriade de procédures dématérialisées, je voudrais rendre hommage à tous les êtres humains, pourtant rouages du système, qui m'ont aidé à les assouplir quand elles étaient de toute évidence trop inadaptées pour pouvoir être appliquées...

Cette version de mon manuscrit de HDR correspond à la version publiée le 15 novembre 2022, après prise en compte des suggestions des différents membres du jury.
--

SOMMAIRE

C.V. (SYNTHÉTIQUE)	3
ACTIVITE DE RECHERCHE	5
DE MES DEBUTS EN RECHERCHE A MON RECRUTEMENT : EMERGENCE ET MATURATION DE MON PROJET DE RECHERCHE	5
THESE : SPECIATION CHEZ LE <i>ZOSTEROPS BORBONICUS</i> : OISEAU ENDEMIQUE DE L'ILE DE LA REUNION (TOULOUSE, FRANCE)	5
POST-DOC 1 & 2 : BIOGEOGRAPHIE ET SYSTEMATIQUE MOLECULAIRE DU GENRE <i>ABUDEFDUF</i> (TAIPEI, TAÏWAN) ET CAUSES EVOLUTIVES DE LA RADIATION ADAPTATIVE DES POISSONS-CLOWNS (LAUSANNE, SUISSE)	7
PROJET DE RECHERCHE PRINCIPAL : LIER GENOTYPES, PHENOTYPES ET FITNESS POUR MIEUX COMPRENDRE LES CAUSES DE LA RADIATION ADAPTATIVE DES ORCHIDEES DU GENRE <i>OPHRYS</i>	9
PROJET D'INTEGRATION A MON EQUIPE D'ACCUEIL	10
ÉTUDE PILOTE N°1 : BIOGEOGRAPHIE DU COMPLEXE D'ESPECES <i>OPHRYS AVEYRONENSIS</i>	16
ÉTUDE PILOTE N°2 : TAXONOMIE INTEGRATIVE DU COMPLEXE <i>OPHRYS CATALAUNICA/O. MAGNIFLORA</i>	21
BILAN INTERMEDIAIRE DES DEUX ETUDES PILOTES	22
LE PROJET « <i>DIVERSIFLY</i> »	23
PROJETS DE RECHERCHE 'ANNEXES'	32
GESTION DITE 'ECO-BIO-SOCIALE' D'UNE INVASION BIOLOGIQUE	32
GENOMIQUE DE LA CONSERVATION DE PLANTES ENDEMIQUES OU EMBLEMATIQUES DES PYRENEES	33
SYSTEMATIQUE ET TAXONOMIE (INTEGRATIVE) D'ORCHIDEES EURO-MEDITERRANEENNES ET D'AUTRES ORGANISMES	39
REFLEXIONS PERSONNELLES POUR UNE RECHERCHE PLUS INTEGRATIVE	41
QUELS MARQUEURS GENETIQUES EMPLOYER ?	43
QUELS TRAITS PHENOTYPIQUES UTILISER ET COMMENT FAIRE LE LIEN AVEC LES GENOTYPES ET LA <i>FITNESS</i> ?	46
BILAN DE L'ACTIVITE D'ENCADREMENT	51
ACTIVITE PEDAGOGIQUE ET D'ENSEIGNEMENT	51
EXPERIENCE D'ENSEIGNEMENT ET D'ENCADREMENT ANTERIEURE A MON RECRUTEMENT EN TANT QUE MAITRE DE CONFERENCES	51
EXPERIENCE D'ENSEIGNEMENT ET D'ENCADREMENT DEPUIS MON RECRUTEMENT EN TANT QUE MAITRE DE CONFERENCES	52
POSTURE D'ENSEIGNEMENT	53
SUSCITER ET ENTRETENIR LA CURIOSITE DES ETUDIANTS	53
FAIRE EN SORTE QUE LES ETUDIANTS SOIENT ACTEURS DE LEUR PROPRE APPRENTISSAGE	53
DEVELOPPER CHEZ LES ETUDIANTS UNE DEMARCHE SCIENTIFIQUE ET DES COMPETENCES TRANSVERSALES	54
BILAN PRELIMINAIRE	55
RÉFÉRENCES	56

C.V. (synthétique)

Joris BERTRAND

Date of birth: 11-oct-85 (36 years old)

Nationality: French



Contact:

✉ Laboratoire Génome & Développement des Plantes - UMR 5096 (UPVD/CNRS) - Université de Perpignan Via Domitia, 58 av. Paul Alduy, Bâtiment T, 66860 Perpignan Cedex 09, France.

✉ joris.bertrand@univ-perp.fr

@ www.jorisbertrand.com

Career

- Since Dec 2017** **Maître de Conférences** – Université de Perpignan Via Domitia (France) - Laboratoire Génome & Développement des Plantes (UMR 5096 UPVD/CNRS)
- Sept. 2016 - Sept. 2017** **First assistant** - University of Lausanne (Switzerland) - Dept. of Ecology and Evolution/ Swiss Inst. of Bioinformatics
- Sept. 2015 - Aug. 2016** **Independent researcher** - Macquarie Uni (Sydney, Australia) - Dept. of Life Sciences
- Jan. 2014 - Jul. 2015** **Post-doctoral fellow** - National Taiwan University (Taipei) - Inst. of Oceanography
- Sept. 2012 - Aug. 2013** **Teaching & research assistant** - University Paul Sabatier, Toulouse III (France) - Laboratoire Evolution & Diversité Biologique (UMR 5174 UPS/CNRS/IRD)
- Nov. 2009 - Aug. 2012** **PhD student & instructor** - University Paul Sabatier Toulouse, III (France) - Laboratoire Evolution & Diversité Biologique (UMR 5174 UPS/CNRS/IRD)

Education

Paul Sabatier University, Toulouse III (Toulouse, France)

- 2013** **Doctorate degree** in “Ecology, Biodiversity & Evolution”
- 2009** **Master degree** in “Biodiversity, Ecology & Evolution” (with honours)
- 2008** “**Maîtrise**” in “Ecology” (with honours)

Jean-François Champollion University (Albi, France)

- 2007** **Bachelor degree** in “Organisms, Populations & Ecosystems Biology” (with honours)

Teaching experience

Completion of > 1000 h of teaching:

- Since 2017** **Maître de Conférences** (>650 h) Université de Perpignan Via Domitia
- 2016** **First assistant** (>30 h) University of Lausanne, (Switzerland)
- 2009 - 2013** **Instructor & teaching Assistant** (> 300 h) Paul Sabatier University, Toulouse III (Toulouse, France)

Research interests (keywords)

- Genomic bases of evolutionary divergence and adaptive radiations
- Biogeography of species and regions
- Conservation genomics and ecological niche modelling of patrimonial, threatened species

Funded research projects as a PI

- 2020** Grant from the Fédération de Recherche en Écologie & Environnement from the UPVD (3,32 k€) – 'Étude intégrative de l'état de santé des populations de Grand Tétrás (*Tetrao urogallus aquitanicus*), une espèce emblématique locale menacée d'extinction' (co-PI : O. Rey & V. Hinoux)
- 2021** Grant from the Agence Nationale pour la Recherche (ANR) Jeune Chercheur Jeune Chercheuse (JCJC) – DiversiFly : Adaptive Diversification of Fly orchids as a new model to investigate ecological speciation and evolutionary radiations (370 k€).
- 2020** Grant from the Fédération de Recherche en Écologie & Environnement from the UPVD (5 k€) – 'Développement d'une approche intégrative pour la conservation d'espèces emblématiques - Étude pilote sur le couple *Ophrys catalaunica*/ *O. magniflora*' (co-PI : C. Bertrand)
- 2020** Funding from the Fondation Biotope (2 k€) - REORDER: Phylogénie moléculaire haute-résolution des *Ophrys* de France - PI (co-PI, E. Véla, E., Delannoy & F. Melki)
- 2019** Inter-LabEx initiative grant funded by the LabExs TULIP & CEMEB (8k€) - 'Pilot study of the disjunct geographic distribution of the endemic and threatened *Ophrys aveyronensis*: causes and implications' - PI (co-PI, B. Schatz)
- 2019** 'Bonus Qualité Recherche (BQR) grant funded by UPVD (14k€) - Study of the genomic bases of adaptive radiation of Euro-Mediterranean *Ophrys* orchids'

Bibliometry, supervision and additional expertise

Orcid/ Researcher ID: 0000-0002-3379-1019 / C-1812-2014

Number of peer-reviewed publications: 24

Number of citations: >410 (Source: Google Scholar)

h-index: 13 (Source: Google Scholar)

Level	Number	Period(s)	Rate of supervision
PhD candidate	1	2019-(2022)	50%
Master Students	6	2019, 2020, 2021, 2021, 2022, 2022, (2022 x 4)	100%, 100%, 50%, 50%, 50%, 25%, (33% x 4)

Since 2020 Member of the scientific commission of the French Orchid Society for (SFO) & Member of the Scientific committee of the Federation of the Catalan Natural Reserves
I was entailed as referee for >20 peer-reviewed research articles and proposals to national calls.

Skills, personal interests & others

- Fieldwork** Sampling experience (birds, fishes and plants): morphological measurements, blood sampling, sample processing and collection management
- Labwork** DNA extraction, quantification, amplification and purification, electrophoresis, Sanger and Next Generation Sequencing, microsatellite and SNPs genotyping, flow cytometry
- Data analyses** Population genetics and Phylogenetics analyses, Biostatistics and bioinformatics: programming skills (R, Unix)
- Languages** French (mother tongue), English (professional, TOEIC score: 900), Spanish (fluent), Occitan (fluent)
- Membership** Ligue pour la Protection des Oiseaux (LPO), Fédération France Orchidées (FFO), Groupe Ornithologique du Roussillon (GOR)
- Personal Interests** Photography, Trekking, Snowboarding (Federal Instructor, Grade II), Running and scuba diving
- Others** Driving licence, First aid certificate

Activité de Recherche

De mes débuts en recherche à mon recrutement : émergence et maturation de mon projet de recherche

Thèse : Spéciation chez le *Zosterops borbonicus* : oiseau endémique de l'île de La Réunion (Toulouse, France)

Contexte Dans la mesure où elles sont dépourvues de toute forme de vie lorsqu'elles émergent, les îles océaniques offrent un contexte biogéographique "simplifié" pour étudier comment la biodiversité colonise de nouveaux territoires et s'y diversifie en lignées endémiques. L'équipe dans laquelle j'ai effectué ma thèse de doctorat (sous la direction de Christophe THÉBAUD et Benoit PUJOL) avait commencé à étudier les causes de variations phénotypiques (coloration du plumage et morphologie) observables chez un oiseau endémique de l'île de La Réunion (*Zosterops borbonicus*). Cette espèce présente différents variants phénotypiques dont les aires de

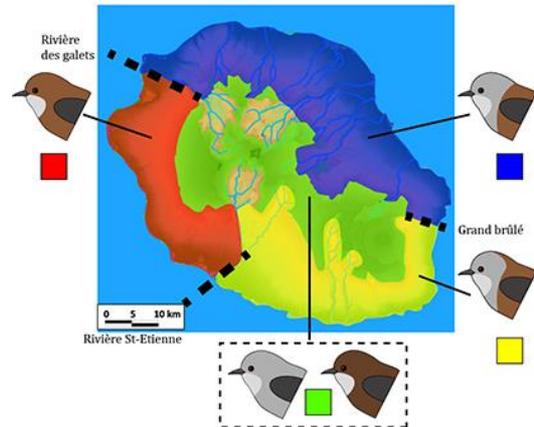


Figure 1 Répartition des variants de *Zosterops borbonicus* sur l'île de La Réunion

distribution géographique apparaissent comme stables dans l'espace et dans le temps (Fig. 1). Différents éléments mis en évidence par des travaux de l'ornithologue Frank GILL (Gill, 1973) ensuite étayés par les premières données génétiques obtenues sur ce projet (Warren *et al.*, 2006; Milá *et al.*, 2010) évoquent un possible cas de spéciation en cours au sein même de l'île. Ce système apparaît donc comme remarquable dans la mesure où cette divergence évolutive s'établirait à une échelle spatiale extrêmement réduite (2512 km²). L'île de La Réunion a été façonnée par une intense activité volcanique et un régime de précipitations torrentiel ayant engendré la formation d'un relief escarpé et une palette d'habitats présentant des conditions écologiques très variées. En s'adaptant à différentes niches écologiques, les populations de *Z. borbonicus* pourraient donc adopter des trajectoires évolutives différentes. Au moment où j'ai intégré ce projet, l'enjeu était de mettre à profit différents outils génétiques afin de décrypter le scénario évolutif sous-jacent.

Objectifs Dans le cadre de ma thèse, je me suis intéressé à caractériser la balance entre plusieurs forces évolutives (sélection, migration, dérive) s'exerçant à différents pas de temps mais toujours à des échelles spatiales très locales au sein de populations naturelles de *Z. borbonicus*. L'objectif était de comprendre en quoi l'environnement, la contingence historique mais aussi les caractéristiques biologiques intrinsèques de l'espèce *Z. borbonicus* ont pu contribuer à façonner la variation observable de nos jours chez ces oiseaux.

Stratégie L'idée première était de confronter des patrons de différenciation génétique (neutre) et phénotypique à ceux de variation environnementale entre populations s'établissant le long d'un gradient environnemental (altitudinal) sur la côte Ouest de l'île où il n'existe pas de barrières physiques évidentes pouvant s'opposer aux échanges entre populations. Pour ce faire, nous avons échantillonné à différentes altitudes, elles-mêmes

correspondant à différentes zones écologiques, plus de 400 individus au niveau de 18 localités. Les oiseaux ont été mesurés (sur 6 traits morphologiques: longueur du tarse, de l'aile, de la queue, du bec; largeur et hauteur du bec) et pesés. Des patchs de plumes ont aussi été prélevés sur différentes parties de leur corps pour quantifier *a posteriori* et de manière objective, la coloration de leur plumage via des analyses spectrales. Les individus ont aussi été génotypés sur 12 marqueurs microsatellites spécifiquement développés pour cette espèce (Bertrand *et al.*, 2012). En parallèle, nous avons obtenu des données relatives à des variables climatiques (température, humidité) et à la végétation (*Normalized Difference Vegetation Index* pour *NDVI*, un indice de « verdeur » de la végétation) sous la forme de couches cartographiques.

Résultats et conclusions Les patrons de différenciation génétique neutre ont permis de confirmer que même dans des environnements écologiquement homogènes, le flux génique entre populations est très faible. Ainsi, les distances de dispersion apparaissent comme extrêmement réduites chez cet oiseau (Bertrand *et al.*, 2014), ce qui est d'autant plus remarquable que les densités de populations sont conséquentes (Bertrand *et al.*, 2016a). Le long du gradient altitudinal, nous avons documenté des différences notables entre les patrons de variation phénotypiques et génétiques (Bertrand *et al.*, 2016b). Alors que plusieurs traits morphologiques semblent contraints par une adaptation aux conditions locales en suivant de manière graduelle la variation environnementale, les données génétiques ont permis de mettre en évidence un cline abrupt ayant son centre à une altitude intermédiaire (autour de 1400 m, Fig. 2). Cet élément, corroboré par de subtiles différences dans la coloration du plumage a révélé la présence d'une zone hybride cryptique entre deux populations (de basse et de haute altitude) ayant été historiquement partiellement isolées l'une de l'autre du point de vue reproductif.

L'exceptionnel degré de différenciation phénotypique et génétique observable chez *Z. borbonicus* serait donc dû à une conjonction très particulière de facteurs. Des différences d'ordre environnemental, très marquées à La Réunion, poussent les phénotypes des individus vers des *optima* changeants d'un endroit à l'autre de l'île et les oiseaux se déplacent de toute façon très peu, ce qui limite l'effet généralement homogénéisant de la migration. A cela s'ajoute le fait que l'île a connu une histoire géologique tourmentée ayant très vraisemblablement affecté à la fois la dynamique démographique et spatiale des populations.

Collaborations J'ai également été associé au volet du projet en lien avec la thèse de mon collègue et ami Yann BOURGEOIS. Ce volet, visait quant à lui à rechercher les bases génétiques du déterminisme de la coloration du plumage afin de comprendre en quoi ce trait pouvait être directement ou indirectement sous sélection (naturelle et/ou sexuelle) et ainsi, d'interpréter son rôle en terme écologique et évolutif. L'idée première était ici de déterminer ce qui varie suivant la coloration en veillant à contrôler les autres facteurs d'ordre géographiques et environnementaux. Nous nous sommes donc focalisés sur les populations d'altitude de *Z. borbonicus* des landes subalpines de l'île qui présentent la particularité de voir des oiseaux bruns et des oiseaux gris vivant en sympatrie. La première étape a consisté à

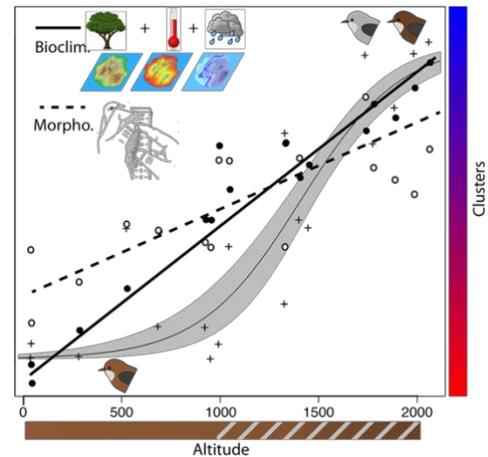


Figure 2 Variation environnementale (trait plein/points noirs), morphologique (pointillés/points blancs) et génétique (sigmoïde, croix) le long du gradient

adopter une approche de type "gènes candidats" afin d'étudier si différents gènes de la voie de bio-synthèse des mélanocortines (e.g. *MC1R*, *Agouti*, *POMC*, *Tyrp 1*) classiquement impliqués dans les variations de coloration mélanique chez les animaux étaient responsables des différences de coloration brun/gris observables chez cette espèce. Aucune des séquences produites n'a présenté de mutation étant associée à ce trait chez *Z. borbonicus* (Bourgeois *et al.*, 2012; 2016). Face à ce résultat "négatif", l'équipe a poursuivi ses efforts en se tournant vers le développement de nouveaux marqueurs afin de travailler à l'échelle génomique (Bourgeois *et al.*, 2013) et a plus récemment identifié un *locus* inédit dans le déterminisme de la coloration mélanique (Bourgeois *et al.*, 2017). Les données génomiques générées ont par la suite été utilisées pour reconstruire l'histoire évolutive de *Z. borbonicus* (Bourgeois *et al.*, 2020) et ont fourni une base solide pour la continuité du projet.

Dans une approche similaire, j'ai activement collaboré avec un autre collègue, Boris DELAHAIE, qui a plus récemment orienté ses travaux de thèse sur l'étude de l'hybridation au niveau des différentes zones de contacts s'établissant entre les trois formes de *Z. borbonicus* de basse altitude (Delahaie *et al.*, 2017).

Post-doc 1 & 2 : Biogéographie et systématique moléculaire du genre *Abudefduf* (Taipei, Taïwan) et causes évolutives de la radiation adaptative des poissons-clowns (Lausanne, Suisse)

Contexte Les sergents-majors (genre *Abudefduf*) comme les poissons-clowns, sont deux représentants des Pomacentridae, une famille de poissons coralliens à la fois très diversifiée et largement distribuée. Ces deux groupes peuvent être vus comme de bons modèles pour étudier par quels mécanismes l'isolement reproductif se met en place, les lignées se diversifient mais aussi maintiennent leur intégrité en cas de remise en contact, et ce, à des échelles trans-océaniques.

Biogéographie & Systématique des Sergents-majors de l'Indo-Pacifique (Pomacentridae: *Abudefduf*)

Objectif et Stratégie Lorsque j'ai commencé mon premier post-doc à l'Institut d'Océanographie de la *National Taiwan University*, les connaissances sur la systématique des sergents-majors demeuraient fragmentaires. Il semblait que le fait d'utiliser des critères phénotypiques pour déterminer la taxonomie de ce groupe posait un certain nombre de problèmes. Sur la base de données génétiques, certains auteurs avaient d'ailleurs suggéré l'existence de lignées cryptiques alors que d'autres venaient de démontrer que certaines espèces d'*Abudefduf* vivant en sympatrie pouvaient produire des hybrides (Quenouille *et al.*, 2004, Coleman *et al.*, 2014). De manière inattendue, certaines espèces se ressemblant fortement pouvaient en fait être assez génétiquement divergentes alors que d'autres, faciles à distinguer d'un point de vue phénotypique n'étaient pas reproductivement isolées les unes des autres.

En se basant sur un échantillonnage géographique et taxonomique étendu, j'ai séquencé près de 500 individus sur différents gènes (mitochondriaux et nucléaires) dont les taux d'évolution respectifs sont censés refléter différents pas de temps évolutifs. Ce jeu de données génétiques devait permettre i) de proposer une nouvelle hypothèse phylogénétique et vérifier la correspondance entre systématique moléculaire et morphologique au sein de ce genre, ii) d'étudier en quoi les barrières biogéographiques pouvaient avoir affecté la diversité

et la différenciation au niveau intra-spécifique et iii) de mettre en évidence et caractériser un hypothétique flux de gènes inter-spécifique.

Résultats S'il existe globalement une bonne correspondance entre phylogénie et systématique morphologique au sein du genre *Abudefduf*, nous avons confirmé qu'une espèce : *A. vaigiensis*, est en réalité polyphylétique et comporte trois lignées qui bien que phénotypiquement très similaires, sont génétiquement assez divergentes alors qu'elles vivent souvent en sympatrie. Au sein de la plupart des espèces, le degré de différenciation génétique est faible hormis chez une autre espèce : *A. sexfasciatus* qui présente une structuration marquée en sous-lignées régionales. De manière intéressante, *A. sexfasciatus* et une des lignées d'*A. vaigiensis* sont indifférenciées sur le *locus* mitochondrial étudié démontrant que les deux espèces présentent un flux de gènes inter-spécifique en dépit de différences phénotypiques marquées (Bertrand *et al.*, 2017, Fig. 3 ; voir aussi Campbell *et al.*, 2018). Ces travaux ont été menés dans l'équipe de Wei-Jen CHEN en collaboration avec Philippe BORSA.

Perspectives Les aspects du modèle biologique *A. sexfasciatus/A. vaigiensis* ayant été révélés au cours de cette étude laissent entrevoir une opportunité prometteuse pour étudier comment les barrières entre espèces peuvent se maintenir malgré la présence de flux de gènes inter-spécifique. Ce phénomène peut même être considéré comme atypique dans la mesure où il a lieu à une échelle spatiale très vaste alors que la plupart des modèles d'étude comparables voient souvent se développer des phénomènes similaires dans les contextes spatiaux beaucoup plus restreints.

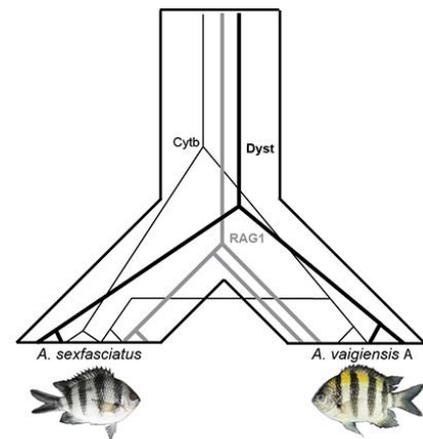


Figure 3 Différents arbres de gènes révèlent l'évolution complexe de deux espèces d'*Abudefduf*

Bases génomiques de la radiation adaptative chez les poissons-clowns (Amphiprioninae)

Objectif et stratégie Trente espèces de poissons-clowns ont été décrites à ce jour et toutes possèdent un ancêtre commun qui aurait vécu il y a moins de 15 millions d'années. Cette diversification à la fois intense et relativement rapide est caractéristique d'une radiation adaptative. Si l'étude de ce type de divergence permet de décrypter les mécanismes impliqués dans leur contexte géographique et historique, les causes de ce phénomène chez les poissons-clowns pourraient en plus avoir été catalysées par différents facteurs: interaction mutualiste avec les anémones de mers (toxiques pour la plupart des autres poissons), hybridation ou encore régime social complexe leur ayant permis d'exploiter des niches écologiques bien spécifiques (Litsios *et al.*, 2012; 2014a, 2014b, 2014c). Dans tous les cas, les bases génomiques de cette radiation adaptative demeurent méconnues.

Mon projet de post-doc visait à découpler les signatures génomiques de l'isolement géographique et de l'adaptation aux conditions écologiques afin d'expliquer la divergence chez trois paires d'espèces sœurs de poissons-clowns qui vivent plutôt dans l'océan Indien

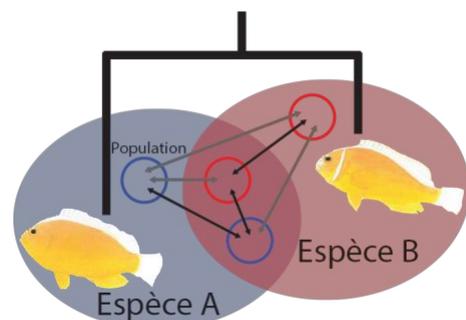


Figure 4 Schéma du design expérimental visant à comparer populations sympatriques et allopatriques de deux espèces sœurs de poissons-clowns

pour l'une, plutôt dans l'océan Pacifique pour l'autre et dont les aires de distribution se chevauchent autour de la barrière Indo-Pacifique (où elles pourraient s'hybrider). Un re-séquençage de génomes complets devait nous permettre de comparer les patrons de diversité et de différenciation génomique entre populations allopatriques et sympatriques afin de mieux appréhender la contribution des facteurs biogéographiques et/ou écologiques à l'œuvre dans ce système (Fig. 4). Ce travail s'intègre à un projet de plus grande ampleur qui vise aussi à caractériser la dynamique de l'interaction poisson-anémone via la quantification de l'expression des gènes (transcriptomique) aussi bien que des composés métaboliques (métabolomique) produits au niveau de la peau et du mucus de l'un et l'autre des deux partenaires. Ces travaux ont été menés dans l'équipe de Nicolas SALAMIN au Département d'Écologie et d'Évolution de l'Université de Lausanne (puis l'équipe a été affiliée au Département de Biologie Computationnelle de la même université). Ma participation active à ce projet a « hélas » été écourtée par mon recrutement à Perpignan. C'est une étudiante en thèse avec qui je collaborais, Anna MARCIONETTI qui a donc pris la suite de ces recherches en parallèle de son projet de thèse.

Projet de recherche principal : lier géotypes, phénotypes et *fitness* pour mieux comprendre les causes de la radiation adaptative des orchidées du genre *Ophrys*

De par mon parcours, le cœur de mon activité de recherche a toujours été principalement orienté vers l'étude des processus à l'œuvre dans la diversification du vivant, et ce, à différentes échelles évolutives, que ce soit en étudiant un phénomène de spéciation naissant chez des oiseaux insulaires (thèse) ou des phénomènes de diversification évolutives à des rangs taxonomiques supérieurs chez des poissons coralliens (post-docs). Malgré un profil plutôt « animaliste », c'est pourtant dans un laboratoire travaillant historiquement sur des modèles exclusivement végétaux que j'ai été recruté. Très vite, les orchidées du genre *Ophrys* (Encadré 1), un genre qui me fascine depuis mon plus jeune âge, m'est apparu comme un modèle de choix pour développer un projet de recherche qui me permettrait de continuer à développer des thématiques qui m'étaient chères tout en respectant les nouvelles contraintes qui s'imposaient à moi, à savoir de mettre un peu de bonne volonté à travailler sur un modèle végétal.

Encadré 1 - Les orchidées du genre *Ophrys*

Au cours des cinq derniers millions d'années, les orchidées du genre *Ophrys* ont connu un phénomène de diversification que l'on pourrait qualifier d'explosif, au point que certaines lignées d'*Ophrys* présenteraient parmi les taux de diversification les plus élevés ayant été documentés à ce jour (Breitkopf *et al.*, 2015). Un consensus scientifique solide s'est peu à peu établi sur le fait que cette diversification serait induite par les pressions de sélection exercées par les insectes pollinisateurs de ces plantes sur les phénotypes floraux. En effet, les *Ophrys* sont connus pour présenter une stratégie de pollinisation bien particulière dans le sens où leurs fleurs font office de leurre sexuel. Concrètement, les fleurs ressemblent à leurs insectes pollinisateurs femelles, ce qui leur permet de duper les mâles et ainsi de leur faire assurer leur pollinisation. En particulier, un des pétales (le labelle) mime la forme, la pilosité et les dimensions de l'insecte femelle. Dans certains groupes, les deux autres pétales sont enroulés sur eux-mêmes, évoquant ainsi des antennes (Fig. 5). Ce mimétisme n'est pas seulement

visuel mais aussi et avant tout olfactif. La fleur émet en effet un bouquet de composés dont la signature chimique s'apparente aux phéromones sexuelles émises par la femelle de pollinisateur réceptive. Le mâle tente littéralement de s'accoupler avec cette « poupée gonflable » que constitue la fleur (c'est la pseudo-copulation) et finit occasionnellement avec du pollen collé au niveau de sa tête ou de son abdomen. En reproduisant l'opération sur une autre plante de la même espèce, il assurera ainsi la pollinisation. Le succès de cette stratégie repose sur un haut degré de spécialisation. Dans un article de synthèse, Joffard *et al.* (2019) ont reporté une moyenne de 1,56 espèce d'insecte pollinisateur connue par espèce d'*Ophrys* (contre 7,44 sur l'ensemble des orchidées euro-méditerranéennes). En l'absence de preuves évidentes de barrières reproductives post-zygotiques au sein de ce genre (voir Cortis *et al.*, 2009 pour une possible exception entre espèces d'*Ophrys* diploïdes : *O. iricolor* et *O. incubacea*), des barrières pré-zygotiques d'ordre éthologique maintenues par l'insecte pollinisateur seraient principalement responsables de l'isolement reproductif entre espèces d'*Ophrys*.

Remarque : Sachant que les *Ophrys* nécessitent d'établir des symbioses mycorrhiziennes pour leur germination, il a pu également être avancé qu'une spécificité à des espèces de champignons bien particulières pouvaient aussi jouer le rôle de barrière à la reproduction au sein de ce système. Au demeurant, cette hypothèse n'a jamais vraiment été étayée du point de vue empirique (e.g. Gervasi *et al.*, 2017).

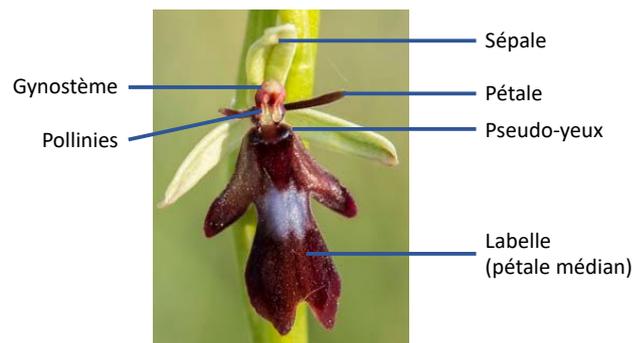


Figure 5 Aperçu d'une fleur mimétique d'*Ophrys insectifera* (Photo J. Bertrand)

J'ai signé ou co-signé, un article de revue (Baguette *et al.*, 2020) ainsi qu'un chapitre d'ouvrage publié en version française et anglaise (Bertrand *et al.*, 2021a; 2021b) consacré au modèle *Ophrys* et ses intérêts en tant que modèle biologique. Ces deux publications font respectivement le point sur les *Ophrys* en tant que candidats pour étudier les radiations adaptatives et sur les défis inhérents à la systématique et à la taxonomie du genre.

Projet d'intégration à mon équipe d'accueil

Obtention de données de génome complet d'*Ophrys* et génomique comparative

Dès 2018, je me suis attelé à répondre à différents appels d'offres pour financer ce projet émergent au sein de mon équipe d'accueil, l'équipe « Analyse du Génome et Évolution », alors animée par Olivier PANAUD. Avec le recul, j'ai très (et sans doute trop) vite voulu en faire un projet d'intégration à mon équipe d'adoption en me lançant sans réserve dans des aspects très axés sur la génomique comparative. Or, il n'y avait alors pas de génome de référence disponible, ni pour le genre *Ophrys*, ni pour aucune Orchidoideae, la sous-famille d'orchidées à laquelle le genre *Ophrys* appartient. L'initiative n'a pas convaincu les différents bailleurs de fonds (ANR et LabEx, notamment). J'ai néanmoins reçu le soutien de l'Université de Perpignan via un financement Bonus Qualité Recherche (BQR) qui m'a entre autres permis d'obtenir un

premier jeu de données génomiques « courtes- » et « longues-lectures » pour trois espèces d'*Ophrys* distinctes, représentatives des trois lignées principales du genre : *Ophrys aveyronensis*, *O. aymoninii* et *O. lutea* (Tab. 1).

Tableau 1 Bilan des données courtes- et longues-lectures générées pour trois espèces d'*Ophrys* : *Ophrys aveyronensis*, *Ophrys aymoninii* et *Ophrys lutea*

Espèce	Données Illumina (« courtes-lectures »)		Données Nanopore (« longues-lectures »)	
	Nombre de lectures	Couverture*	Nombre de nucléotides	Couverture*
<i>O. aveyronensis</i>	319 740 401	8,16 X	9229035263	1,7 X
<i>O. aymoninii</i>	475 551 312	11,1 X	9931337548	1,6 X
<i>O. lutea</i>	610 990 367	17,1 X	5686692253	1,1 X

Les données « courtes-lectures » ont été obtenues via une prestation auprès de Novogene (2x150 pb) mais les données « longues-lectures » ont été générées au sein de l'équipe par notre assistante-ingénieure Christel LLAURO. Ces données Nanopore correspondent à 2 *flowcells* de Minlon (*Oxford Nanopore Technology*) par espèce, soit un total de 6 *flowcells*.

* La couverture a été calculée sur la base des estimations de taille de génome en cytométrie en flux de 5,88 Gpb, 6,41 Gpb et 5,35 Gpb pour *O. aveyronensis*, *O. aymoninii* et *O. lutea* respectivement.

L'équipe possédant une expertise sur les éléments transposables (TEs), j'ai d'abord voulu comparer l'abondance relative des différentes familles de TEs entre espèces d'*Ophrys* pour évaluer de possibles variations structurales au niveau de ces trois génomes. J'ai testé différents outils (e.g REPdevono, *Chu et al.*, 2016 et RepLong, *Guo et al.*, 2018) afin de tenter de reconstruire des éléments répétés sur la base des jeux de données Illumina et Nanopore, respectivement. En parallèle, mes collègues Olivier PANAUD et Moaine EL BAIDOURI ont visé à reconstituer des séquences d'éléments répétés à partir de ces mêmes données brutes. Olivier, ciblait avant tout les rétrotransposons en s'ancrant sur des motifs de transcriptase-inverse dans les lectures alors que Moaine cherchait les éléments répétés en se basant sur une analyse de *k*-mers. Je me suis par ailleurs appuyé sur des distributions de *k*-mers pour estimer la taille du génome de ces trois espèces d'*Ophrys* sur la base des données Illumina (Tab. 2).

Tableau 2 Estimation de la taille du génome sur la base de distribution de *k*-mers (19-mers et 25 mers)

	<i>O. aveyronensis</i>	<i>O. aymoninii</i>	<i>O. lutea</i>
19-mers	6,1	4,5	6,5
25-mers	8,3	7,7	7,0

Si les résultats de ces différentes approches se sont avérés relativement cohérents entre eux, il était illusoire de penser pouvoir être exhaustifs sans génome de référence. Ces travaux ont donc été mis momentanément en *stand-by*. Dans l'attente, cette exploitation préliminaire du jeu de données suggérait déjà i) que le génome d'*Ophrys* est relativement riche en éléments répétés et ii) que les abondances relatives des différents types d'éléments pourraient varier de façon significative d'une espèce à l'autre au sein du genre (un élément corroboré par les différences estimées de taille de génome entre espèces du genre, voire au niveau intra-spécifique). En parallèle, nos résultats montraient que la taille du génome d'*Ophrys* avait probablement été surestimée dans la littérature. En effet, la *Plant DNA C-values database* du *Royal Botanical Garden* de Kew (cvalues.science.kew.org) rapportait jusqu'alors un consensus des estimations obtenues en cytométrie en flux par *Leitch et al.*, (2009) faisant état d'une valeur 1C moyenne de 10,2 pg (9,5 à 10,8 pg sur un échantillon de 4 espèces d'*Ophrys*). La *C-value* correspond à la quantité d'ADN en picogrammes contenue au sein d'un noyau haploïde et peut donc être considéré comme un proxy de la taille du génome,

des radiations adaptatives chez d'autres groupes (voir la notion de *radiating genomes*, Jiggins, 2014; Berner & Salzburger, 2015).

Causes et Conséquence évolutive de la dynamique du génome

N'ayant pas accès au génome de mon modèle biologique principal, j'ai tout de même commencé à m'interroger sur les liens pouvant exister entre dynamique du génome (et plus particulièrement de sa taille) et taux de diversification chez *Ophrys* et les autres orchidées euro-méditerranéennes. La taille de génome varie principalement sous l'effet des phénomènes de polyploïdisation et d'une balance entre accumulation et élimination de séquences répétées (en premier lieu desquelles les rétrotransposons à LTR, pour *Long Terminal Repeats*). Il a ainsi été prouvé que la prolifération d'éléments transposables (éventuellement couplée à un taux d'élimination relativement bas) pouvait expliquer un doublement de la taille du génome (e.g. Piegu *et al.*, 2006; Pellicer *et al.*, 2020). Cela étant dit, les conséquences de telles variations sur l'écologie et l'évolution des organismes demeurent difficiles à appréhender.

Les plantes présentent la gamme de variation la plus importante de tous les organismes vivants en termes de taille de génome (voir Pellicer *et al.*, 2018). A ce jour, le plus gros génome de plante mesuré est celui de l'angiosperme *Paris japonica* (famille des Melanthiaceae), qui avec une taille de 149 milliards de paires de bases (Gb)/1C d'ADN, a un génome environ 2440 fois plus grand que celui de *Genlisea tuberosa* (0,061 Gb/1C), espèce présentant quant à elle le plus petit génome connu chez les plantes. Des génomes de taille très importante ont été documentés principalement chez des angiospermes mais aussi dans une moindre mesure chez des ptéridophytes ou encore chez des gymnospermes ce qui suggère que des génomes de taille imposante ont pu évoluer indépendamment à de multiples occasions dans l'histoire des plantes même si les exemples sont globalement rares. Le mode de la distribution des valeurs de *1C-values* est de 0,587 Gb et la moyenne à 5,020 Gb. On a donc une queue de distribution étalée vers les tailles élevées.

Plusieurs études suggèrent qu'avoir un génome de taille importante serait plutôt un fardeau d'un point de vue écologique et évolutif. Vinogradov (2003) montre que des espèces de plantes à gros génomes tendent à avoir un statut de conservation plus défavorable sur la liste rouge des espèces menacées. Knight *et al.*, (2005) ont trouvé une corrélation négative entre taille de génome et nombre d'espèces chez les angiospermes. Ce patron, ne serait pas si évident chez les animaux même si la corrélation entre taille de génome et richesse spécifique est elle aussi généralement négative (Vinogradov, 2004). Dans un article de revue, Kraaijeveld (2010) mentionne par exemple qu'Olmo (2006) trouve une corrélation inverse entre taille moyenne du génome et nombre d'espèces dans les différentes classes de vertébrés alors que Mank et Avise (2006) mettaient en évidence des corrélations plutôt positives chez différents taxons de poissons.

Parmi les causes possibles résumées par Leitch *et al.*, (2009) ou Pellicer *et al.*, (2018), il apparaît par exemple que les plantes possédant un génome de grande taille ne pourraient s'établir que sur des sols suffisamment riches en azote et en phosphore, deux des constituant majoritaire de la molécule d'ADN. Les espèces possédant un génome de grande taille pourraient aussi être plus sensibles aux radiations, à la pollution par les métaux et auraient des aires de distribution géographique globalement plus réduites que les espèces à génome de taille plus modérée. Les plantes à gros génomes sont par ailleurs souvent des géophytes pérennes. Enfin il existe une corrélation positive entre taille de génome et durée nécessaire pour accomplir le cycle cellulaire ce qui fait que les espèces à gros génomes pourraient avoir

un taux de mutation (et un potentiel évolutif) globalement inférieur à celui des plantes à plus petits génomes (Bromham *et al.*, 2015). Il a même été suggéré que la radiation des angiospermes pourrait avoir été précédée d'une réduction de la taille du génome (Simonin et Roddy, 2018). L'hypothèse qu'avancent ces auteurs serait qu'une réduction de la taille du génome et des cellules aurait permis une optimisation des compartiments cellulaires ayant conféré aux angiospermes un avantage compétitif face aux autres groupes (en premier lieu desquels, les gymnospermes) au moment de leur importante diversification au cours du Crétacé. Si un gros génome ne semble pas être caractéristique des lignées connaissant des taux de diversification importants, il convient néanmoins de bien faire la différence entre taille de génome et taux de variation de tailles de génome.

Hypothèse : un génome considéré comme « dynamique », à savoir présentant des variations de taille (aussi bien des augmentations que des diminutions) conséquentes d'une branche à l'autre dans la phylogénie d'un groupe serait à l'origine de taux de diversification plus élevés.

En des termes plus mécanistiques, on peut postuler que la prolifération d'éléments répétés (en premier lieu desquels les rétrotransposons à LTR, pour *Long Terminal Repeats*) aurait parfois un impact fonctionnel pouvant être à l'origine d'innovations biologiques, favorisant l'apparition de nouvelles lignées évolutives. Le rôle fonctionnel joué par des éléments transposables dans l'apparition de nouveaux phénotypes est documenté de manière croissante dans littérature (e.g. Lisch, 2013; Schrader & Schmitz, 2018). D'autre part si la taille du génome en soi est connue pour être un trait relativement mal conservé dans la phylogénie, une étude empirique relativement récente avait suggéré qu'il existe une corrélation entre taux de variation de taille du génome (et non pas simplement taille du génome ou *C-value*) et taux de spéciation à l'échelle des angiospermes et des autres plantes terrestres (Puttick *et al.*, 2015 voir aussi Ricci *et al.*, 2016). C'est cette hypothèse de travail bien particulière que j'ai voulu tester à l'échelle des orchidées euro-méditerranéennes. La famille des orchidées est en effet connue aussi bien pour être une des, si ce n'est la famille de plantes la plus riche en espèces et elle présente des variations de taille de génomes très importantes entre espèces (*i.e.* de 0,33 pg à 55,4 pg, soit un facteur de 168 entre le plus petit et le plus grand génome d'orchidées connu, Leitch *et al.*, 2009).

Pour être testée, cette hypothèse nécessitait :

- i) une phylogénie (la plus exhaustive possible) datée et
- ii) une estimation de la *C-value* pour les différentes espèces.

Pour ces deux points, des bases de données existent, à savoir GenBank et la *Plant DNA C-value database* (hébergées par le *National Institutes of Health* (NIH) et les *Royal Botanic Gardens* de Kew) mais l'une comme l'autre sont loin d'être exhaustives. Pour la phylogénie, j'ai donc entrepris de générer ma propre hypothèse phylogénétique des orchidées euro-méditerranéennes. Je savais bien sûr d'avance que cela allait être un travail de longue haleine. Il fallait obtenir les échantillons, extraire l'ADN et effectuer le séquençage avant même de pouvoir reconstruire et dater l'arbre phylogénétique à proprement parler. C'est pour ça que j'ai opté sur une stratégie de séquençage Sanger « traditionnelle » suivant laquelle nous allions pouvoir amplifier et séquencer les marqueurs génétiques choisis (à savoir les incontournables *nrITS*, *matK* et *rbcL* plus un marqueur nucléaire supplémentaire, *Xdh*, voir Jin, 2017). Pour l'obtention des échantillons, je pouvais compter sur le soutien d'un réseau de bénévoles experts et passionnés à la tête duquel se trouvait une collègue géologue du CNRS récemment retraitée, présidente de l'association orchidophile locale, Roselyne BUSCAIL. J'ai ainsi pu obtenir en l'espace de quatre ans, les tissus de bien plus d'espèces d'orchidées que je

n'aurais pu espérer en collecter seul en une décennie... En 2019, je recevais également le soutien de Jean-Loup ZITOUN, un étudiant de Master 2 très motivé qui fera faire un grand bond en avant au projet. Il était prévu de postuler à un financement de bourse de thèse région pour que Jean-Loup poursuive l'aventure avec nous. Or, notre demande n'ayant pas été financée, Jean-Loup a poursuivi sa carrière sur un autre projet et ces travaux se poursuivent au ralenti malgré le soutien de notre assistante-ingénieure Christel LLAURO pour le séquençage Sanger. Cette initiative a par ailleurs été perturbée par un détail logistique. En effet, l'appareil de cytométrie en flux du laboratoire sur lequel nous avons commencé à obtenir des *C-values* (avec ma collègue Laëticia BARIAT) a dû être cédé en raison d'une utilisation jugée insuffisante pour pouvoir assurer son entretien. Je me retrouvais donc à devoir envisager une autre option (payante). C'est auprès de la plateforme voisine BioPIC de Banyuls avec David PECQUEUR en tant qu'interlocuteur que je me rends occasionnellement depuis pour poursuivre ce travail.

Un des principaux résultats du travail de Master de Jean-Loup est qu'il existerait effectivement une relation positive entre taux diversification (*i.e.* taux de spéciation – taux d'extinction) inféré sur la base de la phylogénie datée et taux de variation de taille de génome (mais pas entre taux de diversification et taille du génome). Ainsi les *Ophrys* qui présentent parmi les taux de diversification les plus élevés chez les orchidées euro-méditerranéennes présenteraient aussi les taux de variation de taille de génome les plus importants (Fig. 7). Cette corrélation, qui demande certes à être étayée avant publication suggère néanmoins un possible lien entre dynamique du génome et taux de diversification. Dans la version présentée, un alignement des 4 *loci* soit un total de 4058 sites a été aligné à partir de 76 taxons.

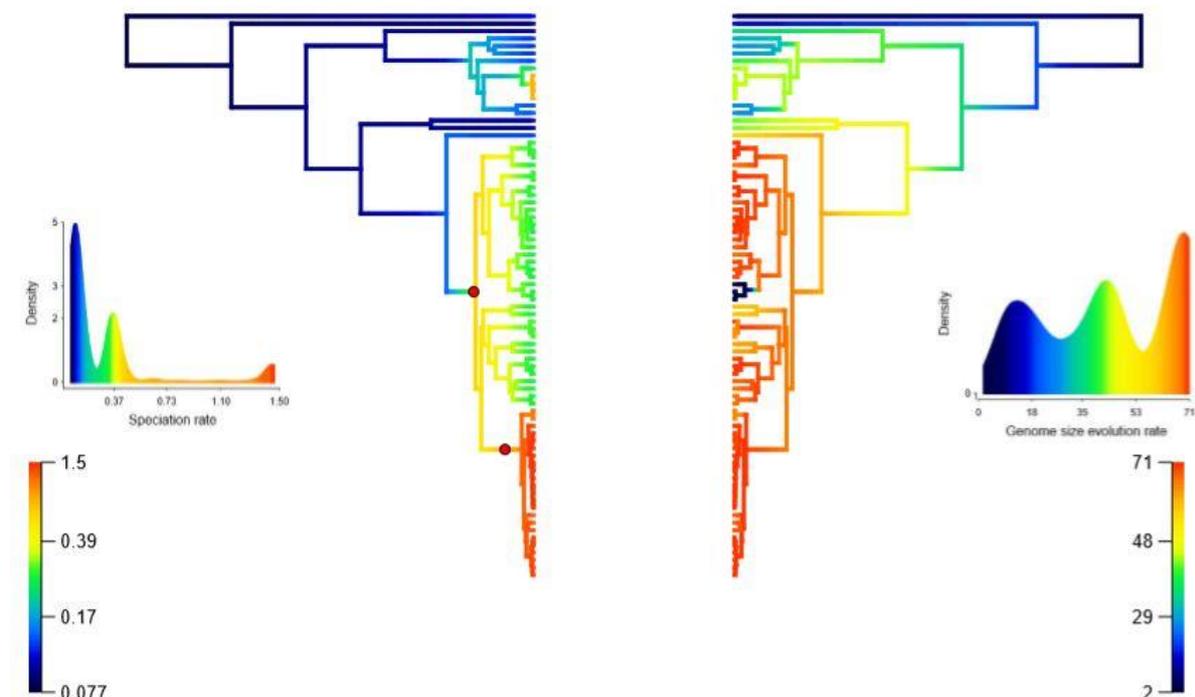


Figure 7 Taux net de diversification ou taux de spéciation (à gauche) et taux de variation de taille du génome (à droite) apparaissent comme positivement corrélés. Les *Ophrys* correspondent au clade situé en bas des deux arbres sur la figure. Les analyses ont été effectuées avec le package Bayesian Analysis of Macroevolutionary Mixture ou BAMM (Rabosky *et al.*, 2014)

Dans l'attente, les données génétiques générées au cours de ce travail pourront trouver leur utilité, ne serait-ce qu'auprès de la communauté des taxonomistes et des

systematiciens. L'hypothèse phylogénétique construite est ainsi la plus exhaustive proposée à ce jour pour les orchidées euro-méditerranéennes. Le fait de rejoindre la commission scientifique de la Fédération France Orchidées (en 2020) ne fera d'ailleurs que me conforter sur ce point.

Étude pilote n°1 : Biogéographie du complexe d'espèces *Ophrys aveyronensis*

Contexte En 2019, j'ai commencé à échanger avec un collègue chercheur montpelliérain, Bertrand SCHATZ. Bertrand travaillait déjà sur le modèle *Ophrys* depuis de nombreuses années et sa recherche me paraissait assez complémentaire de celle que je voulais développer au niveau des approches. En effet, il étudiait le système en se plaçant davantage du point de



Figure 8 Détail d'une fleur d'*Ophrys aveyronensis*, Larzac, Aveyron (Photo J. Bertrand)

vue de l'insecte pollinisateur que je ne l'envisageais et il avait notamment développé une expertise sur l'échantillonnage et l'analyse des odeurs émises par les fleurs. Nous avons donc entrepris de collaborer sur un projet qui nous permettrait de mettre en place une approche intégrative se basant à la fois sur l'analyse des données phénotypiques (morphométrie, photographies, odeurs...) et génétiques. Pour ce faire, nous avons proposé une étude biogéographique sur le complexe *Ophrys aveyronensis* qui a été soutenue par une initiative inter-LabEx (TULIP/CEMEB). La littérature suggérant des degrés de différenciation relativement faibles entre espèces d'*Ophrys*, nous avons opté pour une approche

de génomique des populations dans le but d'obtenir une résolution suffisante pour caractériser la diversité et la différenciation entre populations, de même que pour comparer phénotypes et génotypes. Contrairement à ce que j'avais commencé à aborder dans le cadre de mon post-doc précédent sur les poissons-clowns, l'idée de travailler sur des génomes complets était inenvisageable sur des dizaines d'individus d'*Ophrys*, en raison de la taille de leur génome. Néanmoins, Sedeeck *et al.*, (2014), avaient déjà prouvé le côté prometteur des approches de Génotypage par Séquençage (GBS) chez d'autres *Ophrys*.

Les aires de distributions des deux vicariants du complexe *Ophrys aveyronensis* sont séparées par une distance de plus de 500 km (et le massif des Pyrénées, Fig. 8 et Fig. 9). En bref, le taxon *Ophrys aveyronensis* (ou *Ophrys sphegodes* subsp. *aveyronensis*, Wood, 1983 ; Delforge, 1984) a d'abord été décrit comme endémique de la région des Grands Causses avant que des populations en apparence très similaires soient observées quelques années plus tard dans le Nord de la péninsule ibérique (e.g. Hermosilla & Soca, 1999). Ces deux taxons ont depuis été considérés tantôt comme des espèces (*Ophrys aveyronensis* et *O. vitorica*, Kreutz, 2007), tantôt comme des sous-espèces (*O. aveyronensis* subsp. *aveyronensis* et *O. aveyronensis* subsp. *vitorica*, Paulus, 2017). Sans vouloir épiloguer dans ce débat que nous laissons volontiers aux taxonomistes, il était intéressant d'étudier les causes (et les conséquences) évolutives de ce patron d'aire de distribution géographique disjointe.

L'idée première était donc d'utiliser à la fois des données morphologiques, génétiques, d'odeur et écologiques pour confronter trois hypothèses :

i) L'aire de distribution disjointe du complexe *O. aveyronensis* correspondrait à deux poches relictuelles d'une aire ancestrale s'étant fragmentée et réduite. Si tel était le cas, on devait au moins s'attendre à trouver un degré de différenciation génétique plus important entre populations de part et d'autre des Pyrénées qu'au sein des deux sous-aires.

ii) L'aire de distribution disjointe du complexe ne serait au final que le résultat d'une introduction biologique récente d'une des parties de l'aire vers l'autre. Si tel était le cas on pourrait s'attendre à ce que la diversité génétique dans l'une des deux parties de l'aire s'apparente à un sous-échantillon de ce qu'elle était dans l'autre partie.

iii) Cette configuration correspondrait en fait à des aires de deux taxons ne formant pas nécessairement un clade dans la phylogénie, mais qui auraient pu converger vers des phénotypes similaires en s'adaptant à un pollinisateur commun (comme supposé et récemment soutenu par Benito Ayuso *et al.*, 2019).

En juin 2019, nous avons échantillonné entre 14 et 22 individus au niveau de 6 localités : 3 en France pour *O. a. subsp. aveyronensis* ($n=52$) et 3 en Espagne pour *O. a. subsp. vitorica* ($n=57$). Des mesures morphologiques renseignant sur différents traits phénotypiques de la fleur, ou de la plante dans son ensemble ont été relevés. Nous avons aussi effectué des photographies calibrées de chaque fleur. Des prélèvements de bractées ont aussi été effectués en vue de l'extraction d'ADN et du génotypage sur un total de 86 individus (43 *O. a. subsp. aveyronensis* et 43 *O. a. subsp. vitorica*). Sur un sous-échantillon de 24 individus, les odeurs des labelles ont aussi été échantillonnées grâce à des fibres SPME.

A ce jour, nous n'avons pas encore soumis le manuscrit correspondant à ce travail car certaines analyses restent encore à améliorer ou à faire mais nous livrons ci-après une version en cours de la plupart des résultats préliminaires.

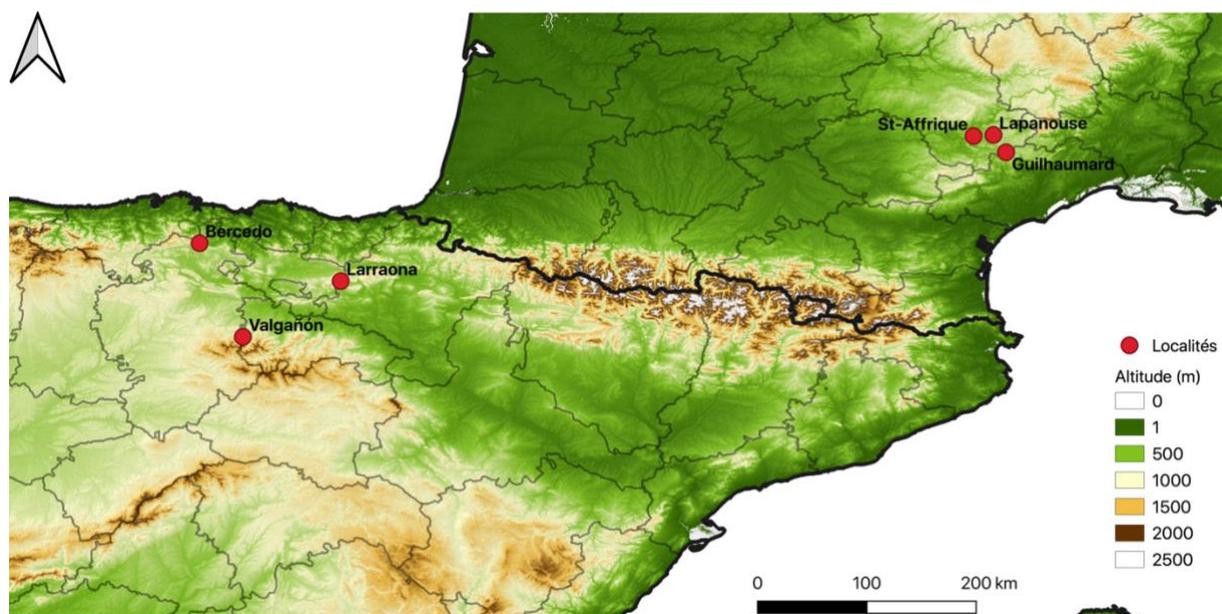


Figure 9 Échantillonnage géographique des 6 localités du complexe *Ophrys aveyronensis*

Résultats préliminaires L'analyse des données morphométriques confirme que les populations des Grands Causses et de la péninsule ibérique sont très peu différenciées sur le plan morphologique (Annexe 1). Pourtant, il est à noter que plusieurs variables : distance de la première à la deuxième fleur sur l'inflorescence, nombre de fleurs/bourgeons, diamètre de la tige, largeur du labelle et largeur du sépale médian présentent des différences significatives (Fig. 10, Annexe 1). Cela suggère que les individus d'*Ophrys a. subsp. aveyronensis* sont en moyenne plus robustes et présentent de légères différences par rapport à leurs homologues ibériques au niveau de la morphologie florale. Par ailleurs, un travail visant à obtenir des informations quantitatives à partir des photos de fleurs avait été initié par Florian LOUTY, un

étudiant de M1. Florian a mis en évidence de subtiles différences au niveau de la coloration et de l'agencement des motifs présent au niveau du labelle (macule). Ces données ont depuis été reprises par Anaïs GIBERT (Ingénieure de Recherche sur le projet ANR JCJC *DiversiFly*, voir ci-après) dont les résultats ont permis d'établir que des modèles de type *random forests* alimentés par ces données ne se trompaient que relativement peu pour différencier les deux taxons, chose pour laquelle même l'œil d'un expert orchidophile commet assez souvent des erreurs!

Remarque : une analyse préliminaire des odeurs laisse entrevoir de possibles différences dans l'identité ou du moins l'abondance relative de certains composés olfactifs. Si cela se confirme, ce résultat pourrait suggérer que contrairement à ce qui était connu jusqu'alors, les deux taxons pourraient avoir deux espèces de pollinisateurs différents et seraient potentiellement engagées dans un processus de spéciation écologique.

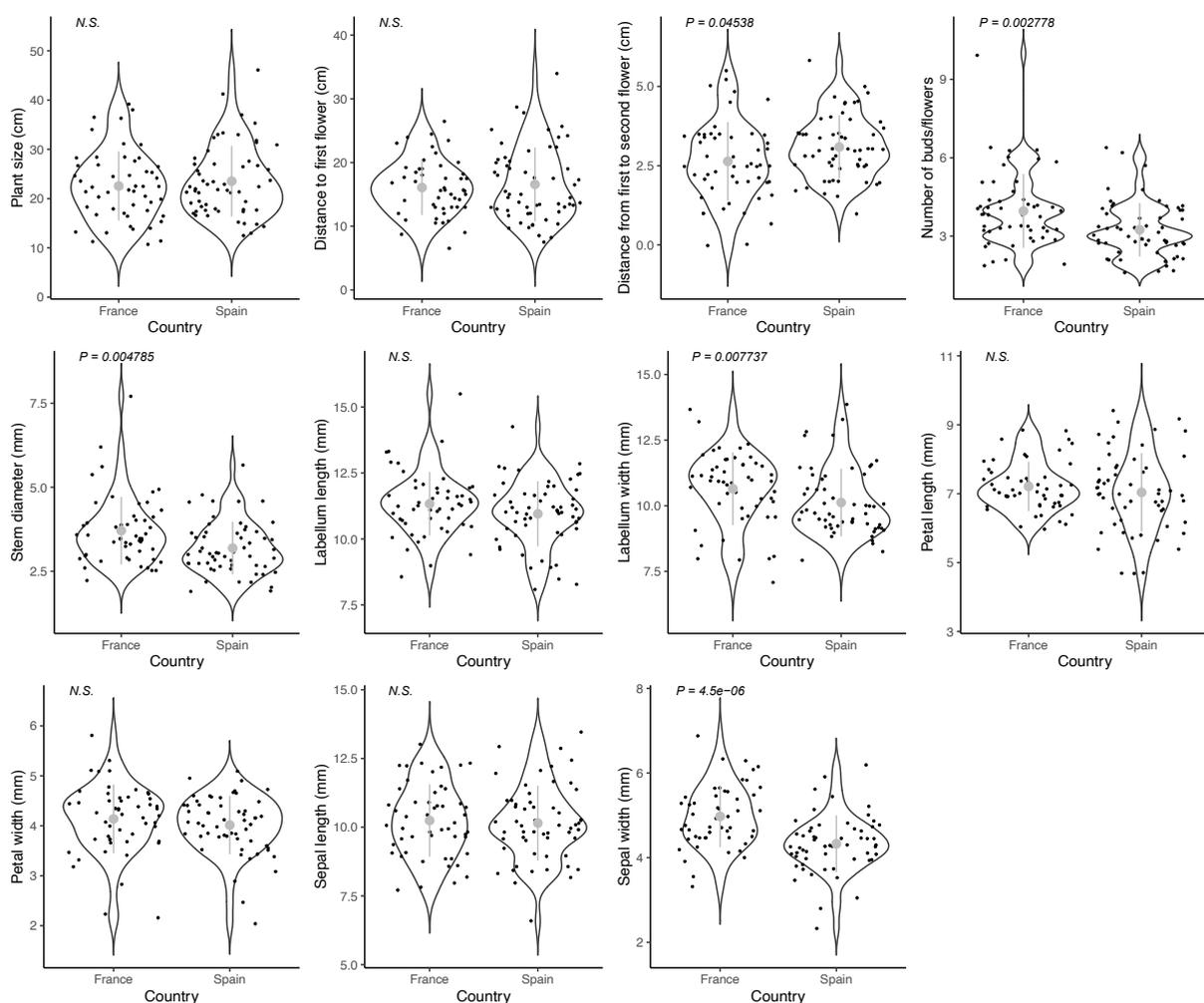


Figure 10 Distribution des valeurs de 11 variables morphométriques et comparaison entre les deux taxons du complexe *O. aveyronensis* (les *p*-values sont indiquées directement sur chaque graphe)

La reconstruction de *stacks* à partir du programme éponyme (Catchen *et al.*, 2011 ; 2013), et ce, en mode *de novo* (c'est-à-dire sans aligner les lectures sur un génome de référence) a permis d'obtenir une matrice de 9231 marqueurs (SNPs) pour étudier la diversité et la différenciation génétique au niveau de ces 6 populations du complexe *O. aveyronensis*. En bref, les populations sont apparues comme relativement structurées (Fig. 11 et 12). Sur

l'ACP, on voit clairement se distinguer les populations d'*O. a. subsp. aveyronensis* (France) et *O. a. subsp. vitorica* suivant l'axe 1 (qui explique 8,53% de la variance). Suivant l'axe 2 (qui explique 3,87% de la variance), on distingue les différentes localités au sein des deux taxons. Ces résultats sont corroborés par le calcul des F_{ST} en montrant que la différenciation entre les deux pays ($0,112 < F_{ST} < 0,148$, $P < 0,01$) est de l'ordre de deux fois plus élevée qu'entre localités au sein des deux zones ($0,048 < F_{ST} < 0,082$, $P < 0,01$). Enfin une analyse d'assignation réalisée avec sNMF (Frichot *et al.*, 2014) confirme que le nombre de partitions le plus adéquat serait $K=5$, groupant ainsi les populations larzacienne du Guilhaumard et de Lapanouse-de-Cernon (Fig. 10). En termes de diversité génétique, toutes les localités présentent un déficit modéré en hétérozygotes (F_{IS} significativement positifs) mais on ne note pas d'écart notable entre populations.



Figure 11 Analyse en Composante Principale (ACP) réalisées sur 86 individus d'*Ophrys aveyronensis* génotypés sur 9231 SNPs. Ici, PC1 (expliquant 8,53% de la variance) et PC2 (expliquant 3,87% de la variance) sont représentés

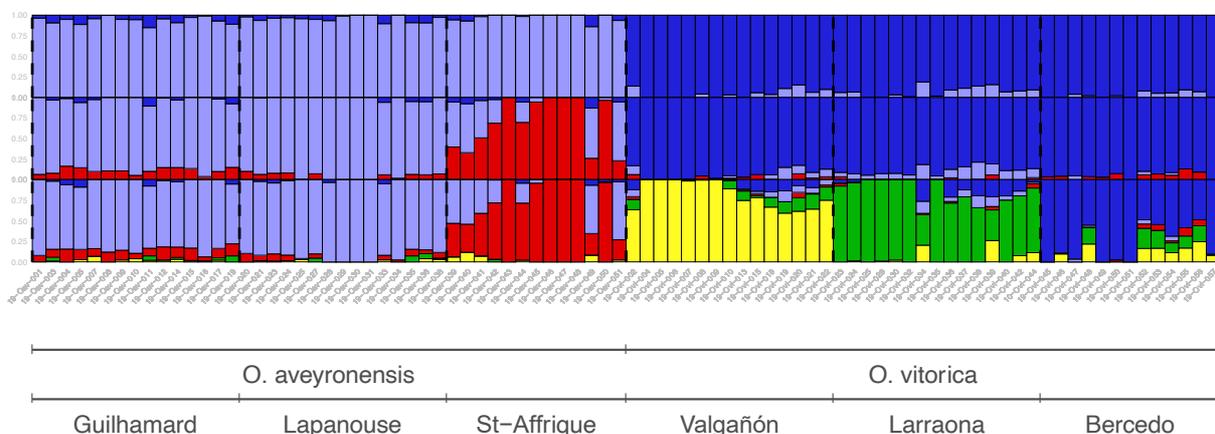


Figure 12 Analyse d'assignation réalisée (avec sNMF) sur 86 individus d'*Ophrys aveyronensis* génotypés sur 9231 SNPs. Les proportions d'assignation aux différents clusters sont détaillées de haut en bas pour $K=2$ à $K=5$

Dans l'attente de la réalisation d'études plus poussées en matière d'inférence et de confrontation de différents scénarios démographiques, ces résultats soutiennent plutôt la première hypothèse suivant laquelle l'aire de distribution disjointe du complexe *O. aveyronensis* correspondrait à deux poches relictuelles d'une aire ancestrale s'étant fragmentée et réduite. Il à noter que ce scénario serait aussi appuyé par une tentative préliminaire de modélisation de niche bioclimatique actuelle et passée que j'ai commencé à réaliser (Fig. 13). Sur la première carte, on voit la projection spatiale de la niche bioclimatique du complexe *O. aveyronensis* telle que modélisée avec le logiciel Maxent (Philips *et al.*, 2017) à partir des données d'occurrences de la base de données gbif (www.gbif.org). J'ai ensuite projeté spatialement cette niche bioclimatique actuelle dans le contexte paléoclimatique inféré pour le dernier maximum glaciaire (ou LGM pour *Last Glacial Maximum*) il y a environ 18 000 ans (Fig. 12b). Sur cette figure, on voit apparaître une aire géographique continue, centrée sur l'Est des Pyrénées. On peut supposer que l'aire ancestrale du complexe *O. aveyronensis* s'est non-seulement scindée en deux mais que les deux sous-aires se sont ensuite déplacées en réponse au réchauffement climatique au cours des millénaires, en longeant le piémont Sud des Pyrénées côté Espagnol et en remontant vers le Nord, jusqu'aux Grands Causses, côtés français. Ce scénario, inféré uniquement sur la base de données climatiques serait au demeurant plausible en considérant les exigences de l'espèce en matière de nature des sols. On trouve en effet des zones calcaires tout le long de ces deux possibles routes de migration.

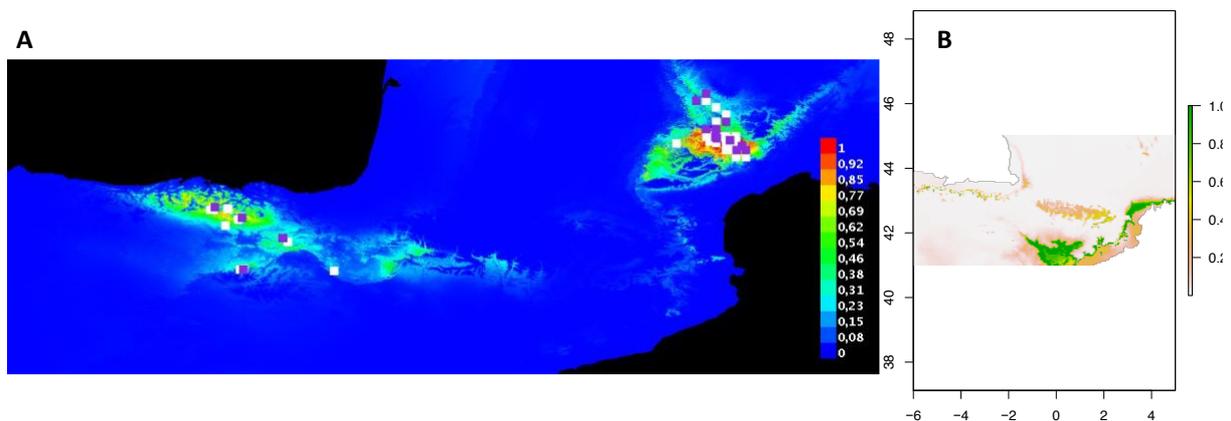


Figure 13 Projection spatiale de la niche bioclimatique d'*Ophrys aveyronensis* telle que modélisée par Maxent A) sur les données actuelles (bioclim) et B) sur les données inférées pour le dernier maximum glaciaire (LGM)

Conclusions & perspectives Bien que non-encore finalisée à ce jour, cette étude m'a rassuré en confirmant l'utilité des protocoles de type *RAD-seq* pour étudier des populations de taxons très faiblement divergents les uns des autres. D'ailleurs, si ce protocole convient pour la réalisation d'une étude de génomique des populations « classique », il a aussi permis d'identifier des *outliers* fortement associés avec la structuration génétique des populations ainsi que des SNPs significativement associés à certains traits phénotypiques. Par exemple, deux méthodes différentes de détection d'*outliers*, l'une basée sur l'analyse sNMF (package LEA, Frichot & François, 2015), l'autre reposant sur une ACP (*pcadat*, Luu *et al.*, 2016), ont permis de mettre en évidence 5 et 7 SNPs significativement associés à la structuration génétique ($K=2$ et PC1, respectivement). De manière intéressante, ces SNPs sont situés sur un même locus nGBS mais un *blast* de ce locus n'a pas permis de mettre en relation cette portion d'ADN d'*Ophrys aveyronensis* avec une séquence connue et annotée (Fig. 14). Compte-tenu du fait que nous pourrons bientôt compter sur un génome de référence de qualité d'une

espèce très proche sur lequel nous pourrions nous aligner, cela ouvre des perspectives prometteuses dans la recherche de régions génomiques candidates dans le déterminisme de l'adaptation et/ou de l'isolement reproductif chez ces orchidées.

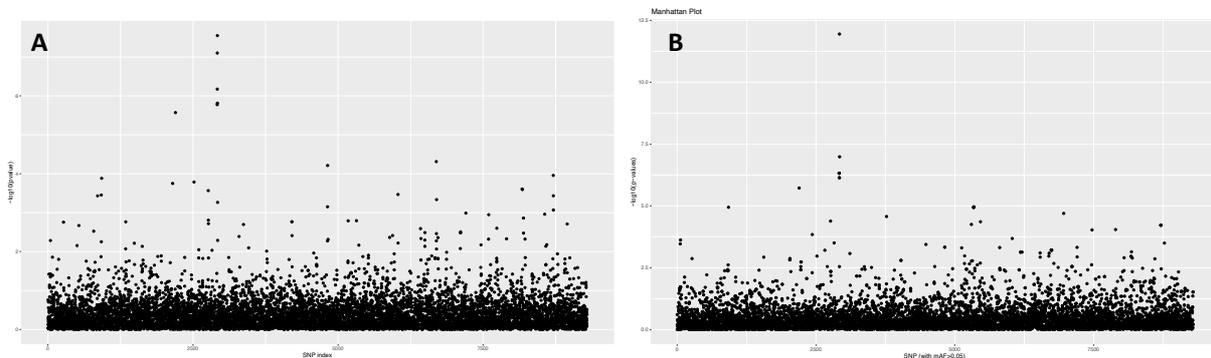


Figure 14 Détection d'outliers à partir A) de l'analyse sNMF et B) d'une analyse *pcadapt* sur 86 individus du complexe *Ophrys aveyronensis* génotypés sur 9231 SNPs

Étude pilote n°2 : Taxonomie intégrative du complexe *Ophrys catalaunica*/*O. magniflora*

Contexte Fin 2020, un nouveau projet pilote soutenu par la Fédération de Recherche locale (la Fédération de Recherche « Énergie Environnement » (FREE FR 2043) allait me permettre de continuer à expérimenter d'autres approches intégratives sur le modèle *Ophrys*. Premièrement, si l'utilisation des approches *RAD-seq* avait donné entière satisfaction lors de l'étude sur le complexe *O. aveyronensis*, d'autres méthodes permettant de sous-échantillonner le complexe génome d'*Ophrys* étaient envisageables. C'est notamment le cas du séquençage de transcriptomes qui permet de cibler les séquences codantes des gènes exprimés (et possédant potentiellement un impact fonctionnel) dans certains organes d'intérêt tels que la fleur. Deuxièmement, des discussions avec Cédric BERTRAND, un collègue chimiste de l'UPVD m'avaient encouragé à essayer de mettre en solution les composés olfactifs des fleurs d'*Ophrys* dans de l'hexane plutôt que d'utiliser des fibres SPME (comme lors de notre étude pilote sur le complexe *O. aveyronensis*). Le complexe d'espèces *Ophrys catalaunica*/*O. magniflora* permettait de faire d'une pierre de coup en proposant de répondre à une problématique biologique pertinente tout en testant les deux alternatives méthodologiques mentionnées ci-dessus.

Le statut taxonomique (espèces distinctes, sous-espèces, voire variétés) d'*Ophrys catalaunica* et d'*O. magniflora* demeure incertain mais ces deux taxons présenteraient des aires de distribution géographique se chevauchant au niveau des départements des Pyrénées Orientales et de l'Aude. Ainsi, *O. catalaunica* serait plutôt endémique du piémont Sud des Pyrénées catalanes alors qu'*O. magniflora* aurait plutôt une distribution centrée sur l'Aude avec un hiatus au niveau de la plaine roussillonnaise. Dans l'arrière-pays, leur ressemblance morphologique rend les individus de la poignée de stations connues délicates à caractériser. La littérature ne reposant que sur des données morphologiques (et le pollinisateur d'*Ophrys magniflora* étant à ce jour inconnu) il nous est apparu intéressant de relever le défi en sachant que ces deux espèces (micro-)endémiques et donc, *de facto*, emblématiques, présentent un degré de menace croissant déjà palpable (*i.e.* destruction des habitats et changement climatique).

Nous avons prévu d'échantillonner au moins deux localités pour chaque taxon suivant un transect transfrontalier orienté Nord/Sud (Fig. 15), mais au printemps 2021, la collecte s'est avérée plus ardue que prévue. Le deuxième confinement et les restrictions de

déplacement en raison du COVID-19 nous ont empêché de nous rendre sur les localités espagnoles prévues pour *O. catalaunica* pour des raisons sanitaires et la floraison a de toute façon été catastrophique pour *O. magniflora* (avec seulement 4 individus trouvés sur une seule localité) pour des raisons météorologiques. La suite du projet s'est tout de même poursuivie même si son aboutissement nécessitera un retour sur le terrain en 2022. En mai 2022, nous avons pu échantillonner une autre localité d'*O. catalaunica* côté ibérique mais le télescopage de ces missions de terrain avec celles de l'ANR qui débute fait qu'il faudra terminer la collecte en 2023, voire plus tard...



Figure 15 Transect initialement proposé pour l'échantillonnage des localités d'*Ophrys catalaunica* (en bleu) et d'*Ophrys magniflora* (en rouge)

Résultats préliminaires Si la robustesse des résultats nécessite d'augmenter l'échantillonnage, le projet reste néanmoins un succès sur le plan technique. Le stockage des fleurs dans une solution de *RNAlater* a permis d'obtenir un ARN en quantité et de qualité suffisante pour séquencer les transcriptomes floraux de 8 individus (4 *O. magniflora* et 4 *O. catalaunica*) et l'immersion de fleurs dans une solution d'hexane s'est avéré être une option pleinement exploitable et qui plus est, économique pour échantillonner les odeurs de la fleur de 20 individus en tout (4 *O. magniflora* et 16 *O. catalaunica*). A l'heure où j'écris ces lignes, les analyses génétiques et d'odeur sont en cours. Une première ACP sur la matrice d'expression des transcrits suggère une différenciation perceptible entre les deux localités/espèces. Pour les analyses d'odeur, je collabore avec mes collègues Cédric BERTRAND et Delphine RAVIGLIONE, l'Ingénieure d'Étude qui est la responsable technique du plateau Métabolites Secondaires Xénobiotiques et Métabolomique Environnementale, MSXM) de la plateforme Bio2mar. Nous avons par ailleurs été appuyé par trois étudiants chimistes : Dominique NGUYEN (étudiante de L3 montpelliéraine en stage hors-cursus) puis Marc RAMOND et Tristan FUENTES (étudiants perpignanais de M1 dans le cadre de leur projet tutoré). Tristan a prolongé l'expérience dans le cadre de son stage de M1 au printemps 2022.

Bilan intermédiaire des deux études pilotes

Je considère ces deux études pilotes comme des réussites dans le sens où elles ont montré qu'on pouvait avec succès utiliser des données de génomique des populations de type *RAD-seq* pour caractériser finement les patrons de diversité et de différenciation génétique entre

populations d'*Ophrys* et même pour mettre en évidence des régions génomiques candidates à l'adaptation et/ou à l'isolement reproductif au sein de ce système. Fort de ces résultats satisfaisants, c'est cette approche qui sera par la suite employée dans d'autres projets dans lesquels j'ai pu m'investir et sur d'autres modèles biologiques (e.g. le Châtaignier (*Castanea sativa*), la Dauphinelle des montagnes (*Delphinium montanum*) ou la Pédiculaire chevelue (*Pedicularis comosa*)). J'ai ainsi pu développer une expertise technique qui a été appliquée sur plusieurs modèles et pour laquelle je possède désormais un certain recul. Ces études m'ont aussi rassuré sur ma capacité à mener un projet sur un modèle plante, moi qui n'avais jusqu'alors travaillé que sur des populations naturelles (et de l'ADN) d'animaux. Couplé à l'exploration d'autres types de données génétiques et génomiques mais aussi phénotypiques, elles m'ont permis de me lancer dans la rédaction d'un projet ANR Jeune Chercheur Jeune Chercheuse en 2020-2021.

Tableau 3 Bilan des jeux de données de génomique des populations générés ou devant être générés prochainement avec une approche nGBS (protocole de type *RAD-seq*)

Espèce	Nombre d'échantillons	Type de lecture	Année
<i>Ophrys aveyronensis</i> (complexe)	86 (+10)	2 x 150 pb	2019
<i>Castanea sativa</i>	120	1 x 75 pb	2019
<i>Delphinium montanum</i>	106	2 x 250 pb	2021
<i>Pedicularis comosa</i>	158	2 x 250 pb	2022
<i>Ophrys</i> groupe <i>insectifera</i>	196 prévus	2 x 250 (?) pb	2022

Le projet « *DiversiFly* »

Ce projet de recherche, intitulé « *Adaptive Diversification of Fly Orchids as a new model to investigate ecological speciation and evolutionary radiations* » (***DiversiFly***) fait l'objet d'un financement de l'Agence Nationale de la Recherche (ANR) via son instrument Jeune Chercheur Jeune Chercheuse (JCJC) pour une durée de quatre ans (sur la période 2022-2025) et constituera donc le cœur de mon activité de recherche actuel et dans un avenir proche.

Objectif général Dans le projet *DiversiFly*, je propose de tirer parti des caractéristiques écologiques et évolutives fascinantes du système *Ophrys* pour contribuer à une meilleure compréhension de la façon dont les populations divergent les unes et autres durant le processus de spéciation écologique, en particulier dans les phénomènes de radiations adaptatives où ce processus semble connaître une sorte d'« emballement ». L'idée consiste à profiter de la relative démocratisation et des dernières avancées des méthodes dites « -omiques » pour développer des approches intégratives permettant désormais d'étudier les causes de la variation phénotypiques, jusqu'à ses bases moléculaires, y compris chez des organismes « non-modèles ». Mais un autre intérêt majeur que je vois dans le système *Ophrys*, c'est cette possibilité d'intégrer ces approches intégratives dans contexte évolutif élargi visant à établir les relations entre génotypes, phénotypes et *fitness*, dans un contexte réaliste (comme explicité dans Byers *et al.*, 2017 et la *Special Issue* de *Molecular Ecology* dans laquelle a été rédigé cet article).

Une question biologique qui m'anime est notamment de comprendre :

Comment une fleur peut s'adapter au point d'imiter son insecte pollinisateur?

État de l'art & objectifs L'architecture génétique qui sous-tend les radiations évolutives pourrait être relativement simple (« génique plutôt que génomique »). En effet des travaux théoriques prédisent qu'un ou quelques *loci* tout au plus peuvent être majoritairement impliqués dans le déterminisme de l'innovation biologique clé étant à l'origine du processus de diversification rapide et intense (voir Orr, 2005; Gavrillets & Losos, 2009 mais voir aussi Dufresnes *et al.*, 2021). Bien que la réalité puisse être plus complexe, ces prédictions ont été largement soutenues par les résultats d'études empiriques (voir Nosil *et al.*, 2017). Chez les angiospermes par exemple, l'architecture de la fleur est principalement déterminée par un nombre restreint de gènes assez bien connus (les gènes MADS box, voir Wang, 2019). Sur cet organe, plusieurs traits sont ainsi susceptibles de se comporter comme des « traits magiques » (« *magic traits* ») i) en faisant à la fois l'objet d'une sélection divergente (induite par les pollinisateurs) et ii) en étant directement impliqués dans l'isolement reproductif (voir Schlüter, 2018). L'étude des traits phénotypiques floraux et leurs bases génétiques apparaît donc comme d'importance capitale pour comprendre l'adaptation et la spéciation chez les plantes. Cependant, une étape critique consiste à évaluer précisément l'impact que la variation au niveau de ces « traits magiques » induit sur la *fitness* individuelle dans les écosystèmes naturels pour appréhender la pertinence de cette variation dans un contexte éco-évolutif réaliste. Parmi ces systèmes, les **zones hybrides** ont été qualifiées de « laboratoires naturels pour études évolutives » ou de « fenêtres sur les processus évolutifs » (voir Abbott, 2017). Les zones hybrides fournissent une opportunité puissante pour prendre en compte l'information géographique et environnementale afin d'identifier les traits pertinents pour l'adaptation et/ou l'isolement reproductif et étudier leurs bases génomiques. Ainsi, au niveau de ces zones hybrides, des traits phénotypiques floraux comme la couleur des fleurs chez *Antirrhinum* (Whibley *et al.*, 2006) ou *Mimulus* (Stankowski *et al.*, 2017) présentent souvent des clines spatiaux cohérents avec les clines de fréquences alléliques qui permettent de pointer des régions génomiques candidates, causales de cette variation phénotypique.

De ce point de vue, certains *Ophrys* (voir encadré 2 et encadré 3), tels ceux du groupe *insectifera* pourraient fournir un système d'étude exceptionnel pour étudier ce genre de questions dans un groupe connaissant une diversification intense et c'est ce que je vise à développer dans ce projet. Comme évoqué dans l'encadré 1, le genre *Ophrys* est de plus en plus considéré comme un modèle de choix pour étudier les phénomènes de divergence évolutives (e.g. Nosil *et al.*, 2017; Nürk *et al.*, 2020) mais il existe également de nombreux facteurs limitants inhérents à ce modèle. Le premier consiste en la difficulté de cultiver ces plantes géophytes pérennes en conditions contrôlées. De plus, il est à noter que les *Ophrys* ont plutôt un cycle de vie relativement long et il faut, dans la nature, plusieurs années (probablement 5 à 7 ans) pour aller de la germination jusqu'à la première floraison. La seconde difficulté consiste en la taille importante du génome d'*Ophrys* (> 5 Gpb) qui a retardé le développement de ressources génomiques chez ce genre et qui continue de rendre relativement coûteuse la production de ce type de données. Au demeurant, travailler en populations naturelles dans un contexte de zone hybride permet en partie de contourner l'étape de culture en permettant d'étudier les traits et gènes responsables du maintien de la barrière entre espèces en dépit d'un flux de gènes potentiellement illimité. Par ailleurs, et comme évoqué précédemment, la publication à venir du premier génome de référence pour

le genre (avec *Ophrys sphegodes*, Schlüter lab) ouvre bien entendu des perspectives prometteuses.

Encadré 2 Le groupe *Ophrys insectifera* en tant que modèle d'étude

Parmi les *Ophrys*, le groupe *insectifera* apparaît comme une opportunité exceptionnelle pour étudier les causes de la divergence évolutive. Ce clade consiste en trois espèces très proches : *Ophrys insectifera*, *Ophrys aymoninii* et *Ophrys subinsectifera*. Ces trois espèces sont pollinisées par des insectes hyménoptères de trois familles phylogénétiquement distinctes : une guêpe (Crabonidae) : *Argogorytes mystaceus* (ou *A. fargeii*) pour *O. insectifera*, une abeille sauvage (Andrenidae) : *Andrena combinata* pour *O. aymoninii* et un symphyte (de la famille des Argidae) : *Sterictiphora gastrica* pour *O. subinsectifera*. Ces différences concernant les insectes pollinisateurs sont associées à des différences phénotypiques qui permettent de distinguer sans ambiguïtés les trois espèces d'*Ophrys* sur le terrain. Par exemple, le labelle est le plus court chez *O. subinsectifera*, de taille intermédiaire chez *O. aymoninii* et le plus long chez *O. insectifera*. Comme souligné par Triponez *et al.*, (2013), cette variation au niveau de la longueur du labelle peut être mise en relation avec la longueur du corps de l'insecte pollinisateur, une observation pouvant être interprétée comme la signature d'une pression de sélection induite par le pollinisateur (car la longueur du labelle contraint l'enlèvement des pollinies par l'insecte lors de la pseudo-copulation). Les représentants du groupe *insectifera* diffèrent également au niveau d'autres traits en lien avec la forme et la coloration du labelle (e.g. présence de marges jaunes chez *O. aymoninii* et *O. subinsectifera*, mais pas *O. insectifera*). La composition du bouquet d'odeur (*i.e.* pour un total allant jusqu'à 107 composés ayant pu être détecté dans le mélange des trois espèces) varie également d'une espèce à l'autre (Joffard *et al.*, 2016, voir aussi Borg-Karlson *et al.*, 1993 et Gervasi *et al.*, 2017). On peut y voir une autre preuve de l'existence d'une sélection induite par le pollinisateur (parce que les mâles des différentes espèces de pollinisateurs sont attirés par différents mélanges de composés biologiquement actifs qui miment les phéromones sexuelles de femelles réceptives). Par exemple, Bohman *et al.*, (2020) ont montré que des (Z)-8-Heptadecenes et des *n*-Pentadecanes étaient à fois présents chez *O. insectifera* et la femelle d'*A. fargeii* et que ces composés étaient électrophysiologiquement actifs chez ces insectes. Mais en dépit de différences phénotypiques claires, ces trois espèces sont génétiquement très peu divergentes, un patron cohérent avec des temps de divergence récents, vraisemblablement associé à une évolution réticulée (Triponez *et al.*, 2013). Les aires de distributions de ces trois espèces varient aussi en fonction de facteurs écologiques. *O. insectifera* possède une vaste aire de répartition en Europe de l'Ouest alors qu'*O. aymoninii* et *O. subinsectifera* sont (micro-)endémiques de la région des Grands Causses (France) et du piémont Sud des Pyrénées, respectivement (Fig. 16).

Là où leurs aires de distribution géographique se chevauchent, des individus présentant des phénotypes intermédiaires sont occasionnellement observables mais une possible hybridation n'a jamais vraiment étudiée d'un point de vue génétique au niveau du groupe *insectifera*. Comme chez d'autres *Ophrys*, il ne semble pas exister de barrières intrinsèques et l'isolement floral dû à l'attraction de différentes espèces de pollinisateurs pourrait seul, expliquer l'isolement reproductif entre ces espèces (Gervasi *et al.*, 2017)

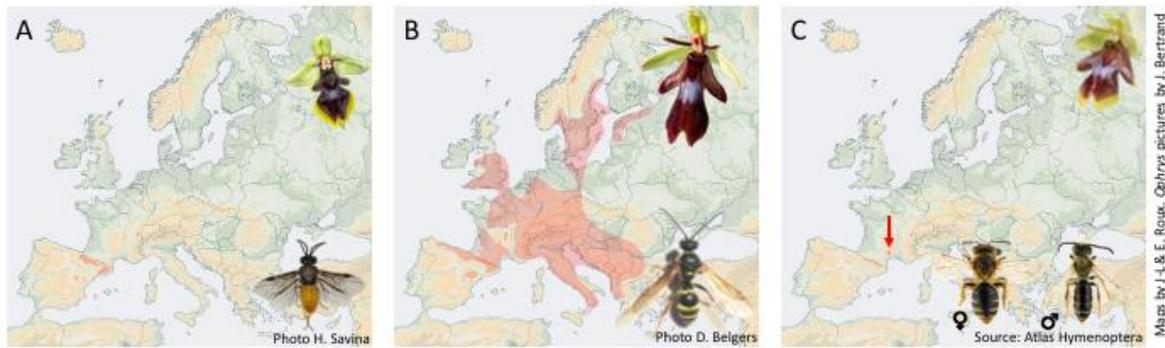


Figure 16 *Ophrys* species of the *insectifera* group: flowers, pollinators and geographic distribution, A) *O. subinsectifera* (pollinator: *Sterictiphora gastrica*), B) *O. insectifera* (poll.: *Argogorytes fargeii*) and C) *O. aymoninii* (poll.: *Andrena combinata*)

Encadré 3 Gènes candidats

Un jeu de gènes candidats pour le déterminisme des phénotypes floraux a déjà été caractérisé chez *Ophrys*. Il a été établi que des homologues de *stearoacyl carrier protein (ACP) desaturase (SAD)* impliqués dans la production de différents alcènes (en particulier *SAD2* et *SAD5*) pourraient constituer des candidats de gènes de spéciation sérieux au niveau du genre car des modifications de leurs niveaux d'expression relatifs modifieraient suffisamment la composition du bouquet d'odeur émis par la fleur pour induire l'attraction d'espèces différentes de pollinisateurs entre deux espèces morphologiquement très similaires : *Ophrys sphegodes* et *O. exaltata* (Schlüter *et al.*, 2011; Sedeek *et al.*, 2016). Des gènes, potentiellement impliqués dans la bio-synthèse d'hydrocarbures et d'anthocyanes ou des facteurs de transcription ont aussi été annotés sur le premier transcriptome d'*Ophrys* publié (Sedeek *et al.*, 2013). Cinq orthologues de gènes MADS box (4 *DIVARICATA*-like et 1 *RADIALIS*-like) ont aussi été identifiées dans une des rares études de transcriptomique comparative portant sur une espèce d'Orchidoideae : *Orchis italica* (Valoroso *et al.*, 2017). Enfin, une approche de type scan génomique a permis d'identifier des *outliers* incluant deux gènes annotés : *VSP45* et *CER1* impliqués dans la circulation des vésicules et la bio-synthèse d'hydrocarbures, respectivement (Sedeek *et al.*, 2014). D'un côté, nous avons donc une liste de *loci* potentiellement ou certainement impliqués dans le déterminisme du phénotype floral des *Ophrys*. D'un autre côté, les liens entre ces gènes et le phénotype n'ont pas encore été démontrés pour *SAD2* et *SAD5* et le rôle qu'ils jouent dans la composition du bouquet d'odeurs chez deux espèces. Bien que les voies de bio-synthèse de ces odeurs puissent être similaires chez d'autres *Ophrys*, nous manquons d'études empiriques pour confirmer ou infirmer cette hypothèse. De plus, la relation entre gènes MADS box et le spectaculaire phénotype floral des *Ophrys* demeure à ce jour largement sous-explorée.

Dans *DiversiFly*, je propose de tirer profit de ce contexte de zone hybride pour étudier ces questions qui sont structurées en deux *Work Packages* (WP1 & WP2)

WP1 Génomique des Populations – Étude des traits et des régions génomiques impliquées dans l'adaptation et/ou l'isolement reproductif au niveau des zones hybrides naturelles que forment localement *O. insectifera* et l'un ou l'autre des *Ophrys* endémiques du groupe (*Ophrys aymoninii* et *Ophrys subinsectifera*)

- **Hypothèse 1** : Si on considère comme hypothèse de travail que les bases de la divergence sont « géniques plutôt que génomiques » chez *Ophrys* (*sensu* Sedeek *et al.*, 2014) alors les

traits faisant l'objet de pressions de sélection divergentes entre espèces (et les fréquences des allèles qui leur sont associés) devraient présenter des clines plus marqués que la moyenne au niveau des zones hybrides. On peut aussi espérer que leur signature génomique soit détectable via des approches de détection d'*outlier*, de scans génomiques et/ou de GWAS (Fig. 17).

- **Hypothèse 2** : Les individus hybrides devraient pouvoir être identifiés car ils présenteraient des valeurs intermédiaires au niveau de certains traits. Cependant, il a été prouvé que des hybrides entre *Ophrys arachnitiformis* et *O. lupercalis* pouvaient aussi émettre des composés olfactifs inédits, en tout cas non émis par les populations parentales (Vereecken *et al.*, 2010). Conformément au modèle de zone de tension, on peut aussi supposer que la *fitness* des hybrides (à évaluer) sera inférieure à celle des individus des taxons parentaux.

- **Hypothèse 3** : Comme les traits impliqués dans la pollinisation par leurre sexuel font l'objet de fortes pressions de sélections exercées par le pollinisateur sur la plante, on peut s'attendre à ce qu'ils présentent une variation inter-individuelle (et inter-annuelle) limitée. Au demeurant et comme nous l'avons proposé dans Baguette *et al.*, (2020), on peut aussi s'attendre à ce que la plasticité de certains traits (notamment des odeurs) ait pu être positivement sélectionnée. En effet, la compétition inter-individuelle pour l'accès aux pollinisateurs est forte chez *Ophrys* et émettre des odeurs nouvelles pourrait s'apparenter à une sorte de « prime à l'innovation » en permettant d'attirer une nouvelle espèce d'insecte pollinisateur, et en évitant la compétition intra-spécifique entre individus d'*Ophrys*.

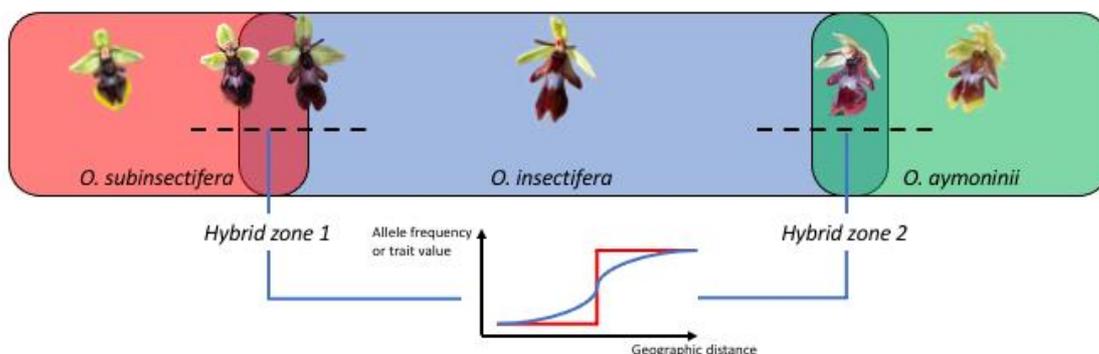


Figure 17 Les patrons de variation clinale seront étudiés pour différents traits phénotypiques et caractéristiques génomiques au niveau de transects entre paires d'espèces

Ce WP1 se décline en quatre tâches afin de tester les hypothèses explicitées ci-dessus.

Tâche n°1 - Marquage et phénotypage

Il est prévu que jusqu'à 96 individus soient échantillonnés le long de transects au travers des deux zones hybrides (*i.e.* jusqu'à 192 plantes). Afin de pouvoir suivre un même individu d'une année sur l'autre, une pointe métallique sera plantée à proximité de chaque pied étudié et les coordonnées géographiques de la plante seront enregistrées avec un GPS de précision centimétrique. En parallèle, nous visons à prendre plusieurs photos calibrées de la fleur (*i.e.* avec un objectif macro et un rapport de grandissement fixé, une échelle, un flash annulaire censé délivrer une température de couleur homogène d'une prise de vue à l'autre et une charte des gris pour pouvoir ajuster *a posteriori* la balance des blancs). Des données quantitatives seront dérivées des photos et permettront d'effectuer des mesures

morphométriques et géomorphométriques, afin de quantifier les couleurs et l'agencement spatial des motifs au moyen de *packages* R : *geomorph* (Adams, 2021), *colordistance* (Weller & Westneat, 2019) ou *patternize* (Van Belleghem *et al.*, 2018) ou de *plugins* ImageJ : *PAT-GEOM* (Chan *et al.*, 2019) dédiés. En parallèle, nous mesurerons aussi au moins les 4 traits représentatifs de la plante entière évoqués dans l'étude pilote n°1 (hauteur de la plante, nombre de bourgeons/fleurs, distance du sol à la première plante, et de la première à la deuxième plante) et les 6 traits floraux (à savoir les longueurs et largeurs des pétales (dont le labelle) et les sépales). En ce qui concerne les odeurs, nous avons initialement tablé sur l'emploi de fibres SPME pour l'échantillonnage mais le succès du protocole employé dans l'étude pilote n°2 et les suggestions faites par un des *reviewers* du projet ANR nous incitent désormais à immerger les fleurs dans de l'hexane. En plus d'être moins onéreuse, cette option permet lors des analyses d'identifier la position exacte des doubles liaisons au niveau de certaines molécules et ainsi de les caractériser plus finement. Pour finir, une fleur sera immergée dans une solution de *RNAlater* puis stockée à -20°C jusqu'à extraction de l'ARN et l'équivalent de 6 à 10 confettis de feuille seront stockés dans du Silica Gel ou de l'éthanol à 95% jusqu'à extraction de l'ADN.

Tâche n°2 – Séquençage de transcriptomes floraux et génotypage nGBS

Comme vu précédemment, le séquençage de transcriptomes est un moyen efficace de sous-échantillonner le génome en se focalisant sur les régions codantes, avec un possible impact fonctionnel. De plus le transcriptome permet à la fois de travailler sur les variations de séquences et de niveaux d'expression. Au demeurant ne serait-ce que parce que les régions génomiques d'intérêt ne se limitent pas nécessairement à l'espace génique, tous les individus seront aussi génotypés suivant le protocole nGBS que nous avons déjà expérimenté avec succès chez *Ophrys* et d'autres espèces de plantes. La panoplie de méthodes standard sera utilisée pour estimer les niveaux d'expression et appeler les variants afin d'évaluer la variation des séquences aux travers des deux zones hybrides. Nous allons ensuite rechercher des *outliers*. Comme différentes méthodes de détection peuvent parfois mener à différents résultats (e.g. voir Dalongeville *et al.*, 2018), nous confronterons plusieurs approches : chercher les *loci* associés à de fortes valeurs de *F*-statistiques (F_{ST} ou F_{CT}) comme dans OutFLANK (Whitlock & Lotterhos, 2015), utiliser des méthodes se basant sur les *Latent Factors Mixed Models* (LFMM) telles qu'implémentées dans le package R LEA (Frichot & François, 2015) ou des analyses multivariées telles qu'implémentées dans *pcadapt* (Luu *et al.*, 2017). De plus, des approches GWAS permettront de détecter des SNPs associés aux valeurs de certains traits phénotypiques.

Tâche n°3 – Analyse de clines

Au niveau de chaque zone hybride nous étudierons si des traits phénotypiques présentent un patron de variation sigmoïde en ajustant différents modèles d'équilibre de clines avec *hzar* (Derryberry *et al.*, 2014). Les caractéristiques de ces traits peuvent être intégrées de façon brutes (par exemple des fréquences alléliques ou la quantité absolue d'un composé organique pour une odeur) ce qui peut être intéressant pour tester une variation clinale sur des traits ou des *loci a priori* identifiés comme candidats. Cependant, ce type d'analyse peut aussi être mené sur des scores d'analyses multivariées (pour la morphologie, la coloration et l'odeur) ou sur des proportions d'admixture, comme nous l'avons déjà fait dans Bertrand *et al.*, (2016). Par ailleurs, Stankowski *et al.*, (2017) ont utilisé les 1% de valeurs les plus élevées de F_{CT} obtenues d'une AMOVA pour se concentrer sur les *loci* présentant une différenciation

maximale entre populations de *Mimulus* dont la couleur des fleurs variait. Les paramètres des clines, notamment leur centre *c*, leur largeur *l* pourront être comparés entre traits et au travers des deux zones hybrides.

Tâche n°4 – Suivi temporel de la variation phénotypique, du taux de pollinisation (succès reproductif mâle), du *fruit set* (succès reproductif femelle) et du taux de « survie »

Pour chaque individu, le taux de pollinisation (*i.e.* le nombre de fleurs pour lequel les pollinies sont absentes divisé par le nombre total de fleur sur une inflorescence donnée) sera évalué à la fin de la session de marquage et de génotypage. A la fin de la période de floraison (*i.e.* environ un mois plus tard), nous retournerons sur le terrain pour mesurer le *fruit set* (*i.e.* le nombre de capsules s'étant développées divisé par le nombre total de fleurs sur une inflorescence donnée). Cette étape est cruciale pour mettre en évidence de possible différences de *fitness* entre espèces et entre espèces parentales et hybrides (de même qu'entre individus tout court). Cette opération (à l'exception du marquage et du génotypage nGBS) sera répétée chaque année sur les quatre ans du projet pour estimer le taux de « survie » (ou du moins la probabilité de fleurir d'une année sur l'autre) mais aussi pour estimer la variabilité inter-individuelle des traits phénotypiques (morphologie, coloration, odeur) et des niveaux d'expression des gènes.

WP2 Transcriptomique comparative - Rechercher et valider les gènes clés dans la radiation évolutive des *Ophrys*

- Hypothèse 4 : A l'échelle du genre, les gènes impliqués dans le déterminisme du phénotype floral (fleurs évoquant un insecte) présentent probablement des signes de sélection positive marqués (*i.e.* un excès de mutation non-synonyme par rapport aux mutations synonymes) par rapport à d'autres orchidées (Fig. 18a).

- Hypothèse 5 : Comme les voies génétiques de la production d'odeur pourraient être similaires pour bon nombre d'*Ophrys* (Bohman *et al.*, 2020), les gènes sous-jacents sont particulièrement susceptibles de présenter des différences de niveau d'expression, même entre espèces phylogénétiquement proches mais pollinisées par des espèces d'insectes attirés par des odeurs différentes (comme c'est le cas pour le clade *insectifera*). Inversement deux espèces d'*Ophrys* pollinisées par des insectes aux préférences très proches sont censés montrer des niveaux d'expression similaires, même si elles sont relativement éloignées du point de vue phylogénétique (Fig. 18b).

Nous avons déjà des données préliminaires pour initier ce WP2. Le financement BQR dont j'ai bénéficié en 2018 m'avait en effet permis de séquencer et d'assembler 11 transcriptomes d'*Ophrys* (de 8 des 10 lignées principales) : *O. bombyliflora*, *O. catalaunica*, *O. exaltata marzuola*, *O. insectifera*, *O. lutea*, *O. speculum*, *O. marmorata*, *O. scolopax*, *O. tenthredinifera*, *O. apifera* et *O. subinsectifera*. Cela s'ajoute aux transcriptomes de quelques *Ophrys* et une vingtaine d'autres orchidées déjà publiés (e.g. Piñeiro Fernández *et al.*, 2019). Sur ces données, j'ai déjà pu i) identifier des candidats tels que des homologues de gènes MADS-box et de gènes codant pour des *stearoyl-acyl-carrier-protein desaturase* (*SAD*) impliqués dans la production d'odeur chez *Ophrys* and ii) mettre en évidence un jeu de transcrits différentiellement (et de façon congruente) exprimés chez différentes espèces d'*Ophrys* comparé à *Orchis italica*, considéré comme groupe externe (Fig. 19).

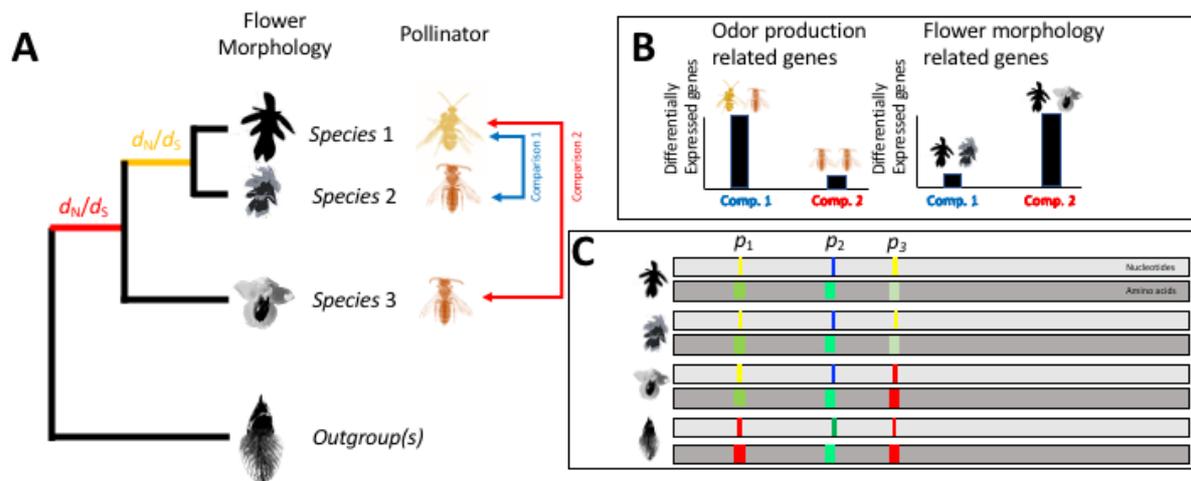


Figure 18 Aperçu des hypothèses en lien avec l'étude de la variation des séquences et des niveaux d'expression des transcriptomes. **A)** Arbre de gènes représentant les relations phylogénétiques entre espèces orthologues. **B)** Le nombre de gènes différemment exprimés devrait être plus important entre gènes codant pour différentes valeurs à un phénotype donné quelles que soient les relations phylogénétiques (arbre d'espèces) entre les espèces. **C)** Dans les alignements de séquences nucléotidiques (en gris clair) et d'acides aminés (gris foncé), les gènes positivement sélectionnés présentent un excès de mutations non-synonymes (d_N) comme dans p_1 et p_3 comparé aux mutations synonymes (d_S). La mutation en p_1 augmente le ratio d_N/d_S sur la base à la base des *Ophrys* (en rouge) alors que celle en p_3 l'augmente sur la branche menant au clade *insectifera* (en orange).

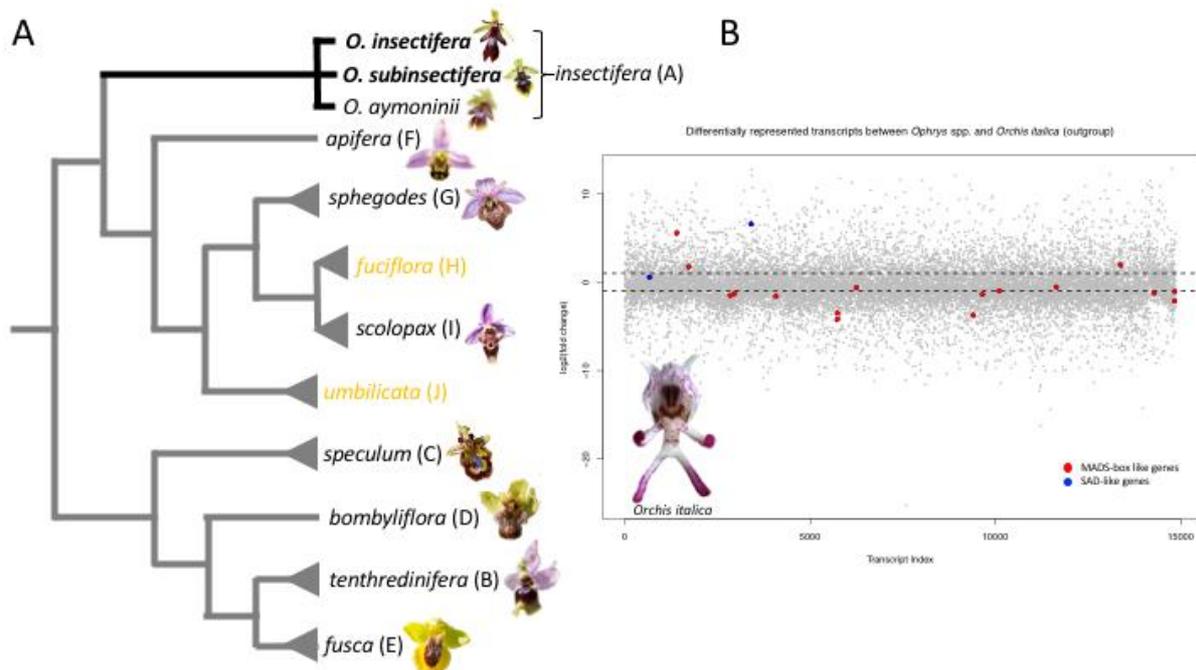


Figure 19 A) Arbre phylogénétique (consensus) du genre *Ophrys*, mettant en exergue le groupe *insectifera* (les lignées pour lesquelles nous ne disposons pas encore de transcriptomes sont indiquées en orange). **B)** Aperçu des différences de niveaux d'expression entre transcrits chez les 11 espèces d'*Ophrys* et *Orchis italica* (groupe externe). Les homologues de SAD (production d'odeur) et gènes MADS box (morphogénèse florale) sont indiqués en bleu et rouge, respectivement.

Ce WP2 se décline en trois tâches afin de tester les hypothèses explicitées ci-dessus.

Tâche n°5 – Obtenir des transcriptomes floraux représentatifs du genre *Ophrys* et d'autres orchidées

Une fois de plus, l'idée n'est pas de mener une étude RNAseq *per se* mais d'utiliser les transcriptomes floraux comme un moyen de « sous-échantillonner » l'espace génique de représentants des 9-10 lignées principales (parfois appelées macro-espèces ou espèces tête de groupe) d'*Ophrys* (et d'autres orchidées comme groupes externes). En plus des 11 espèces représentatives de 8 des 9/10 lignées, l'idée est de compléter le jeu de données jusqu'à une trentaine d'espèces en sachant que deux espèces d'*Ophrys* (*O. aveyronensis* et *O. corbariensis*) et deux d'espèces d'*Himantoglossum* (*H. robertianum* et *H. hircinum*) ont été récoltées en 2021 et que nous pouvons compter sur les 4 transcriptomes floraux d'*O. catalaunica* et 4 transcriptomes d'*O. magniflora* générés en 2021 pour l'étude pilote n°2.

Tâche n°6 – Rechercher les gènes sous sélection positive

Cette recherche sera menée à la base des *Ophrys* comparé aux groupes externes et au sein des *Ophrys* (en particulier au niveau du groupe *O. insectifera*). En utilisant une approche phylogénétique employée avec succès par Anna MARCIONETTI et l'équipe de Nicolas SALAMIN dans laquelle j'ai effectué mon post-doc pour identifier des gènes impliqués dans la relation mutualiste poisson-clown/anémone de mer (Marcionetti *et al.*, 2019), nous nous baserons sur les transcriptomes floraux pour inférer les Groupes Hiérarchiques Orthologues (HOGs). Nous alignerons les HOGs résultants et utiliserons ces alignements pour reconstruire les arbres de gènes. Pour les orthologues « *single-copy* » (1-to1 OGs), la sélection positive (*i.e.* excès de substitutions non-synonymes sur synonymes, avec $d_N/d_S > 1$) sera évaluée i) à la base des *Ophrys* et ii) entre lignées d'*Ophrys* au sein du genre (e.g. groupe *insectifera* vs. le reste). Une approche similaire sera menée pour les gènes dupliqués et d'une manière générale, les « *multi-copy* HOGs ». Une annotation fonctionnelle sera menée sur chaque gène candidat ; Bien sûr nous pourrons et envisageons d'utiliser cette approche pour étudier les gènes déjà connus pour être impliqués dans le développement floral (ainsi que pour les candidats qui auraient pu être identifiés dans le WP1).

Tâche n°7 – Rechercher les gènes différenciellement exprimés chez *Ophrys*

En suivant une autre approche employée par Martha SERRANO et l'équipe de Nicolas SALAMIN pour identifier les gènes impliqués dans des adaptations plantes/pollinisateurs (Serrano-Serrano *et al.*, 2019) nous comparerons les lignées d'*Ophrys* pour chercher des gènes dont les niveaux d'expression sont cohérents i) entre lignées (e.g. groupe *insectifera*) ou ii) entre groupes d'*Ophrys* (tous deux considérés comme des « conditions ») et partageant des caractéristiques particulières (e.g. « floraison précoce », ou « pollinisation par des abeilles (de la famille des Apidae) » qui différenciellement exprimés entre conditions. La matrice normalisée d'expression de gènes pour les 1-to-1 OGs sera utilisée pour effectuer une analyse multivariée (ACP), des analyses de *clustering* (*k*-mean) et une analyse d'expression différentielle de gènes entre conditions. Là encore, une annotation fonctionnelle des candidats sera effectuée et nous nous servirons également de ce cadre pour ré-évaluer le statut de tous les candidats connus ou révélé dans le WP1. Pour résumer et comme l'indique

le schéma ci-dessous, les deux WP2 sont donc complémentaires mais pas nécessairement interdépendants (Fig. 20).

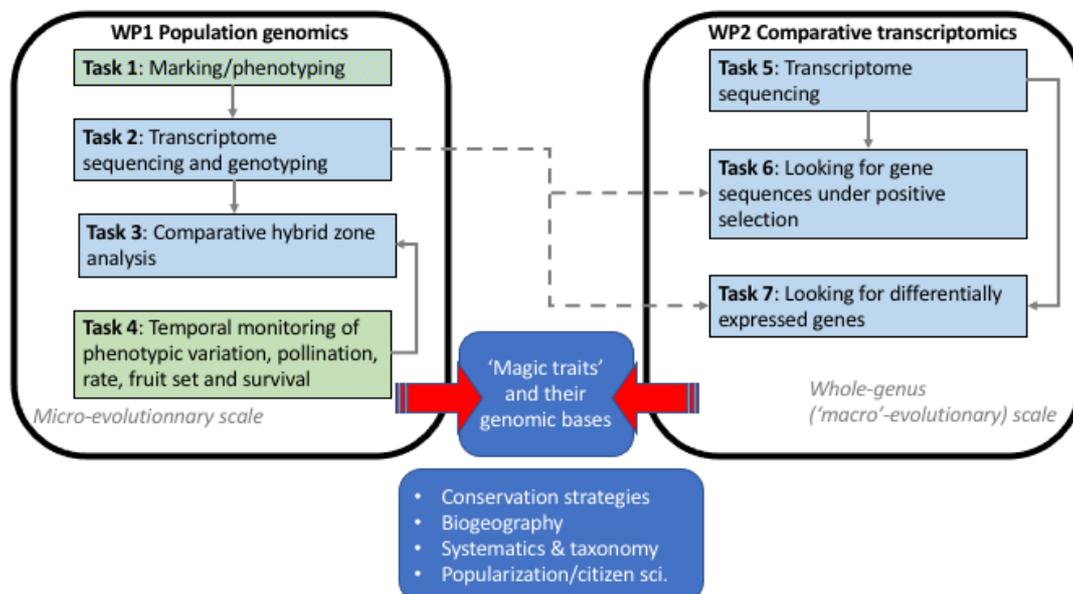


Figure 20 Aperçu du workflow du projet *DiversiFly*

En plus des attendus explicités ci-dessus, les résultats du projet *DiversiFly* pourront sans doute être valorisés dans d'autres disciplines telles que la biologie de la conservation (notamment pour les deux espèces (micro)-endémiques) la biogéographie, la systématique et la taxonomie. Le modèle *Ophrys* se prête aussi bien à la vulgarisation en raison de la stratégie de pollinisation originale (et au-dessous de la ceinture d'un point de vue très anthropocentré) mais aussi parce que le réseau d'orchidophile bénévole très développé pourrait permettre de mettre en place des actions de science participative.

Projets de recherche 'annexes'

Ces projets de recherche dans lesquels j'ai été impliqué depuis mon recrutement sont présentés dans l'ordre chronologique.

Gestion dite 'éco-bio-sociale' d'une invasion biologique

Présentation du projet Fin 2018, le dépôt d'un dossier pour la création de la Fédération de Recherche « Énergie Environnement » (FREE FR 2043) auprès des tutelles (UPVD et CNRS) nous a encouragé à discuter entre collègues perpignanais dans le but de faire émerger, d'abord localement, une recherche résolument plus intégrative. Un collègue modélisateur, Sébastien GOURBIÈRE, a ainsi été à l'initiative d'un projet visant à implémenter une approche dite 'éco-bio-sociale' pour comprendre l'invasion biologique des populations de châtaigniers (*Castanea sativa*) du département des Pyrénées-Orientales par un hyménoptère parasite (*Dryocosmus kuriphilus*) venu d'extrême Orient. Ce petit insecte induit la formation de galles au niveau des feuilles et provoque une perte de production de biomasse parfois jusqu'à entraîner un dépérissement de l'arbre. En 2019, le projet a obtenu un premier financement de la Mission pour les Initiatives Transverses et Interdisciplinaires (MITI) du CNRS dans le cadre

du 'Défi Modélisation du vivant' (21 k€) et a reçu le soutien financier de la Région Occitanie dans l'attribution d'un financement doctoral (2019-2022). L'idée de ce projet était de combiner des compétences reposant sur l'acquisition de données biologiques et écologiques (*i.e.* taux d'infestation, prévalence, taux de prédation des galles par des hyperparasites, caractérisation des peuplements forestiers...) et génétiques pour alimenter différents modèles visant à caractériser et prédire l'évolution de ce phénomène d'invasion biologique. Mon rôle était plus particulièrement d'apporter mon expertise en génomique des populations pour caractériser les patrons de diversité et de différenciation des populations de châtaigniers à l'échelle de la zone d'étude, puis de voir en quoi certaines caractéristiques génétiques de l'arbre hôte pouvaient être associées à des phénotypes de résistance (*i.e.* via des approches de type GWAS) et ainsi jouer un rôle dans un contrôle de type '*bottom up*' de l'invasion par la ressource. En plus de l'échantillonnage des châtaigniers sur le terrain, j'ai été impliqué dans l'estimation des taux d'infestation et des taux de prévalence sur la première session de terrain (été 2019). Ce projet m'a par ailleurs permis de participer à l'encadrement de la thèse de doctorat de Jean-Loup ZITOUN (à hauteur de 50%). Ce projet sera par la suite financé par la FREE en 2020 (4,5 k€).

Génomique de la conservation de plantes endémiques ou emblématiques des Pyrénées

Présentation du projet Dans le cadre d'un projet financé par le Fonds Européen de Développement Régional (FEDER) via l'instrument Interreg-POCTEFA pour la période 2020-2022 (Coût total estimé : 625,3 k€ / coût FEDER approuvé : 394,2 k€), j'ai pris part à un projet de coopération transfrontalière (France, Espagne et Andorre). Ce projet intitulé 'FLORALAB' (www.floralab.eu) est porté par la Fédération des Réserves Naturelles Catalanes (FRNC). C'est un projet de préservation et de valorisation de la flore patrimoniale transfrontalière de l'Est des Pyrénées (départements des Pyrénées Orientales et de l'Ariège en France, Catalogne (Espagne) et Andorre) dont les objectifs sont :

- i) d'améliorer les connaissances sur la flore Est-pyrénéenne rare et menacée,
- ii) de favoriser une gestion concertée, transfrontalière et interprofessionnelle de cette flore et de ses habitats,
- iii) de former étudiants et acteurs du territoire à la connaissance de cette flore et aux enjeux de sa conservation et,
- iv) de rendre accessible ces informations au grand public.

Le projet FLORALAB a pour ambition de s'appuyer sur la recherche scientifique pour atteindre ces objectifs et l'UPVD en est un partenaire bénéficiaire (académique). FLORALAB cible 10 taxons représentatifs d'un patrimoine floristique rare et menacé de l'Est des Pyrénées. Quatre de ces taxons ont une répartition eurosibérienne (et atteignent dans les Pyrénées la limite méridionale de leur aire de distribution) : le botryche à feuilles de matricaires (*Botrychium matricariifolium*), la dorine à feuilles alternes (*Chrysosplenium alternifolium*), le maïanthème à deux feuilles (*Maianthemum bifolium*) et la gentiane des marais (*Gentiana pneumonathe*). Trois autres taxons ont une répartition plus globale mais présentent de rares populations isolées dans la chaîne pyrénéenne : l'orchis à corne courte (*Orchis spitzelii*), le saule des Lapons (*Salix lapponum*) et l'astragale à fleurs pendantes (*Astragalus penduliflorus*). Enfin, trois taxons sont endémiques de l'Est des Pyrénées : le persil d'isard (*Xatardia scabra*), la **dauphinelle des montagnes (*Delphinium montanum*)** et la **pédiculaire fausse asperge (*Pedicularis comosa* subsp. *asparagoides*)**.

En raison de leur endémisme local et/ou parce qu'elles montrent des signes alarmant de déclin, ces populations ont été pressenties pour la réalisation d'une étude génétique et l'expertise de ma collègue Valérie HINOUX et moi-même a été sollicitée pour mener à bien sa réalisation. Il est à noter que si le financement de l'étude génétique a été assuré par FLORALAB pour la dauphinelle des montagnes (*Delphinium montanum*), c'est ensuite le commissariat de massif qui a pris en charge celui de l'étude relative à la pédiculaire fausse asperge (*Pedicularis comosa* subsp. *asparagoides*). Nous avons pu compter sur le soutien précieux du réseau d'acteurs de FLORALAB pour nous assister dans l'échantillonnage en milieu naturel ainsi que sur le recrutement d'une étudiante en master, Pascaline SALVADO qui a assuré l'analyse des données au cours de sa formation. Pascaline s'est consacrée à l'étude sur la dauphinelle des montagnes lors de sa première année de Master (2020-2021) et travaille désormais sur la pédiculaire fausse asperge (2021-2022).

Peu d'espoir pour la dauphinelle des montagnes (*Delphinium montanum*) : éléments combinés obtenus sur la base d'approches de génomique des populations et de modélisation de niche écologique

Le manuscrit correspondant a été publié dans *Ecology & Evolution* (Salvado *et al.*, 2022a) et un *preprint* a été déposé sur bioRxiv.

Contexte À ce jour, deux études génétiques avaient été réalisées sur la dauphinelle des montagnes (Simon *et al.*, 2001, Lopéz-Pujol *et al.*, 2007). Toutes deux suggéraient déjà un déficit en hétérozygotie de même qu'un degré de différenciation relativement important entre populations. Au demeurant, ces résultats avaient été obtenus sur un nombre limité de marqueurs enzymatiques (allozymes) et/ou de populations. De plus, plusieurs éléments concordants soutenaient que la dauphinelle des montagnes (*Delphinium montanum*) était une espèce auto-tétraploïde, un résultat que nous avons pu corroborer via une estimation de la valeur 1C de l'espèce en cytométrie en flux. Avec cette 1C-value de 10,32 pg, nous nous devons donc de composer avec une espèce polyploïde dont la taille du génome rendait peu envisageable une étude de génomique des populations sur la base de génomes complets. A cela s'ajoutait le fait que le *corpus* méthodologique existant pour analyser des données de génomique des populations s'avère plutôt restreint pour étudier des organismes polyploïdes.

Objectifs & Hypothèses L'idée première était donc de ré-évaluer les patrons de diversité et de structuration de la diversité génétique neutre avec une résolution « génomique » pour mieux appréhender le statut actuel de la métapopulation de dauphinelle des montagnes (*Delphinium montanum*) et proposer des solutions pour la conservation de l'espèce. Nous avons opté pour une approche de génomique des populations, toujours avec le protocole nGBS que j'avais déjà utilisé avec succès sur *Ophrys aveyronensis* ainsi que sur le châtaignier (*Castanea sativa*) et qui nous avait permis d'obtenir respectivement 9231 et 31736 marqueurs de type SNPs de qualité satisfaisante. L'échantillonnage a été réalisé au niveau de 9 localités représentatives de l'aire de distribution de l'espèce (Fig. 20).

En parallèle, nous avons cherché à modéliser la niche écologique actuelle de la dauphinelle des montagnes (*Delphinium montanum*) dans le but de déterminer à quel point le changement climatique pouvait constituer une menace pour l'espèce et de prédire l'évolution de son aire de distribution dans le futur. Pour ce faire, nous avons téléchargé des données correspondant à 19 variables bioclimatiques « actuelles » (*i.e.* interpolées à une

maille de 30 secondes d'arc sur la base d'observations réalisées par un réseau de stations météorologiques sur la période 1981-2010. Ces différentes variables sont en lien avec les températures et les précipitations. Les prédictions de différents GCM (pour *Global Circulation Models*) ont ensuite permis d'inférer la dynamique spatiale de la niche écologique de l'espèce à l'horizon 2035 (sur la base des données de la période 2011-2040), 2055 (2041-2070), et 2085 (2071-2100).

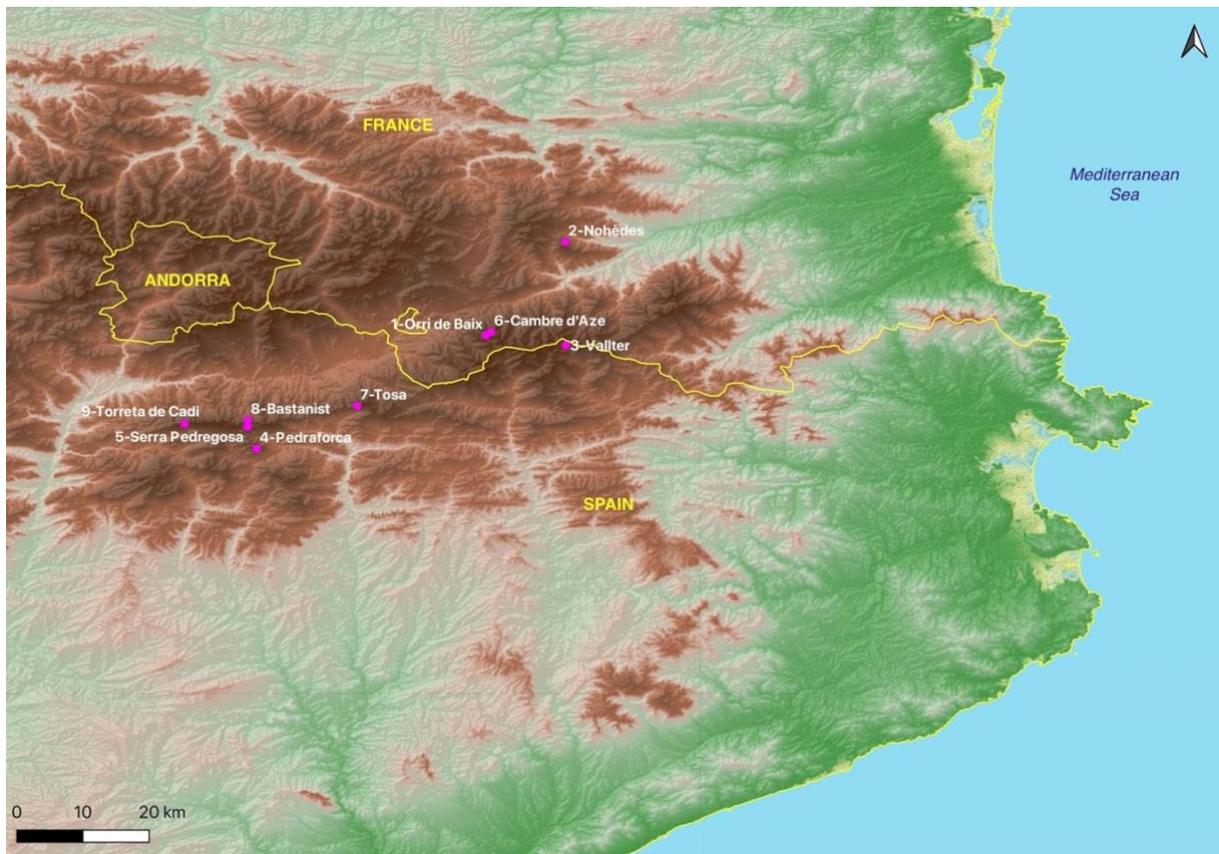


Figure 20 Échantillonnage géographique des 9 localités de *Delphinium montanum* (d'après Salvado et al., 2022a)

Résultats Les indicateurs d'écart à la panmixie (G_{IS}) et de différenciation génétique (e.g. G_{ST}), le patron clair d'isolement par la distance, de même qu'une analyse d'assignation (sNMF) confirment que les différentes populations de dauphinelle des montagnes (*Delphinium montanum*) se comportent comme des entités souffrant de consanguinité sur leurs massifs montagneux s'apparentant à de véritables îles émergeant d'une matrice d'habitats défavorables (« sky islands ») très peu connectées entre elles (Fig. 21).

L'utilisation de données bioclimatiques actuelles et les projections à différents horizons dans le futur laisse présager une diminution drastique de l'aire de distribution de l'espèce au cours du XXI^{ème} siècle et ce, quel que soit le scénario suivi (Fig. 22). En effet, nous avons effectué une prédiction de la niche suivant différents RCP (pour *Representative Concentration Pathways*) censés représenter différentes trajectoires d'augmentation des gaz à effet de serre (i.e. forçage radiatif) plutôt optimiste (+2,6 W/m²), « réaliste » (+7 W/m²) et relativement pessimiste (+8,5W/m²). Tous, suggèrent que l'aire de distribution de la dauphinelle des montagne (*Delphinium montanum*) va se réduire comme peau de chagrin, parfois jusqu'à disparaître totalement.

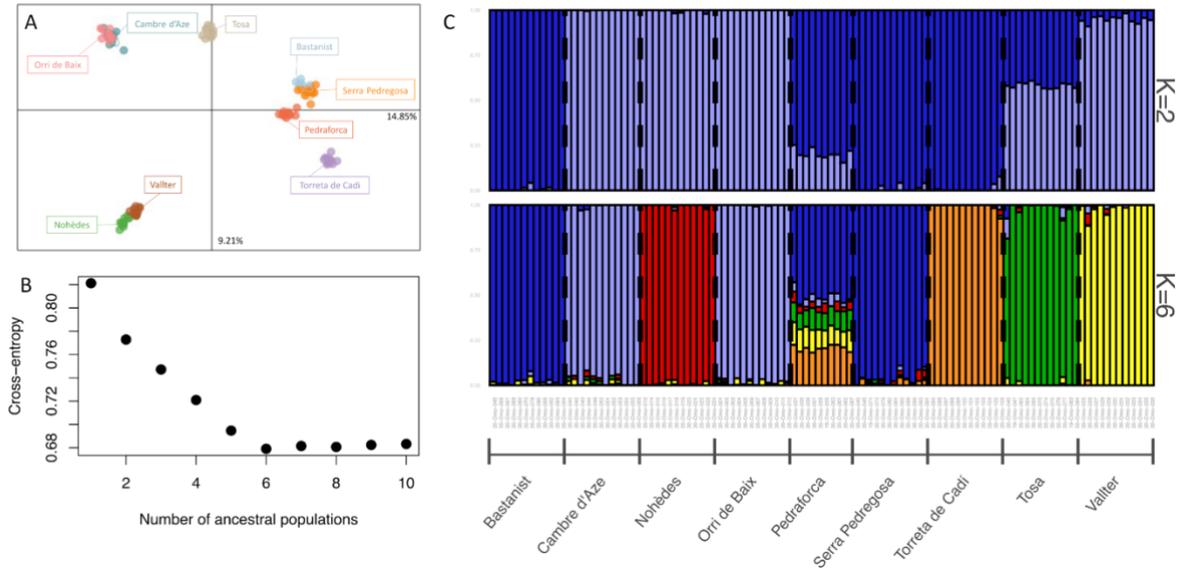


Figure 21 Diversité et structuration génétique sur la base de 106 individus de *Delphinium montanum* génotypés sur 5095 SNPs **A)** ACP suivant les deux premiers axes (PC1 et PC2) représentant respectivement 14,85% et 9,21% de la variance génétique totale. **B)** Valeurs de critère de cross-entropie pour un nombre de groupes allant de $K=1$ à $K=10$ (10 répliques sNMF pour chaque valeur de K). La valeur optimale considérée a été $K=6$. **C)** Diagrammes en barres représentant les *ancestry coefficients* obtenus grâce à sNMF à $K=2$ et $K=6$ (d'après Salvado *et al.*, 2022a)

En outre, cette étude a permis de générer des données de séquençage (Nanopore et Illumina) nous ayant permis d'assembler et d'annoter le génome chloroplastique de *D. montanum* (Salvado *et al.*, 2022b).

Conclusion et perspectives S'il est en l'état actuel des choses, difficile de conclure quant à l'impact réel de la consanguinité et de l'absence de connectivité sur les populations de dauphinelles des montagnes (*Delphinium montanum*), nos résultats suggèrent que la principale menace pesant sur l'espèce pourrait être le changement climatique. Le degré de différenciation observé à l'échelle de la métapopulation montre que toute mesure de conservation telle que le renforcement, la translocation ou la migration assistée aurait très probablement un impact notable sur l'augmentation des effectifs, l'augmentation de la diversité génétique et de l'aire de distribution, respectivement. Plusieurs études suggèrent d'ailleurs qu'une politique « interventionniste » est généralement plus bénéfique qu'une vision plus « conservatrice » pour sauvegarder des populations d'espèces menacées. Si des croisements en milieu contrôlé pourraient permettre d'évaluer un éventuel phénomène de dépression d'allogamie, nos résultats suggèrent aussi que toutes ces mesures pourraient s'avérer vaines si la niche écologique (et l'habitat) de l'espèce venait de toute façon à disparaître. Des tests préliminaires menés par des acteurs du réseau FLORALAB ont aussi montré que la dauphinelle des montagnes (*Delphinium montanum*) était difficile à cultiver en dehors de son habitat naturel. Au demeurant, persévérer dans cette voie permettrait peut-être d'élucider des phénomènes tels que l'absence récurrente de floraison chez les individus de la population de Nohèdes (massif du Madres-Coronat). Sur ce point particulier donc, continuer à développer les ressources génomiques chez cette espèce pourrait permettre d'assurer une collaboration pertinente entre scientifiques et gestionnaires ainsi que la mise en place d'une recherche intégrative visant à lier génomique, biologie du développement et écologie.

Quand la couleur (de la fleur) ne reflète pas les relations systématiques (et la taxonomie) entre populations Est-pyrénéennes de Pédiculaires chevelue (*Pedicularis comosa*) mais révèle une zone hybride cryptique

Contexte et hypothèse Deux sous-espèces de pédiculaires chevelues (*Pedicularis comosa*) ont été décrites sur la base de critères morphologiques, à savoir la couleur des fleurs. La sous-espèce *Pedicularis comosa* subsp. *comosa* possède des fleurs jaunes et se rencontre dans plusieurs massifs montagneux du Sud de l'Europe alors que la sous-espèce *Pedicularis comosa* subsp. *asparagoides* possède des fleurs roses et serait endémique du massif des Albères (s'étendant de manière anecdotique jusqu'au massif du Canigou). C'est donc pour évaluer le statut taxonomique de cette « sous-espèce » et envisager des mesures de gestion appropriées pour ce taxon que la Fédération des Réserves Naturelles Catalanes nous avait sollicité pour une étude génétique par l'intermédiaire du projet FLORALAB. A l'instar de ce qui avait été fait sur la dauphinelle des montagnes (*Delphinium montanum*), nous avons utilisé l'approche nGBS pour étudier les patrons de diversité et de différenciation au niveau de 11 localités (pour un total de 158 individus génotypés) dans la partie orientale des Pyrénées.

Résultats préliminaires Sur la base des 4427 SNPs retenus, il apparaît que la structuration génétique des populations ne reflète pas la différence de coloration de la fleur. En effet, elle serait plutôt cohérente avec la géographie (et notamment les massifs montagneux : Albères, Canigou/Cerdagne, Madres/Coronat/Donezan) et les deux sous-espèces telles que décrites sur la base de critères morphologiques ne formeraient pas des groupes réciproquement monophylétiques! Entre Albères et Cerdagne, les analyses d'assignation suggèrent l'existence d'une zone hybride cryptique relativement étroite centrée sur l'Ouest des Albères (au niveau du pic des Salines). Sur le terrain, néanmoins, nous avons noté des fleurs d'un rose pâle atypique au niveau du massif du Canigou (vers le refuge de Mariailles). Bien que cela nécessite confirmation, il pourrait donc y avoir un décalage entre cline phénotypique et cline génomique. La recherche de loci *outliers* possiblement impliqués dans la variation de coloration ou dans d'autres traits potentiellement sous sélection sont en cours de même que la modélisation et la confrontation de différents scénarios démographiques qui nous permettront de mieux comprendre l'histoire évolutive de ce système. Bien que les résultats obtenus ne soient *a priori* pas conformes à ceux « attendus » par nos collègues des réserves, ce modèle biologique ouvre donc des perspectives prometteuses sur des aspects plus fondamentaux en évoquant notamment un « négatif » du modèle *Antirrhinum majus* possédant lui aussi deux « sous-espèces », l'une à fleurs jaunes (subsp. *striatum*), l'autre à fleur rose (subsp. *pseudomajus*), formant une zone hybride ayant déjà été exhaustivement étudiée (e.g. Whibley *et al.*, 2006; Khimoun *et al.*, 2011; 2013 ; Bradley *et al.*, 2017) dans le même contexte géographique. La première différence évidente avec *Pedicularis* tient au fait que la zone hybride serait moins exubérante et que c'est l'aire de distribution de la forme à fleur jaune qui englobe celle à fleur rose chez ce dernier.

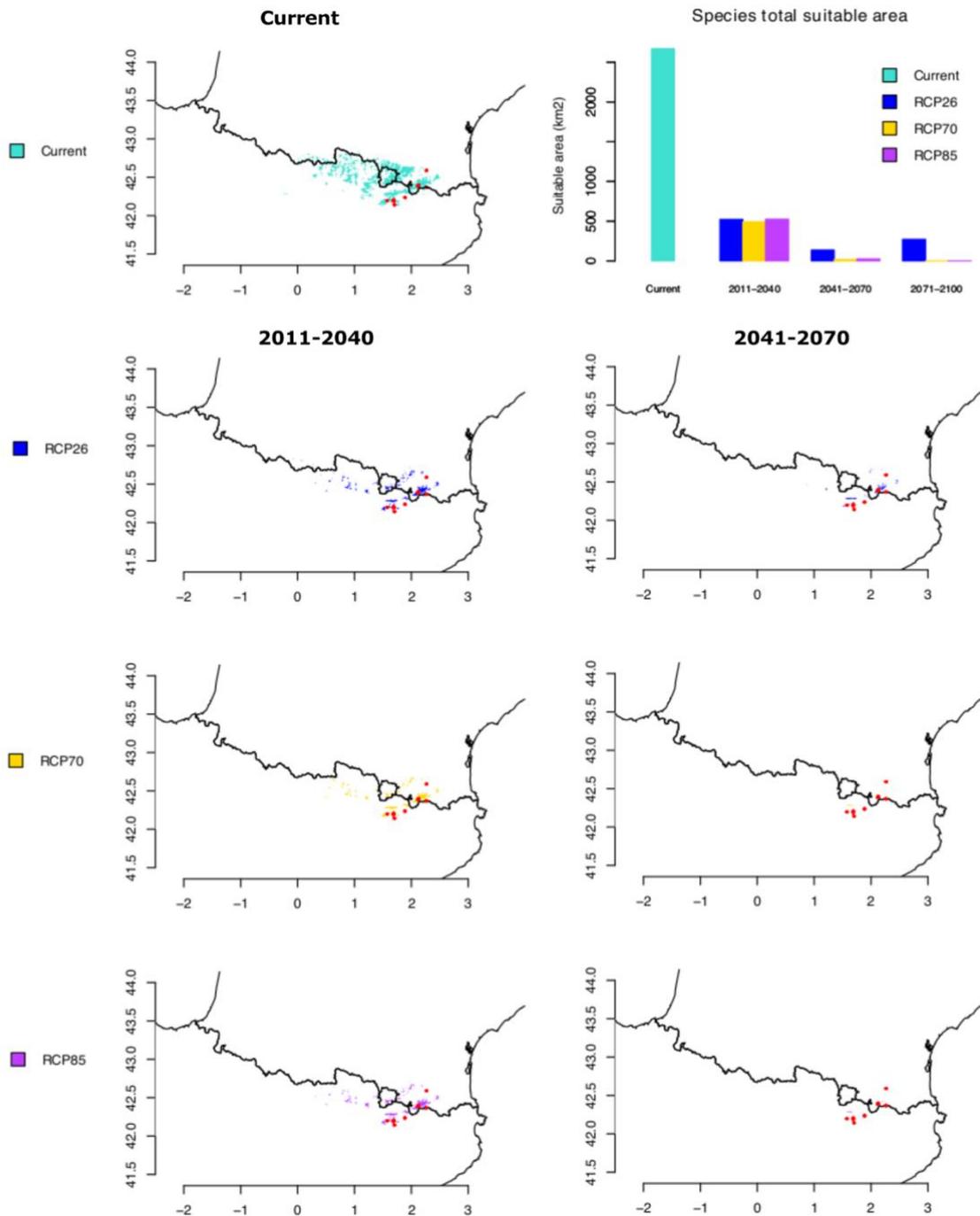


Figure 22 Projection spatiales des habitats considérés comme favorables à l'établissement de *Delphinium montanum* sur la base de la modélisation de niche environnementale réalisées à partir de variables bioclimatiques. Ces projections sont faites sur la base des données actuelles (moyenne 1981-2010) et futures avec un consensus obtenu à partir de 5 *Global Climate Models* (CGMs) pour différentes périodes : 2011-2040 et 2041-2070 et suivant différents *Representation Concentration Pathways* (RCPs) : 2,6, 7,0 et 8,5 (d'après Salvado *et al.*, 2022a).

Le tutorat du Junior Lab: *Genetic approaches to assess the decline of the Western Capercaillie (Tetrao urogallus subsp. aquitanicus) as a new way to improve its conservation in the Catalan natural Reserves (GenyWesNCo)*

Dans un contexte socio-géographique similaire, je contribue actuellement à l'encadrement de quatre étudiantes de M1: Céleste PARIS, Manon CANAL, Morgane TASSAINT et Pauline BUSO dans le cadre d'un *Junior Lab* du master *Functional Biology & Ecology* de la *TULIP Graduate School*. Sur proposition d'un acteur socio-économique, ici la Fédération des Réserves

Naturelles Catalanes et plus particulièrement Sylvain DANIELO, conservateur de la réserve naturelle de Py-Mantet, les étudiantes doivent s'emparer d'une problématique et d'y apporter une réponse par l'élaboration et la réalisation d'un projet scientifique. Les étudiantes travaillent sur la conservation des populations locales de Grand Tétrás (*Tetrao urogallus* subsp. *aquitanicus*) en développant des approches de génétique des populations, de *metabarcoding* (diète et charge parasitaire), de dosage d'hormones révélant le stress des individus, le tout sur la base de technique d'échantillonnage non-invasives. Elles sont co-encadrées par Valérie HINOUX, Olivier REY et moi-même. En plus de mon intérêt pour les modèles ornithologiques et les problématiques de conservation, le développement d'approches non-invasives et l'acquisition de compétence en *metabarcoding* pourront sans aucun doute être mises à profit dans mes autres projets, par exemple pour mieux retracer les relation plantes-insectes chez les *Ophrys*.

Systématique et taxonomie (intégrative) d'orchidées euro-méditerranéennes et d'autres organismes

Si la systématique et la taxonomie sont, dans le cadre de mes recherches, plutôt un sous-produit de projets de recherche plus fondamentaux qu'un objectif en soi, j'ai néanmoins contribué à plusieurs projets plus ou moins centrés sur la systématique et la taxonomie. Pour certaines espèces notamment, l'échantillonnage sur le terrain est compliqué, soit parce que l'habitat de ces espèces est difficile d'accès, soit parce que l'état de conservation de leurs populations rend le recours à des méthodes « classiques » peu envisageables.

Muséomique des pigeons couronnés de Nouvelle Guinée et mitogénomique des Clupéiformes

J'ai ainsi été impliqué dans un projet de recherche visant à mettre à profit des collections de spécimens disponibles en musée. L'objectif de ce projet, dirigé par mon directeur de thèse Christophe THÉBAUD et Guillaume BESNARD, était d'obtenir de manière non-destructive, des données génétiques pour étudier l'histoire évolutive de trois espèces menacées ou quasi-menacées de pigeons couronnés de Nouvelle-Guinée (appartenant au genre *Goura*). Avec mes collègues Yann BOURGEOIS et Boris DELAHAIE, nous avons prélevé de petits échantillons sur les coussinets plantaires d'oiseaux entreposés dans les collections du *Natural History Museum* de Tring (Royaume-Uni). Bon nombre d'individus avaient été collectés il y a plus de 150 ans (certains par Alfred Russel WALLACE en personne). Un séquençage Illumina de l'ADN extrait à partir de ces échantillons a ensuite été réalisé (via un protocole de type *shotgun*). En dépit d'un ADN fortement dégradé et très faiblement concentré, nous avons pu reconstituer avec succès les génomes mitochondriaux complets de trois espèces (quelques gènes nucléaires d'intérêt ont aussi été récupérés). Les séquences obtenues ont ensuite été alignées avec d'autres séquences de Colombiformes disponibles dans la littérature. Nous avons pu proposer une hypothèse phylogénétique informative et bien résolue de ce groupe (Besnard *et al.*, 2016) et ces travaux ont constitué le point de départ de la thèse de doctorat de Jade BRUXAUX.

Les compétences acquises en assemblage de mitogénomes m'ont par la suite amené à collaborer avec des collègues ichtyologues (notamment Sébastien LAVOUÉ) pour assembler des génomes mitochondriaux et les utiliser en tant que données génétiques pour inférer la phylogénie de Clupéiformes. Ce groupe constitué par les anchois, les harengs et les sardines est bien connu pour son intérêt économique mais recèle une diversité génétique parfois

cryptique. Or, la prise en compte de cette diversité (cryptique) apparaît comme cruciale, ne serait-ce que pour aider à établir des politiques durables de gestion des stocks. Trois manuscrits en lien avec ce travail ont depuis été publiés (Lavoué *et al.*, 2017a; 2017b; 2019).

Depuis mes travaux de Master, j'ai toujours été sensible à la confrontation des signaux fournis par les génomes nucléaires et mitochondriaux lors de l'étude la biogéographie des espèces et des régions. Sur les poissons coralliens des genres *Abudefduf* et *Amphiprion* de la famille des Pomacentridae que j'ai eu l'opportunité d'étudier dans le cadre des deux post-docs successifs, un examen comparatif des deux types d'information a ainsi permis de révéler aussi bien des lignées cryptiques *a priori* impossible à distinguer les unes des autres sur la plan morphologique mais présentant des taux de divergences pouvant atteindre jusqu'à 10% entre elles sur les séquences de certains gènes, de même que des phénomènes de capture mitochondriale (voir Bertrand *et al.*, 2017).

Fort de ces expériences, j'ai par la suite toujours été attentif à l'étude des génomes mitochondriaux (chez les animaux) et chloroplastiques chez les plantes (voir Marcionetti *et al.*, 2018 pour la publication du premier génome de poisson-clown). C'est notamment le cas sur le manuscrit à propos de la biogéographie d'un clade de poissons-clowns, le groupe *Amphiprion akallopisos* qui est en cours de finalisation et pour lequel nous mettons aussi en évidence des phénomènes d'incongruences mito-nucléaires (Marcionetti *et al.*, *in prep.*).

Taxonomie et systématique intégrative des genres *Ophrys* et *Allium*

En 2021, j'ai également initié une collaboration avec un collègue montpelliérain (Errol VÉLA) avec qui nous avons co-encadré un étudiant de Master 1 (Hugo BELLAVOIR) pour proposer une hypothèse phylogénétique multi-locus visant à placer dans l'arbre du vivant des taxons du genre *Allium* récoltés il y a plusieurs années dans des zones aujourd'hui très peu accessibles d'Irak. Le nombre d'individus repoussant années après années s'amenuisant, Errol m'avait proposé de collaborer pour extraire l'ADN, amplifier un certain nombre de gènes par PCR puis les séquencer en Sanger afin de reconstruire la phylogénie, déterminer la position des différents individus afin de vérifier la cohérence entre cette position phylogénétique et l'identification qui avait été faite par Errol et ses confrères botanistes sur la base de critères morpho-anatomiques. Les premiers résultats suggèrent que si globalement la taxonomie telle que définie sur des critères morphologiques est cohérente avec notre phylogénie moléculaire, il est envisageable de décrire plusieurs nouveaux taxons sur la base de la systématique moléculaire.

En dehors de ces différentes contributions, mon intégration au sein de la commission scientifique de la Fédération France Orchidées (FFO) m'a incité à candidater à des sources de financement pour utiliser des données génomiques dans le but de faire avancer les connaissances sur des genres d'orchidées dont la systématique et la taxonomie demeurent problématiques, en premier lieu desquels les *Ophrys*. Avec mes collègues Etienne DELANNOY, Errol VÉLA et Frédéric MELKI, nous avons ainsi obtenu un premier financement de la Fondation Biotope pour un projet baptisé 'REORDER' (2,5 k€) en 2020. L'idée est d'employer une approche de type *RAD-seq* pour obtenir une matrice de plusieurs milliers de SNPs sur un échantillon exhaustif des taxons d'*Ophrys* présents en France. Pour ce faire, nous pouvons compter sur le réseau de cartographes bénévoles de la FFO pour nous assister dans la collecte. Bien que le travail de terrain soit ambitieux et ait demandé d'être étalé sur plusieurs années (*a fortiori* en raison des restrictions en lien avec la crise sanitaire du COVID 19), l'objectif et d'inclure les premiers résultats de cette étude lors de la sortie de la troisième édition de l'ouvrage *Orchidées de France, Belgique et Luxembourg* prévue pour le printemps 2023.

Tableau 4 Bilan des projets financés dans lesquels j'ai été impliqué depuis mon recrutement en tant que Maître de Conférences

Type de projet	Nom du projet	Période	Porteurs (si partenaire)	Montant
Bonus Qualité Recherche UPVD	Étude des bases génomiques de la radiation adaptative des orchidées euro-méditerranéennes du genre <i>Ophrys</i>	2019		14,2 k€
Initiative Inter-LabEX (TULIP-CEMEB)	Étude pilote de la distribution géographique disjointe de l'orchidée endémique et menacée <i>Ophrys aveyronensis</i> : causes et implications	2019	Co-porté avec B. Schatz	8,35 k€
Projet Interreg-POCTEFA	<i>Floralab</i>	2020-2022	UPVD partenaire bénéficiaire	394,2 k€
Fédération de Recherche en Écologie & Environnement	Développement d'une approche intégrative pour la conservation d'espèces emblématiques	2020	Co-porté avec C. Bertrand	4,46 k€
Fondation Biotope	<i>REORDER</i>	2020-2011	Co-porté avec E. Véla, F. Melki et E. Delannoy	2,5 k€
Fédération de Recherche en Écologie & Environnement	Étude intégrative de l'état de santé des populations de Grand Tétras (<i>Tetrao urogallus aquitanicus</i>), une espèce emblématique locale menacée d'extinction	2021	Co-porté avec O. Rey & V. Hinoux	3,3 k€
ANR JCJC	<i>DiversiFly</i>	2022-2025 (48 mois)		371,6 k€
Projet BiodivOc	<i>DevOCGen</i>	2022-2024 (36 mois)	Porté par S. Boitard et R. Lebblois	329,4 k€

Réflexions personnelles pour une recherche plus intégrative

Fort de l'expérience présentée ci-dessus, je présente ci-après quelques pistes qui permettraient la mise en place d'une recherche plus intégrative. A mon échelle, j'essaie de ne pas auto-censurer mes questions de recherche par des verrous biologiques et techniques et essaie de toujours garder en ligne de mire cet objectif ultime de lier génotypes, phénotypes et *fitness*. Cette démarche, explicitée par exemple par Byers *et al.*, (2017) et la *special issue* de *Molecular Ecology* dans laquelle cet article a été publié ou Schlüter (2018) est schématisée de façon simplifiée dans la figure suivante (Fig. 23).

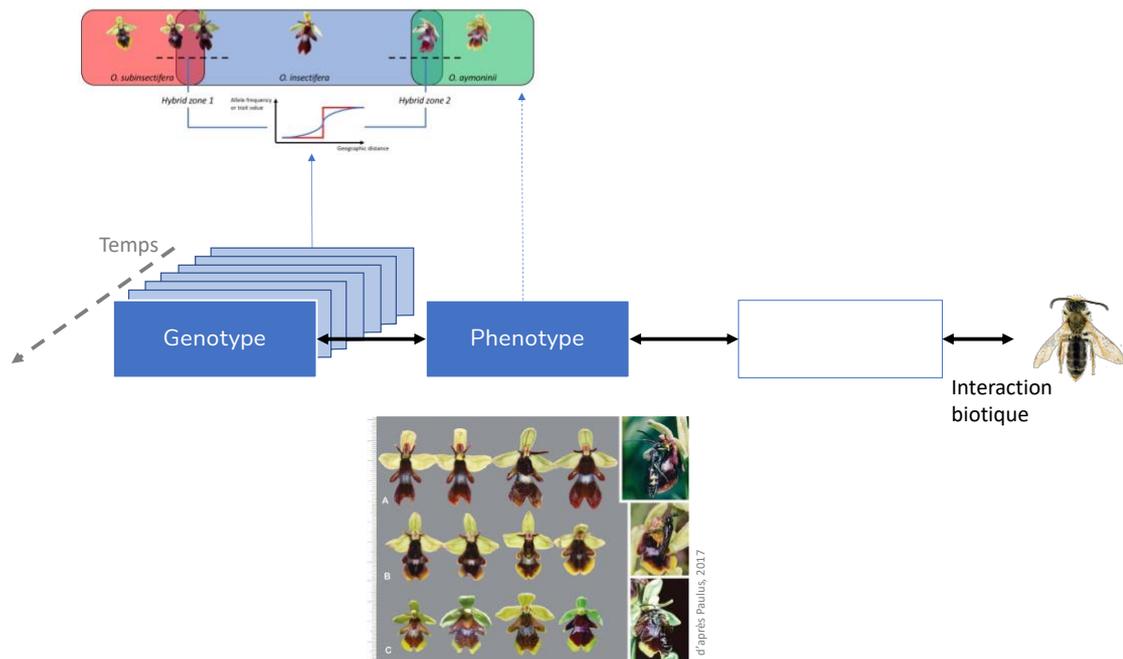


Figure 23 Schéma de l'approche intégrative visée sur la base de mes projets en cours, illustré avec l'exemple des *Ophrys*

Je me suis spécialisé dans l'utilisation de marqueurs génétiques pour tenter d'expliquer la mise en place de patrons de diversité et de différenciation phénotypique. J'ai donc utilisé les génotypes (séquences, microsatellites ou SNPs) pour retracer l'histoire biogéographique de populations mais je n'ai, au final, que peu creusé les liens entre génotypes et phénotypes et entre phénotypes et *fitness*. Les approches de type P_{ST} - F_{ST} (e.g. comme employées dans Bertrand *et al.*, 2016) permettent par exemple, en comparant degrés de différenciation génétiques et phénotypiques, de mettre en évidence des traits potentiellement sous sélection. Au demeurant, elles ne permettent pas d'élucider les bases génétiques d'un trait d'intérêt, et donc, de faire le lien entre génotype et phénotype dans le sens, étude des bases génétiques de la variation. C'est une densité conséquente en marqueurs génétiques et des approches de type scan génomiques, GWAS, détection d'*outliers* et lorsque le contexte s'y prête, clines, qui permettent plutôt de franchir le pas en populations naturelles. Cela ne dispense pas d'études ultérieures plus approfondies du mécanisme sous-jacent pour aller vers une validation fonctionnelle, mais c'est souvent une étape incontournable pour la caractérisation de régions génomiques candidates (en populations naturelles).

Pour ce faire, les approches de génomique des populations via le re-séquençage de génomes complet (et éventuellement l'alignement sur une référence) seraient désormais comparables en termes de coûts à des approches de type *RAD-seq* et *Poolseq* chez bon nombre d'organismes (Lou *et al.*, 2021). Toutefois, même un re-séquençage à faible profondeur demeure coûteux à tous les points de vue pour une espèce non-modèle telle qu'*Ophrys* (> 300€/ind. pour une profondeur de 2X). Reste alors les méthodes permettant de réduire la complexité génomique (autrement dit de sous-échantillonner le génome).

Quels marqueurs génétiques employer ?

Sous échantillonner le génome aux moyens d'approches de réduction de la complexité génomique de type *RAD-seq* : avantages et limitations

Les approches de la famille *RAD-seq* sont plus exhaustives (en termes de représentativité du génome) et au final plus faciles à mettre en œuvre qu'un jeu de marqueurs microsatellites sur lesquels j'avais travaillé en début de carrière. Elles sont entièrement sous-traitables, ne nécessitent pas vraiment de phase de mise au point des conditions d'amplification et d'optimisation des *multiplex*. L'obtention des génotypes est aussi exempte de la subjectivité de l'expérimentateur lors de la phase d'interprétation des profils, propres à des marqueurs tels que les microsatellites (mais voir aussi Shafer *et al.*, 2016 pour une étude montrant comment les *pipelines* et les paramètres choisis peuvent influencer l'analyse des données et les interprétations qui en découlent). En contrepartie, le protocole est dépendant de l'action d'enzymes de restriction et d'une éventuelle étape de sélection de taille des fragments (sans compter la présence éventuelle d'une étape de normalisation). Les *pipelines* d'analyses bioinformatiques des données brutes évoluent aussi au point que se pose la question de la répétabilité du génotypage. Une chose est sûre, un protocole *RAD-seq* peut difficilement être considéré autrement que comme une approche « *one shot* » générant un jeu de données qui pourra difficilement être étoffé au fur et à mesure de l'accumulation des échantillons. Ainsi, combiner différents jeux de données, même générés avec des protocoles *a priori* identiques (méthode de préparation des banques et séquençage) mais de *runs* réalisés à différentes dates est sans doute à proscrire. A cela s'ajoute le fait que l'approche en elle-même demeure relativement coûteuse (de l'ordre de 70 à 80 € par échantillon) ce qui limite sa démocratisation, par exemple lorsqu'il s'agit de convaincre des partenaires non-académiques dans le cadre d'études de biologie de la conservation.

Dans les cas particulier des espèces dont la taille de génome est relativement importante, comme c'est le cas pour *Ophrys*, *Pedicularis comosa* ou *Delphinium montanum*, la part de génome réellement couverte par ce type de données est dérisoire (<1% sur une matrice de SNPs filtrée) ce qui apparaît comme peu crédible quand on prétend s'intéresser à la recherche de régions génomiques candidates à l'adaptation et/ou à l'isolement reproductif, quand bien même le déséquilibre de liaison serait conséquent au niveau des populations étudiées. Pour nuancer ce point, notons toutefois que ces données permettent d'obtenir des informations avec une résolution suffisante pour appréhender la diversité, la structuration et dans une certaine mesure, l'histoire démographique des populations. Nous pouvons aussi assurer qu'elles nous ont permis de mettre en évidence des *outliers* « fiables » qui ne demandent sans doute qu'à être confirmés.

La transcriptomique, vue comme une approche pour sous-échantillonner le génome: avantages et limitations

Parmi les approches de réduction de la complexité génomique, les transcriptomes permettent de cibler les régions codantes de gènes exprimés dans les tissus à partir desquels les ARN sont extraits. Travailler sur des transcriptomes possède donc le double avantage de permettre à la fois d'analyser les niveaux d'expression des gènes mais aussi les variations de séquence au niveau de régions génomiques possédant potentiellement un impact fonctionnel. Pour ce faire, différentes approches peuvent être employées. Les transcriptomes peuvent ainsi être assemblés à partir d'un génome ou d'un transcriptome de référence, ou *de novo* à partir d'outils spécifiques tels que Trinity (Grabherr *et al.*, 2011). C'est cette dernière approche qui est actuellement employée par Lucas VANDENABEELE, étudiant de M1 que je co-encadre avec

Hervé PHILIPPE, Michel BAGUETTE et Bertrand SCHATZ. L'idée est d'assembler (et d'annoter) les transcriptomes afin d'inférer les groupes d'orthologues avant d'employer des méthodes de reconstruction phylogénétiques. Sur la base de centaines de *loci*, il devrait donc être possible de confronter arbres de gènes et arbres d'espèces et de mieux appréhender le degré de réticulation des *Ophrys* à une échelle génomique. Il a ainsi été montré que la phylotranscriptomique pourrait être une alternative crédible à la phylogénomique lorsque le traitement des données suit un *pipeline* rigoureux (Cheon *et al.*, 2020). En outre, ce type d'approche pourrait permettre d'isoler des *barcodes* permettant de distinguer génétiquement les *Ophrys* même à une échelle taxonomique infra-spécifique. Parmi les applications envisagées, il serait par exemple possible de se servir de ces *barcodes* pour identifier à l'espèce auxquelles appartiennent des pollinies d'*Ophrys* collectées sur un insecte pollinisateur (chose impossible sur la base de marqueurs disponibles à l'heure actuelle, car trop peu résolutifs). Il est à noter que certains outils tels KisSplice (Sacamoto *et al.*, 2012) permettent quant à eux d'identifier des SNPs et d'employer des approches de génomique des populations sur la base des transcriptomes. Bien sûr, travailler sur des transcriptomes n'a pas que des avantages. L'ARN est une molécule relativement instable comparée à celle d'ADN et elle peut-être délicate à obtenir en quantité et en qualité satisfaisante à partir d'échantillons obtenus en conditions de terrain. Le signal phylogénétique délivré par les parties codantes transcrites des ARN messagers peut aussi apparaître comme trop peu variable et de toute façon biaisé par la sélection naturelle. Au demeurant, les parties 5'-UTR et 3'-UTR sont des portions transcrites (mais pas traduites) situées en amont du codon d'initiation et en aval du codon stop qui ne sont pas censées être soumises au même régime de pressions sélectives et sont plus variables, et potentiellement plus informatives. C'est en tout cas ce qu'il ressort des premiers alignements effectués sur *Ophrys*.

L'*haplotagging* : une option prometteuse pour reconstituer les haplotypes en début d'un séquençage de faible profondeur

Dans le cadre du projet BiodivOc (financé par la Région Occitanie) intitulé *DevOCGen* porté par deux collègues montpelliérains du laboratoire CBGP: Simon BOITARD et Raphaël LEBLOIS, il est prévu que 200 individus d'*Ophrys aymoninii* échantillonnés dans le cadre du projet ANR JCJC *DiversiFly* soient séquencés en suivant un protocole innovant : l'*haplotagging* (Meier *et al.*, 2021). Cette approche se basant sur un séquençage courte-lecture de type (Illumina) consiste à marquer des *loci*, physiquement proches sur le génome pour séquencer et assembler des longs fragments permettant par la suite de reconstituer les haplotypes. Cette approche est relativement peu coûteuse et elle serait simple à mettre en œuvre. L'idée de ce projet est d'inférer l'histoire démographique récente (~ 100 dernières générations) de populations d'espèces emblématiques de la région Occitanie. Parmi elles, *Ophrys aymoninii* constitue un défi technique car la taille de son génome fait que i) malgré l'effort de séquençage apporté, la profondeur sera probablement plus faible pour *Ophrys* que pour les autres organismes étudiés dans cette étude et ii) théoriquement, il est nativement possible de marquer jusqu'à environ 85 millions de fragments, ce qui pourrait s'avérer être insuffisant pour un génome > 5 Gbp. Si *Ophrys (aymoninii)* fait donc office de test en conditions extrêmes de cette approche dans le cadre de ce projet, les données obtenues pourront être confrontées à celles obtenues sur un lot commun d'individus dans le cadre du projet ANR JCJC *DiversiFly* avec des données nGBS et transcriptomiques.

Alternatives crédibles aux SNPs obtenus via des méthodes de réduction de la complexité génomique

Ne pouvant pas se satisfaire d'une phylogénie se basant sur le séquençage Sanger d'une poignée de *loci* pour étudier la systématique et la taxonomie de taxons ayant connu une diversification à la fois rapide et intense, se pose aussi la question de l'emploi de protocoles faisant appel à des méthodes de séquençage haut-débit (et se basant généralement sur la technologie Illumina) afin de séquencer simultanément des centaines de *loci*. A ce titre, j'avais commencé à suivre avec intérêt les avancées des protocoles ciblant des régions génomiques dont les régions flanquantes étaient suffisamment conservées à l'échelle du rang taxonomique étudié (e.g. une famille), mais qui étaient-elles-même suffisamment variables pour apporter une information utile pour la reconstruction phylogénétique. Dans l'équipe de Nicolas SALAMIN au sein de laquelle j'ai effectué mon dernier post-doc, une doctorante, Oriane LOISEAU, s'était ainsi attelée à adapter des kits pré-existants pour étudier la systématique des Geonomateae (une tribu de palmiers des Néotropiques), ce qui lui avait permis de pouvoir se baser sur près de 4000 régions génomiques séquencées avec succès sur la majorité des individus inclus dans le jeu de données (Loiseau *et al.*, 2019).

Cette approche pourrait donc s'avérer fructueuse pour faire progresser la systématique des orchidées euro-méditerranéennes, voire pour tenter de résoudre les relations phylogénétiques au sein de genre problématiques comme *Ophrys* (et *Allium*) sur lesquels j'ai commencé à me pencher. Le problème, c'est que de tels kits n'ont que très récemment été publiés pour les Orchidacées et ne sont tout simplement pas disponibles pour les Amaryllidacées. La première solution consisterait donc à se lancer dans le développement d'un tel kit. Pour *Ophrys* par exemple, je dispose désormais du jeu de données de transcriptomes floraux, représentatifs de la diversité phylogénétique du genre à partir duquel il serait possible d'isoler des régions candidates. C'est en suivant une démarche similaire que Breitkopf *et al.* (2015) avait définis les marqueurs nucléaires inédits qui leurs avaient servi à produire une phylogénie multi-locus du genre, en se basant sur le transcriptome d'*Ophrys sphegodes* publiée en 2013 (Sedeek *et al.*, 2013). Alternativement, un kit « universel » permettant de séquencer avec succès un panel de 353 régions génomiques à l'échelle des angiospermes a également été publié il y a quelques années (Johnson *et al.*, 2019). Se posait alors la question de savoir si ce kit universel allait permettre d'obtenir une résolution nécessaire pour travailler à des rangs taxonomiques inférieur que celui de l'ensemble des plantes à fleur. Yardeni *et al.*, (2021) viennent d'apporter des éléments de réponse en utilisant en parallèle le fameux *Angiosperms353kit* et un kit (*Bromeliad1776*) plus spécifique de la famille des Broméliacées sur laquelle les auteurs travaillaient. Comme on pouvait s'y attendre, les résultats sont meilleurs avec le kit spécifique à la famille mais les performances globales du kit universel seraient relativement satisfaisantes et mèneraient notamment à des résultats comparables en termes de reconstruction phylogénétique. Récemment, Pérez-Escobar *et al.* (2021) ont mené un travail similaire en utilisant le kit *Angiosperms353* chez des orchidées et un kit spécifique orchidées a d'ailleurs été publié la même année (Eserman *et al.*, 2021). Les choses évoluent très vite au point que nous envisageons désormais une approche « kit » (délaisant ainsi l'approche *RAD-seq* initialement envisagée) pour mener à bien le projet REORDER sur la systématique moléculaire des *Ophrys* présenté précédemment. Pour ce faire, nous avons envisagé de collaborer avec une collègue, Louise BROUSSEAU, récemment recrutée à l'IRD dans la même unité qu'Errol VÉLA avec qui je porte le projet.

Quels traits phénotypiques utiliser et comment faire le lien avec les génotypes et la *fitness* ?

Chez *Ophrys*, le mimétisme évident entre la fleur et « son » espèce d'insecte pollinisateur pousse à une réévaluation de la prise en compte des phénotypes floraux en essayant de ne pas se limiter à des mesures morphométriques (longueurs, largeurs, ratio...) et à tendre vers la quantification la plus objective possible d'autres traits phénotypiques, en lien avec les formes, les couleurs ou les patrons de coloration, tels les motifs plus ou moins complexes présents sur le labelle. Si ces traits sont généralement utilisés en taxonomie pour délimiter les espèces, leur utilisation dans des clés d'identification ne peut généralement pas être incluse autrement qu'au moyen d'une description fastidieuse et souvent très relative. Des mesures pseudo-quantitatives telles que le spectre maculaire (Ring, 2008) ont pu être proposées par certains orchidophiles mais leur utilisation demeure difficile à mettre en œuvre pour un non spécialiste et délivre des données de portée limitée pour tester des hypothèses scientifiques. L'idée était donc d'explorer des outils permettant d'extraire des informations quantitatives à partir d'images. Une fois de plus, ce sont mes expériences passées sur les modèles animaux qui

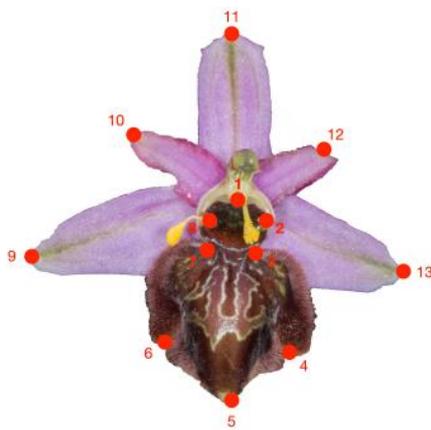


Figure 24 Position des 13 *landmarks* sur une fleur d'*Ophrys aveyronensis*.

m'ont suggéré l'idée d'utiliser trois *packages* R : *geomorph* (Adams *et al.*, 2021), *colordistance* (Weller & Westneat, 2019) et *patternize* (Van Belleghem *et al.*, 2018).

Le *package geomorph* permet, comme son nom l'indique, de faire de la **géomorphométrie**. Cette approche consiste à positionner des points de repères (*landmarks*) au niveau de régions qui sont homologues d'une image (et donc d'un individu) à l'autre. Chez les plantes, cette opération peut être compliquée par le fait que certains organes comme la fleur sont composés de tissus mous contrairement à une coquille de mollusque ou un squelette de vertébré. Au pire, l'information sera bruitée mais pas nécessairement biaisée ce qui ne me paraît donc pas un frein pour

envisager une telle approche chez les plantes. Une fois l'ensemble des *landmarks* positionnées, l'étape suivante consiste à effectuer une analyse procrustéenne qui par un jeu de centrage, de rotation et d'homothétie va maximiser la superposition des *landmarks*. C'est donc bel et bien sur des formes et non pas des mensurations qu'on travaille à l'issue de cette transformation. Suite à cette première étape, des analyses multivariées et la batterie de tests statistiques associés permettent d'étudier de possibles différences de formes entre individus, populations, espèces...

Remarque : Notons que rien n'empêche de relever des valeurs de distances, d'angles ou de surfaces entre *landmarks* avant cette transformation. On retombe alors dans une démarche de **morphométrie** « classique » mais se faisant sur photo plutôt que sur la plante vivante. Cette approche de morphométrie sur image peut permettre de corroborer, voire de compléter des mesures qui pourraient avoir été effectuées sur le terrain.

Le *package colordistance* a été conçu pour quantifier et évaluer les différences de coloration entre poissons coralliens. Typiquement, une image numérique est constituée de plusieurs millions de pixels colorés qui peuvent être placés dans un repère en trois dimensions dont les axes représentent par exemple la proportion de Rouge, de Vert et de Bleu qui constitue leur couleur (RVB). Sur des photos détournées (dépourvues d'arrière-plan), il serait

donc intuitif de vouloir comparer les valeurs de coloration de pixels entre images. Or, cela pose problème pour plusieurs raisons. La première c'est qu'un pixel donné d'une première image a peu de chance d'être « homologue » avec un pixel donné de l'image 2. De plus, le nombre pixels total peut ne pas être rigoureusement le même entre deux images ayant été détourées. La seconde, c'est que des comparaisons de pixels deux à deux mèneraient à des temps de calculs trop important. Pour reprendre la métaphore de l'auteur principale de ce *package*, l'idée est donc de grouper au préalable les pixels en un certain nombre de catégories (ou *bins*) un peu à la manière dont les fabricants de chaussures nous proposent de faire rentrer l'infinie gamme de variation de la taille de nos pieds dans un nombre limité de pointures commercialisées. Ce sont ensuite ces *bins* qui seront comparées d'une image à l'autre en sachant qu'ils sont constitués soit en rangeant les pixels dans des catégories fixes pour chaque image, soit redéfinis sur chaque image via une approche de type *k-means*. Sur cette base, des distances peuvent enfin être calculées entre chaque paires d'images pour donner des matrices de dissimilarité.

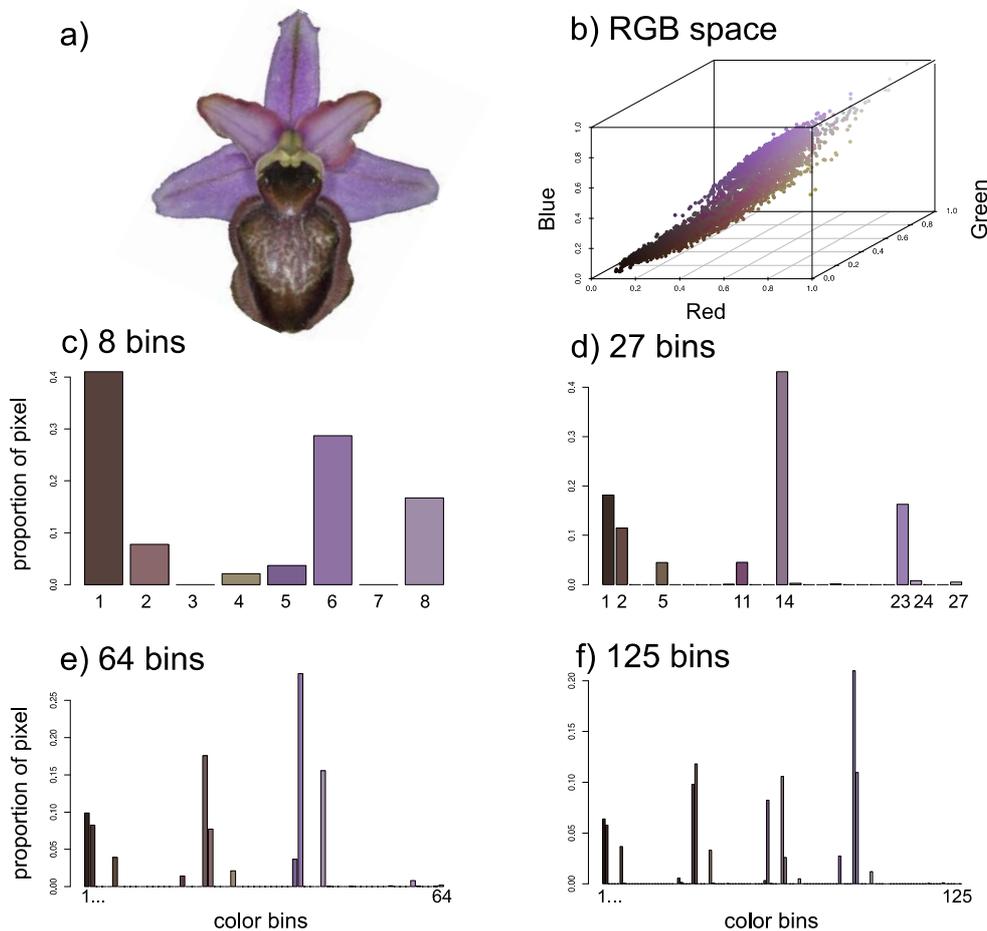


Figure 25 Binning d'une image d'*Ophrys aveyronensis*. (a) image de la fleur, (b) décomposition de cette image en pixels dans l'espace RVB. (c) Histogramme montrant les proportions de pixels assignés à chacun des 8 *bins*, (d) 25 *bins*, (e) 64 *bins* et (f) 125 *bins* prédéfinis par l'utilisateur (d'après Gibert *et al.*, soumis).

Le *package patternize* va plus loin en considérant chaque image comme une grille *raster* dans laquelle les pixels sont géoréférencés suivant leurs coordonnées dans un repère en deux dimensions. Cela exige donc que les images soient alignées entre elles mais ce n'est donc pas simplement la couleur mais bien son agencement spatial qui est évalué ici. Pour définir le patron d'intérêt, l'utilisateur peut soit i) définir une couleur intérêt (en entrant sa

composante RVB), ii) soit via une approche *k*-means (en définissant un nombre de *clusters* que le programme se chargera ensuite de définir de façon optimale sur l'image), soit iii) via une approche *watershed* dans laquelle l'utilisateur pointe une zone d'intérêt qui est ensuite étendue par un algorithme pour tenter de définir le patron précisément.

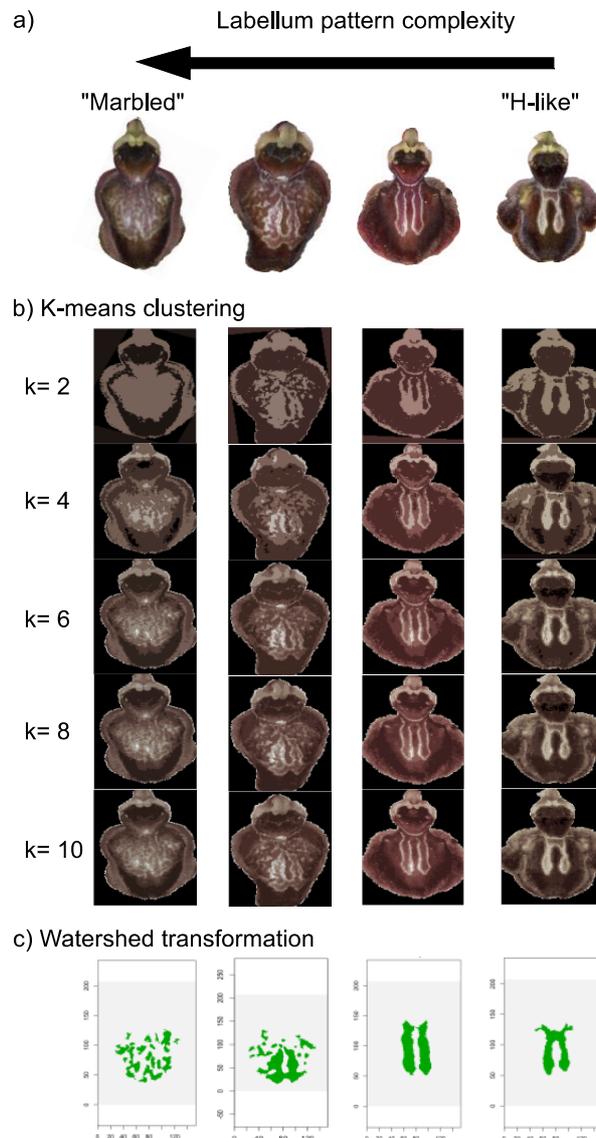


Figure 26 Comparaison de patrons extraits avec *patternize* à partir d'une assignation par *k*-means et une transformation *watershed* de quatre spécimens d'*Ophrys aveyronensis* (a) photographies des labelles, (b) assignation par *k*-means pour $k=2, 4, 6, 8$ et 10 et (c) transformation *Watershed*.

D'autres *package* R (e.g. *Momocs*, Bonhomme *et al.*, 2014) mais aussi des *plugins* ImageJ (e.g. *pat-geom*, Chan *et al.*, 2019) n'ont à ce jour pas été essayés dans ce projet mais pourraient avantageusement compléter des approches déjà utilisées. Par ailleurs, je suis conscient qu'une prise en compte rigoureuse de la coloration nécessitera de ne pas se restreindre à notre point de vue anthropocentré. Les pollinisateurs des *Ophrys* perçoivent certainement une partie du spectre UV (comme l'abeille domestique et bien d'autres insectes) et nous devons sans doute envisager des approches se basant sur les spectres pour aller plus loin dans la compréhension des liens entre phénotype, pollinisateur, et donc *fitness*.

A ce jour, nous avons mis en application ces approches sur les images obtenues sur 109 individus du complexe *Ophrys aveyronensis* afin de tester s'il était possible i) d'obtenir des variables quantitatives fiables à partir de photos ii) de tenter de discriminer les deux

taxons de ce complexe d'espèces (*O. a.* subsp. *aveyronensis* et *O. a.* subsp. *vitorica*) et iii) de caractériser de possibles nouveaux traits phénotypiques impliqués dans l'adaptation et/ou l'isolement reproductif.

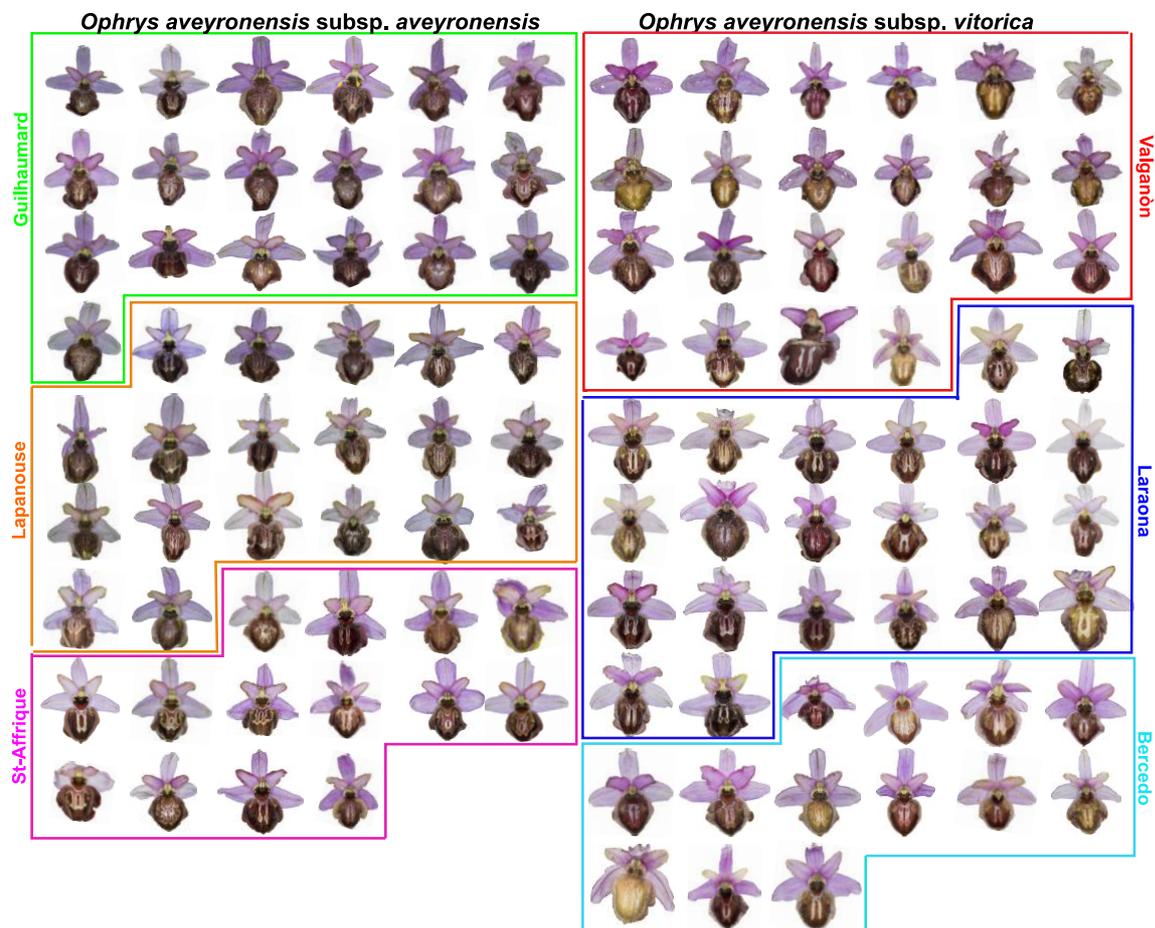


Figure 27 Photographies des 109 individus du complexe *Ophrys aveyronensis*: 52 *O. a.* subsp. *aveyronensis* et 57 *O. a.* subsp. *vitorica* et provenant de 6 localités, trois en France : Guilhaumard, Lapanouse-de-Cernon et Saint-Affrique et trois en Espagne : Valgañón, Larraona et Bercedo, utilisés pour l'étude. (d'après Gibert *et al.*, soumis)

En bref, l'utilisation conjointe des *packages geomorph*, *colordistance* et *patternize* couplée à une approche de *machine learning* de type *random forest* (implémentée dans la *package R* éponyme, Liaw *et al*, 2002) nous a permis de discriminer les deux taxons avec une précision allant jusqu'à 95% (soit bien mieux que l'œil humain, même expert). On voit donc déjà le potentiel de ce type d'approche en taxonomie. Parmi les variables permettant le mieux d'expliquer ces subtiles différences entre taxons, on trouve des variables morphométriques telles la distance entre les pseudo-yeux qui peut intuitivement jouer un rôle clé dans l'attraction et le positionnement du pollinisateur sur la fleur d'*Ophrys* lors de la pseudo-copulation. Cependant d'autres variables morphométriques telles que la largeur du sépale (un trait étant déjà apparu comme le plus fortement différencié entre taxons lors de nos mesures sur le terrain) et surtout des différences de coloration (*bins* 3 et 5 tels que déterminés par *colordistance*) ou d'agencement spatial de cette coloration (scores du premiers axe d'ACP dérivés des *K=6 clusters* obtenus avec *patternize*) apparaissent comme encore plus important. Ce type d'approche sans *a priori* sur les traits à analyser permettrait donc peut-être de mettre en évidence des traits pouvant passer inaperçus mais étant pertinents d'un point de vue éco-évolutif.

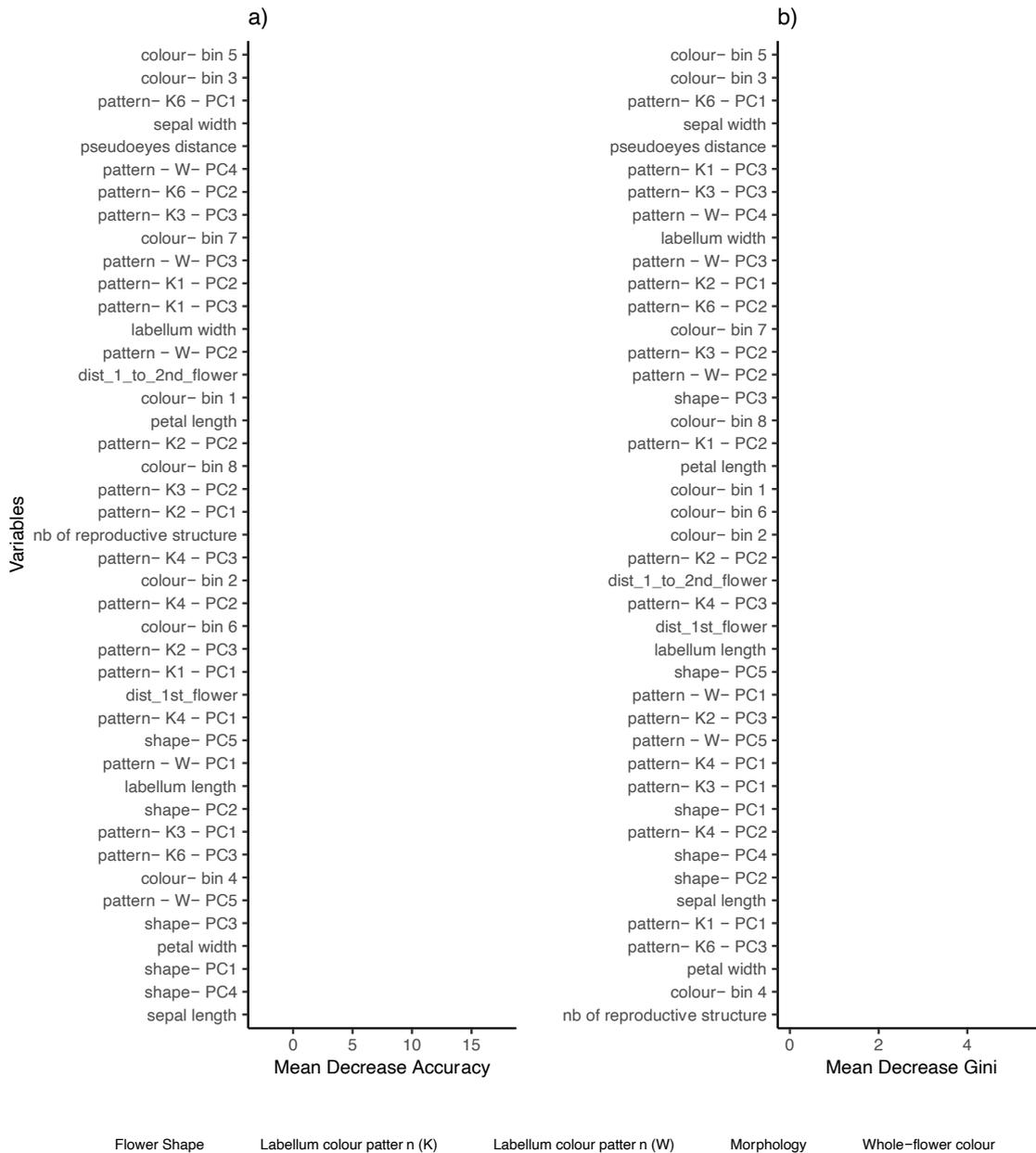


Figure 28 Importance des variables dans la distinction d'*Ophrys aveyronensis* subsp. *aveyronensis* et *O. a.* subsp. *vitorica* sur la base de 1000 arbres avec une approche *random forest* (d'après Gibert *et al.*, 2022, soumis)

Le manuscrit correspondant à cette étude vient d'être accepté dans *Diversity* avec la référence suivante:

Gibert A., Louty F., Buscail R., Baguette M., Schatz B. & Bertrand J.A.M. Extracting meaningful quantitative information from images taken in the wild: a case study of two vicariants of an orchid taxon (*Ophrys aveyronensis* species complex).

Bilan de l'activité d'encadrement

Tableau 5 Bilan des activités d'encadrement d'étudiants réalisées depuis mon recrutement en qualité de Maître de Conférences

Période	Licence 3	Master 1	Master 2	Doctorat
2019			1 (100%)	1 (50%)
2020		1 (100%)		
2021	1 (50%)*	2 (50%)		
2022		4 (33%) + 2 (50%)	1 (50%)	

* stage réalisé hors-cursus.

Activité pédagogique et d'enseignement

Je possède une expérience en enseignement de plus de 1200 heures détaillée dans le tableau ci-dessous et explicitées dans la section ci-après (voir aussi [Annexe 2](#) pour un détail exhaustif de mes enseignements avant mon recrutement : c'est-à-dire en qualité de Doctorant Chargé d'Enseignement, d'ATER et de Maître de Conférences)

Tableau 6 Bilan des enseignements réalisés à ce jour

Poste	Année universitaire	Heures (éq. TD)	Sous-total (éq. TD)	Total (éq. TD)
Doctorant Chargé d'Enseignement	2009-2010	71,5	206,5	1211,25
	2010-2011	64		
	2011-2012	71		
ATER	2012-2013	100	100 (123 en réel)	
Premier Assistant	2016-2017	34,5	34,5	
Maître de Conférences	2017-2018	122,5	870,25	
	2018-2019	173,75		
	2019-2020	182,25		
	2020-2021	192,25		
	2021-2022	199,5		

Expérience d'enseignement et d'encadrement antérieure à mon recrutement en tant que Maître de Conférences

J'ai d'abord enseigné (> 300 h) en qualité de Doctorant Chargé d'Enseignement et en tant qu'Attaché Temporaire d'Enseignement et de Recherche (½ poste, voir détail en [Annexe 2](#)). Puis, en tant que Premier Assistant à l'Université de Lausanne (Suisse), jusqu'à 30% de mon temps de travail pouvait être dédié à l'enseignement. Au cours de ces différentes expériences, j'ai été amené à enseigner en français et anglais, aussi bien en TD qu'en TP mais aussi à évaluer des étudiants de différents niveaux du cursus universitaire (de la L1 au M2) et dans des disciplines variées. Cette expérience, enrichissante à plus d'un titre, m'a notamment permis de consolider mon socle de connaissances, mais aussi d'améliorer mon aptitude à la pédagogie et, plus généralement, à la communication. Cette activité m'a aussi confronté aux difficultés inhérentes aux missions de l'enseignant et a nourri un dialogue formateur avec les étudiants ainsi que les membres des différentes équipes pédagogiques.

Les travaux menés au cours de mon master, de mon doctorat et de mes post-docs m'ont aussi donné l'opportunité de contribuer à l'encadrement plus ou moins officiel d'au moins 7 étudiants : Gaston CHASSAGNARD (2010), Lucie GARNIER et Roxanne LEBERGER

(2011), Myriam GOUDET et Boris DELAHAIE (2012) à Toulouse, puis Caroline COSNARD et Alberto GARCÍA JIMÉNEZ (2018) à Lausanne. Cet encadrement a principalement consisté en leur formation au travail de laboratoire (extraction, dosage et amplification de l'ADN), à l'acquisition ainsi qu'à l'analyse des données (propriétés des séquences, phylogénie, phylogéographie, génétique des populations, biostatistiques...) et à l'assistance dans la rédaction du mémoire de stage. L'encadrement s'est accompagné de multiples échanges m'ayant notamment permis de mieux appréhender la gestion d'un projet de recherche et le travail en équipe depuis le *design* expérimental jusqu'à la publication d'un travail achevé.

Expérience d'enseignement et d'encadrement depuis mon recrutement en tant que Maître de Conférences

Ma prise de fonction au 01/12/2017 fait que je n'ai pas enseigné au premier semestre de l'année universitaire 2017-2018. J'ai enseigné 122,5 h (éq. TD) en 2018, 173,75h (éq. TD) en 2018-2019, 182,25h (éq. TD) en 2019-2020 et 192,25h (éq. TD) en 2020-2021. Au cours de ces quatre années universitaires, j'ai principalement été amené à enseigner en Biologie Moléculaire (L1 et L2), en Génétique et en Dynamique des Populations (L3), en Outils Moléculaires d'Analyse de la Biodiversité (M1) en Modélisation des Systèmes d'Information Géographique (M1) et en Introduction à l'Écologie et l'Évolution (L1). Mon arrivée dans une maquette déjà en cours avec des enseignements déjà en place a fait qu'il a parfois été difficile d'assurer un service complet (de 192h éq. TD) et d'avoir un volume horaire reportable et reporté d'une année à l'autre.

Sur la maquette d'enseignement en cours (depuis la rentrée de septembre 2021), je suis désormais responsable, ou co-responsable de 5 U.E. Cette nouvelle offre de formation m'a globalement permis de recentrer mes enseignements sur des disciplines dans lesquelles je me sens plus à l'aise et qui sont en tout cas davantage en adéquation avec mes activités de recherche. Par ailleurs, je suis membre du jury restreint du Master Biologie, Écologie, Évolution parcours Biodiversité et Développement Durable en sachant que j'avais déjà intégré le comité de sélection des candidatures en juin 2021.

Tableau 7 Enseignements dispensés sur la maquette en cours

Unité d'enseignement	Code U.E.	Niveau/ Semestre	Volume horaire
Introduction à l'évolution	TLPV1IE1	L1/S1	4,5h CM
R et informatique pour les scientifiques : introduction	TLNF2RI1	L1/S2	27h TD
Introduction à l'écologie	TLPC2IE1	L1/S2	18h TD
Écologie des interactions biotiques	TLSC5IT1	L3/S5	8h TP
Projet intégratif en écologie	TLSC6PE1	L3/S6	3h CM + 10,5h TP
Écologie moléculaire & Génétique des Populations	TLSC6EM1/TLSC6ET1	L3/S6	3h CM + 3 x 3h TP
Évolution/Génomique évolutive & génomique des populations	TMBD1U21/ TMBD1GE1	M1/S7	3h CM + 1,5h TD
Systèmes d'Information Géographique (SIG)	TMBD1SI1	M1/S7	2 x 24h TP
Junior Lab	TMTU1JL1	M1/S8	13h TD
Structure & Dynamique des génomes	TMBI2DG1	M1/S8	6,5h TP
Analyse de la diversité intraspécifique	TMBD2AG1	M1/S8	3h CM + 3h TD + 2 x 3h TP
Suivi temporels des populations	TMBD2ST1	M1/S8	4,5h CM + 15h TP

Posture d'enseignement

J'ai vécu le cursus universitaire français, d'abord en tant qu'étudiant, puis en tant qu'enseignant. Mes expériences à l'étranger m'ont également permis de confronter ma vision à celle de pays de culture asiatique (Taiwan), anglo-saxonne (Australie) et certes plus européenne (Suisse) mais tout de même assez différente du système français. Se faisant, j'ai muri une posture d'enseignement personnelle. D'un point de vue historique, l'université a toujours eu pour mission de dispenser des savoirs fondamentaux en formant par, et souvent, pour la recherche. Cette composante reste à mon sens à la base de l'enseignement universitaire mais celui-ci se doit également de prendre en compte un certain nombre de facteurs additionnels afin de conserver tout son sens dans le contexte actuel. Il convient par exemple de considérer que tous les étudiants n'aspirent pas à faire de la recherche. Il serait aussi illusoire de penser que tous les étudiants du cursus de licence et dans une moindre mesure, de celui Master, ont une idée claire et réfléchie de leur propre orientation. Mais cela ne m'apparaît pas comme un problème si on considère qu'en complément d'un enseignement de savoirs fondamentaux, ce qu'enseigne l'université, c'est l'acquisition d'une démarche intellectuelle. Cette démarche s'accompagne d'un développement de compétences faisant notamment intervenir la rigueur, l'autonomie ou encore la capacité à construire un raisonnement étayé, autant de qualités indéniables dans toute sphère professionnelle et même personnelle. En tant qu'enseignant, j'aspire donc plus à être un accompagnant avec lequel on dialogue plutôt qu'une autorité dont le rôle est de dispenser une éducation de manière unilatérale. La posture d'enseignement à laquelle j'adhère s'articule notamment autour de trois points principaux.

Susciter et entretenir la curiosité des étudiants

Bien des étudiants arborent au premier abord une attitude qu'on pourrait qualifier de passive. Une timidité naturelle, une incertitude quant à leur orientation, ou un besoin d'adaptation au mode de fonctionnement de l'université sont autant de causes possibles pour expliquer ces comportements, somme toute normaux. Cette attitude aura d'ailleurs souvent tendance à s'estomper au fur et à mesure que l'étudiant progresse dans son cursus. C'est en particulier au début du cursus de licence que l'enseignant doit susciter puis entretenir la curiosité de son auditoire. En sciences naturelles, les exemples concrets abondent mais gagneraient certainement à être davantage encore expliqués avec des perspectives plus écologiques et évolutives (en salle et sur le terrain). Ainsi, je crois que des exemples judicieusement choisis peuvent être successivement traités par les différents acteurs de l'équipe pédagogique, en biologie animale ou végétale, puis en écologie fonctionnelle et évolutive pour décroiser et favoriser les passerelles entre les différentes matières dans l'esprit de l'étudiant. Cela lui permet ainsi de mieux intégrer d'entrée les différentes échelles du vivant (e.g. gène, individu, population, écosystème...) en apprenant à les considérer simultanément avec différents points de vue.

Faire en sorte que les étudiants soient acteurs de leur propre apprentissage

Mon expérience m'a aussi permis de constater que les étudiants sont plus confiants et motivés sitôt qu'ils sont acteurs de leur propre apprentissage. Cet élément revêt donc une importance capitale pour entretenir la curiosité évoquée dans le paragraphe précédent. Le fait d'impliquer

l'étudiant dans cette tâche est aussi un bon moyen de favoriser le dialogue avec l'enseignant. A ce titre, se baser sur les exemples vus en cours pour les approfondir *via* une pédagogie par projets est un bon moyen de favoriser les échanges entre l'étudiant et l'enseignant-chercheur. Ces projets peuvent s'articuler autour de travaux de recherche en leur rajoutant éventuellement une dimension en lien avec des enjeux sociétaux actuels (e.g. agroécologie et agriculture productiviste, conservation des grands prédateurs dans un contexte socio-économique tendu...). Ce qui me paraît intéressant dans cette démarche, c'est que l'étudiant apprend à synthétiser et communiquer autour d'un consensus scientifique pour le confronter à un contexte socio-politique réel dans lequel des compromis doivent être trouvés. Tout au long de ce travail, les étudiants sont amenés à faire état de l'avancement de leurs projets *via* des présentations devant leurs camarades et l'équipe enseignante. Cet exercice les fait donc passer de l'autre côté de la barrière en les mettant en situation de communicants (cohérent avec une pédagogie dite "inversée", voire "renversée").

Développer chez les étudiants une démarche scientifique et des compétences transversales

En développant leurs projets, les étudiants acquièrent des compétences diverses. D'abord ils développent une démarche scientifique. Accompagnés par l'enseignant, ils dégagent une problématique et définissent la stratégie à adopter. Ils doivent ensuite réaliser des recherches bibliographiques, s'adresser à des personnes susceptibles de les renseigner et éventuellement collecter des données et mener leur propre expérimentation. Au cours de ce projet, les étudiants travaillent par petits groupes et développent leur capacité à coopérer. Puis, les données devront être analysées, les résultats synthétisés et le travail doit se concrétiser sous la forme d'un rapport écrit et d'une présentation orale finale. En outre, il est bénéfique d'envisager la valorisation des projets sous d'autres formats et en employant d'autres médias de communication (e.g. clip vidéo ou exposition, blogs et diffusion sur les réseaux sociaux). Au final et en parallèle de l'acquisition d'un savoir, les étudiants développent donc un certain nombre de compétences transversales qui leur seront utiles dans une carrière de chercheur, d'enseignant ou dans toute autre voie professionnelle. Les aspects ayant trait à la professionnalisation peuvent d'ailleurs être concertés en amont entre l'équipe enseignante et différents acteurs non-académiques. Les enseignants peuvent ainsi se charger de consulter les intervenants extérieurs afin de faire remonter des sujets d'étude à proposer en début d'année.

De ce que j'ai perçu, la majorité des étudiants ne se lance pas dans un cursus de sciences naturelles par dépit, mais bel est bien animée par une fibre naturaliste préalable qui ne demande qu'à être affirmée. Les jeunes générations sont aussi de plus en plus habituées à jongler avec différents médias de communications reposant sur une interaction entre l'image et les réseaux sociaux. Cela ouvre de nouvelles perspectives quant à la valorisation des travaux tout en visant à optimiser un couple clarté/esthétisme pour gagner en efficacité. Cet aspect est souvent perçu comme facultatif voire superflu par bon nombre de nos confrères. A titre personnel, j'y vois une opportunité intéressante pour dépoussiérer l'image que le grand public peut avoir du milieu académique. J'essaye donc toujours de travailler et améliorer les supports visuels que ce soit en tant que chercheur lors de la présentation et la communication autour de mon travail, d'enseignant lors de la présentation des exemples traités et même dans des sphères plus personnelles en tant que "blogueur-photographe" et communicateur scientifique.

Bilan préliminaire

Au terme de mes cinq premières années universitaires passées à enseigner au rythme relativement soutenu qui est celui des maître de conférences dans le système français, je me permets une première prise de recul sur cette posture d'enseignement que j'avais exposé à peu près en ces termes lors de mes candidatures à plusieurs postes de maître de conférences, dont celui sur lequel j'ai été recruté. J'ai fait de mon mieux pour mettre en œuvre ma posture d'enseignement en essayant de décloisonner mes enseignements de deux façons, l'une assez conceptuelle, l'autre plus méthodologique. Ainsi, j'ai mis à profit mes enseignements de M1 en « Modélisation des Systèmes Biologiques » et en « Systèmes d'Information Géographique (SIG) » pour faire travailler les étudiants en petits groupes de 3 ou 4 sur des projets de modélisation de niche écologique en utilisant des outils communs (QGIS et R) avant d'effectuer une restitution à l'oral. Cette initiative a été bien perçue car elle aidait à décloisonner les connaissances théoriques acquises dans différentes matières. C'est en tout cas ce qui en a été dit par les représentants des étudiants lors du conseil de perfectionnement de la formation. D'une manière plus générale, une utilisation de R dans des disciplines autres que les biostatistiques, par exemple en génétique des populations, permet de limiter l'utilisation de logiciels très (voire trop) spécifiques et parfois obsolètes (e.g. Arlequin et autres applets java en génétique des populations). Le succès de ces initiatives est cependant à nuancer. D'abord parce que si j'ai pu les proposer à des effectifs de M1 d'une vingtaine d'étudiants, ces enseignements seraient dans les faits beaucoup plus difficiles à mettre en place de façon aussi efficace en cursus de Licence. Ensuite, parce ces changements conceptuels et méthodologiques se heurtent généralement au conservatisme de certains collègues qui ne veulent pas, ou ne partagent pas notre enthousiasme vis-à-vis de cette façon d'enseigner.

Les restrictions liées à la crise sanitaire du COVID-19 a aussi catalysé le passage à différentes pédagogies dites innovantes. En effet, les périodes de confinement nous ont obligé à repenser nos enseignements de sorte qu'ils puissent être dispensés en tout distanciel, puis en co-modal avant de revenir en présentiel. Pour ce faire, je me suis, comme bon nombre de mes collègues largement appuyé sur des outils permettant d'assurer des enseignements en visioconférence (e.g. Zoom). La plateforme Moodle de l'espace numérique de travail a aussi permis non-seulement de partager les supports pédagogiques, mais aussi de permettre la remise de travaux en ligne de même que d'élaborer des quizz de révisions, voire des examens en ligne. Depuis le retour à la normale relatif, j'essaye de continuer sur cette lancée. A l'initiative de Martin ROSALIE, un collègue du laboratoire enseignant au département de mathématiques-informatique, j'ai pu par exemple enseigner dans un module de L1 de découverte de R pour les scientifiques dans lesquels les TD sont entièrement effectués sur Moodle. De la liste de présence à la réponse aux questions de l'énoncé, les étudiants répondent en ligne en entrant directement leurs lignes de code dans la console. Ces réponses aux questions sont ensuite automatiquement corrigées (tout ça via le *plug-in CodeRunner*) ce qui permet un maximum d'interactivité avec le langage de programmation pour les étudiants, et un gain de temps considérable pour les enseignants. Certes, des améliorations restent à apporter mais les premiers retours des étudiants sont encourageants. Par ailleurs, on peut citer l'utilisation d'outils comme *kahoot!* ou *wooclap* que je n'ai pas encore personnellement employé en cours magistraux mais qui permettent aux étudiants (même les plus réservés) de répondre de façon ludique à des questions via leur Smartphone, avant de procéder à une évaluation collective (mais potentiellement anonyme) des réponses.

Références

- Abbott R.J. (2017) Plant speciation across environmental gradients and the occurrence and nature of hybrid zones. *Journal of Systematics & Evolution*, **55**, 238-258.
- Abreu J.A., Hawkins J.A., Cotrim H., Fay M.F., Hidalgo O. & Pellicer J. (2017) *Ophrys fusca* and *Ophrys dyris* (Orchidaceae) – constancy of tetraploidy amongst populations in Central Portugal. *New Journal of Botany*, **7**, 94-100.
- Adams D., Collyer M., Kaliontzopoulou A. & Baken E. (2021) *Geomorph*: software for geometric morphometric analyses. R package version 4.0: <https://cran.rproject.org/package=geomorph>
- Baguette M., **Bertrand J.A.M.**, Stevens V. & Schatz B. (2020) Why are there so many bee-orchid species? Adaptive radiation by intra-specific competition for mnesic pollinators *Biological Reviews*, **95**, 1630-1663.
- Bateman R.M. Rudall P.J., Murphy A.R.M., Cowan R.S., Devey D.S. & Pérez-Escobar O.A. (2021) Whole plastomes are not enough: phylogenomic and morphometric exploration at multiple demographic levels of the bee orchid clade *Ophrys* sect. *Shegodes*. *Journal of Experimental Botany*, **72**, 654-681.
- Benito Ayuso J. Estudios sobre polinización en el género *Ophrys* (Orchidaceae), I. *Flora Montiberica*, **74**, 32-37.
- Berner D. & Salzburger W. (2015) The genomics of organismal diversification illuminated by adaptive radiation. *Trends in Ecology & Evolution*, **31**, 491-499.
- Bertrand J.A.M.**, García-Jiménez R., Bourgeois Y., Duval T., Heeb P., Thébaud C. & Milá B. (2012) Isolation and characterization of twelve polymorphic microsatellite loci for investigating an extreme case of microgeographical variation in an island bird (*Zosterops borbonicus*). *Conservation Genetics Resources*, **4**, 323-226.
- Bertrand J.A.M.**, Bourgeois Y.X.C., Delahaie B., Duval T., García-Jiménez R., Cornuault J., Heeb P., Milá B., Pujol B. & Thébaud C. (2014) Extremely reduced dispersal and gene flow in an island bird. *Heredity*, **112**, 190-196.
- Bertrand J.A.M.**, Bourgeois Y.X.C. & Thébaud C. (2016) Population density of the Réunion Grey White-eye (*Zosterops borbonicus*) within the summit ecosystems of Réunion (Mascarene Islands). *Ostrich*, **87**, 85-88.
- Bertrand J.A.M.**, Delahaie B., Bourgeois Y.X.C., Duval T., García-Jiménez R., Cornuault J., Pujol B., Thébaud C. & Milá B. (2016b) The role of selection and historical factors in driving population differentiation along an elevational gradient in an island bird. *Journal of Evolutionary Biology*, **29**, 824-836.
- Bertrand J.A.M.**, Gibert A., Llauro C. & Panaud O. (2019). Characterization of the complete plastome of *Ophrys aveyronensis*, a Euro-Mediterranean orchid with an intriguing disjunct geographic distribution. *Mitochondrial DNA Part B*, **4**, 3256-3257.
- Bertrand J.**, Baguette M., Joffard N. & Schatz B. (2021) Les défis inhérents à la systématique et à la taxonomie de genres ayant connu une radiation explosive récente: le cas des orchidées du genre *Ophrys*. Dans *Systématique et exploration du vivant* (eds. M.-C. Maurel & P. Grandcolas). ISTE, Paris, pp 115-135.
- Bertrand J.**, Baguette M., Joffard N. & Schatz B (2021) Challenges inherent in the systematics and taxonomy of genera that have recently experienced explosive radiation: the case of Orchids of the genus *Ophrys*. In *Systematics and the Exploration of Life* (eds. M.-C. Maurel & P. Grandcolas). ISTE/(Wiley), Paris, pp 113-134.

- Bertrand J.A.M.**, Gibert A., Llauro C. & Panaud O. (2021) Whole plastid genome-based phylogenomics supports an inner placement of the *O. insectifera* group rather than a basal position in the rapidly diversifying *Ophrys* genus (Orchidaceae). *Botany Letters*, **168**, 452-457.
- Besnard G., **Bertrand J.A.M.**, Delahaie B., Bourgeois Y.X.C., Lhuillier E. & Thébaud C. (2016) Valuing museum specimens: high-throughput DNA sequencing on historical collections of New Guinea crowned pigeons (*Goura*). *Biological Journal of the Linnean Society*, **117**, 71-82.
- Bohman B., Weinstein A.M., Mozuraitis R., Flematti G.R. & Borg-Karlson A.-K. (2020) Identification of (Z)-8-Heptadecene and n-Pentadecane as electrophysiologically active compounds in *Ophrys insectifera* and its *Argogorytes pollinator*. *International Journal of Molecular Sciences*, **21**, 620.
- Bonhomme V., Picq S., Gaucherel C. & Claude J. (2014) Momocs: Outline analysis using R. *Journal of Statistical Software*, **56**, 1-24.
- Borg-Karlson A.-K., Groth I., Ågren L. & Kullenberg B. (1993) Form-specific fragrances from *Ophrys insectifera* L. (Orchidaceae) attract species of different pollinator genera. Evidence of sympatric speciation? *Chemoecology*, **4**, 39-45.
- Bou Dagher-Kharrat M., Abdel-Samad N., Douaihy B., Bourge M, Fridlender A., Siljak-Yakovlev S. & Brwon S.C. (2013) Nuclear DNA C-values for biodiversity screening: Case of the Lebanese flora. *Plant Biosystems*, **147**, 1228-1237.
- Bourgeois Y.X.C., **Bertrand J.A.M.**, Thébaud C. & Milá B. (2012) Investigating the role of the Melanocortin-1 receptor gene in an extreme case of microgeographical variation in the pattern of melanin-based plumage pigmentation. *PLoS ONE*, **7**: e50906.
- Bourgeois Y.X.C., Lhuillier E., Cézard T., **Bertrand J.A.M.**, Delahaie B., Cornuault J., Duval T., Bouchez O., Milá B. & Thébaud C. (2013) Mass production of SNP markers in a non-model passerine bird species through RAD-sequencing and contig mapping to the zebra finch genome. *Molecular Ecology Resources*, **13**, 899-907.
- Bourgeois Y.X.C., **Bertrand J.A.M.**, Delahaie B., Cornuault J., Duval T., Milá B. & Thébaud C. (2016) Candidate gene analysis suggests untapped genetic complexity in melanin-based pigmentation in birds. *Journal of Heredity*, **107**, 327-335.
- Bourgeois Y.X.C, Delahaie B., Gautier M., Lhuillier E., Malé P.-J.G., **Bertrand J.A.M.**, Cornuault J., Wakamatsu K., Bouchez O., Mould C., Bruxaux J., Holota H., Milá B. & Thébaud C. (2017) A novel locus on chromosome 1 underlies the evolution of a melanic plumage polymorphism in a wild songbird. *Royal Society Open Science*, **4**: 160805.
- Bourgeois Y.X.C., **Bertrand J.A.M.**, Delahaie B., Holota H., Thébaud C. & Milá B. (2020) Differential divergence in autosomes and sex chromosomes is associated with intra-island diversification at a very small spatial scale in a songbird lineage. *Molecular Ecology*, **29**, 1137-1153.
- Bradley D., Xu P., Mohorianu I.-I., Whibley A., Field D., Tavares H., Couchman M., Copsey L., Carpenter R. Li M., Li Q. Xue Y., Dalmay T. & Coen E. (2017) Evolution of flower color pattern through selection on regulatory small RNAs. *Science*, **358**, 925-928.
- Breitkopf H., Onstein R.E., Cafasso D., Schlüter P.M. & Cozzolino S. (2015) Multiple shifts to different pollinators fuelled rapid diversification in sexually deceptive *Ophrys* orchids. *New Phytologist*, **207**, 377-389.
- Bromham L., Hua, X., Lanfear R. & Cowman P.F. (2015) Exploring the relationships between mutation rates, life history, genome size, environment, and species richness in flowering plants. *American Naturalist*, **185**, 507-524.

- Byers K.J.R.P., Xu S. & Schlüter P.M. (2017) Molecular mechanisms of adaptation and speciation: why do we need an integrative approach? *Molecular Ecology*, **26**, 277-290.
- Campbell M.A., Ross Robertson D., Vargas M.I., Allen G.R. & McMillan W.O. (2018) Multilocus molecular systematics of the circumtropical reef-fish genus *Abudefduf* (Pomacentridae): history, geography and ecology of speciation. *PeerJ*, **6**, e5357.
- Catchen J., Amores A., Hohenlohe P., Cresko W. & Postlethwait K. (2011) *Stacks*: building and genotyping *loci de novo* from short-read sequences. *G3: Genes, Genomes Genetics*, **1**, 171-182.
- Catchen J., Hehenlohe S., Amores A. & Cresko W. (2013) *Stacks*: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124-3140.
- Chan IZW, Stevens M & Todd PA (2019) PAT-GEOM: a software package for the analysis of animal patterns. *Methods in Ecology & Evolution*, **10**, 591-600.
- Cheon S., Zhang J. & Park C. (2020) Is phylotranscriptomics as reliable as phylogenomics? *Molecular Biology & Evolution*, **37**, 3672-3683.
- Chu C., Nielsen R. & Wu Y. (2016) REPdenovo: Inferring *De Novo* repeat motifs from short sequence reads. *Plos ONE* **11**, e0150719.
- Coleman R.C., Gaither M.R., Kimokeo B., Stanton F.G., Bowen B.W. & Toonen R.J. (2014) Large-scale introduction of the Indo-Pacific damselfish *Abudefduf vaigiensis* into Hawai'i promotes genetic swamping of the endemic congener *A. abdominalis*. *Molecular Ecology*, **23**, 5552-5565.
- Cortis P., Vereecken N.J., Schiestl F.P., Barone Lumaga M.R., Scrugli A. & Cozzolino S. (2009) Pollinator convergence and the nature of species' boundaries in sympatric Sardinian *Ophrys* (Orchidaceae). *Annals of Botany*, **104**, 497-506.
- Dalongeville A., Benestan L., Mouillot D., Lobreaux S. & Manel S. (2018) Combining six genome scan methods to detect candidate genes to salinity in the Mediterranean striped red mullet (*Mullus surmuletus*). *BMC Genomics*, **19**, 217.
- Delahaie B., Cornuault J., Masson C., **Bertrand J.A.M.**, Bourgeois Y.X.C., Milá B. & Thébaud C. (2017) Narrow hybrid zones in spite of very low population differentiation in neutral markers in an island bird species complex. *Journal of Evolutionary Biology*, **30**, 2132-2145.
- Delforge P. (1984) L'Ophrys de l'Aveyron. *L'Orchidophile*, **15**, 577-583.
- Derryberry E.P., Derryberry G.E., Maley J.M. & Brumfield R.T. (2014) HZAR: hybrid zone analysis using an R software package. *Molecular Ecology Resources*, **14**, 652-663.
- Dufresnes C., Brelsford A., Jeffries D.L., Mazepa G., Suchan T., Canestrelli D., Nicieza A., Fumagalli L., Dubey S., Martínez-Solano I., Litvinchuk S.N., Vences M., Perrin N. & Crochet P.-A. (2021) Mass of genes rather than master genes underlie the genomic architecture of amphibian speciation. *Proceedings of the National Academy of Science of the USA*, **118** e2103963118.
- Eserman L.A., Thomas S.K., Coffey E.E.D & Leebens-Mack J.H. (2021) Target sequence capture in orchids: developing a kit to sequence hundreds of single-copy loci. *Applications in Plant Sciences*, **9**, e11416.
- Frichot E., Mathieu F., Trouillon T., Bouchard G. & François O. (2014) Fast and efficient estimation of individual ancestry coefficients. *Genetics*, **196**, 973-983.
- Frichot E. & François O. (2015) LEA: An R package for landscape and ecological association studies. *Methods in Ecology & Evolution*, **6**, 925-929.
- Gavrilets S. & Losos J.B. (2009) Adaptive radiation: contrasting theory with data. *Science*, **6**, 732-737.

- Gervasi D.D.L., Selosse, M.-A., Sauve M., Francke W., Vereecken N.J., Cozzolino S. & Schiestl F.P. (2017) Floral scent and species divergence in a pair of sexually deceptive orchids. *Ecology and Evolution*, **7**, 6023-6034.
- Gill F.B. (1973) Intra-island variation in the Mascarene White-Eye *Zosterops borbonica*. *Ornithological Monographs*, **12**, 1-66.
- Grabherr M.G., Hass, B.J., Yassour M., Levin, J.Z., Thompson D.A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z., Mauveli E., Hacohen N., Gnirke A., Rhind N., Di Palma F., Birren B.W., Nusbaum C., Linblad-Toh K., Friedman N. & Regev A. (2011) Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnology*, **29**, 644-652.
- Guo R., Li Y.-R., He S., Ou-Yang L., Sun Y. and Zhu Z. (2018) RepLong: de novo repeat identification using long read sequencing data. *Bioinformatics*, **34**, 1099-1107.
- Hermosilla C.E. & Soca R. (1999) Distribution of *Ophrys aveyronensis* (J.J. Wood) Delforge and survey of its hybrids. *Caesiana*, **13**, 31–38.
- Jiggins C.D. (2014) Evolutionary biology: radiating genomes. *Nature*, **513**, 318-319.
- Jin W.-T., Schuiteman A., Chase M.W., Li J.-W., Chung S.-W., Hsu T.-C., Jin X.-H. (2017) Phylogenetics of subtribe Orchidinae s.l. (Orchidaceae; Orchidoideae) based on seven markers (plastid *matK*, *psaB*, *lrbcl*, *ltnrL-F*, *trnH-psba* and nuclear nrITS, *Xdh*): implications for generic delimitation. *BMC Plant Biology*, **17**, 222.
- Joffard N., Buatois B. & Schatz B. (2016) Integrative taxonomy of the fly orchid group: insights from chemical ecology. *The Science of Nature*, **103**, 77.
- Joffard N., Massol F., Grenié M., Montgelard C. & Schatz B. (2019) Effects of pollination strategy, phylogeny and distribution on pollination niches of Euro-Mediterranean orchids. *Journal of Ecology*, **107**, 478-490.
- Johnson M.G., Pokorny L., Dodsworth S., Botigué L.R., Cowan R.S., Devault A., Eiserhardt W.L., Epitawalage N., Forest F., Kim J.T., Leebens-Mack J.H., Leitch I.J., Maurin O., Soltis D.E., Soltis P.S., Wong G. K.S., Baker W.J. & Wickett N.J. (2019) A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematics Biology*, **68**, 594-606.
- Khimoun A., Burrus M., Andalo C. Liu Z.-L., Vicédo-Cazettes C. Thébaud C & Pujol. B. (2011) Locally asymmetric introgressions between subspecies suggest range expansion at the *Antirrhinum majus* global scale. *Journal of Evolutionary Biology*, **24**, 1433-1441.
- Khimoun A., Burrus M., Cornuault J., Burrus M., Pujol B., Thébaud C. & Andalo C. (2013) Ecology predicts parapatric distributions in two closely related *Antirrhinum majus* subspecies. *Evolutionary Ecology*, **27**, 51-64.
- Knight C.A., Molinari N.A. & Petrov D.A. (2005) The large genome constraint hypothesis: evolution, ecology and phenotype. *Annals of Botany*, **95**, 177-190.
- Kraaijeveld K. (2010) Genome size and species diversification. *Evolutionary Biology*, **37**, 227-233.
- Kreutz C.A. (2007) Beitrag zur Taxonomie und Nomenklatur europäischer, mediterraner, nordafrikanischer und vorderasiatischer Orchideen. *Ber. Arbeitskrs. Heim. Orchid.*, **24**, 77-141.
- Lavoué S. **Bertrand J.A.M.**, Chen, W.-J., Ho, H.-C., Motomura H., Sado T. Miya M. (2017) Phylogenetic position of the rainbow sardine *Dussumieria* (Dussumieriidae) and its bearing on the early evolution of the Clupeoidei. *Gene*, **623**, 41-47.

- Lavoué S., **Bertrand J.A.M.**, Wuang H.-Y., Chen W.-J., Ho H.-C. H., Motomura H., Hata H., Sado T. & Miya M. (2017) Molecular systematics of the anchovy genus *Encrasicholina* in the Northwest Pacific. *PLoS ONE*, **12**, e0181329.
- Lavoué S., **Bertrand J.A.M.**, Shen K.-N., Ratmuangkhwang S., Sado T., Miya M. and Mood Nor S.A. (2019). Deep genetic differentiation between two morphologically similar species of wolf herrings (Teleostei, Clupeoidei, Chirocentridae). *Journal of Applied Ichthyology*, **35**, 693-700.
- Leitch I.J., Kahandawala I., Suda J., Hanson L., Ingrouille M.J., Chase M.W. and Fay M.F. (2009) Genome size diversity in orchids: consequences and evolution. *Annals of Botany*, 469-481.
- Liaw A. & Wiener M.R. (2002) Classification and regression by random forest. *R News*, **2**, 18-22.
- Lisch D. (2013) How important are transposons for plant evolution? *Nature Reviews Genetics*, **14**, 49-61.
- Litsios G., Sims C.A., Wüest R.O., Bearman P.B., Zimmermann N.E. & Salamin N. (2012) Mutualism with sea anemones triggered the adaptive radiation of clownfishes. *BMC Evolutionary Biology*, **12**, 212.
- Litsios G., Bearman P.B., Lanterbecq D., Toulou N. & Salamin N. (2014) The radiation of the clownfishes has two geographical replicates. *Journal of Biogeography*, **41**, 2140-2149.
- Litsios G., Kostikova A. & Salamin N. (2014) Host specialist clownfishes are environmental niche generalists. *Proceedings of the Royal Society B*, **281**, 20133220.
- Litsios G. & Salamin N. (2014) Hybridisation and diversification in the adaptive radiation of clownfishes. *BMC Evolutionary Biology*, **14**, 245.
- Loiseau O., Olivares I., Paris M., de La Harpe M., Weigand A., Koubínová D., Rolland J., Bacon C.D., Balslev H., Borchsenius F., Cano A., Couvreur T.L.P., Delnatte C., Fardin F., Gayot M., Mejía F., Mota-Machado T., Perret M., Roncal J., Sanin M. J., Stauffer F., Lexer C. Kessler M. & Salamin N. (2019) Targeted capture of hundreds of nuclear genes unravels phylogenetic relationships of the diverse Neotropical palm tribe Geonomateae. *Frontiers in Plant Science*, **10**, 864.
- López-Pujol J., Orellana M.R., Bosch M., Simon J. & Blanché C. (2007) Low genetic diversity and allozymic evidence for autopolyploidy in the tetraploid Pyrenean endemic larkspur *Delphinium montanum* (Ranunculaceae) *Biological Journal of the Linnean Society*, **155**, 211-222.
- Lou R.N., Jacobs A., Wildr A.P. & Overgaard Therkildsen N. (2021) A beginner's guide to low-coverage whole genome sequencing for population genomics. *Molecular Ecology*, **30**, 5966-5993.
- Luu K., Bazin E. & Blum M.G.B. (2016) *pcadapt*: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, **17**, 67-77.
- Mank J.E. & Avise J.C. (2006) Cladogenetic correlates of genomic expansions in the recent evolution of actinopterygian fishes. *Proceedings of the Royal Society of London B*, **273**, 33-38.
- Marcionetti A., Rossier V., **Bertrand J.A.M.**, Litsios G. & Salamin N. (2018) First draft genome assembly of an iconic clownfish species (*Amphiprion frenatus*). *Molecular Ecology Resources*, **18**, 1092-1101.

- Marcionetti A., Rossier V., Roux N., Salis P., Laudet V. & Salamin N. (2019) Insights into the genomics of clownfish adaptive radiation: genetic basis of the mutualism with sea anemones. *Genome Biology & Evolution*, **11**, 869-882.
- Meier J.I., Salazar P.A., Kučka M., William Davies R., Dréau A., Aldás I., Box Power O., Nadeau N.J., Bridle J.R., Rolian C., Barton N.H., McMillan W. O., Jiggins C.D. & Chan Y.F. (2021) Haplotype tagging reveals parallel formation of hybrid races in two butterfly species. *Proceedings of the National Academy Of Sciences of the USA*, **118**, e201500518.
- Milá B., Warren B.H., Heeb P. & Thébaud C. (2010) The geographic scale of diversification on islands: genetic and morphological divergence at a very small spatial scale in the Mascarene grey white-eye (*Aves: Zosterops borbonicus*). *BMC Evolutionary Biology*, **10**, 158.
- Nosil P., Feder J.L., Flaxman S.M. & Gompert Z. (2017) Tipping points in the dynamics of speciation. *Nature in Ecology & Evolution*, **1**, 0001.
- Nürk N.M., Linder H.P., Onstein R.E., Larcombe M.J., Hughes C.E., Piñeiro Fernández L., Schlüter P.M., Valente L., Beierkuhnlein C., Cutts V., Donoghue M.J., Edwards E.J., Field R., Flantua S.G.A., Higgins S.I., Jentsch A., Liede-Schumann S. & Pirir M.D. (2020) Diversification in evolutionary arenas – Assessment and synthesis. *Ecology & Evolution*, **10**, 6163-6182.
- Olmo E. (2006) Genome size and evolutionary diversification in vertebrates. *Italian Journal of Zoology*, **73**, 167-171.
- Orr H.A. (2005) The genetic theory of adaptation: a brief history. *Nature Reviews Genetics*, **6**, 119-127.
- Paulus H.F. (2017) Bestäubungs-biologie *Ophrys* in Nordspanien. *Journal Europäischer Orchideen*, **49**, 427-471.
- Pellicer J., Hidago O., Dodsworth S. & Leitch I.J. (2018) Genome size diversity and its impact on the evolution of land plants. *Genes*, **9**, 88, doi.org/10.3390/genes9020088.
- Pellicer J., Fernández P, Fay M.F., Michálková E & Leitch I.J. (2021) Genome size doubling arises from the differential repetitive DNA dynamics in the genus *Heloniopsis* (Melanthiaceae). *Frontiers in Genetics*, **12**, 726211.
- Pérez-Escobar O.A., Dodsworth S., Bogarín D., Bellot S., Balbuena J.A., Schley R.J., Kikuchi I. A., Morris S.K., Epitawalage N., Cowan R., Maurin O., Zuntini A., Arias T., Serna-Sánchez A., Gravendeel B., Torres Jimenez M.F., Nargar K., Chomicki G., Chase M.W., Leitch I.J., Forest F. & Baker W.J. (2021) Hundreds of nuclear and plastid loci yield novel insights into orchid relationships. *American Journal of Botany*, **108**, 1-15.
- Phillips S.J., Anderson R.P., Dudík M, Schapire R.E. & Blair M.E. (2017) Opening the black box: an open-source release of Maxent. *Ecography*, **40**, 887-893.
- Piegu B., Guyot R., Picault N., Roulin A., Saniyal A., Kim H., Collura K., Brar D.S., Jackson S., Wing R.A. & Panaud O. (2006) Doubling genome size without polyploidization: Dynamics of retrotransposition-driven genomic expansion in *Oryza australiensis*, a wild relative of rice. *Genome Research*, **16**, 1262-1269.
- Piñeiro Fernández L., Byers K.J.R.P., Cai J., Sedeek K.E.M., Kellenberger R.T., Russo A., Qi W., Aquino Fournier C. & Schlüter P.M. (2019) A phylogenomic analysis of the floral transcriptomes of sexually deceptive and rewarding European orchids, *Ophrys* and *Gymnadenia*. *Frontiers in Plant Science*, **10**, 1553.
- Puttick M.N., Clark J. & Donoghue C.J. (2015) Size is not everything: rates of genome size evolution, not C-value, correlate with speciation in angiosperms. *Proceedings of the Royal Society B*, **282**, 20152289.

- Quenouille B., Hubert N., Bermingham E. & Planes S. (2011) Speciation in tropical seas: allopatry followed by range change. *Molecular Phylogenetics & Evolution*, **58**, 546-552.
- Rabosky D.L., Grudler M., Anderson C., Title P., Shi J.J., Brown J.W., Huang H. & Larson J.G. (2014) BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology & Evolution*, **5**, 701-707.
- Ricci M., Peona V., Guichard E., Taccioli C. & Boattini A. (2018) Transposable elements activity is positively related to rate of speciation in Mammals. *Journal of Molecular Evolution*, **86**, 303-310.
- Ring J.-P. (2008) Le spectre maculaire, une autre approche de la taxonomie des *Ophrys*. Bulletin de la Société Française d'Orchidophilie Poitou-Charentes-Vendée (SFO PCV). 33 p.
- Roma L., Cozzolino S., Schlüter P.M., Scopece G. & Cafasso D. (2018) The complete plastid genomes of *Ophrys iricolor* and *O. sphegodes* (Orchidaceae) and comparative analyses with other orchids *Plos ONE*, **13**, e204174.
- Sacamoto G.A.T., Kielbassa J., Chikhi R., Uriacaru R., Antoniou P., Sagot P.-F., Peterlongo P. & Lacroix V. (2012) *BMC Bioinformatics*, **13**, S5.
- Salvado P., Aymerich Boixader P., Perera J., Vila Bonfill A., Quélenec C., Lewin, J.-M., Delorme-Hinoux V. & **Bertrand J.A.M.** (2022) Little hope for the polyploid endemic Pyrenean Larkspur (*Delphinium montanum*): Evidences from population genomics and Ecological Niche Modeling. *Ecology & Evolution*, **12**, 8711.
- Salvado P., Lauro C. Carpentier M.-C., Delorme-Hinoux V. & **Bertrand J.A.M.** (2022) Characterization of the complete plastome of *Delphinium montanum*, a polyploid, endemic and endangered Pyrenean Larkspur. *Mitochondrial DNA Part B*, **7**, 590-592.
- Shafer A.B.A, Peart C.R., Tusso S. Maayan I., Brelsford A., Wheat C.W. & Wolf J.B.W. (2016) Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology & Evolution*, **8**, 907-917.
- Schlüter P.M., Xu S., Gagliardini V., Whittle E., Shankin J., Grossniklaus B. & Schiestl F.P. (2011) Stearoyl-acyl carrier protein desaturases are associated with floral isolation in sexually deceptive orchids. *Proceedings of the National Academy of Sciences of the U.S.A.*, **108**, 5696-5701.
- Schlüter P.M. (2018) The magic of flowers or: speciation genes and where to find them. *American Journal of Botany*, **105**, 1957-1961.
- Shrader L. & Smith J. (2018) The impact of transposable elements in adaptive evolution *Molecular Ecology*, **28**, 1537-1549.
- Sedeek K.E.M., Qi W., Schauer M.A., Gupta A.K., Proveda L., Xu S., Liu Z.-J., Grossniklaus U., Schiestl F.P. & Schlüter P.M. (2013) Transcriptome and proteome data reveal candidate genes for pollinator attraction in sexually deceptive orchids. *PLoS One*, **8**, e64621.
- Sedeek K.E.M., Scopece G., Staedler Y.M., Schönenberger J., Cozzolino S., Schiestl F.P. & Schlüter P.M. (2014) Genic rather than genome-wide differences between sexually deceptive *Ophrys* orchids with different pollinators. *Molecular Ecology*, **23**, 6192-6205.
- Sedeek K.E.M., Whittle E., Guthörl, Grossniklaus U., Shanklin J. & Schlüter P.M. (2016) Amino acid change in an orchid desaturase enables mimicry of the pollinator's sex pheromone. *Current Biology*, **26**, 1505-1511.
- Serrano-Serrano M.L., Marcionetti A., Perret M. & Salamin N. (2019) Convergent changes in gene expression associated with repeated transition between hummingbird and bee pollinators. *bioRxiv*, DOI:10.1101/701627.

- Simon J., Bosch M., Molero J. & Blanché C. Conservation biology of the Pyrenean larkspur (*Delphinium montanum*): a case of conflict of plant versus animal conservation? *Biological Conservation*, **98**, 305-314.
- Simonin K.A. & Roddy A.B. (2018) Genome downsizing, physiological novelty and the global dominance of flowering plants. *PLoS Biology*, **16**, e2003706.
- Stankowski S., Sobel, J. & Streisfeld M.A. (2017) Geographic cline analysis as a tool for studying genome-wide variation: a case study of pollinator mediated divergence in a monkeyflower. *Molecular Ecology*, **26**, 107-122.
- Triponez Y., Arrigo N., Pelissier L., Schatz B. & Alvarez N. (2013) Morphological, ecological and genetic aspects associated with endemism in the Fly Orchid group. *Molecular Ecology*, **22**, 1431-1446.
- Valoroso M.C., De Paolo S., Lazetti G. & Aceto S. (2017) Transcriptome-wide identification and expression analysis of *DIVARICATA*- and *RADIALIS*-like genes of the Mediterranean Orchid *Orchis italica*. *Genome Biology & Evolution*, **9**, 1418-1431.
- Van Belleghem SM, Papa R, Ortiz-Zuazaga H, Hendrickx F, Jiggins CD, McMillan WO & Counterman BA (2018) *patternize*: An R package for quantifying colour pattern variation. *Methods in Ecology Evolution*, **9**, 390-398.
- Vereecken N.J., Cozzolino S. & Schiestl F.P. (2010) Hybrid floral scent novelty drives pollinator shift in sexually deceptive orchids *BMC Evolutionary Biology*, **10**, 103.
- Vinogradov A.E. (2003) Selfish DNA is maladaptive: evidence from the plant Red List. *Trends in Genetics*, **19**, 609-614.
- Vinogradov A.E. (2004) Genome size and extinction risk in vertebrates. *Proceedings of the Royal Society of London B*, **271**, 1701-1705.
- Wang W., Schalamun M., Morales-Suarez A., Kainer D., Schwessinger B. & Lanfear R. (2018) Assembly of chloroplast genomes with long- and short-read data a comparison of approaches using *Eucalyptus pauciflora* as a test case. *BMC Genomics*, **19**, 977.
- Wang S.-L., Viswanath K.K., Tong C.-G., An H.R., Jang S. & Chen F.-C. (2019) Floral induction and flower development of Orchids. *Frontiers Plant Science*, **10**, 1258.
- Warren B.H., Bermingham E., Prys-Jones R.P. & Thébaud C (2006) Immigration, species radiation and extinction in a highly diverse songbird lineage: white-eyes on Indian Ocean islands. *Molecular Ecology*, **15**, 3769-3786.
- Weller H.I. & Westneat M.W. (2019) Quantitative color profiling of digital images with earth mover's distance using the R package *colordistance*. *Peer J*, **7**, e6398.
- Whibley A.C. Langlade N., Andalo C., Hanna A.I., Bangham A., Thébaud C. & Coen E. (2006) Evolutionary paths underlying flower color variation in *Antirrhinum*. *Science*, **313**, 963-966.
- Whitlock M.C. & Lotterhos K.E. (2015) Reliable detection of *loci* responsible for local adaptation: inference of a null model through trimming the distribution of F_{ST} . *American Naturalist*, **186**, S24-S36.
- Wick R.R., Judd L.M., Gorrie C.L. & Holt K.E. (2017) Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *Plos Computational Biology*, **13**, e1005595.
- Wood J.J. (1983) Eine neue subspecies von *Ophrys sphegodes* Mill. aus Süden Zentralfrankreichs: *Ophrys sphegodes* Mill. subsp. *aveyronensis*. *Die Orchidee*. **34**, 105-109.
- Yardeni G., Viruel J., Paris M., Hess J., Groot Crego C., de La Harpe M., Rivera N., Barfuss M.H.J., Till W., Guzmán-Jacob V., Krömer T., Lexer C., Paun O. & Leroy T. (2021) Taxon or

universal? Using target capture to study the evolutionary history of rapid radiations.
Molecular Ecology Resources, **22**, 927-945.

Annexe 1

PCA – Biplot

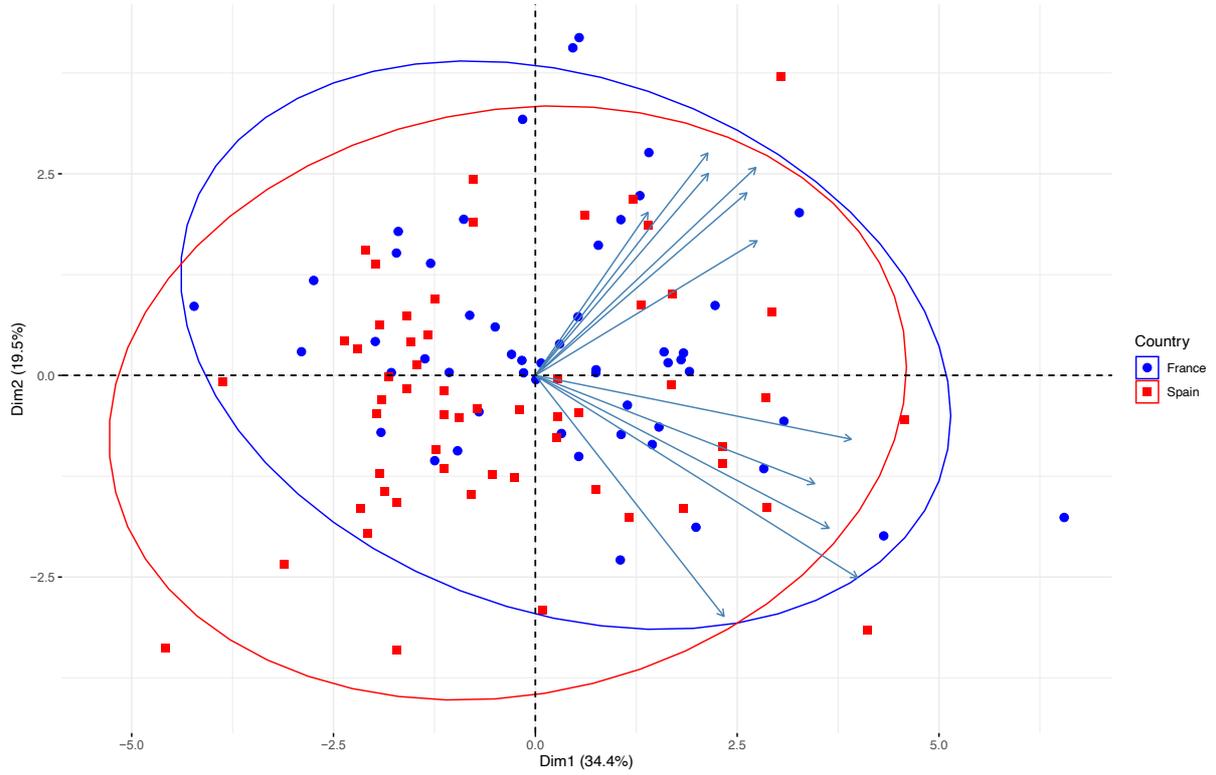


Figure 1 Analyse en Composante Principale (ACP) réalisée sur 11 variables morphométriques sur 109 individus du complexe *O. aveyronensis*

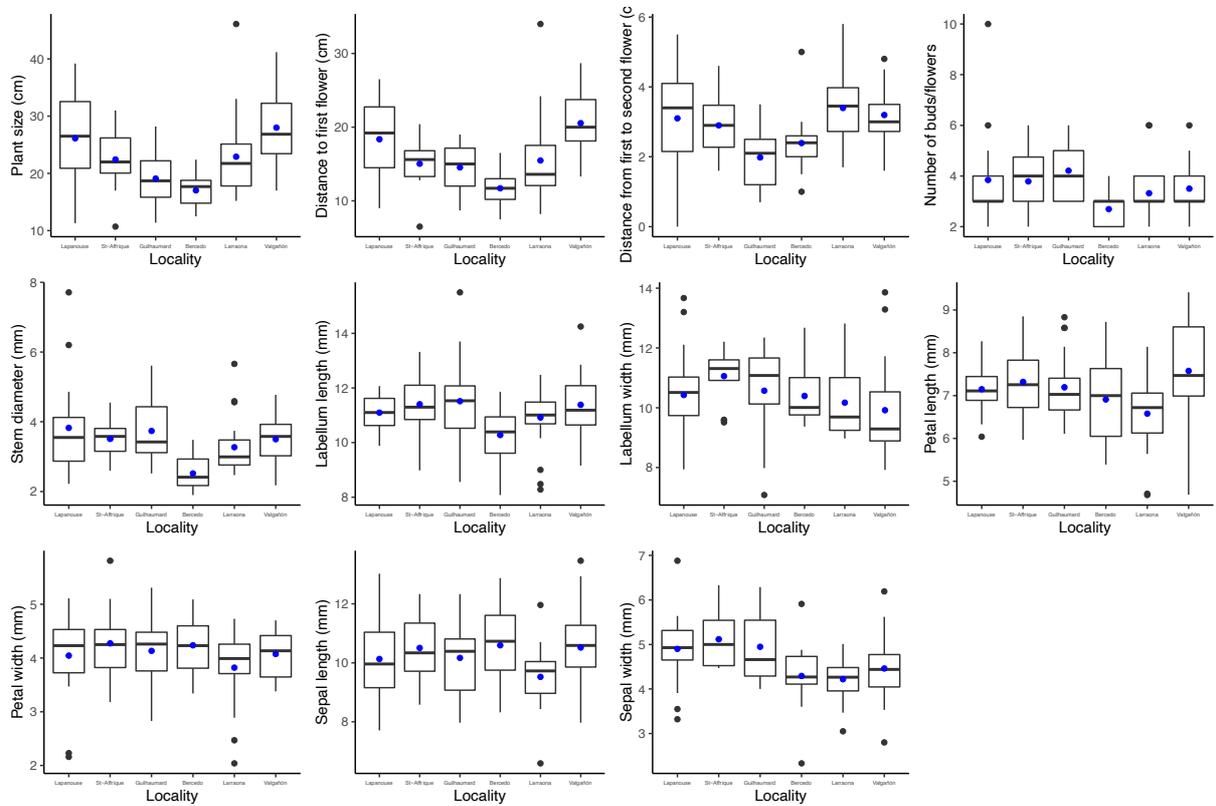


Figure 2 Distribution des valeurs de 11 variables morphométriques et comparaison entre les deux taxons du complexe *O. aveyronensis*

Annexe 2

J'ai effectué des enseignements à l'Université Paul Sabatier Toulouse III en tant que Doctorant Chargé d'Enseignement sur la période 2009-2012 puis en tant qu'ATER (½ poste) sur la période 2012-2013 (total de 329,5 h soit 306,5 h éq. TD). Ces enseignements sont détaillés dans les tableaux ci-après :

2009 - 2010 (71,5 h) :

Formation	Libellé de l'U.E	Code UE	TD/TP	Volume
L1	Organisation du monde vivant	DL1T461	TD	8 h
L1 (BBcP)	Biologie animale	DL2T931	TD	6 h
L1 (BG)	Biologie animale	DL2T811	TP	2,5 h
L1 (PCAg)	Biologie animale	DC2T731	TP	15 h
L1 (PCAg)	Biologie végétale	DL2T932B	TP	12 h
L2 (BOPE)	Traitement des données biologiques – méthodologie expérimentale	3L3BP04M	TD	4 h
L3 (BOPE)	Diversité, écologie et évolution des oiseaux	3L6UO6M	TP	12 h
L3 (BOPE)	Biologie évolutive et génomique	3L5EB1M	TD	6 h
M1 (Écologie)	Écologie évolutive	3M7EL4M	TP	4 h
M1 (Écologie)	Génétique & dynamique des populations	3M7EL2M	TD	2 h

Abréviations : BBcP (Biologie, Biochimie et Physiologie), BG (Biologie Générale), PCAg (Préparation aux concours « B » type Agro), BOPE (Biologie des Organismes des Populations et des Écosystèmes)

2010 - 2011 (64 h) :

Formation	Libellé de l'U.E	Code UE	TD/TP	Volume
L1	Organisation du monde vivant	DL1T461	TD	34 h
L1 (PCAg)	Biologie végétale	DL2T932B	TP	12 h
L3 (BOPE)	Diversité, écologie et évolution des oiseaux	3L6UO6M	TP	3 h
L3 (BOPE)	Biologie évolutive et génomique	3L5EB1M	TD	6 h
L3 (BOPE)	Biogéographie évolutive	3L5UO1M	TD	9 h

Abréviations : PCAg (Préparation aux concours « B » type Agro), BOPE (Biologie des Organismes des Populations et des Écosystèmes)

2011 - 2012 (71 h) :

Formation	Libellé de l'U.E	Code UE	TD/TP	Volume
L1 (PCAg)	Biologie végétale	EDS2BPC2	TD/TP	12 h
L1	Biodiversité 1	EDST1BBM	TD	8 h
L2/L3 (BOPE)	Biogéographie évolutive	ED4BOPJ2	TP	6 h
L3 (BOPE)	Biologie évolutive et génomique	EL5BOPAM	TD	6 h
L3 (BOPE)	Diversité, écologie et évolution des oiseaux	EL6BOP1	TD/TP	8 h
L3 (MABS)	Biologie des populations	EL6BIOJ12	TD	12 h
M1 (MABS)	Évolution moléculaire	EM7BMAAM	TP	8 h
M1 (Écologie)	Introduction à la génétique des populations	EM7BECBM	TD	10 h
M1 (Écologie)	Phylogénie et méthodes comparatives	EM7BECGM	TD	1 h

Abréviations : PCAg (Préparation aux concours « B » type Agro), BOPE (Biologie des Organismes des Populations et des Écosystèmes), MABS (Microbiologie-Agrobiosciences Bioinformatiques et Biologie des Systèmes)

2012 - 2013 (100 h) :

Formation	Libellé de l'U.E	Code U.E	TD/TP	Volume
-----------	------------------	----------	-------	--------

M1 (Écologie)	Phylogénie & méthode comparative	EM7BECGM	TD	2 h
M1 (Écologie)	Introduction à la dynamique des populations	EM7BECCM	TP	20 h
M1 (Écologie)	Biodiversité et dynamique des écosystèmes tropicaux, ethnoécologie	EM7BECFM	TD/TP	6 h
M1 (Écologie)	Introduction à la génétique des populations	EM7BECBM	TD	8 h
M1 (Écologie)	Phylogénie & méthode comparative	EM7BECGM	TD	1 h
M1 (Écologie)	Génétique des populations (avancé)	EM7BECPM	TD	4 h
L3	Zoologie	EL5BOPBM	TP	12 h
L2	Biologie animale	ED3BOPC1	TP	15 h
L1	Biodiversité 1	EDST1BBM	TD	8 h
M1 (Écologie)	Macroécologie et Biogéographie	EM7BECMM	TD	2 h
L1 (PCAg)	Biologie Animale	EDS2BPC1	TP	4 h
L3 (BOPE)	Diversité, évolution et écologie des oiseaux	EL6BOPT1	TP	2 h
L1	Approches expérimentales en biodiversité	EDST2BHM	TD/TP	8 h
M1 (Écologie)	Génomique des populations & Évolution moléculaire	EM8BECFM	TP	4 h
L2/L3 (BOPE)	Biogéographie évolutive de terrain	ED4BOPJ2	TD	6 h
L3 (BOPE)	Diversité, évolution et écologie des oiseaux - Terrain	EL6BOPT2	TD	12 h

Abréviations : PCAg (Préparation aux concours « B » type Agro), BOPE (Biologie des Organismes des Populations et des Écosystèmes)

Tuteur pédagogique: Alexandre RIBÉRON, Maître de Conférences à l'Université Paul Sabatier, Toulouse III

Les enseignements en lien avec la première maquette d'enseignement sur laquelle je suis intervenu en tant que Maître de Conférences de l'année universitaire 2017-2018 à 2020-201 sont détaillés ci-après :

2017-2018 (122,5 h)

Formation	Libellé de l'U.E	Code U.E	CM/TD/TP	Volume
L1 Biologie Écologie	Biologie moléculaire 1 : Concepts et Méthodes	TLSV2BM1	CM/TP	56,25 h
L2 Biologie Écologie	Biologie Moléculaire 2 : Méthodes et Techniques	TLBI4MO1	TP	15 h
L3 Biologie Écologie	Dynamique des Populations	TLSV6DP1	CM/TD/TP	31,5 h
L3 Biologie Écologie	Génétique des Populations	TLSV6GP1	TD/TP	19,5 h

2018-2019 (173,75 h)

Formation	Libellé de l'U.E	Code U.E	CM/TD/TP	Volume
L1 Biologie Écologie	Biologie moléculaire 1 : Concepts et Méthodes	TLSV2BM1	CM/TP	56,25 h
L2 Biologie Écologie	Biologie Moléculaire 2 : Méthodes et Techniques	TLBI4MO1	TP	15 h
L3 Biologie Écologie	Dynamique des Populations	TLSV6DP1	CM/TD/TP	22,5 h
M1 Biodiversité & Développement Durable	Outils moléculaires d'analyse de la biodiversité 1	TSBE1OM1	TP	24 h
M1 Biodiversité & Développement Durable	Modélisation des Systèmes Biologiques	TSBE1MS1	CM/TP	52,5 h
M1 Biodiversité & Développement Durable	Immersion en milieu professionnel	TSBE2ST1	TP	2h
M1 Biodiversité & Développement Durable	Préparation au stage 2	TSBE4PS1	TD	1,5h

2019-2020 (182,25 h)

Formation	Libellé de l'U.E	Code U.E	CM/TD/TP	Volume
L1 Biologie Écologie	Biologie moléculaire 1 : Concepts et Méthodes	TLSV2BM1	CM/TP	56,25 h
L1 Biologie Écologie	La cellule : division et différenciation	TLSV1CE1	TP	6h
L1 Biologie Écologie	Introduction à l'Écologie et à l'Évolution	TLSC2IE1	TD	9h
L2 Biologie Écologie	Biologie Moléculaire 2 : Méthodes et Techniques	TLBI4MO1	TP	30 h
L3 Biologie Écologie	Dynamique des Populations	TLSV6DP1	CM/TD/TP	24 h
L3 Biologie Écologie	Génétique des Populations	TLSV6GP1	TP	27 h
M1 Biodiversité & Développement Durable	Systèmes d'Information Géographique (SIG)	TSBE1MS1	CM/TP	28,5 h
M1 Biodiversité & Développement Durable	Préparation au stage 2	TSBE4PS1	TD	1,5h

2020-2021 (192,25 h)

Formation	Libellé de l'U.E	Code U.E	CM/TD/TP	Volume
L1 Biologie Écologie	Biologie moléculaire 1 : Concepts et Méthodes	TLSV2BM1	CM/TP	44,25 h
L2 Biologie Écologie	Biologie Moléculaire 2 : Méthodes et Techniques	TLBI4MO1	TP	30 h
L3 Biologie Écologie	Dynamique des Populations	TLSV6DP1	CM/TD/TP	31,5 h
M1 Biodiversité & Développement Durable	Systèmes d'Information Géographique (SIG)	TSBE1OM1	TP	28,5 h
M1 Biodiversité & Développement Durable	Modélisation des Systèmes Biologiques	TSBE1MS1	CM/TP	52,5 h
M1 Biodiversité & Développement Durable	Immersion en milieu professionnel	TSBE2ST1	TP	2h
M1 Biodiversité & Développement Durable	Préparation au stage 2	TSBE4PS1	TD	1,5h