Etude fonctionnelle et évolutive de la voie de l'acide rétinoique et de la phosphorylation des récepteurs chez le poisson zèbre

Eric Samarut

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE

IGBMC – IGFL – ENS de Lyon

THÈSE

présentée par

Eric SAMARUT

soutenue le 16 décembre 2013

pour obtenir le grade de

Docteur de l'université de Strasbourg

Sciences du Vivant – Aspects Moléculaires et Cellulaires de La Biologie

ETUDE FONCTIONNELLE ET EVOLUTIVE DE LA VOIE DE L'ACIDE RETINOIQUE ET DE LA PHOSPHORYLATION DES RECEPTEURS CHEZ LE POISSON ZEBRE

Soutenance publique devant le jury composé de

DIRECTEURS DE THESE
Dr Cécile Rochette-Egly
(Université de Strasbourg)
Pr Vincent Laudet
(ENS de Lyon)

RAPPORTEURS
Denis Duboule
(Ecole Polytechnique Fédérale de Lausanne)
Frédéric Michon
(University of Helsinki)

EXAMINATEUR
Julien Vermot
(Université de Strasbourg)

A mes poissons morts pour la Science,
A Maurice Star,

Remerciements

Bien évidemment, je tiens en tout premier lieu à remercier mes deux directeurs de thèse avec qui cela a toujours été un plaisir de discuter science, de confronter des idées, et d'imaginer la suite. Cette codirection a été une très bonne balance de « liberté cadrée » entre accompagnement et autonomie, suivi et confiance, choucroute et gratin dauphinois ! Merci au dynamisme et à la réactivité de Cécile, merci à la souplesse et à la confiance de Vincent.

« et pof ça phosphoryle, et clac ça interagit… » Cécile R-E.

Cette codirection a aussi été l'occasion pour moi de découvrir Strasbourg et l'IGBMC. Je ne serai jamais assez reconnaissant de l'accueil qui m'a été réservé à mon arrivée au laboratoire. Un merci particulier à la dinde Christine et au grand steak Sébastien pour leurs conseils techniques, leurs recommandations culinaires et leur bonne-humeur. J'ai passé une première année très agréable au sein de l'équipe aux côtés de Régis, Vanessa, Nathalie, Aleksandr et Gabi. Un grand merci à Lady Samia et Ziad avec qui j'ai passé une grande partie de ma seconde année strasbourgeoise. Les pauses cafés (p'tit caf ?) sur des fonds de discussions plus ou moins scientifiques, les soirées arrosées au Barco Latino et au Seven me resteront en souvenir ! Je remercie également les Vermot et particulièrement Emilie et Stéphane pour m'avoir fourni en poissons et m'avoir fait une petite place à l'animalerie ! Merci aussi à tous les membres des services communs de l'IGBMC dont Mustapha et ses anticorps, Betty et ses cellules, Pascal et ses peptides et plus particulièrement La Blonde et ses vacances (*Dear everyone*…).

« Si je ne lis pas ton papier d'ici ce soir, je t'offre une caisse de vin! » Vincent L.

« euh…cellules COS c'est pour cosinus?? » Samia G.

« Titafait, titafait… » Ziad A T.

Après deux ans passés au grand Nord, c'est le retour sur Lyon. Merci au bureau de m'avoir fait une place et de m'avoir accueilli au sein de leurs délires de chef ligoté et autres élans déchaînés ! J'ai pris un plaisir tout particulier à parler le vieux avec Florent ou encore réapprendre l'espagnol avec Juli avec quelques verres dans le nez ! Un immense merci à Laure le pétoncle, la sirène de l'animalerie, la poissonnière de l'IGFL pour sa gentillesse, sa compréhension, sa naïveté et les passe-droits qu'elle a pu m'accorder ! Un merci tout particulier à Cyril le *crazy dude* qui m'a été d'une *RAR* inspiration ! Grâce à lui, je sais que

« Qui a dessiné des cerises sur mon agenda?? » Laure B.

mon Mac peut parler, et ce fût un plaisir de passer une grande partie de mon temps au sein de l'équipe GOSAM©…je reste optimiste pour qu'elle se monte un jour dans un labo junior inoccupé! Le déménagement dans le nouvel IGFL a été l'occasion de nouveaux rapprochements, en particulier avec les Volfettes, ex-réfugiées du froid sous-sol. Merci à Domi et sa douceur à toute épreuve ainsi qu'à Dédé, son rire ravageur et sa gentillesse authentique! J'ai beaucoup apprécié les moments en votre compagnie au sein de la DreamTeam avec Roro le puriste et les membres de l'équipe Aïe-Troudp(o)ut' ! Sandrine, merci d'avoir toujours modéré nos propos masculins, Benj et Romain, merci de toujours les enrichir de nouvelles données (*quoi ? un hareng géant ?*). Merci également pour cet engouement sportif dans lequel vous avez réussi à m'enliser !

"Go of GOSAM and proud to be!" Cyril G.

"Ah mais j'tai pas raconté… (Rires)" Delphine G.

"J'en ai une bonne… Tu connais la différence…" Romain G.

"L'Amphigatocut (…) à jamais gravé dans ma mémoire…" Benjamin G.

Je souhaite remercier la sympathie et l'aide de nombreuses autres personnes au sein de l'équipe mais aussi des équipes environnantes. En particulier, merci à Marie qui a aidé SamSam dans l'analyse de ses données. Je remercie les Viriot (Manu et al..), les Samarut (Damien, Karine et al.), les Mérabet (Marilyne, Manon et al.), les Flamant (Fabrice et al.), les Vanacker (Violaine et al.), les Volff (Fred, Emilien, Magalie et al.) et les Ruggiero (Elise et al.). Un immense merci à nos chefs financières qui sont Sonia et Fabienne : merci d'avoir accepté avec gentillesse et bonne humeur mes demandes d'achat improbables entre grillage anti-moustique pour fabriquer des pondoirs, à des lots de poissons exotiques dont même le CNRS doutait de leur utilisation à des fins scientifiques !

Enfin, je souhaite remercier mes amis et proches pour m'avoir soutenu et pour être encore aux côtés d'un fou qui travaille sur… …des poissons-zèbres ! Merci à mes parents de m'avoir initié à la science dès mon plus jeune âge en observant le développement d'œufs de grenouille pêchés dans une marre autour de Saint-Christophe, à Sylvain pour m'avoir donné le goût de l'aquariophilie et à Camille qui s'intéresse à mon travail en ressassant des vieux mots de TP de S.V.T !

"Bec-bunsen…, erlenmeyer…, ADN!" Camille S.

Merci aux belles-sœurs et beaux-frères de tous horizons (Marion, Marlène, Chloé, Marine, Arthur, Tiébo) ainsi qu'à ma nouvellement belle-famille Dufour qui se sont toujours interrogés sur le pourquoi du comment je travaille sur le poisson-zèbre. Un merci tout particulier à ma femme Liselotte qui m'a supporté dans les durs moments de désespoir de la thèse et pour sa relecture attentive de ce manuscrit ! Nous aurons finalement réussi à combiner avec brio un mariage, une thèse et un départ prochain dans les grands froids canadiens ! Merci pour ton enthousiasme et pour m'avoir toujours soutenu et suivi dans ces projets d'éloignements géographiques ! Merci également à mes amis, les futurs Enderlin, Alcaras, Galopin, Sanchiz et Brunon pour leur présence et pour s'être toujours intéressés de près ou de loin à ce que je pouvais bien « rechercher »…

"Non mais sérieux, tu vas à un meeting QUE sur le zebrafish???" *Marlène P.*

"Et en fait ils font comment tes poissons sans nageoires? Ils coulent?" *Marine E.*

"C'est moche un amphioxus!" *Liselotte S.*

"J'aimerais bien voir tes poissons-tigres!" *Clémence C.*

"Tu bosses sur la toporination des centrales nucléaires c'est ca?" *Loïc A.*

Un grand merci à JazzRadio reprise, à ChromaRadio, à Illy, à Coca-Cola Zéro et à Dinausorus, qui m'ont soutenu au sens propre du terme tout au long de cette rédaction !

Pour finir, je remercie par avance tous ceux que j'aurais pu oublier, dans la précipitation du dernier jour de rédaction, de m'excuser de cette étourderie !

Publications

Samarut E, Fraher D, Laudet V, Gibert Y. (2014) zebRA : an overview of retinoic acid effects during zebrafish development. **Soumis, Biochim Biophys Acta**.....page 66

Samarut E*, Gaudin C*, Hughes S, Gillet B, De Bernard S, Jouve P.E, Buffat L, Alliot A, Lecompte O, Rochette-Egly C, Laudet V. (2013) RAR-subtype specific transcriptotypes in early zebrafish embryos. **Soumis, Mol Endo.** .....page 93

Samarut E*, Gibert Y*, Pasco-Viel E, Bernard L, Schulte-Merker S, Labbe C, Viriot L, Laudet V. (2013) Retinoic acid controls tooth morphology and number in Cypriniformes. **Soumis, Cell reports**.....page 177

Tohmé M, Prud'homme S.M, Boulahtouf A, Samarut E, Brunet F, Bernard L, Gibert Y, Balaguer, P and Laudet V. (2013) ERRγ is an in vivo receptor of bisphenol A. **En révision, FASEB J.**

Seritrakul P, Samarut E, Lama TT, Gibert Y, Laudet V, Jackman WR. (2012) Retinoic acid expands the evolutionarily reduced dentition of zebrafish. **FASEB J.** 26(12):5014-24.....page 166

Samarut E, Rochette-Egly C. (2012) Nuclear retinoic acid receptors: Conductors of the retinoic acid symphony during development. **Mol Cell Endocrinol.** 348(2):348-60.....page 52

Ferry C, Gaouar S, Fischer B, Boeglin M, Paul N, Samarut E, Piskunov A, Pankotai-Bodo G, Brino L, and Rochette-Egly C. (2011) Cullin 3 mediates SRC-3 ubiquitination and degradation to control the retinoic acid response. **PNAS** 108(51):20603-8. .....page 221

Lalevée S, Anno Y. N, Chatagnon A, Samarut E, Poch O, Laudet V, Benoit G, Lecompte O and Rochette-Egly C. (2011) Genome-wide in silico identification of conserved and functional DR5 retinoic acid receptors response elements. **J Biol Chem.** 286(38):33322-34. .....page 207

Samarut E, Amal I, Markov GV, Stote R, Dejaegere A, Laudet V, Rochette-Egly C. (2011) Evolution of nuclear retinoic acid receptor alpha phosphorylation sites. Serine gain provides fine-tuned regulation. **Mol Biol Evol.** 28(7):2125-37. .....page 142

Grijota-Martínez C, Samarut E, Scanlan TS, Morte B, Bernal J. (2011) In vivo activity of the thyroid hormone receptor beta- and _-selective agonists GC-24 and CO23 on rat liver, heart, and brain. **Endocrinology** 152(3):1136-42.

Lalevée S, Bour G, Quinternet M, Samarut E, Kessler P, Vitorino M, Bruck N, Delsuc MA, Vonesch JL, Kieffer B, Rochette-Egly C. (2010) Vinexinß, an atypical "sensor" of retinoic acid receptor gamma signaling:union and sequestration, separation, and phosphorylation. **FASEB J**. 24(11):4523-34. .....page 195

Introduction – Etat de l'Art 1

Liste des abréviations

A/P: Antéro-Postérieur
ABP: Actin Binding Protein
ADH: Alcohol Dehydrogenase
AF-1: Activation Fonction 1
AF-2: Activation Fonction 2
Akt: Protéine kinase B
AR: Acide Rétinoique
ARNm: Acide RiboNucléique messager
CAK: CDK-Activating Kinase
Cdk7: Cyclin-dependant kinase 7
Cdk8: Cyclin-dependant kinase 8
Cellule ES: Cellules Embryonnaires Souches
ChIP-seq: Chromatin ImmunoPrecipitation sequencing
ChIP: Chromatin ImmunoPrecipitation
CoR-NR: CoRepressor Nuclear Receptor
CRABPII: Cytoplasmic Retinoic Acid Binding Protein II
CRBPI, II, III: Cellular Retinol Binding Protein I, II, III
CYP1A1, B1: Cytochrome P450 1A1, B1
CYP26: Cytochrome P450 26
D/V: Dorso-Ventral
DAD: Deacetylase Activating Domain
DBD: DNA Binding Domain
DBX: Developing Brain, homeoboX
DEAB: 4-DiEthylAminoBenzaldehyde
DLX: Distal-Less homeoboX
DR5: Direct Repeat 5
EDN: EnDotheliN
EDNRA: EnDotheliN Receptor type A
ERK: Extracellular signal Regulated Kinase
EVX: Even-skipped homeoboX
FGF: Fibroblast Growth Factor
Fli1: Flightless 1
FSGD: Fish-Specific Genome Duplication
Gli3: GLI-Kruppel family member 3
GluR1,2: Glutamate Receptor 1, 2
HAT: Histone Acetyl-Transferase
HDAC: Histone DeACetylase
HOX: HomeOboX
Hpf: heures post-fertilisation
IUD: Intrinsically Unstructured Domain
JNK: c-Jun N-terminal Kinases
KO: Knock-Out
LBD: Ligand Binding Domain
LBP: Ligand Binding Pocket
LCoR: Ligand-dependent CoRepressor
MAPK: Mitogen Activated Protein Kinase
MAT1: Ménage A Trois 1
MEIS: Myeloid Ecotropic viral Integration Site
MSK1: Mitogen and Stress Activated Kinase 1

NCC: Neural Crest Cells
NCoR: Nuclear hormone receptor Corepressor
NID: Nuclear receptor Interaction Domain
NLS: Nuclear Localization Signal
NTD: N-Terminal Domain
NVP: Nodal Vesicular Parcels
Oca2: Oculocutaneous albinism II
OTX: OrThodenticle homeoboX
P160: Protéine coactivatrice de 160kDa
PAX: PAired boX gene
PBX: Pre-B-cell leukemia homeoboX
PI3K: PhosphatidylInositol 3 Kinase
PIC: Pre-Initiation Complex
PITX: Paired-like homeodomain transcription factor
PKA: Protein Kinase A
PKC: Protein Kinase C
PPAR: Peroxisome Proliferator Activated Receptor
PRAME: Preferentially expressed Antigen in Melanoma
PRM: Proline-Rich Motif
PSM: PreSomitic Mesoderm
PTM: Post-Translational Modification
RAE: Retinoic Acid Embryopathy
RALDH: RetinALdehyde DeHydrogenases
RAR: Récepteur de l'Acide Rétinoique
RARE: Retinoic Acid Response Element
RIP140: Receptor Integrating Protein of 140 kDa
RN: Récepteur Nucléaire
RNA pol II: RNA polymerase II
RXR: Récepteur X des Rétinoides
SDR: Short-chain Dehydrogenases/Reductases
SH3: Src Homology 3
SHH: Sonic HedgeHog
SMRT: Silencing Mediator of Retinoid and Thyroid hormone receptors
SNC: Système nerveux Central
SRC: Steroid Receptor Coactivator
SUG1: Suppressor for Gal1
TACC1: Transforming Acidic Coiled-Coil 1
TBLR1: Transducin Beta-Like Related Protein 1
TFIIH: Transcription Factor II H
TR: Thyroid Hormone Receptor
TSS: Transcription Start Site
UICPA: Union internationale de chimie pure et appliquée
VAD: Vitamin A Deficiency
VDR: Vitamin D Receptor
WGD: Whole Genome Duplication
Wnt: Wingless-related integration site
WW: Tryptophane/Tryptophane
XPD: Xeroderma Pigmentosum group D-complementing protein
ZFN: Zinc Finger Nuclease
Zic2: Zic family member 2

Liste des figures

Liste des annexes

Avant-Propos

Mon travail de thèse s'inscrit dans ma volonté de connecter plusieurs aspects de la biologie, en reliant des observations moléculaires au développement et à la physiologie et de conférer à la génomique fonctionnelle une dimension évolutive. Relier la structure à la fonction d'une protéine et l'étudier dans un contexte développemental et évolutif est ainsi à la base de mes travaux de recherche. Durant mon master, j'ai étudié la voie de l'acide rétinoïque (AR) chez le poisson-zèbre ce qui m'a naturellement conduit à vouloir développer ce sujet en accord avec mes envies de « biologie multidimensionnelle ». Le projet initial de mon doctorat était d'apporter une dimension *in vivo* au rôle de la phosphorylation des récepteurs de l'acide rétinoïque (RAR) en utilisant le poisson-zèbre comme modèle de biologie du développement. Je liais ainsi l'expertise en biologie moléculaire du laboratoire de Cécile Rochette-Egly à la biologie du développement au sein de l'équipe de Vincent Laudet. Au cours des années de doctorat, mes travaux ont intégrés une dimension évolutive à laquelle j'ai pris goût, tout en conservant mon désir premier de génomique fonctionnelle. Ainsi, mon travail durant ces quatre années de thèse se décompose en trois parties majeures :

- Une étude fonctionnelle de l'activité transcriptionnelle des différents sous-types de RAR dans l'embryon précoce de poisson-zèbre : La biologie moléculaire et la génomique fonctionnelle s'y entremêlent.
- Une étude d'Evo-Fun qui corrèle l'**Evo**lution, la structure et la **Fon**ction des RAR en s'intéressant au rôle des phosphorylations dans l'évolution de la régulation de leur activité.
- Une étude d'Evo-Devo qui met en avant le rôle de la voie de l'AR dans le **Dév**eloppement des dents et l'**Evol**ution de la denture chez les poissons.

Après un état de l'art qui se veut complet mais concis sur les aspects moléculaires de la voie de l'AR et de son rôle dans le développement, j'ai divisé ce manuscrit en trois chapitres. Chacun est composé d'une brève introduction retraçant le contexte scientifique dans lequel ont été effectué mes recherches, des résultats publiés ou en cours de soumission, et d'une discussion ouvrant sur les perspectives à venir.

Très bonne lecture.

Introduction – Etat de l'Art

De la molécule d'acide rétinoïque à son rôle dans le développement et l'évolution

A. ASPECTS MOLECULAIRES DE LA VOIE DE L'ACIDE RETINOIQUE

I. METABOLISME DE L'ACIDE RETINOIQUE

L'acide rétinoïque (AR) est le dérivé actif majeur de la vitamine A, avec le rétinal *11-cis*, et existe sous forme de différents isomères dont les principaux sont l'AR 9-*cis* et *all-trans*. La forme majoritaire dans l'organisme et celle dont la majorité des effets est décrite est la forme *all-trans* alors que la détection de l'isomère 9-*cis* n'a que rarement été décrite *in vivo* jusqu'à maintenant (Kane et al., 2010). L'acide rétinoïque présent dans l'organisme provient exclusivement de l'alimentation (vitamine A ou pro-vitamine A) et ne peut être synthétisé *de novo*. De ce fait, la vitamine A doit être convertie via plusieurs réactions enzymatiques (Gutierrez-Mazariegos et al., 2011; Maden, 2007; Theodosiou et al., 2010). Brièvement, et comme illustré dans la figure 1, la voie de synthèse classique de l'AR correspond à l'oxydation du rétinol en rétinal, par des enzymes de la familles des ADHs et SDRs (*Alcohol dehydrogenase* et *short-chain dehydrogenases/reductases*). Le rétinal est par la suite oxydé de manière irréversible en acide rétinoïque sous l'action des enzymes de la famille des RALDH (ou ALDH1A) (*retinaldehyde dehydrogenases*), ou plus rarement par certaines enzymes mono-oxygénases de la famille des cytochromes p450 comme CYP1B1 ou CYP1A1 (Chambers et al., 2007; Chen et al., 2000). Toutefois, une voie de synthèse annexe peut être assurée par le clivage des béta-carotènes (pro-vitamine A) en rétinal par les enzymes BCO-I (*beta-carotene-15,15'-oxygenases*). Concernant son catabolisme, l'acide rétinoïque est dégradé en dérivés oxydés comme l'acide 4-oxo ou 4-hydroxy rétinoïque par l'action des enzymes de la famille des cytochromes p450 (CYP26) (Figure 1) (Simoes-Costa et al., 2008).

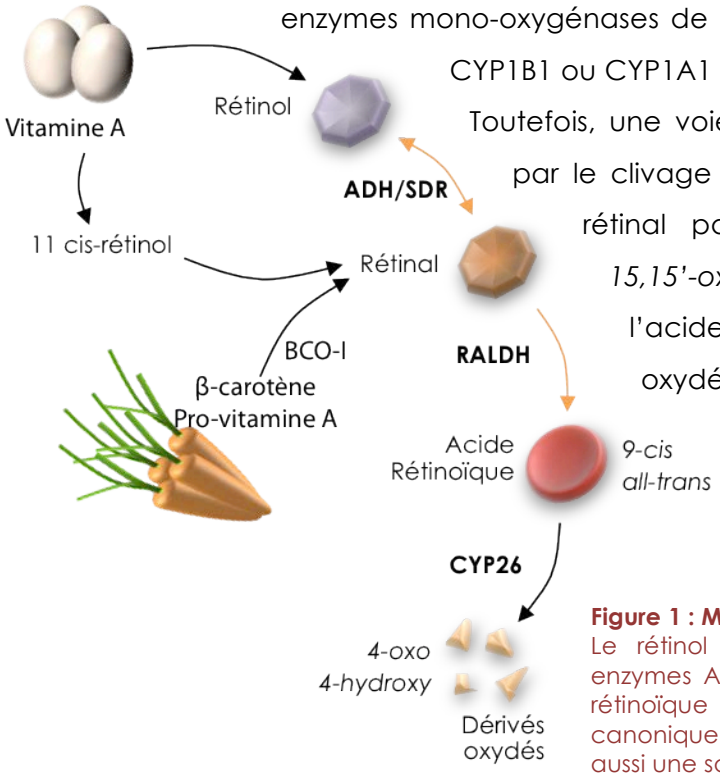


Figure 1 : Métabolisme de l'AR.
Le rétinol est transformé en rétinal par l'action des enzymes ADH/SDR de manière réversible, puis en acide rétinoïque sous l'action des RALDH (voie de synthèse canonique en orange). Le clivage des béta-carotènes est aussi une source de rétinal. Les enzymes CYP26 dégradent l'AR en métabolites polaires oxydés.

Ainsi, le terme « rétinoïdes » regroupe à la fois le rétinol, l'AR, les dérivés métaboliques et les composés synthétiques actifs (Sporn et al., 1994).

Les rétinoïdes sont des substances de petite taille, hydrophobes et liposolubles, ce qui leur permet de traverser facilement la bicouche lipidique membranaire. Une fois dans le cytoplasme, ils sont pris en charge par les protéines CRBP (*Cellular retinol binding protein*) et CRABP (*Cellular retinoic acid binding protein*) qui régulent leur métabolisme soit en les stockant, soit en facilitant leur présentation aux différentes enzymes pour des réactions d'estérification ou d'hydrolyse (Theodosiou et al., 2010).

II. RAR et RXR: UNE STRUCTURE MODULAIRE

Les Récepteurs de l'Acide Rétinoïque (RAR) ainsi que les Récepteurs X des Rétinoïdes (RXR) appartiennent à la superfamille des Récepteurs Nucléaires (RN) (Laudet and Gronemeyer, 2002). Au moins 48 membres ont été identifiés dans le génome humain (Robinson-Rechavi et al., 2001) et peuvent être regroupés en plusieurs sous-familles incluant les récepteurs aux hormones thyroïdiennes (TR), à l'acide rétinoïque (RAR), à la vitamine D (VDR), à l'ecdysone (EcR), les récepteurs contrôlant l'activation de la prolifération des peroxysomes (PPAR), les récepteurs aux rétinoïdes (RXR), les récepteurs aux hormones stéroïdiennes telles que les glucocorticoïdes (GR), les androgènes (AR), les minéralocorticoïdes (MR), les progestatifs (PR), les œstrogènes (ER), ainsi que les récepteurs orphelins comme ERR (*Estrogen-Related Receptor*) pour lesquels aucun ligand n'a encore été décrit (Laudet and Gronemeyer, 2002). Ils présentent tous une structure modulaire en six domaines (de A à F) commune aux différents RN comme décrite ci-après (Figure 2).

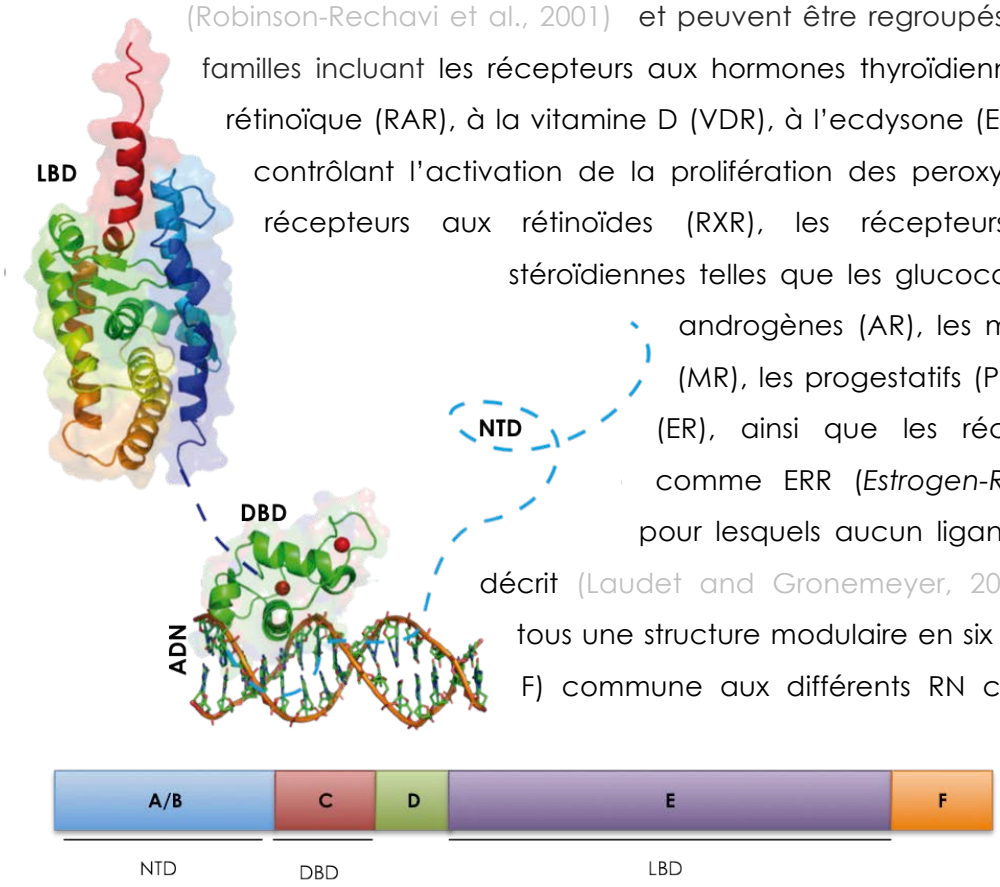


Figure 2 : Organisation modulaire des RAR. Les RAR sont composés de trois principaux domaines que sont le NTD, le DBD et le LBD. La région D est une région charnière et la région F correspond à la partie C-terminale non structurée

Les RAR sont des facteurs de transcription dépendants du ligand qui agissent sous forme d'hétérodimères avec les RXR pour transduire le signal de l'AR. Trois sous-types de RAR codés par des gènes distincts sont présents chez les mammifères : α (NR1B1), β (NR1B2) et γ (NR1B3) (Chambon, 1996; Germain et al., 2006a) ainsi que pour RXR : α (NR2B1), β (NR2B2), γ (NR2B3)(Germain et al., 2006b). Pour chaque sous-type, différentes isoformes différant seulement dans leur région N-terminale peuvent être transcrits par l'utilisation de différents promoteurs et par épissage alternatif.

1. LE DOMAINE DE LIAISON A L'ADN (DBD)

Ce domaine, aussi dénommé région centrale ou région C (Figure 2) correspond à la région la plus conservée au sein des RN et entre les sous-types de RAR. Le noyau du DBD est composé de deux motifs en doigt de zinc organisés en deux hélices alpha (trois pour RXR) qui se croisent perpendiculairement pour former une structure globulaire (Freedman et al., 1988). Seulement trois résidus présents sur l'hélice α, définissant la boite P, déterminent la spécificité de liaison entre le récepteur et l'ADN au niveau des éléments de réponse (Lee et al., 1993) (Figure 3).

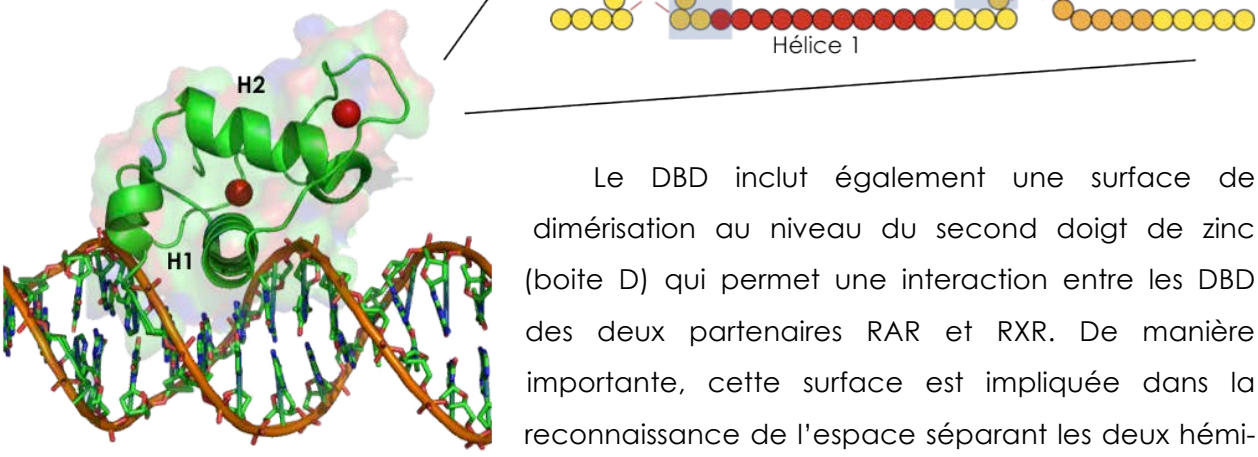


Figure 3: Organisation et structure 3D du DBD de RAR sur l'ADN. Représentation schématique du DBD et des doigts de zinc (Zn-I-II) comprenant les boites P et D. Les hélices 1 (H1) et 2 (H2) sont indiquées.

Le DBD inclut également une surface de dimérisation au niveau du second doigt de zinc (boite D) qui permet une interaction entre les DBD des deux partenaires RAR et RXR. De manière importante, cette surface est impliquée dans la reconnaissance de l'espace séparant les deux hémi-sites au sein des éléments de réponse (Mangelsdorf and Evans, 1995; Umesono et al., 1991a) (voir section A-III-2, page 12.)

2. LE DOMAINE DE LIAISON DU LIGAND (LBD)

Le LBD fait partie de la région E, en position C-terminale par rapport au DBD (Figure 2) et est impliqué dans la reconnaissance et la fixation du ligand. C'est un domaine fonctionnel complexe puisqu'il comprend la poche de fixation du ligand (LBP), le domaine principal de dimérisation, un sillon hydrophobe impliqué dans l'association avec les corégulateurs ainsi qu'une hélice H12 C-terminale renfermant la fonction de transactivation dépendante du ligand (AF-2) (Figure 4). Des données de structures ont permis de définir une structure canonique composée de douze hélices alpha (H1 à H12) et d'un feuillet beta, le tout formant un domaine globulaire replié en sandwich selon trois feuillets d'hélices antiparallèles (Bourguet et al., 2000; Renaud and Moras, 2000; Renaud et al., 1995).

Brièvement, la poche du ligand, située dans la partie inférieure du LBD, définie la spécificité et l'affinité du ligand de par sa taille et sa forme ainsi que via des interactions spécifiques entre le ligand et des résidus polaires du fond de la poche (Klaholz et al., 2000). Chez les mammifères, seuls trois résidus polaires de la LBP impliqués dans la liaison de l'AR différent entre les sous-types de RAR et ces différences sont exploitées pour le développement de ligands synthétiques spécifiques d'un sous-type donné (Gronemeyer et al., 2004).

Après fixation du ligand (conformation holo), la poche est refermée par l'hélice H12 du LDB assurant ainsi la stabilité du ligand. En effet, cette hélice est exposée au solvant en absence de ligand (conformation apo), mais la fixation de ce

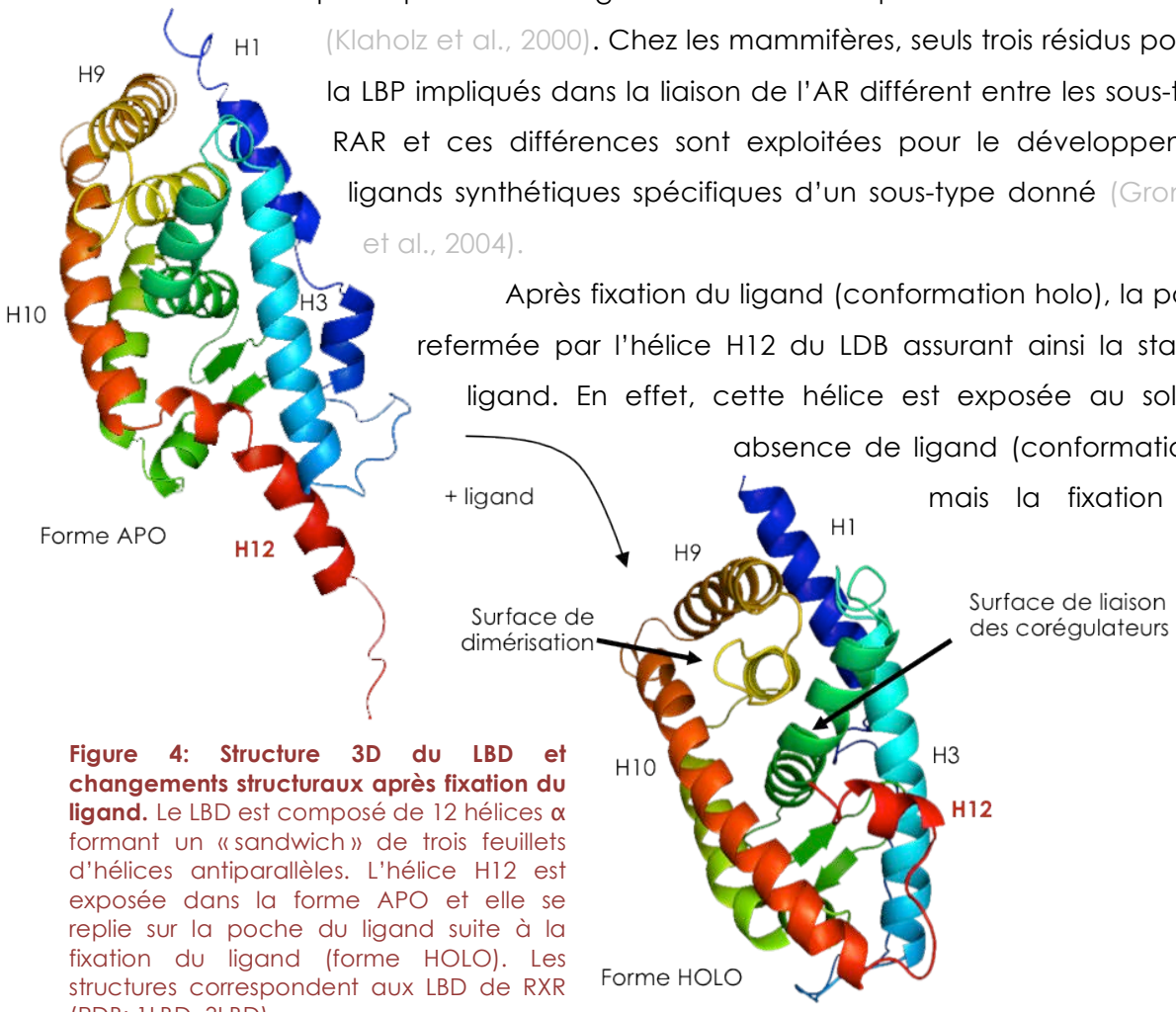


Figure 4: Structure 3D du LBD et changements structuraux après fixation du ligand. Le LBD est composé de 12 hélices α formant un « sandwich » de trois feuillets d'hélices antiparallèles. L'hélice H12 est exposée dans la forme APO et elle se replie sur la poche du ligand suite à la fixation du ligand (forme HOLO). Les structures correspondent aux LBD de RXR (PDB: 1LBD, 3LBD).

dernier mène à des repositionnement d'hélices au sein du LBD (transconformation (Moras and Gronemeyer, 1998)), en particulier l'alignement des hélices H10 et H11, qui résulte *in fine* au basculement de l'hélice H12 sur la LBP, la refermant comme un piège à souris (Figure 4)(Moras and Gronemeyer, 1998; Renaud et al., 1995). Dans cette conformation, des ponts salins au sein du LBD stabilisent et rigidifient la structure (Bourguet et al., 1995) et de nouveaux résidus sont exposés à la surface du LBD formant un sillon hydrophobe (domaine AF-2) propice à l'interaction avec des coactivateurs. Enfin, il a été montré que la fixation du ligand induit également des changements allostériques entre la LBP et la surface de dimérisation, conférant une stabilité de l'hétérodimère RAR-RXR (Brelivet et al., 2004; Pogenberg et al., 2005).

En conclusion, le LBD est le domaine majeur du récepteur lui conférant sa spécificité de ligand. C'est également ce qui le distingue d'autres facteurs de transcription en conférant à son activité, une dépendance au ligand. C'est un domaine complexe au sein duquel de nombreuses régions sont énergétiquement liées favorisant ainsi des changements allostériques subtils et importants permettant une adaptation du récepteur en fonction de son environnement moléculaire. Ces communications intramoléculaires sont particulièrement discutées dans le chapitre 2, page 156.

3. LE DOMAINE N-TERMINAL (NTD)

Le NTD correspond aux deux régions A et B et contient une fonction de transactivation indépendante du ligand AF-1 (Nagpal et al., 1993; Nagpal et al., 1992). Il est le domaine qui diffère le plus en séquence et en longueur au sein des RN mais aussi entre les sous-types de RAR et RXR. De plus pour chaque sous-type, il existe plusieurs isoformes qui diffèrent au niveau de la région A suite à des épissages alternatifs ou à l'utilisation de promoteurs différents (Chambon, 1996). A l'inverse, la région B est très conservée entre les différents sous-types de RAR. A lui seul, le NTD est capable d'activer la transcription de gènes cibles en absence de ligand du fait de sa fonction AF-1. Bien que cette fonction redevienne dépendante du ligand dans le contexte du récepteur entier, cela suggère de fortes relations synergiques entre le NTD et le LBD des RAR (Taneja et al., 1997).

Du fait de sa structure naturellement désordonnée (IUD : *Intrinsically Unstructured Domain*) (McEwan et al., 2007), caractérisée par l'absence d'une structure tertiaire stable, le NTD est très sensible à la protéolyse et aucune donnée structurale n'a jamais été obtenue

concernant ce domaine des RAR. Cependant, ces IUD peuvent adopter des états conformationnels pseudo-structurés transitoires suite à l'interaction avec d'autres protéines ou après liaison à l'ADN (Dyson and Wright, 2005; Lavery and McEwan, 2005; Wright and Dyson, 2009). De plus, le NTD des RAR possède un motif riche en proline contenant un motif consensus de phosphorylation au niveau d'une serine (voir section A-IV-2.1, page 18.) qui est susceptible de former des pseudo-structures hélicoïdales transitoires (Bielska and Zondlo, 2006; Rochette-Egly et al., 1997). Enfin ces domaines riches en prolines (PRM : *Proline rich motif*) ont la capacité d'interagir avec des protéines à domaines SH3 ou WW (Reimand et al., 2012).

Récemment, plusieurs études ont révélé des interactions entre le NTD et certaines protéines comme HACE1, Acinus-S' ou la vinexine beta réprimant l'activité transcriptionnelle des RAR (Bour et al., 2005b; Vucetic et al., 2008; Zhao et al., 2009). De manière intéressante, de par la proximité du NTD avec le DBD, des interactions protéiques au niveau des régions A et B seraient susceptibles de réguler la liaison du récepteur à l'ADN (Lalevee et al., 2010). Cette hypothèse est renforcée par la localisation du domaine A/B proche de l'ADN révélée par des données récentes de structure du complexe RXR/VDR lié à l'ADN (Orlov et al., 2012).

En conclusion, bien que le NTD soit encore très peu étudié de par sa nature désordonnée à l'inverse du LBD et du DBD, son implication fonctionnelle dans l'activité transcriptionnelle des RAR est de plus en plus reconnue. De plus, le fait que ce domaine diffère grandement entre les différents sous-types de RAR permet de spéculer sur son implication dans la spécificité d'action de chaque sous-type de RAR (discuté en chapitre 1, page 134).

4. LA REGION CHARNIERE D

Cette région qui relie le DBD au LBD est très peu conservée au sein des RN bien qu'elle contienne la séquence de localisation nucléaire (NLS : Nuclear Localization Signal) (Hamy et al., 1991). Cette région souple et flexible permettrait une adaptation de l'hétérodimère à des

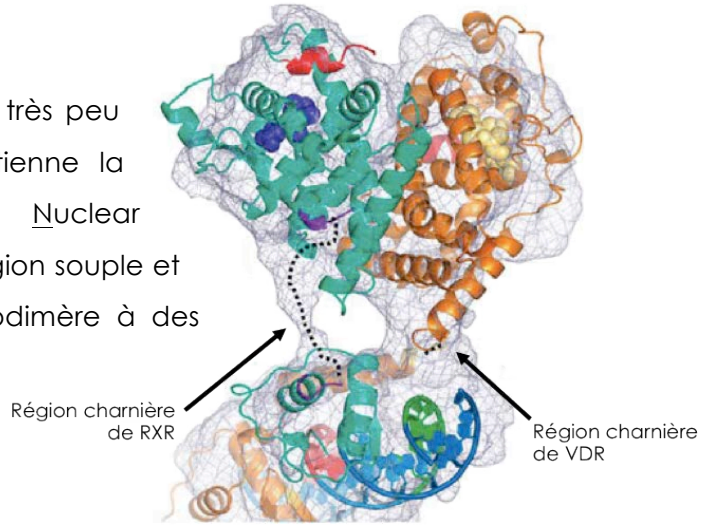


Figure 5: Structure 3D de l'hétérodimère RXR-VDR. Cette structure publiée par Orlov *et al.*, en 2012 montre pour la première fois la région charnière reliant le LBD et le DBD de VDR.

éléments de réponses différents au niveau de l'ADN (voir section A-III-2, page 12.). Bien qu'elle ne soit pas toujours considérée comme une entité à part entière, la région D a été mise en évidence par des données structurales récentes au sein du complexe RXR/PPAR ou RXR/VDR (Chandra et al., 2008; Orlov et al., 2012). Ces études décrivent un rôle clé de cette région charnière dans la conformation et la stabilisation de l'hétérodimère sur l'ADN (Figure 5).

5. LA REGION F C-TERMINALE

La région F est très peu conservée en longueur et en séquence entre les sous-types de RAR et est absente chez la plupart des RN dont RXR. Aucune structure n'est connue et sa fonction dans l'activité du récepteur reste floue. Etant positionnée en continuité de l'hélice H12 du LBD, elle stabiliserait son exposition au solvant en conformation apo et favoriserait ainsi l'interaction avec les corépresseurs (Farboud and Privalsky, 2004). Enfin, cette région C-terminale est le siège de nombreuses phosphorylations (Rochette-Egly et al., 1997; Srinivas et al., 2006)(voir section A-IV-2.1, page 18.) et pourrait servir d'ancrage pour certains ARN messagers (Poon and Chen, 2008) (cf section A-IV-3.1, page 23.)

III. VOIE CLASSIQUE DE TRANSDUCTION DU SIGNAL DE L'AR : LA VOIE GENOMIQUE

1. LE MODELE BINAIRE

Les RAR sont des facteurs de transcription dépendants du ligand. En effet, en présence de celui-ci, ils transduisent son signal en activant l'expression de gènes cibles. Cependant, ils ont aussi un rôle répresseur en absence de ligand et répriment activement l'expression des gènes cibles en contrôlant l'état de compaction de la chromatine. Ils agissent donc via un modèle binaire ON/OFF décrit ci-après (Figure 6).

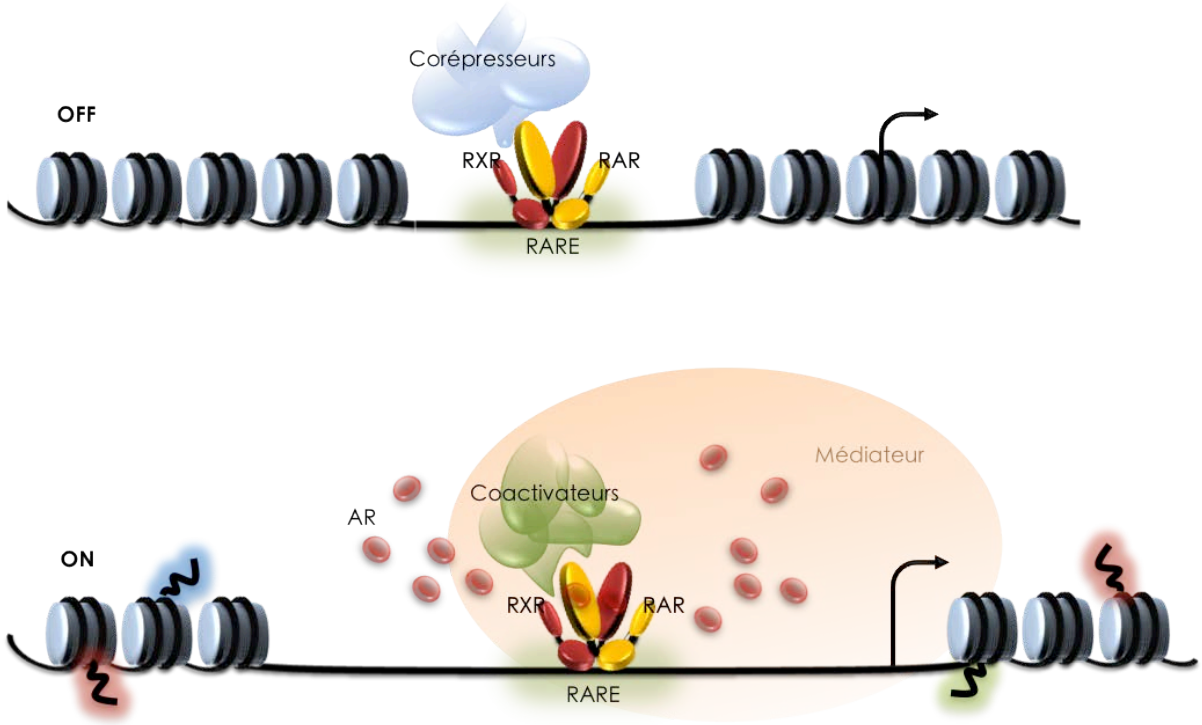


Figure 6: Modèle d'action binaire classique des RAR. En absence de ligand, l'hétérodimère RAR/RXR est lié à sont élément de réponse et interagit avec des corépresseurs qui maintiennent activement la chromatine compactée. Après fixation du ligand, des changements conformationnels au niveau du LBD induit l'interaction avec des coactivateurs qui détendent la chromatine pour l'initiation de la transcription.

1.1. OFF : Répression active en absence de ligand.

Selon le modèle classique, les hétérodimères RAR/RXR sont associés à leurs éléments de réponse au niveau de l'ADN dans le promoteur des gènes cibles et interagissent avec des complexes corépresseurs qui maintiennent activement la chromatine environnante compactée et donc impropre à la transcription (Chakravarti, 2009; Jones and Shi, 2003) (Figure 6). Parmi les corépresseurs les plus connus et les premiers identifiés, on trouve entre autres les protéines SMRT (*Silencing Mediator of Retinoid and Thyroid hormone receptors*) et N-CoR (*Nuclear hormone receptor Corepressor*). Ces deux corépresseurs présentent une forte conservation de séquence et des domaines en commun leur procurant des fonctions répressives similaires (Privalsky, 2004). Les différents domaines sont détaillés dans la figure 7, dont le domaine N-terminal DAD (*Deacetylase Activating Domain*) qui participe au recrutement et à l'activation des HDAC (*Histone DeACetylase*) ou les domaines NID (*Nuclear receptor Interaction Domain*) du côté C-terminal qui participent à l'interaction avec les RN (Cohen et al., 2001; Guenther et al., 2000). Au sein des NID, on retrouve des

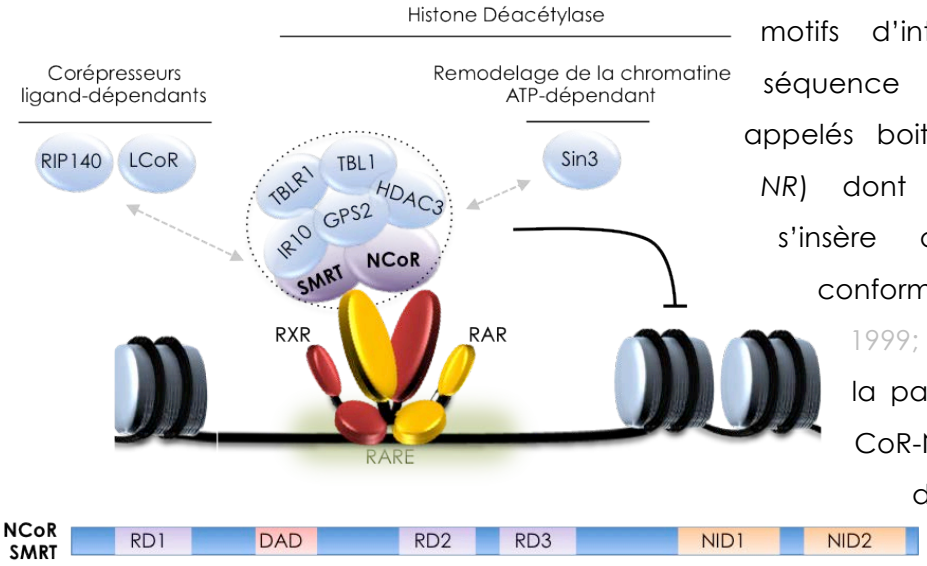


Figure 7: Corépresseurs et complexes associés. Présentation du « cœur » du complexe corépresseur (en pointillé) et des complexes associés. L'organisation modulaire commune des corépresseurs NcoR et SMRT indiquant les domaines RD = Domaine de répression (3 dans NCoR, 4 dans SMRT) et NID = domaine d'interaction avec les RN (3 dans NCoR et 2 dans SMRT)

motifs d'interaction définis par la séquence protéique LxxI/HIxxxI/L appelés boites CoR-NR (*CoRepressor-NR*) dont la structure hélicoïdale s'insère au sein du LBD en conformation apo (Hu and Lazar, 1999; Perissi et al., 1999). De plus, la partie N-terminale des boites CoR-NR masque la surface d'interaction de l'hélice H12 rendant ainsi exclusive l'interaction des corépresseurs au dépend des coactivateurs (Nagy et al., 1999).

Ces corépresseurs sont dépourvus d'activité enzymatique et ne constituent que des plateformes de recrutement pour de gros complexes multi-protéiques possédant des activités enzymatiques intrinsèques comme les HDAC (Guenther et al., 2001) (Figure 7). C'est principalement via la désacétylation des queues d'histones environnantes que les complexes corépresseurs maintiennent la chromatine compactée et répriment ainsi la transcription (Vermeulen et al., 2004). De plus, la formation du complexe de pré-initiation à la transcription (PIC) est activement inhibée par la présence de SMRT et N-CoR, assurant ainsi une inhibition complète de la transcription (Wong and Privalsky, 1998).

1.2. ON : Activation de la transcription en présence du ligand.

La fixation du ligand est l'élément clé de l'activation du récepteur en menant à des changements conformationnels au sein du LBD qui sont à l'origine de la dissociation des complexes corépresseurs, et exposent ainsi de nouvelles surfaces permettant le recrutement de complexes coactivateurs. Le changement majeur au cours de cette transconformation (Moras and Gronemeyer, 1998) induite par la fixation du ligand, conduit au basculement de l'hélice H12 sur la LBP (Figure 4). Cela permet l'exposition du domaine AF-2 pour des interactions avec des complexes coactivateurs. Ceux-ci possèdent un domaine NID

composé de boites NR au motif protéique LxxLL qui est nécessaire et suffisant pour l'interaction avec le sillon hydrophobe du domaine AF-2 (Chang et al., 1999; Darimont et al., 1998; Heery et al., 1997). Parmi ces coactivateurs, on trouve classiquement la famille protéique p160 dont les SRC (SRC-1, SRC-2, SRC-3, *Steroid Receptor Coactivator*) et la famille de protéines p300. Ils renferment une activité HAT (*Histone Acetyl-Transferase*) intrinsèque et agissent en synergie pour remodeler la chromatine et recruter la machinerie transcriptionnelle (Chakravarti et al., 1996; Demarest et al., 2002; Dilworth and Chambon, 2001). Cependant, ils servent aussi de plateforme de recrutement pour de nombreux complexes multi-protéiques aux activités enzymatiques diverses permettant un remodelage de la chromatine propice à la transcription (McKenna et al., 1999). Un aperçu de ces complexes et de leurs activités est présenté en figure 8.

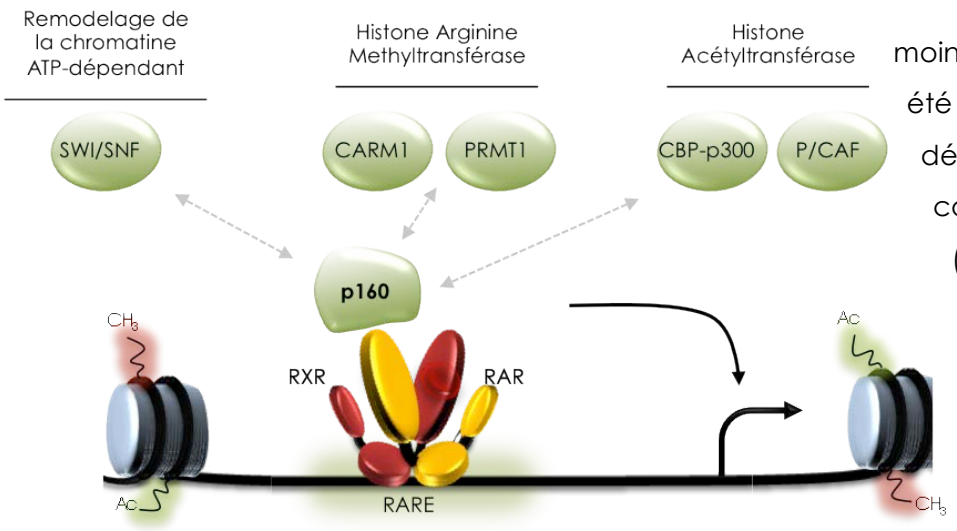


Figure 8: Coactivateurs et complexes associés. Après fixation du ligand, l'hétérodimère RAR-RXR recrute les coactivateurs dont les p160 qui jouent un rôle de plateforme pour le recrutement d'autres complexes coactivateurs à activité enzymatique.

D'autres coactivateurs moins conventionnels ont aussi été mis en évidence. Certains, dépourvus du motifs LxxLL comme la cycline H, TACC1 (*Transforming Acidic Coiled-Coil 1*) ou encore SUG1 (*Suppressor for Gal1*) ont été montrés comme interagissant avec les RAR et régulant positivement leur activité transcriptionnelle (Bour et al., 2005a; Ferry et al., 2009; Gianni et al., 2002; Guyot et al., 2010). A l'inverse, et de manière intéressante, certaines protéines possédant des motifs LxxLL ont été répertoriées comme corépresseurs, car bien qu'interagissant avec les RAR ayant liés l'AR, elles recrutent des complexes répresseurs de la transcription (Gurevich et al., 2007; White et al., 2004). Citons par exemple RIP140 (*Receptor Integrating Protein of 140 kDa*) (Farooqui et al., 2004; Heery et al., 2001), PRAME (*Preferentially expressed Antigen in Melanoma*) (Epping et al., 2005) ou encore LCoR (*Ligand-dependent CoRepressor*) (Fernandes et al., 2003).

1.3. Au delà du modèle classique.

Bien que le modèle classique ON/OFF que nous venons de décrire ait été validé expérimentalement, il apparaît que l'activation de l'expression des gènes cibles contrôlée par les RAR peut aussi dépendre de programmes différents. Ces différences dans le mode d'action des RAR semblent grandement dépendre du contexte chromatinien des promoteurs des gènes cibles régulés. Il a par exemple été montré que le complexe Médiateur de transcription ainsi que la RNA Pol II peuvent avec RARα déjà occupés les promoteurs de certains gènes même en absence de ligand (Flajollet et al., 2006; Pavri et al., 2005). De plus, tous les promoteurs des gènes cibles ne sont pas systématiquement occupés par les RAR en absence de ligand comme suggéré par le modèle classique. En effet, les éléments de réponse ne sont pas toujours disponibles pour accueillir l'hétérodimère RAR/RXR et certains remodelages chromatiniens induits par l'AR via des effets non génomiques peuvent être nécessaires pour permettre cette interaction (discuté en section A-IV-2.2-iii, page 22.) (Bruck et al., 2009).

En conclusion, l'activité transcriptionnelle des RAR, qu'elle soit répressive en absence de ligand, ou activatrice après exposition au ligand, est un phénomène extrêmement dynamique qui nécessite un recrutement coordonné et combiné de nombreux complexes protéiques (Rochette-Egly and Germain, 2009).

2. L'ADN COMME LIGAND ALLOSTERIQUE DES RAR

Les éléments de réponses des RAR (RARE) sont également des acteurs majeurs de leur action. En effet, ces sites lorsqu'ils sont accessibles, sont reconnus par les hétérodimères RAR/RXR et contrôlent ainsi leur recrutement spécifique au niveau d'éléments régulateurs de leurs gènes cibles. Ainsi, leur présence dans les promoteurs ou autres séquences *cis*-régulatrices de certains gènes leur confère généralement une régulation dépendante des RAR et donc de l'AR.

2.1. Les RARE canoniques

Les éléments de réponse classiques sont composés d'un motif nucléotidique hexamèrique de séquence RGKTSA (R = A ou G ; K = G ou T ; S = C ou G d'après la convention de l'UICPA) répété directement et séparé par 5 nucléotides (DR5)(Germain et

al., 2003; Leid et al., 1992; Umesono et al., 1991b). Plusieurs RARE de type DR5 ont été caractérisés dans les promoteurs de gènes cibles canoniques comme *RARb2* (de The et al., 1990), *cyp26a1* (Loudig et al., 2000) ou certains gènes *hox* (Dupe et al., 1997). La recherche *in silico* à l'échelle du génome entier de motifs DR5 conservés entre espèces de vertébrés a permis d'élargir la liste de RARE de type DR5 fonctionnels dans les promoteurs de gènes cibles pour lesquels aucun élément de réponse n'était encore connu ((Lalevee et al., 2011) publication en annexe, page 207).

Toutefois, l'hétérodimère RAR/RXR peut aussi se fixer à des RARE de type DR2 ou DR1 avec une affinité décroissante (Figure 9). Des éléments DR2 ont été identifiés dans les promoteurs des gènes CRBPI et CRABPII (Durand et al., 1992; Smith et al., 1991) et un élément DR1 dans le promoteur du gène de rat de CRBPII (Mangelsdorf et al., 1991). Les différences d'espacement entre les DR contrôlent l'orientation de l'hétérodimère sur l'ADN. Des expériences *in vitro* ont montré que pour les DR2 et DR5, le RAR occupe le demi-site en 3' et le RXR lie le demi-site en 5', alors que cette orientation est inversée dans le cas du DR1. Ces différences d'orientations induites par l'espacement entre les demi-sites influent sur la conformation générale de l'hétérodimère et donc sur son activité transcriptionnelle. Ainsi,

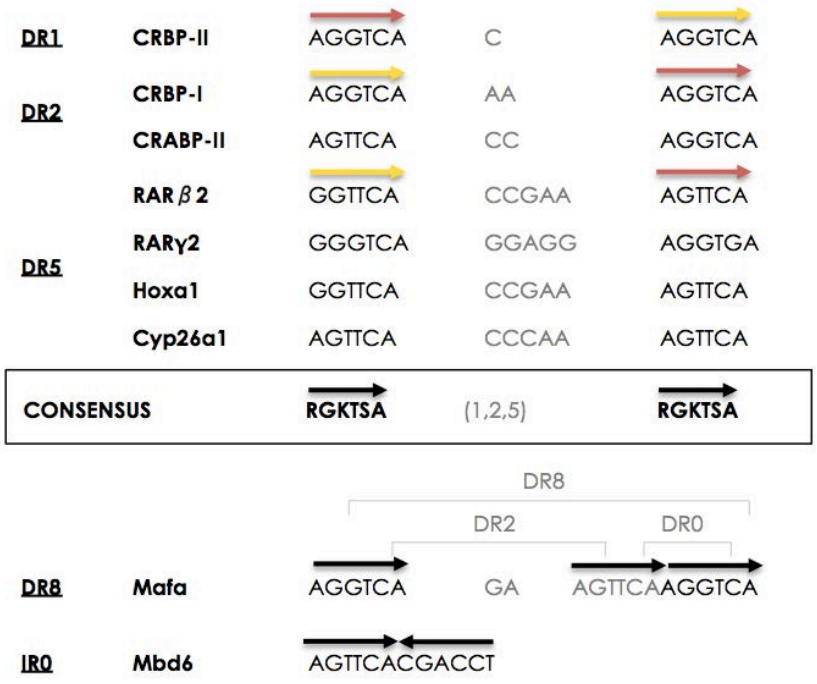


Figure 9: Les différents types de RARE. L'hétérodimère RAR-RXR peut se fixer sur des RARE avec différents espacements. Les RARE de type DR1, DR2 et DR5 ont été les premiers décrits et définissent la séquence consensus. Récemment, des concaténations de DR0 et DR2 en DR8 ainsi que des IR0 ont été mis en évidence. Les séquences données correspondant aux RARE trouvés dans les gènes de souris correspondants. Les demi-sites occupés par RAR sont en rouge et ceux par RXR en jaune

alors que la fixation de RAR/RXR sur un DR5 a un rôle activateur, l'hétérodimère devient répresseur sur un élément de type DR1 du fait d'une interaction constitutive avec des corépresseurs (Zamir et al., 1997). En plus de l'espacement, il a été montré dans le cas du récepteur aux glucocorticoïdes (GR), que la séquence nucléotidique de l'élément de réponse pouvait également jouer un rôle sur la conformation du DBD (Meijsing et al., 2009). Ainsi, des éléments de réponse de même conformation mais de séquences variables peuvent réguler différemment l'activité du récepteur.

L'ADN peut donc être considéré comme un ligand allostérique des RAR car sa liaison aux récepteurs régule la dynamique d'association-dissociation avec des corégulateurs et influence ainsi leur activité transcriptionnelle au même titre que leur ligand moléculaire qu'est l'AR (Kurokawa et al., 1995; Meijsing et al., 2009).

2.2. Un éventail plus large d'éléments de réponse

Des études récentes de séquençage haut débit après immunoprécipitation de la chromatine (ChiP-seq) dans des cellules indifférenciées de souris (cellules F9 ou corps embryonnaire de cellules souches) ont permis de mettre en évidence la gamme d'éléments de réponse reconnus effectivement par RAR. En effet dans ces modèles cellulaires, les RAR occupent un large répertoire de sites DR0, DR8 et IR0 (répétition inversée, figure 9) et l'affinité du récepteur pour ces RARE est aussi forte que celle pour les éléments classiques de type DR5 ou DR2 (données *in vitro*) (Moutier et al., 2012). Le nombre de sites occupés arborant un DR0 est même grandement supérieur à ceux comprenant un DR5. Cependant, contrairement aux DR8 et IR0, le motif DR0 ne semble pas pouvoir à lui seul réguler l'expression du gène qui lui est associé. Toutefois, et de manière intéressante, il est souvent observé une concaténation d'éléments DR0 et DR2 en DR8 (figure 9). Ce composite DR8 est fonctionnel et est retrouvé fréquemment dans les nouvelles données de ChIP-seq comme liant RAR/RXR, il est alors proposé de considérer les composites DR8 comme de nouveaux éléments de réponse pour les RAR. Concernant les sites liant RAR/RXR mais ne permettant pas de régulation transcriptionnelle, comme les DR0, ils pourraient être impliqués dans des occupations compétitives d'éléments normalement lié par d'autres facteurs de transcription (Gu et al., 2005).

En plus de l'espacement, la séquence consensus des demi-sites ne semble pas toujours répondre strictement au motif RGKTSA décrit jusqu'alors. En effet, de nombreux RARE dégénérés, c'est à dire différant par un ou deux nucléotides dans au moins un de leurs

demi-sites, ont été identifiés (Delacroix et al., 2010). Alors que les sites les plus fortement occupés par RAR/RXR dans des cellules souches de souris sont composés de RARE canoniques, les sites dégénérés correspondent à des sites moins fréquemment occupés. Cela suggère qu'il existe un continuum de différents RARE (en séquence et en espacement) qui lient l'hétérodimère avec des affinités différentes et dans différents contextes. Ainsi, en fonction du contexte cellulaire (quiescence, prolifération, différentiation) ou développemental, les sites occupés par les RAR pourraient être variables du fait des différences d'organisation de la chromatine, et ainsi participer à la régulation fine de l'activité des RAR (Delacroix et al., 2010; Lalevee et al., 2011).

Ces résultats issus de nouvelles techniques révèlent une grande diversité de topologies et d'espacement des RARE. Cette diversité semble être spécifique des RAR puisque d'autres RN comme VDR ou PPAR ont un répertoire de site de liaison beaucoup plus restrictif (Nielsen et al., 2008; Ramagopalan et al., 2010). La dégénérescence des RARE pourrait participer à réguler finement la liaison du récepteur à l'ADN en engendrant des niveaux d'affinité différents. Enfin, le répertoire de RARE occupés par les RAR diffèrent d'un type cellulaire à l'autre suggérant que la fonctionnalité de ces sites dépend du contexte cellulaire.

2.3. Distribution distale des éléments de réponse

Par des expériences de ChIP-seq, il a été montré que les éléments de réponse des RN au niveau de l'ADN sont souvent lointains des séquences codantes des gènes régulés (jusqu'à 200kb). De plus, la majorité de ces sites se retrouvent dans des régions inter-géniques ou introniques très distales des sites d'initiation de la transcription (TSS). Cela remet en question le modèle classique qui suggère une action proximale des RN comme de simples facteurs de transcription. En effet, les RN pourraient agir sur l'expression de gènes en

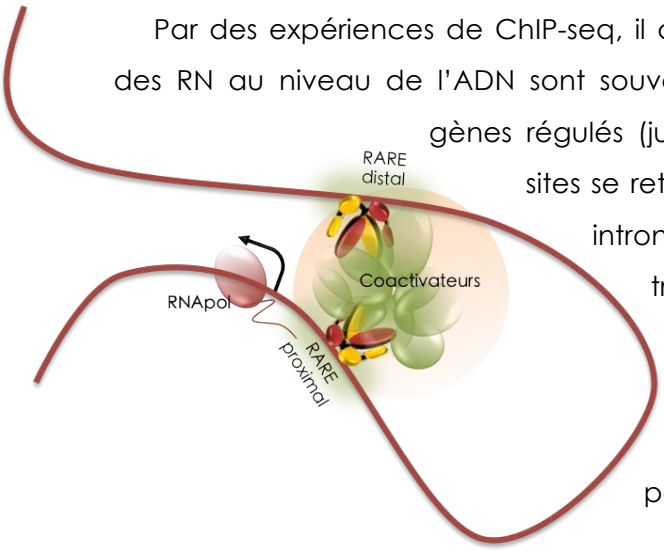


Figure 10: Illustration du rôle synergique des RARE proximaux et distaux. Dans le cas du gène *cyp26a1*, il a été montré un pontage entre le RARE distal et proximal. Cela implique la formation de longues boucles de chromatine pouvant rapprocher des éléments de plusieurs centaines de kilobases au niveau du promoteur. Le rapprochement de plusieurs RARE pourrait renforcer le recrutement de la machinerie transcriptionnelle pour initier la transcription.

tant qu'enhancers à longue portée (*long-range enhancers*) de par leur position distale par rapport aux promoteurs des gènes cibles (Carroll et al., 2006; Moutier et al., 2012; Nielsen et al., 2008; Welboren et al., 2009). De ce fait, la régulation de ces gènes ferait intervenir de longues boucles chromosomiques regroupant ainsi des éléments de réponse lointains au niveau des promoteurs des gènes régulés (Figure 10) (Biddie et al., 2010; Carroll et al., 2005). En particulier, des études récentes du laboratoire de Cécile Rochette-Egly ont montré que les deux éléments de réponse du gène *cyp26a1* de mammifère (DR5 distal et proximal) sont rapidement pontés après fixation des récepteurs à l'ADN et coopèrent ainsi pour l'initiation de la transcription induite par l'AR (Bruck et al., 2009).

2.4. Dialogue avec l'épigénome

L'épigénome qui constitue l'ensemble des modifications épigénétiques d'une cellule, joue également un rôle important dans la liaison des récepteurs à l'ADN. En effet, en plus de la séquence brute des éléments de réponse, les modifications épigénétiques de la chromatine vont, entre autres, déterminer leur accessibilité (Biddie et al., 2010). Comme vu précédemment, les RN recrutent des complexes protéiques à activité enzymatique qui modifient la chromatine environnante pour la garder compacte ou pour la rendre accessible (voir section A-III-1, page 8). Ainsi, de nombreuses modifications post-traductionnelles de queues d'histone favorisant l'accessibilité de la chromatine ont été corrélées à la liaison des RN au niveau des éléments de réponse (Kininis et al., 2007; Lupien et al., 2009). Cependant, il a été montré que la liaison même du récepteur est sensible à l'environnement épigénétique et que ces modifications jouent un rôle primordial dans le recrutement du récepteur au niveau de l'élément de réponse (Eeckhoute et al., 2006; John et al., 2008; Martens et al., 2011). En d'autres termes, certains prérequis épigénétiques sont nécessaires pour permettre l'accessibilité des éléments de réponse et donc la liaison du récepteur à l'ADN. Ces observations remettent en question le rôle des RN comme facteurs premiers permettant le remodelage chromatinien mais suggèrent plutôt un dialogue étroit avec d'autres voies de signalisation cellulaire qui régulent l'épigénome en amont. Dans le cas des RAR, il a été montré que certaines kinases activées par l'AR participent à l'accessibilité de la chromatine au niveau des RARE (Bruck et al., 2009) (voir section A-IV-2.2-iii, page 22.).

IV. VOIE NON GENOMIQUE DE L'AR : NOUVEAUX CONCEPTS

De nombreuses études, certaines issues du laboratoire de Cécile Rochette-Egly, ont montré avec évidence qu'en plus des effets génomiques classiques de l'AR via l'activité transcriptionnelle des RAR, le signal de l'AR était également transduit par des voies non génomiques moins conventionnelles, mais faisant également intervenir les RAR (Al Tanoury et al., 2013).

1. ACTIVATION DE VOIES KINASIQUES ET RAR MEMBRANAIRES

Dans différentes lignées cellulaires, il a été montré que l'AR active rapidement la protéine kinase PKC (Kambhampati et al., 2003) ainsi que la voie de signalisation PI3K/Akt (Bastien et al., 2006; Masia et al., 2007). L'AR active aussi la voie des MAPK/ERK dans différentes lignées de cellules neuronales lors de la différentiation des cellules souches embryonnaires en neurones (Miloso et al., 2004; Stavridis et al., 2010). Enfin, l'AR active la voie des p38MAPK et de la kinase MSK1 en aval dans plusieurs lignées cellulaires de mammifères (Alsayed et al., 2001; Gianni et al., 2002; Ren et al., 2007). De par l'activation extrêmement rapide de cette voie après traitement à l'AR (quelques minutes) (Bruck et al., 2009), ces effets ne peuvent être qu'indépendants de tous phénomènes transcriptionnels. Toutefois, il a été montré que l'activation de cette voie kinasique nécessite RARα puisque l'activation des p38MAPK n'est plus observée dans des cellules mutées (knock-out) pour ce gène (Bruck et al., 2009). Alors que les mécanismes moléculaires d'une telle activation via des processus non génomiques restaient peu compris, des études récentes du laboratoire de Cécile Rochette-Egly ont montré qu'une fraction du pool cellulaire de RARα est localisée

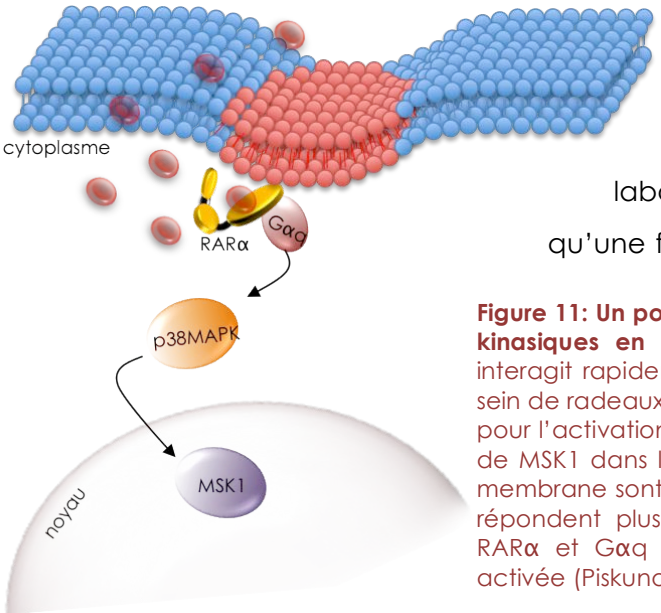


Figure 11: Un pool de RARα membranaire permet l'activation de voies kinasiques en réponse à l'AR. Dans des cellules humaines, RARα interagit rapidement avec la protéine G Gαq en réponse à l'AR au sein de radeaux lipidiques (en rouge). Cette interaction est nécessaire pour l'activation de la voie des p38MAPK et l'activation subséquente de MSK1 dans le noyau. Les mécanismes d'ancrage de RARα à la membrane sont inconnus. Dans des cellules de cancer du sein qui ne répondent plus à l'effet antiprolifératif de l'AR, l'interaction entre RARα et Gαq n'est pas observée et la voie p38MAPK n'est pas activée (Piskunov and Rochette-Egly, 2012).

dans la membrane cellulaire au niveau de radeaux lipidiques. C'est cette fraction membranaire de récepteur qui, via l'interaction avec la protéine G Gαq, permet l'activation de la voie des p38MAPK (Piskunov and Rochette-Egly, 2011b) (Figure 11). Contrairement aux autres récepteurs des stéroïdes, les RAR n'ont pas de site de palmitoylation qui faciliterait leur ancrage à la membrane (Pedram et al., 2007). Cependant, la partie N-terminale (région A/B) est requise pour la localisation de RARα au niveau des radeaux lipidiques membranaires suggérant que ce domaine non conservé entre les différents sous-types de RAR joue un rôle pour la localisation extranucléaire (Piskunov and Rochette-Egly, 2011b). A noter que d'autres RN, comme ER ou GR sont aussi capables d'activer la voie des MAPK via un pool de récepteurs membranaires (Le Romancer et al., 2011; Marquez et al., 2006; Matthews et al., 2008; Piskunov and Rochette-Egly, 2011a).

2. INTEGRATION NUCLEAIRE DES ACTIVATIONS KINASIQUES

Dans plusieurs types cellulaires, il a été montré qu'une fois activées par l'AR dans le cytosol, les kinases sont rapidement transloquées dans le noyau où elles activent d'autres kinases en aval. Dans le cas de l'activation des p38MAPK par l'AR dans des cellules de mammifères, c'est la kinase MSK1 qui est finalement activée dans le noyau (Bruck et al., 2009). Là, elle va transduire le signal de l'AR en phosphorylant de nombreuses protéines de manière coordonnée.

2.1. Phosphorylation des RAR

i. Les RAR sont des phosphoprotéines

Différents sites de phosphorylations ont été décrits au sein des domaines des RAR et ont été associés à des fonctions spécifiques (récapitulés dans la figure 12.) De nombreux signaux, en activant des kinases cytosoliques comme Akt, PKC ou JNK qui transloquent ensuite dans le noyau, peuvent induire la phosphorylation des RAR à diverses positions et ainsi réguler certaines de leurs activités et leur localisation (Figure 12) (Rochette-Egly and Germain, 2009). Nous nous intéresserons par la suite, plus particulièrement aux résidus qui sont phosphorylés en réponse à l'AR et qui modulent l'activité transcriptionnelle des RAR.

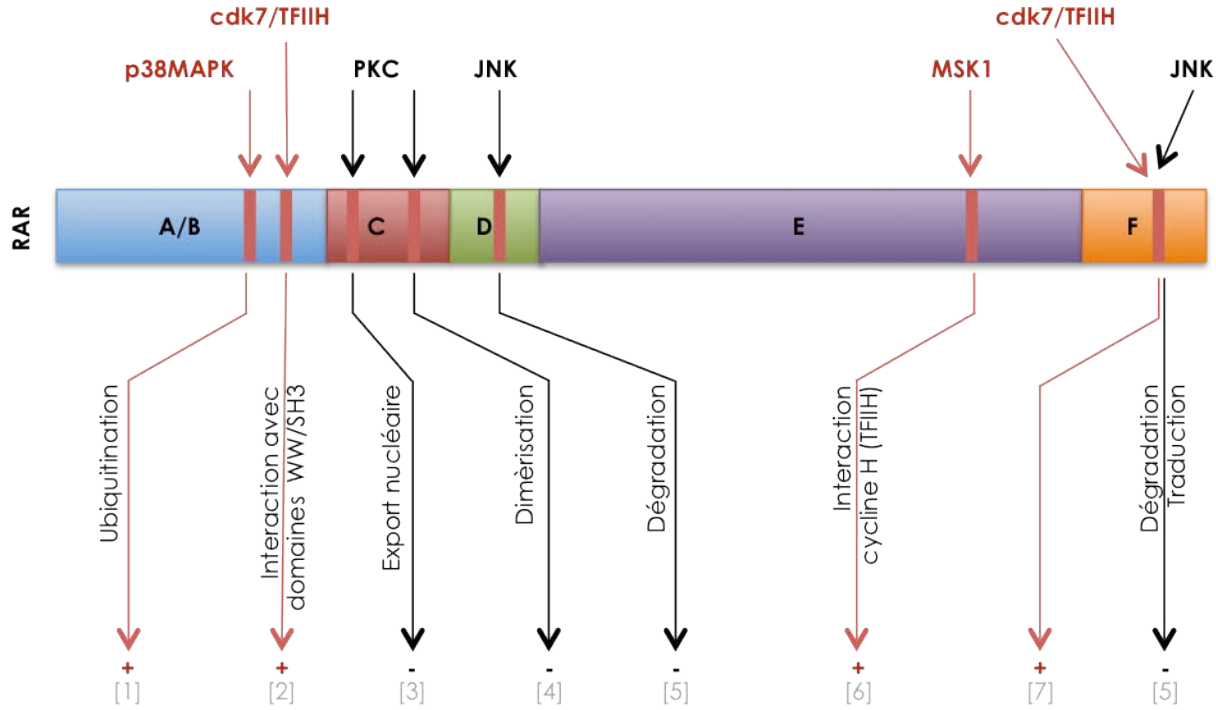


Figure 12: Principales phosphorylations de RAR. Récapitulation des voies de signalisation impliquées dans la phosphorylation des RAR et de leurs effets. Les voies en rouge sont celles induite par l'AR. La conséquence transcriptionnelle (positive ou négative) des phosphorylations est indiquée par un « + » ou un « – ». *[1]Gianni et al., 2002; [2]Bour et al., 2005; [3]Sun et al., 2007; [4]Delmotte et al., 1999; [5]Srinivas et al., 2005; [6]Bruck et al., 2009; [7]Rochette-Egly et al., 1997; [8]Poon and Chen, 2008.*

ii. Cascade de phosphorylations des RAR en réponse à l'AR

Des études récentes du laboratoire de Cécile Rochette-Egly ont montré que RARα était rapidement phosphorylé en réponse à l'AR au niveau de deux sérines : l'une située dans le LBD (S(LBD) en position 369 chez RARα de mammifère) et une autre située dans le NTD (S(NTD) en position 77 chez RARα de mammifère). Ces sites de phosphorylation sont conservés entre les sous-types de RAR (Figure 13). La sérine S(LBD) est localisée au sein d'un motif riche en arginine/lysine et qui correspond à un motif consensus de phosphorylation pour différentes kinases dont PKA et MSK1 (Bruck et al., 2009; Gaillard et al., 2006). Cette

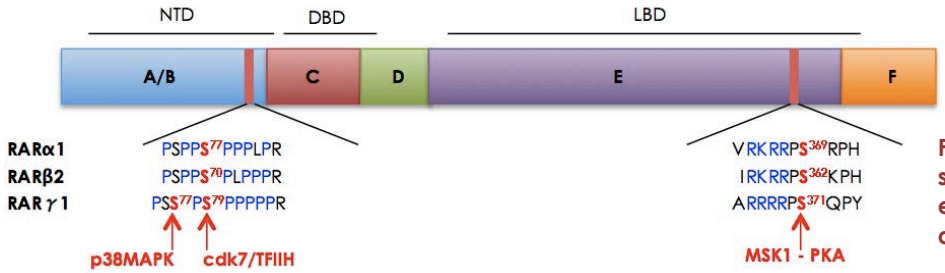


Figure 13: Conservation des sites de phosphorylation entre les sous-types de RAR de mammifères.

sérine est située sur une boucle entre les hélices 9 et 10 du LBD et est donc fortement exposée au solvant. L'autre sérine modifiée en réponse à l'AR, S(NTD), est située dans le NTD au niveau d'un motif riche en proline et est phosphorylée par la kinase cdk7. Cette kinase fait partie du facteur général de transcription TFIIH et son activité dépend de son association avec la cycline H et MAT1 au sein du sous-complexe CAK. De manière intéressante, il a été montré que la phosphorylation de S(LBD) est pré-requise pour la phosphorylation en aval de S(NTD) (Gaillard et al., 2006). En effet, la phosphorylation de S(LBD) est nécessaire au recrutement de la cycline H au niveau de la boucle entre les hélices 8 et 9 du LBD (site d'ancrage de la cycline H ou *cyclin H docking site*). C'est cette interaction qui permet le recrutement simultané de la kinase cdk7 (qui interagit avec la cycline H au sein du complexe CAK) et mène donc à la phosphorylation en aval de S(NTD) par cdk7 (Figure 14). Le recrutement de la cycline H au niveau de la boucle L8-9 induit par la phosphorylation de la S(LBD) au niveau de la boucle L9-10 suggère des changements allostériques subtiles induit par l'AR au niveau du LBD. Mes travaux de thèse ont permis d'élucider et de modéliser ces changements (voir Chapitre 2).

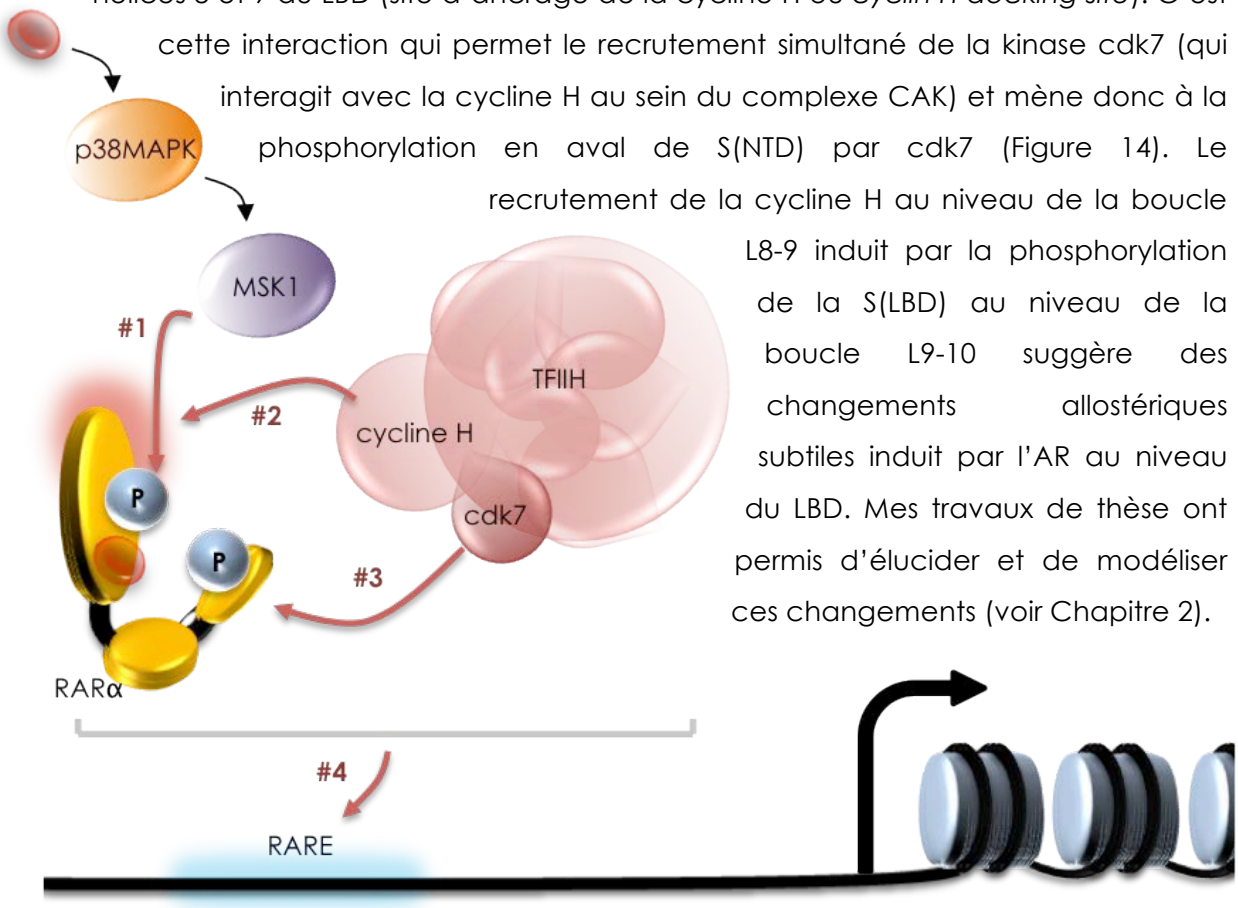


Figure 14: Cascade de phosphorylation coordonnée de RARα. MSK1 est rapidement activée par l'AR via des mécanismes non génomiques impliquant la voie des p38MAPK. **#1**: MSK1 phosphoryle RARα au niveau de S(LBD). **#2**: La phosphorylation de récepteur permet l'interaction avec la cycline H (halo rouge). **#3**: l'interaction RARα-cyclineH permet le bon positionnement de cdk7 pour phosphoryler RARα au niveau de S(NTD). **#4**: Tout le complexe est recruté au niveau du RARE dans le promoteur du gène cible.

D'un point de vue fonctionnel, la phosphorylation des RAR par TFIIH est requise pour la régulation correcte de l'expression des gènes cibles de l'AR dans des cellules de mammifères (Bastien et al., 2000; Rochette-Egly et al., 1997). De plus, il a été montré que

dans des cellules de patients atteint de la maladie *Xeroderma pigmentosum* où le gène d'une sous-unité du facteur TFIIH est muté (sous-unité XPD) , RARα n'est plus phosphorylé correctement en réponse à l'AR et la réponse transcriptionnelle de l'AR est anormale (Keriel et al., 2002). Plus récemment, il a été montré par immunoprécipitation de chromatine (ChIP) que la phosphorylation de S(NTD) de RARα par cdk7 est nécessaire pour son recrutement à l'ADN au niveau des RARE du gène *cyp26a1* (Bruck et al., 2009). Ainsi, des modifications post-traductionnelles au sein du NTD semblent réguler la dynamique d'interaction du récepteur avec l'ADN. Des études récentes de structures sur des récepteurs nucléaires stéroïdiens ont montré que le NTD est très proche du DBD et de l'ADN (Orlov et al., 2012), ainsi des protéines interagissant avec le NTD pourraient gêner physiquement la liaison du DBD à l'ADN. Aussi, la phosphorylation des RAR au niveau de S(NTD) régule l'association/dissociation de certains corégulateurs non conventionnels comme la vinexine beta (Lalevee et al., 2010), publication en annexe page 195. Ainsi, il est envisageable que la régulation de ces interactions par la phosphorylation de S(NTD) soit à l'origine du contrôle du recrutement du récepteur à l'ADN.

2.2. Phosphorylation des partenaires

i. RXR

Le partenaire d'hétérodimérisation RXR est aussi la cible de kinases et présente plusieurs sites de phosphorylation au niveau du NTD, du DBD et du LBD (Figure 15). Au sein du NTD, la phosphorylation constitutive de la sérine en position 22 par des Cdk est nécessaire pour l'action antiproliférative de l'AR (Bastien et al., 2002). Trois autres sérines du NTD sont phosphorylées en réponse à l'AR et sont importantes pour la transactivation des

Récepteur	Région	Résidus	Kinases	Références
RXRα1	NTD	S22	cdk/cyc	Bastien et al., 2002
		S32	JNK	Mann et al., 2005
		S61, S75 JNK		Adam-Stitah et al., 1999
		T87	JNK	Bruck et al., 2006
	DBD	T162	PKC	Sun et al., 2007
	LBD	S265	JNK	Adam-Stitah et al., 1999 Bruck et al., 2006 Zimmerman et al., 2006
		Y248, Y397	MSK4/SEK1	Lee et al., 2000

Figure 15: Domaines et site de phosphorylation des RXR et kinases associées.

gènes cibles de l'AR (Adam-Stitah et al., 1999; Gianni et al., 2003). Au niveau du DBD, la thréonine en position 162 est phosphorylée par PKC et semble impliquée dans la localisation cytoplasmique de RXR à la fin de son processus transcriptionnel (Sun et al., 2007). Enfin, trois résidus sont phosphorylés dans le LBD (S265, Y248, Y397) et semblent être associés à l'export nucléaire et à l'inhibition de l'activité de RXR après un stress cellulaire (Lee et al., 2000; Zimmerman et al., 2006).

ii. Corégulateurs

Les corégulateurs, répresseurs ou activateurs sont aussi sujets à des phosphorylations qui participeraient à leur interaction dynamique avec les RAR. Par exemple, la protéine TBLR1 associée au complexe corépresseur est phosphorylée en réponse à l'AR et cette modification permet le recrutement du protéasome pour la dégradation des protéines du complexe comme SMRT ou NCoR (Perissi et al., 2008). Cela suggère une coopération entre la dissociation des corépresseurs suite à la fixation du ligand au niveau du LBD et leur dégradation induite par ces phosphorylations en réponse à l'AR.

Le coactivateur SRC-3 est également phosphorylé en réponse à l'AR par la kinase p38MAPK. Cette modification favorise aussi sa dissociation des RAR puis sa dégradation par le protéasome pour permettre le recrutement d'autre corégulateurs (Ferry et al., 2011), publication en annexe page 221. Ainsi, tout au long du processus transcriptionnel induit par l'AR, les phosphorylations réguleraient la dynamique et la coordination du recrutement des différents partenaires des RAR

iii. Histones

Dans des expériences de traitement de cellules de mammifères avec des agonistes des rétinoïdes, il a été montré que les queues des histones H3 étaient phosphorylées, spécifiquement au niveau des promoteurs de gènes cibles de l'acide rétinoïque comme *RARb2* (Lefebvre et al., 2002). Selon le code histone, la phosphorylation au niveau de la sérine 10 des histones H3 est associée a une décompaction de la chromatine et une activation de la transcription (Vicent et al., 2006). De plus, la kinase MSK1 activée par l'AR est recrutée au niveau des RARE du gène *cyp26a1* et est associée à la phosphorylation de H3S10 (Bruck et al., 2009). Ainsi, il est suggéré que MSK1, en plus de cibler les RAR, prépare la chromatine pour l'initiation de la transcription des gènes cibles de l'AR (Figure 17, page 25).

3. NOUVEAUX PARTENAIRES DES RAR

3.1. Interaction avec des ARNm

Récemment, un nouveau rôle indépendant de son activité transcriptionnelle a été mis en évidence pour RARα. Dans des cultures de neurones, l'AR joue un rôle dans la dynamique de libération et d'actions de certains neurotransmetteurs au niveau des synapses et cet effet est dépendant de RARα (Chen and Napoli, 2008; Maghsoodi et al., 2008). En effet, celui-ci est localisé dans le noyau et dans le cytoplasme (dendrites) de ces cellules au sein de granules composées d'ARNm ainsi que de facteurs nécessaires à leur traduction. Là, RARα interagit directement via son domaine C-terminal avec les ARNm codants pour les récepteurs du glutamate GluR1 et GluR2. Cette interaction séquestre ces ARNm et empêche leur traduction (Chen et al., 2008; Poon and Chen, 2008). Suite à la fixation du ligand et aux changements conformationnels au sein du LBD induit par celui-ci, les ARNm sont relâchés et rapidement traduits (Chen and Napoli, 2008) (Figure 16). Ainsi, via des effets non génomiques dans le cytoplasme, RARα et l'AR seraient impliqués dans la plasticité synaptique en régulant finement la traduction d'ARNm (Aoto et al., 2008).

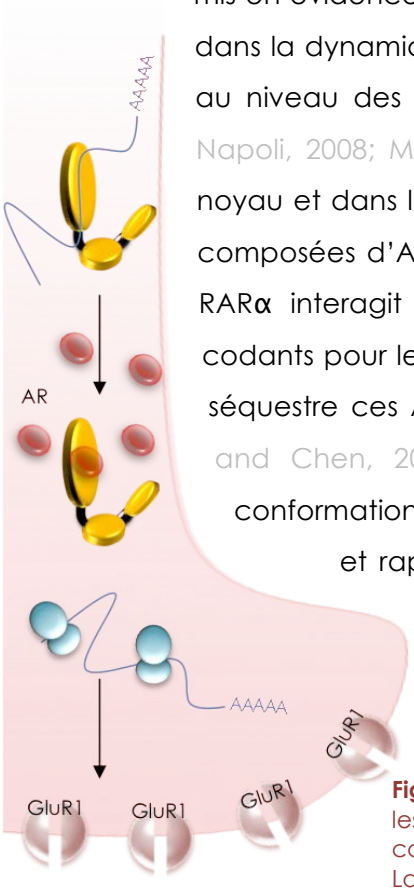


Figure 16: Régulation de la traduction des ARNm GluR1 par RARα. Dans les neurones, RARα interagit au sein de granules d'ARN avec l'ARNm codant pour le récepteur du glutamate GluR1 et réprime sa traduction. La fixation de l'AR permet la dissociation de l'ARNm et sa traduction rapide. Ainsi, en réponse à l'AR, les récepteurs GluR1 sont rapidement synthétisés et exprimés à la synapse.

3.2. Interaction avec des protéines liant l'actine (ABP)

Une des principales protéines de la matrice nucléaire, l'actine, a été identifiée comme essentielle pour la formation du complexe de pré-initiation de la transcription et régule ainsi de nombreux processus transcriptionnels (Percipalle and Visa, 2006). Son rôle dans la transcription est associé à de nombreuses autres protéines capables de lier l'actine, dénommées ABP (*Actin Binding Protein*) (Zheng et al., 2009). Parmi ces protéines, la vinexine beta a été identifiée comme interagissant avec RARγ au niveau du motif riche en proline du NTD. La liaison de la vinexine beta au récepteur empêche son recrutement au niveau des promoteurs de gènes cibles. La phosphorylation du NTD déstabilise cette association et

permet à RARγ de lier les RARE des gènes régulés. Ainsi la vinexine beta est une ABP considérée comme corépresseur des RAR (Lalevee et al., 2010).

De nombreuses autres ABP interagissent avec d'autres récepteurs nucléaires telles que la gelsoline, Fli1, Actinine 2 ou encore la cofiline (Gettemans et al., 2005; Huang et al., 2004; Lee et al., 2004; Ting and Chang, 2008). Ces résultats confirment l'implication de l'actine et de ses protéines associées dans la régulation de la transcription. De plus, ils élargissent la liste de protéines régulatrices des RN et suggèrent un rôle important de la matrice nucléaire dans leur activité transcriptionnelle.

En conclusion, l'ensemble de ces résultats suggère une coopération des effets génomiques et non génomiques de l'AR pour initier la transcription de gènes cibles (Figure 17). De plus, il apparaît que ce processus transcriptionnel est extrêmement dynamique avec des recrutements et des échanges rapides et coordonnés de nombreux corégulateurs et autres partenaires. Les phosphorylations apparaissent comme des modifications majeures qui, de par leur caractère rapide et réversible, agissent comme de véritables horloges moléculaires régulant les interactions protéiques au cours du processus transcriptionnel (Rochette-Egly and Germain, 2009; Samarut and Rochette-Egly, 2012). De manière intéressante, mes travaux de thèse ont permis de mettre en évidence que les sites de phosphorylation des RAR en réponse à l'AR ne sont pas tous conservés au sein des vertébrés et que l'acquisition de résidus phosphorylables au cours de l'évolution aurait permis l'ajout d'un mécanisme de régulation fine de l'activité des RAR (Chapitre 2, page 139).

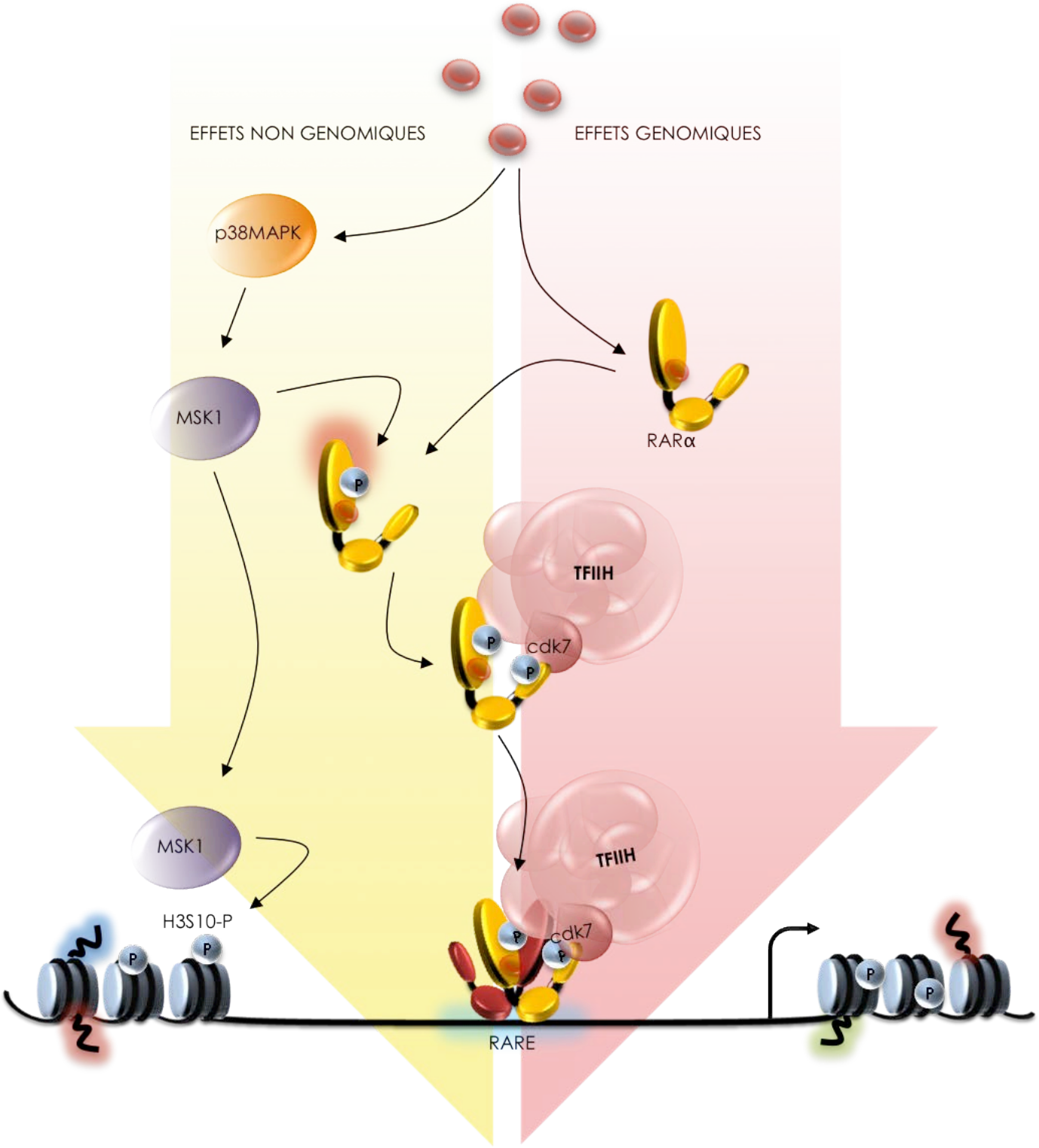


Figure 17: Coopération des effets génomiques et non génomiques de l'AR. L'AR active la voie de p38MAPK et la kinase MSK1 qui phosphoryle le récepteur au niveau de S(LBD), permettant le recrutement de TFIIH via la cycline H. La kinase cdk7 phosphoryle le récepteur au niveau de S(NTD). En parallèle, la kinase MSK1 est recrutée dans la région promotrice du gène cible et phosphoryle la S10 des histones 3 permettant la décompaction de la chromatine. Ainsi, le complexe RAR-TFIIH est recruté à la chromatine au niveau des RARE pour initier la transcription.

B. EFFETS DEVELOPPEMENTAUX DE L'AR ET IMPLICATIONS EVOLUTIVES

I. ORIGINE ET EVOLUTION DE LA VOIE DE L'AR

1. ORIGINE EVOLUTIVE DE LA VOIE DE L'AR

Bien qu'il soit connu depuis maintenant longtemps que la voie de l'acide rétinoïque n'est pas limitée aux vertébrés, les études fonctionnelles du rôle de l'AR chez les invertébrés restent rares (Denuce, 1991; Escriva et al., 2002; Holland and Holland, 1996; Schubert et al., 2005). Cependant, certains travaux ont montré la présence de composants de la voie de l'AR dans le génome de quelques invertébrés non chordés comme les échinodermes ou les hémichordés suggérant ainsi que l'origine de la voie de l'AR serait antérieure aux chordés (Campo-Paysaa et al., 2008; Cañestro et al., 2006; Simoes-Costa et al., 2008). En effet, des analyses *in silico* ont révélé l'existence de gènes homologues aux *rar*, *cyp26* et *raldh* chez l'oursin de mer (ambulacraires) repoussant ainsi l'existence putative de la voie de l'AR à l'ancêtre commun de tous les deutérostomes (Figure 18) (Campo-Paysaa et al., 2008; Howard-Ashby et al., 2006; Ollikainen et al., 2006). Enfin, des membres de la voie de l'AR dont les gènes *rar* et *rxr* ont été identifiés chez certains protostomes comme des mollusques ou des annélides (lophotrochozoaires) suggérant que la voie de l'AR n'est pas spécifique des deutérostomes (Figure 18) (Campo-Paysaa et al., 2008). De plus, des travaux en cours dans le laboratoire de Vincent Laudet ont mis en évidence que les RAR de lophotrochozoaires peuvent lier l'AR et

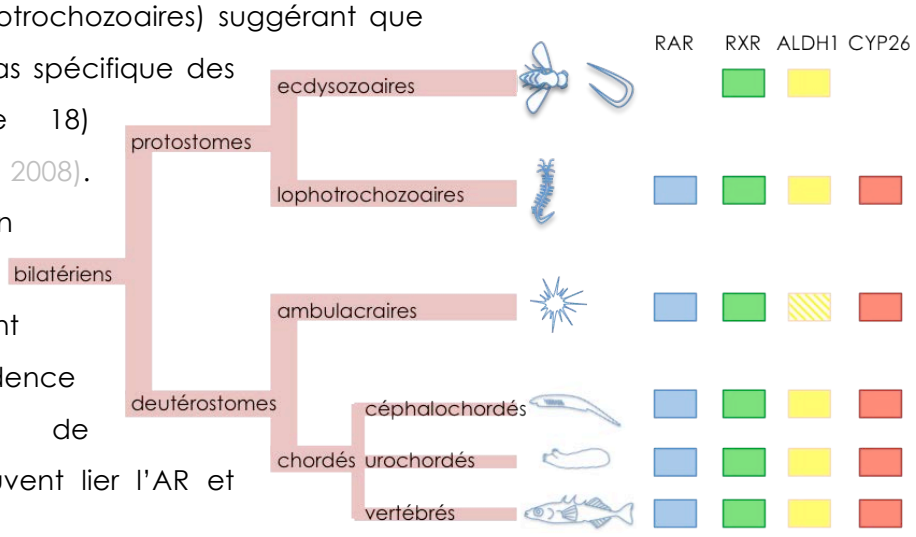


Figure 18: Gènes de la voie de l'AR chez les métazoaires. Figure adaptée de Campo-Paysaa et al., 2008. La présence des gènes de la voie de l'AR chez différentes espèces a été déterminée par BLAST et est indiquée par un carré de couleur. Les carrés rayés signifie que le gène n'a pas été retrouvé chez l'espèce en question mais est présent chez une autre espèce du phylum.

d'autres dérivés de la vitamine A et peuvent transactiver l'expression de gènes rapporteurs sous le contrôle d'un RARE, confirmant ainsi que l'homologie inférée grâce a l'analyse des séquences s'étend également aux aspects fonctionnels (Gutierrez-Mazariegos et al., en préparation). Ces données impliquent une origine ancienne de la voie de l'AR chez l'ancêtre commun des bilatériens et une perte subséquente chez les ecdysozoaires (Figure 18).

2. EVOLUTION DES RAR CHEZ LES VERTEBRES

Comme décrit précédemment, trois gènes paralogues de *rar* sont présents chez les mammifères, codants pour chacun des sous-types α, β et γ de RAR alors qu'une seule copie du gène orthologue des *rar* est présente chez l'amphioxus, codant pour un RAR unique nommé AmphiRAR (Escriva et al., 2002). L'amphioxus est un céphalochordé ayant divergé de la lignée des vertébrés avant l'occurrence des évènements de duplication de génome (*whole genome duplication*, WGD) (Figure 19). En effet, deux évènements majeurs de duplication de génome ont eu lieu au cours de l'évolution au sein de la lignée des vertébrés (Dehal and Boore, 2005; Holland et al., 2008; Taylor and Raes, 2004). Ainsi, pour chaque groupe de gènes paralogues présents chez les vertébrés (comme les trois sous-types de RAR), l'amphioxus ne contient en général qu'une copie du gène orthologue en question (AmphiRAR) (Garcia-Fernandez and Holland, 1994). Ainsi, l'étude comparative de la régulation et des fonctions d'un gène entre une espèce de vertébré et l'amphioxus permet d'appréhender la question du rôle des duplications de génome dans l'acquisition de nouveaux traits phénotypiques au cours de l'évolution (voir section B-III-1.2, page 40).

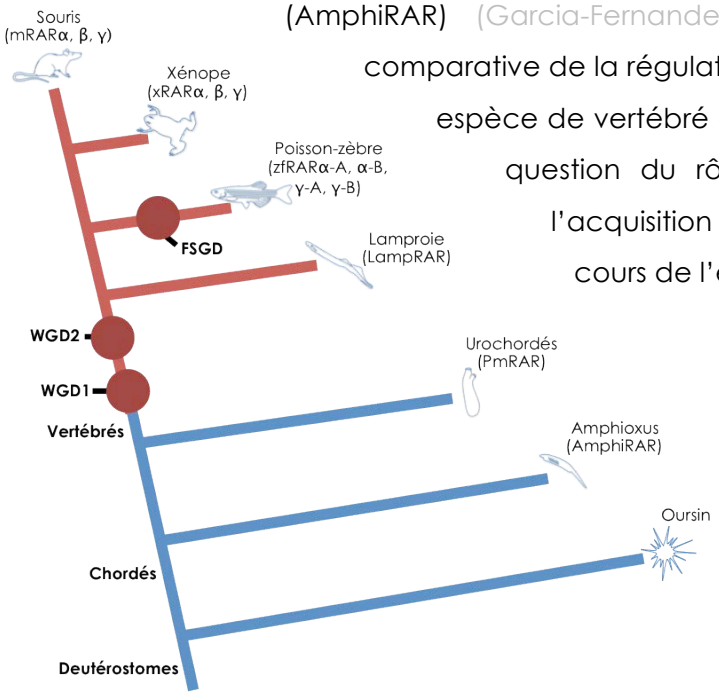
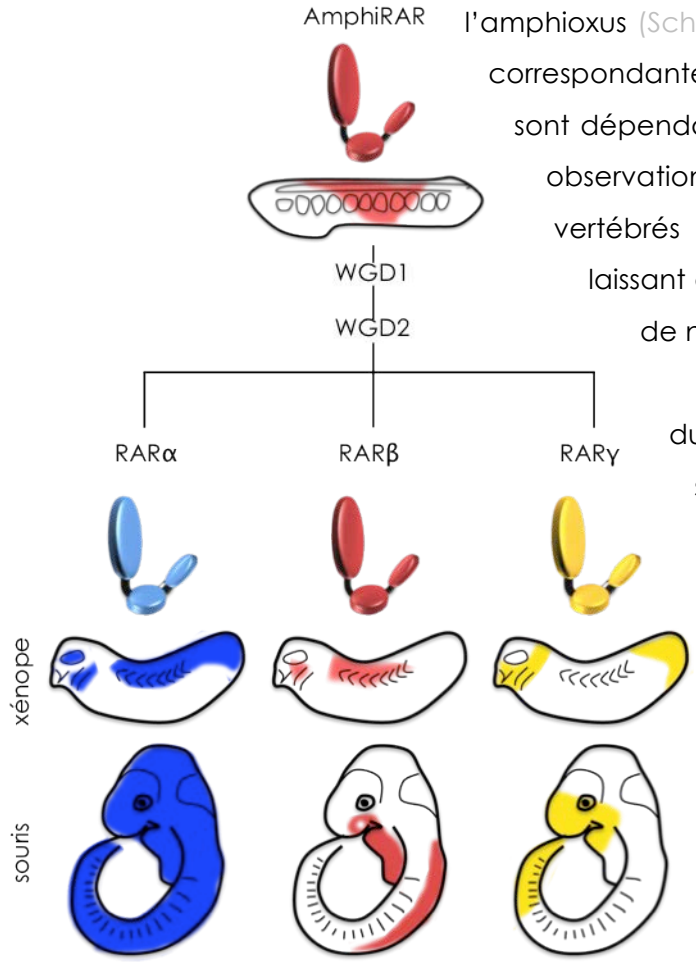


Figure 19: Phylogénie des Deutérostomes. A noter la position basale de l'amphioxus au sein des chordés. Les deux évènements de duplication de génome (WGD) sont indiqués ainsi que la duplication de génome spécifique des poissons (FSGD).

En particulier, une étude du laboratoire de Vincent Laudet s'est intéressée à la régulation de l'expression des sous-types de RAR ainsi qu'à leur fonction respective dans une étude comparative au sein

des chordés (Escriva et al., 2006). De manière intéressante, ils ont montré que les RAR de chordés ont évolués à partir d'un récepteur ancestral qui liait l'AR et dont la LBP ressemble fortement à celle de RARβ de mammifère. A l'inverse, les sous-types α et γ des RAR de mammifères semblent avoir accumulé certaines mutations au sein de leur LBP leur conférant de subtiles différences structurales qui influencent leur affinité pour les rétinoïdes. D'autre part, leur patron d'expression et leur fonction au cours du développement sont en accord avec cette hypothèse d'une néo-fonctionnalisation des sous-types α et γ. En effet, le patron d'expression d'AmphiRAR au niveau du tronc de l'embryon d'amphioxus ressemble à celui de RARβ dans un embryon de souris. A l'inverse, RARα et RARγ de souris semblent avoir acquis une expression dans de nouveaux territoires à savoir dans le cerveau antérieur et dans la queue respectivement (Figure 20). Enfin, l'AmphiRAR est impliqué dans le développement de l'endoderme pharyngale chez l'amphioxus (Schubert et al., 2005) et les structures orthologues correspondantes chez les mammifères (région branchiale) sont dépendante de RARβ (Matt et al., 2003). Cette ultime observation confirme que le sous-type β des RAR de vertébrés a conservé ses caractéristiques ancestrales laissant ainsi « champ-libre » aux sous-types α et γ pour de nouvelles fonctions.



A noter qu'un troisième épisode de duplication de génome a eu lieu spécifiquement chez les poissons téléostéens comme le poisson zèbre (FSGD, (Jaillon et al., 2004; Meyer and Van de Peer, 2005). Ainsi, le nombre de RAR paralogues a doublé en passant de trois à six chez la plupart des poissons téléostéens comme le medaka *Oryzias latipes* par rapport aux vertébrés. Cependant, plusieurs études de génomique comparative effectuées avec

Figure 20: Néofonctionnalisation des RAR chez les vertébrés. Adapté d'Escriva et al., 2006: L'expression et les caractéristiques de liaison du ligand d'AmphiRAR chez l'amphioxus ressemblent à celles du sous-type beta des vertébrés (expression dans la région du tronc en rouge). Les sous-types alpha (bleu) et gamma (jaune) issus des duplications de génome (WGD) ont acquis de nouvelles expressions dans d'autres régions de l'embryon.

les RAR du poisson zèbre et d'autres vertébrés suggèrent que le poisson zèbre a totalement perdu les sous-types issus de la duplication de l'orthologue RARβ et ne possède que quatre sous-types de RAR (α-A, α-B, γ-A, γ-B) (Hale et al., 2006; Tallafuss et al., 2006). De manière intéressante, l'invalidation de RARα-A chez le poisson zèbre génère des anomalies du développement du cerveau postérieur et du pharynx similaires à la perte de RARβ chez la souris (Linville et al., 2009). De plus, c'est le seul sous-type de RAR non exprimé dans la nageoire en développement tout comme RARβ qui est le seul RAR qui n'est pas exprimé dans les membres antérieurs (Bertrand et al., 2007; Mollard et al., 2000). Ces résultats suggèrent que le sous-type RARα-A récemment dupliqué chez le poisson zèbre pourrait avoir acquis des caractéristiques ressemblantes au sous-type beta des mammifères et pourrait ainsi compenser la perte des sous-type beta (Linville et al., 2009).

En conclusion, l'ensemble de ces données suggère que la duplication des gènes codants pour les RAR au cours de l'évolution des vertébrés pourrait avoir participé à la mise en place de nouvelles spécifications du développement propres aux vertébrés. Cependant, aucune donnée à l'échelle transcriptionnelle n'a jamais décrit une activité spécifique de chaque sous-type de RAR. Mes travaux de thèse ont permis de mettre en évidence qu'il existe des répertoires de gènes différentiellement régulés par les différents sous-types de RAR chez l'embryon de poisson zèbre (Chapitre 1, page 89).

II. EFFETS PLEIOTROPES DE L'AR AU COURS DU DEVELOPPEMENT DES CHORDES

L'AR est une molécule pléiotrope qui est impliquée dans de nombreux processus développementaux et plusieurs articles de revue ont entrepris de récapituler ses effets au cours du développement de différentes espèces de vertébrés (Mark et al., 2009; Niederreither and Dollé, 2008; Rhinn and Dollé, 2012). Dans cette partie introductive, je ne souhaite pas faire une liste exhaustive de tous ces effets, ce qui serait utopiste au vu de la masse d'informations disponible, ainsi que redondant avec les revues récentes à ce sujet, mais plutôt donner un aperçu général des effets de l'AR sur les trois principaux feuillets embryonnaires (ectoderme, mésoderme et endoderme) et leurs structures dérivées (Figure 21). Elle se veut principalement illustrative de la nature pléiotrope du rôle de l'AR au cours

du développement précoce chez les vertébrés. En effet, il régule à la fois la spécification des grands axes de l'embryon dans tous les feuillets embryonnaires, mais est aussi requis pour le développement de nombreux organes dérivés. De plus, il agit principalement comme un morphogène en fournissant différentes identités positionnelles selon sa concentration locale en activant des gènes homéotiques. Cette action repose sur la mise en place d'un gradient qui est établit par la dynamique de synthèse et de dégradation de l'AR, elle-même régulée par l'expression des enzymes du métabolisme de l'AR à des points clés de l'embryon. Cela illustre également qu'il existe des zones répondantes ou non au signal de l'AR, déterminées entre autres par l'expression localisée des RAR. Enfin, ce sont les nombreuses interactions avec d'autres voies de signalisation qui permettent à l'AR d'avoir des effets variés et qui assurent la coordination temporelle et spatiale de ces effets au cours du développement.

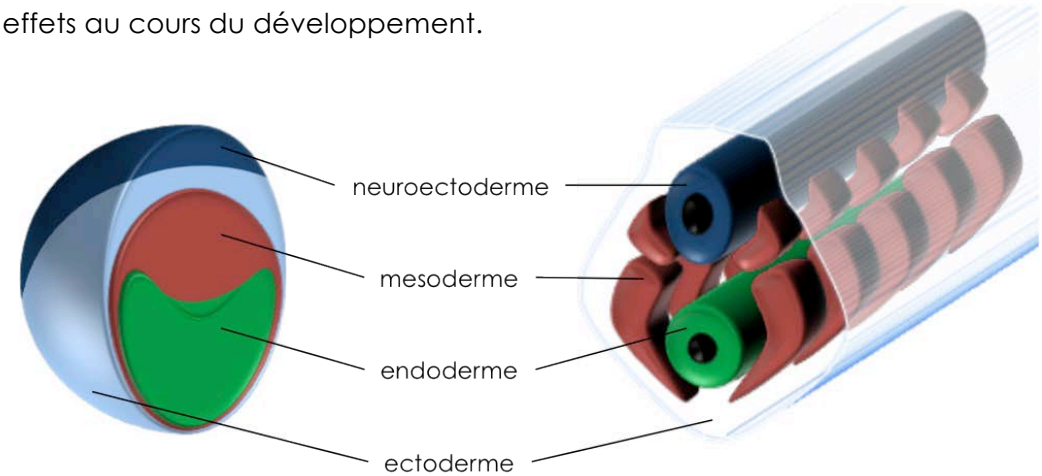


Figure 21: Illustration des différents feuillets embryonnaires. Les feuillets embryonnaires se forment lors de la gastrulation (à gauche) et se mettent en place selon le plan d'organisation de l'embryon via de multiples processus (neurulation, somitogénèse). Le neuroectoderme est à l'origine des structures nerveuses, l'ectoderme donnera les structures épidermiques, Le mésoderme donnera principalement les tissus musculaires alors que l'endoderme est à l'origine du tube digestif et des glandes annexes. A noter que les cellules de crêtes neurales ne sont pas représentées mais sont considérées comme un 4ème feuillet embryonnaire.

Il est également intéressant de noter que de nombreux effets développementaux de l'AR sont retrouvés chez des espèces d'invertébrés suggérant ainsi une origine ancienne de la voie de l'AR chez les métazoaires. D'autre part, certains rôles de l'AR semblent être spécifiques des vertébrés et pourraient correspondre à des innovations (Albalat, 2009). Cela suggère que la voie de l'AR via l'apparition de nouvelles régulations développementales, pourrait contribuer à l'acquisition de nouveautés morphologiques au cours de l'évolution.

1. NEUROECTODERME, SYSTEME NERVEUX CENTRAL ET DEVELOPPEMENT NEURAL

1.1. Développement du cerveau postérieur

Le rôle de l'AR au cours de la mise en place du système nerveux central (SNC) des vertébrés a été principalement étudié dans le contexte de la mise en place de l'axe antéro-postérieur (A/P). En effet, l'AR a des actions spécifiques sur différentes zones du SNC lors de sa mise en place, tout particulièrement dans le cerveau postérieur. Là, il agit comme un morphogène en spécifiant différentes régions du cerveau postérieur, les rhombomères, selon sa concentration suivant un gradient antéro-postérieur (Godsave et al., 1998; Niederreither et al., 2000). Il semblerait que l'action de l'AR soit double au cours du développement du cerveau postérieur en induisant en premier lieu la segmentation physique des rhombomères puis, plus tardivement, en spécifiant l'identité de chacun d'eux (Marshall et al., 1992; Wood et al., 1994). L'identité de chacun des rhombomères est définie par l'expression spécifique de différents facteurs de transcription dont les gènes *pbx*, *teashirt*, *meis* ou encore *hox* (Figure 22) (Erickson et al., 2011; Lumsden and Krumlauf, 1996; Moens and Prince, 2002). L'expression des gènes *hox* est colinéaire le long de l'axe A/P et cette colinéarité est fortement régulée par l'AR (Glover et al., 2006; Murakami et al., 2004; Wada et al., 1999). En effet, l'altération de la voie de l'AR au cours de la mise en place du cerveau postérieur mène à des transformations homéotiques au niveau des rhombomères (Dupe and Lumsden, 2001; Marshall et al., 1992). En particulier, un excès d'AR provoque une extension antérieur du cerveau postérieur et inhibe la formation de structures cérébrales antérieures (Durston et al., 1989). Au niveau moléculaire, l'AR active l'expression de gènes exprimés dans le cerveau postérieur comme *hoxb3*

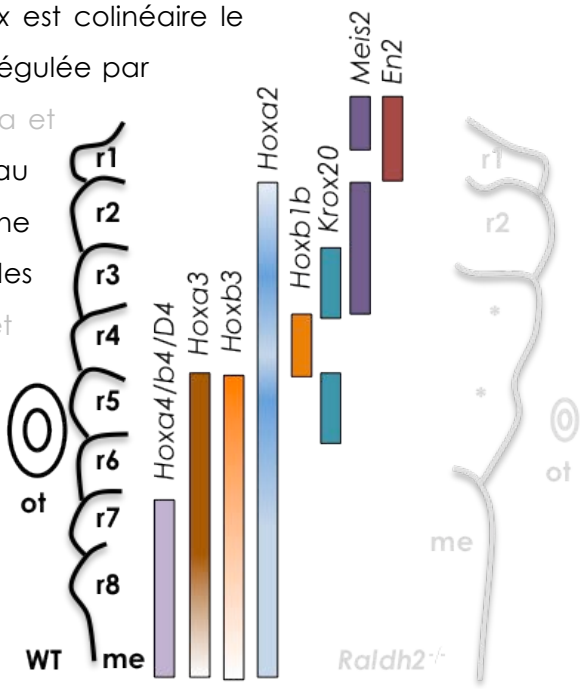


Figure 22: Identité moléculaire des rhombomères lors du développement du cerveau postérieur. Adaptée de Niederreither et al., 2000. L'expression des marqueurs dans le cerveau postérieur murin de 8.5 dpc est indiqué par des carrés colorés, les nuances d'intensité reflètent le niveau d'expression. La morphologie des rhombomères chez un embryon sauvage est à gauche et les altérations d'identités chez un mutant raldh2-/- sont illustrées à droite (r1-r8: rhombomères; me: moelle épinière; ot: vésicule otique; *: identité non clairement définie).

ou *hoxb4* et réprime des marqueurs de structures plus antérieures comme *otx2* ou *engrailed-2* (Godsave et al., 1998). Ainsi, une concentration forte d'AR spécifie une identité postérieure alors qu'un niveau plus faible d'AR détermine les rhombomères antérieurs. Ces différences d'activité dépendantes de la concentration en AR nécessitent l'expression dynamique des enzymes de synthèse et de dégradation de l'AR dans l'embryon. En particulier, il apparaît que l'expression des enzymes CYP26 est finement régulée pour permettre le contrôle robuste de la concentration en AR tout au long de l'axe A/P (Abu-Abed et al., 2001; Emoto et al., 2005; Hernandez et al., 2007; White et al., 2007). La mise en place du gradient A/P de l'AR par l'expression des enzymes du métabolisme de l'AR est détaillée dans la section D de l'introduction, page 65. L'action de l'AR dans la mise en place du cerveau postérieur dépend également d'autres voies de signalisation, comme la voie des FGF (*Fibroblast Growth Factor*) qui permet, entre autres de définir clairement la limite entre le cerveau postérieur et cerveau moyen (*hindbrain/midbrain boundary*) en antagonisant les effets postériorisants de l'AR (Schilling, 2008). La voie des Wnt semble aussi impliquée dans la régulation de l'expression de *cyp26a1* et donc dans le contrôle du gradient d'AR (Kudoh et al., 2002).

Des études effectuées sur le céphalochordé amphioxus et chez certaines espèces d'urochordés ont montré que l'expression colinéaire des gènes *hox* dans le cerveau postérieur est également régulée par l'AR (Katsuyama and Saiga, 1998; Nagatomo and Fujiwara, 2003; Schubert et al., 2006). De plus, les mêmes effets postériorisants de l'AR y ont été décrits ce qui laisse penser que la régulation de l'expression des marqueurs du cerveau postérieur par l'AR est conservée au sein des chordés (Wada et al., 2006).

1.2. Spécification neuronale et neurogénèse

Un autre rôle de l'AR a été mis en évidence dans la spécification neuronale le long de l'axe dorso-ventral (D/V) dans la moelle épinière antérieure (Maden, 2002). En effet, l'AR est capable de spécifier *in vitro* la différentiation de la plaque neurale en un certain type d'inter-neurones, exprimant des marqueurs qui leur sont propres comme certains gènes *dbx* et *evx* (Pierani et al., 1999) et cette population d'inter-neurones est absente chez des embryons de cailles déficients en AR (Wilson et al., 2004). De plus, il a été noté chez ces mutants une expansion des populations ventrales de neurones aux dépends des populations dorsales et des observations similaires ont été effectuées chez l'amphioxus (Schubert et al., 2006). Ainsi l'AR aurait un rôle conservé au sein des chordés dans

l'expression des gènes neuronaux dorsaux nécessaires au développement des inter-neurones (Figure 23). D'autre part, l'AR a aussi été mis en évidence comme étant nécessaire pour la genèse des neurones primaires chez le xénope. Ces neurones sont les premiers à se former à partir de la plaque neurale et sont cruciaux car coordonnent les premiers mouvements échappatoires de l'embryon. L'AR régule le nombre de neurones primaires qui se développent et un excès d'AR mène a une plus grande quantité de neurones primaires et à leur présence dans des régions plus antérieures (Sharpe and Goldstone, 2000). Il a été montré que l'AR active des gènes régulant positivement la neurogénèse comme *Gli3* et inhibe les régulateurs négatifs tel *Zic2* (Maden, 2007). De manière intéressante, la voie Shh semble antagoniser les effets de l'AR sur la neurogénèse des neurones primaires.

Au sein de la moelle épinière, l'AR joue également un rôle dans le développement et la spécification des motoneurones. En effet, il est requis pour leur développement et des traitements à l'AR peuvent changer la destinée de motoneurones branchiaux en motoneurones somatiques (Novitch et al., 2003). Enfin, l'AR promeut l'excroissance des neurites *in vitro* dans des cellules neuronales de poulet et de souris (Corcoran et al., 2000; Wuarin et al., 1990) ainsi que les projections axonales des motoneurones *in vivo* chez le rat et la caille (Maden et al., 1996; White et al., 2000). Des traitements à l'AR dans des cellules neuronales de mollusque (*L. stagnalis*) induisent des excroissances de neurites et prolongent leur survie *in vitro* suggérant ainsi une conservation de cette fonction de l'AR chez les bilatériens (Dmetrichuk et al., 2006).

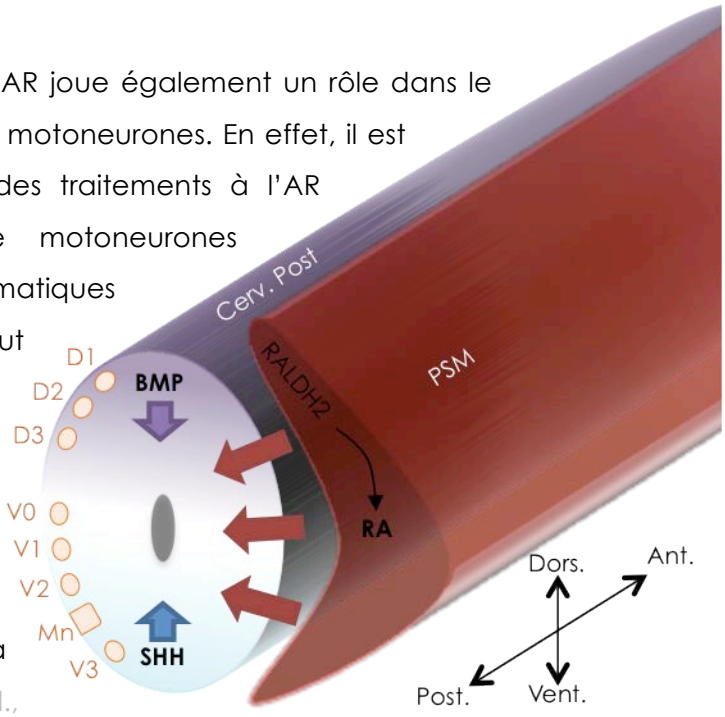


Figure 23: Spécification dorso-ventrale des neurones au sein du cerveau postérieur. Adaptée de Maden et al., 2007. Dans le PSM, RALDH2 exprimée postérieurement synthétise l'AR et CYP26C1 le dégrade plus antérieurement créant ainsi un gradient A/P d'AR qui diffuse depuis le PSM vers le cerveau postérieur adjacent. Là, la voie BMP en provenance des régions dorsales, la voie SHH depuis la région ventrale et l'AR originaire du PSM latéral spécifient les différents types de neurones (D1-3, V0-3) le long de l'axe D/V au sein du cerveau postérieur dont les motoneurones (Mn) en position ventrale.

2. MESODERME, SOMITOGENESE ET ASSYMETRIE GAUCHE/DROITE

Le mésoderme présomitique (PSM) est une des sources les plus importantes d'AR dans l'embryon de vertébré. Il exprime fortement l'enzyme RALDH2 qui synthétise l'AR nécessaire pour la formation de nombreuses structures dérivées dont les somites (Dubrulle and Pourquie, 2004; Niederreither et al., 1997).

2.1. Somitogénèse

La somitogénèse correspond à la formation d'unités mésodermiques symétriques le long de l'axe A/P qui mèneront entre autres à la formation de nombreux muscles et des os du squelette axial (Wellik, 2007). Le mécanisme de segmentation des somites est finement contrôlé par une horloge moléculaire régulant la périodicité temporelle du processus et par un front de détermination qui assure sa périodicité spatiale (Baker et al., 2006; Cooke and Zeeman, 1976; Dubrulle et al., 2001; Kerszberg and Wolpert, 2000). Ce sont principalement les voies de signalisation de Notch, Wnt et FGF qui sont impliquées dans l'horloge de segmentation en agissant en tant qu'oscillateurs transcriptionnels (Goldbeter and Pourquie, 2008) et la voie de l'AR est aussi fortement impliquée dans la dynamique du front de détermination (Dubrulle and Pourquie, 2004). En effet, ce front est déterminé à la limite entre deux gradients antagonistes au niveau du PSM : un gradient antéro-postérieur d'AR et un gradient postéro-antérieur de FGF. Ces deux gradients opposites sont maintenus par des inhibitions d'expression réciproques entre l'AR et *fgf8* et entre FGF8 et *raldh2* (Figure 24) (Diez del Corral et al., 2003). Il est généralement admis que le signal FGF en postérieur conserve les propriétés souches des cellules alors que le signal de l'AR en antérieur conduit à leur différenciation. La voie Wnt semble impliquée dans la régulation de cet équilibre entre les voies FGF et de

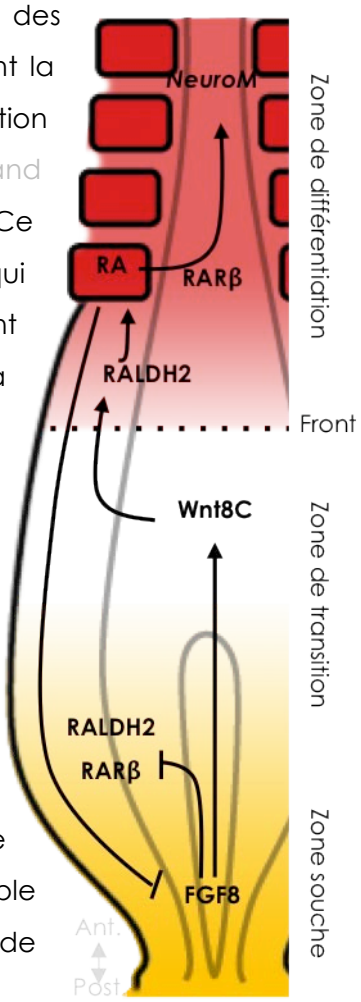


Figure 24: Réseau de signalisation impliqué dans la dynamique de somitogénèse. Adaptée de Olivera-Martinez and Storey, 2007. L'élongation de l'axe embryonnaire le long de l'axe A/P est contrôlée par un double gradient FGF-RA. En postérieur, FGF8 inhibe l'expression de *raldh2* et *rarβ* et maintient le caractère « souche » des cellules. FGF8 active Wnt8C dans la zone de transition entre les deux gradients permettant l'activation subséquente de *raldh2* dans les régions antérieures définissant le front de détermination. L'AR est ainsi synthétisé dans les somites en formation et active, via l'action de RARβ, l'expression de gènes de différentiation neuronale comme *NeuroM*.

l'AR en jouant un rôle de médiateur moléculaire entre ces deux signaux au niveau de la zone de transition du PSM (Figure 24) (Olivera-Martinez and Storey, 2007).

En plus de réguler la dynamique du front de détermination, les voies antagonistes AR/FGF contrôlent également l'expression différentielle dès gènes *hox* au sein du PSM le long de l'axe A/P. Cela va permettre la spécification de l'identité des structures issues du PSM ce qui est essentiel pour la future identité des vertèbres par exemple. Alors que FGF8 active préférentiellement les gènes *hox* postérieurs situés en 5' du cluster génique des *hox*, l'AR permet l'expression de gènes *hox* antérieurs situés en 3' du cluster. Ainsi, ces voies de signalisation assurent la coordination de la segmentation des somites ainsi que la spécification de leur identité le long de l'axe A/P (Dubrulle et al., 2001; Zákány et al., 2001). Les mécanismes de la somitogénèse semblent très bien conservés chez différentes espèces de vertébrés, cependant l'implication de l'AR dans ce processus en dehors des vertébrés reste incertaine. Il n'existe pas de PSM chez l'amphioxus et les somites se forment par bourgeonnement à partir du bourgeon caudal. Ce processus ne semble pas requérir l'AR, cependant il est important pour le positionnement des somites le long de l'axe A/P (Schubert et al., 2006). Ainsi, le rôle de l'AR dans la segmentation des somites semble être apparu chez les vertébrés parallèlement à l'acquisition du PSM.

2.2. Symétrie Gauche/Droite (G/D)

Au cours de la somitogénèse, l'AR régule également la formation synchrone et bilatérale des somites de chaque côté de la notocorde le long de l'axe G/D. En effet, l'altération de la synthèse endogène d'AR chez des mutants de souris, de poulet ou encore de poisson-zèbre résulte en la production de somites asymétriques et de manière désynchronisée (Kawakami et al., 2005; Vermot et al., 2005). Il est admis que l'AR produit antérieurement dans le PSM tamponne des signaux d'asymétrie G/D émanant des zones postérieures du PSM et qui sont requis pour le développement asymétrique de certains organes internes comme les viscères (Vermot and Pourquie, 2005).

Par ailleurs, la mise en place de cette asymétrie globale au sein de l'embryon est également dépendante de la voie de l'AR. Ce processus repose sur l'établissement d'un courant au niveau du nœud embryonnaire, appelé flux nodal, qui est généré par un mécanisme hydrodynamique. En effet, les cellules ventrales du nœud sont ciliées et la rotation synchronisée des cils permet la formation du flux de la droite vers la gauche chez la souris (Figure 25) (Hirokawa et al., 2009). Le flux nodal permet entre autres le transport de

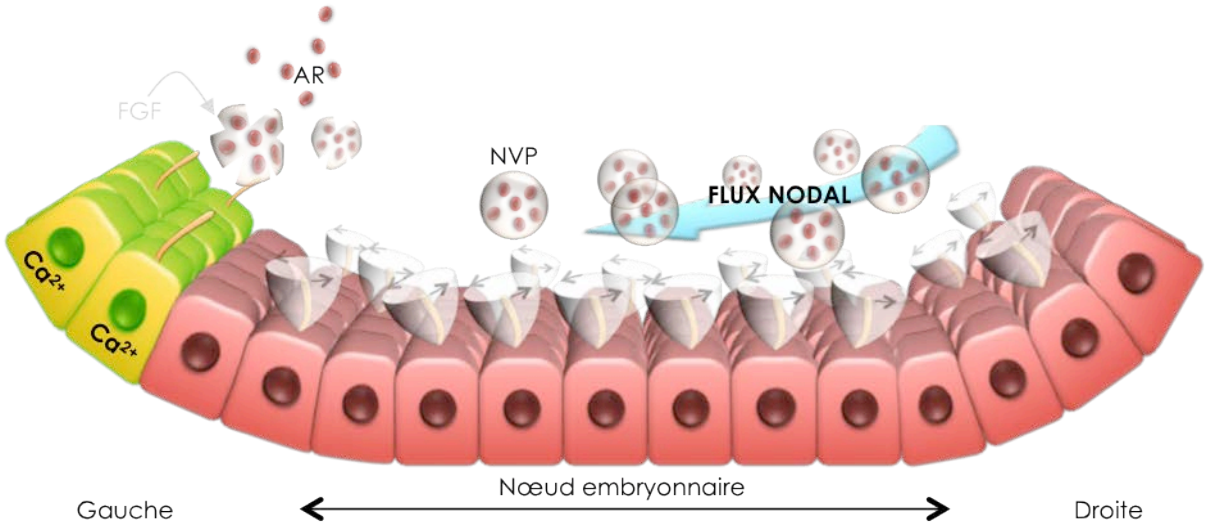


Figure 25: Etablissement du flux nodal et asymétrie gauche/droite. Adaptée de Hirokawa et al., 2009. Les cellules ventrales du nœud embryonnaire sont pourvues de cils mobiles en rotation. La coordination de la rotation des cils génère un flux de la droite vers la gauche chez la souris appelé « flux nodal ». Les NVP renfermant de l'AR sont soumises à ce flux et se concentrent dans la partie gauche du nœud. Via un mécanisme dépendant de la voie des FGF et au contact des cils non mobiles des cellules à la gauche du nœud, les NVP éclatent et libèrent l'AR localement. S'en suit l'activation unilatérale par l'AR de gènes importants dans la mise en place de l'asymétrie G/D. A noter une augmentation de l'influx de Ca2+ dans les cellules activées à gauche du nœud.

vésicules lipidiques chargées d'AR et de SHH (NVP : *Nodal Vesicular Parcels*) du côté gauche de l'embryon de souris. L'AR est ensuite libéré par un mécanisme dépendant des FGF, et va réguler l'expression unilatérale de gènes impliqués dans la mise en place de l'asymétrie G/D comme *lefty*, *pitx* ou *nodal* (Sirbu and Duester, 2006; Tanaka et al., 2005; Tsukui et al., 1999). Chez des invertébrés, l'asymétrie G/D semble être gouvernée par les mêmes hiérarchies génétiques et l'expression des gènes homologues à *shh*, *pitx* et *nodal* sont exprimés de manière asymétrique durant le développement de l'embryon d'amphioxus (Duboc et al., 2005; Schubert et al., 2005; Yu et al., 2007). Toutefois, chez ce dernier ou chez certains ascidiens, des traitements à l'AR n'altèrent pas la mise en place de l'asymétrie G/D, ainsi l'action de l'AR dans ce processus pourrait être spécifique des vertébrés (Hinman and Degnan, 1998; Schubert et al., 2005).

2.3. Autres dérivés mésodermiques

Enfin, en plus de son rôle dans la spécification des grands axes de l'embryon, l'AR est aussi impliqué dans l'organogénèse de nombreux dérivés mésodermiques dont les reins (Wingert et al., 2007), les bourgeons de membres (Duboc and Logan, 2011), les structures cutanées (Prin and Dhouailly, 2004), les muscles (Bentzinger et al., 2012) ou encore le cœur

(Collop et al., 2006). De nombreuses études sur les vertébrés dont le poisson-zèbre ont permis de décrypter la mécanistique précise des effets de l'AR ainsi que les nombreuses interactions avec d'autres voies de signalisation pour le développement de ces organes. Quelques exemples sont présentés dans la revue de la section D de l'introduction, page 65. A l'inverse, rares sont les données du rôle de l'AR dans l'organogénèse des invertébrés, ce qui ne permet donc aucune spéculation sur l'origine évolutive du rôle de l'AR sur le développement de ces structures.

3. ENDODERME, PHARYNX ET ORGANES INTERNES

3.1. Pharynx

L'AR est impliqué dans la spécification des structures endodermiques le long de l'axe A/P au sein de l'endoderme pharyngien. En effet, chez la caille et le poisson-zèbre, des mutants de la voie de synthèse de l'AR présentent une expansion de l'endoderme pharyngien au dépend des structures de l'intestin antérieur (Grandel et al., 2002; Quinlan et al., 2002). A l'inverse, un excès d'AR provoque une postériorisation de ces structures en transformant par exemple les structures antérieures de l'intestin en foie ou en pancréas (Stafford and Prince, 2002). Au niveau des arcs branchiaux, une concentration élevée d'AR est nécessaire pour le développement des arcs postérieurs 3 à 6 chez l'embryon de souris, alors que la différentiation du second arc requiert une dose plus faible d'AR (Mark et al., 2004). En effet, chez des embryons de souris ou de poisson-zèbre où la voie de l'AR est altérée, les arcs 3 à 6 sont manquants et l'expression de certains gènes *hox* et *pax* est modifiée. Ces gènes étant importants dans la différentiation des arcs branchiaux postérieurs (Wendling et al., 2000), cela suggère que l'AR agit au niveau des cellules de l'endoderme pharyngien en leur pourvoyant une identité positionnelle via l'expression de certains gènes de détermination. A noter que le sous-type beta des RAR est spécifiquement requis pour l'effet tératogénique de l'AR au niveau des arcs branchiaux (Matt et al., 2003). En conclusion, la régulation fine de la structuration des différents arcs branchiaux semble être corrélée à des zones d'activité de l'AR très localisées au sein de l'endoderme pharyngien. Chez l'amphioxus également, la structuration de l'endoderme pharyngien est dépendante de l'AR. (Escriva et al., 2002). En effet, des traitements à l'AR mènent à la perte de certaines structures pharyngiennes et postériorisent l'expression de marqueurs comme les gènes *pax* ou *otx* (Schubert et al., 2005). Les mêmes effets ont aussi été observés chez certaines

ascidies ce qui suggère que l'AR était déjà impliqué dans la spécification de l'endoderme pharyngien chez l'ancêtre commun de tous les chordés.

A noter que l'action de l'AR sur l'endoderme pharyngien est indépendante de l'action des cellules de crêtes neurales (NCC) (Veitch et al., 1999) bien que certaines NCC craniales migrent dans les arcs branchiaux et participent activement à leur morphogénèse (Santagati and Rijli, 2003). En effet, chez l'amphioxus, l'AR joue un rôle dans la spécification du pharynx bien que ce chordé soit dépourvu de NCC (Escriva et al., 2002). Toutefois, l'AR a un rôle important dans l'induction et la migration des NCC ainsi que dans la morphogénèse des arcs branchiaux. La revue de la section D de l'introduction, page 65, fournie plus de détails sur les effets de l'AR sur les NCC.

3.2. Organes internes

L'AR synthétisé dans le mésoderme via l'activité de RALDH2 est impliqué dans l'organogénèse de certaines structures dérivées de l'endoderme adjacent (Molotkov et al., 2005). Par exemple, la différentiation du foie et du pancréas nécessite des signaux de l'AR depuis le mésoderme des plaques latérales (Kumar et al., 2003; Kumar and Melton, 2003). De manière générale, il apparaît que comme dans le mésoderme et dans le CNS, l'AR agit comme agent postériorisant dans l'endoderme. En effet, des traitements antagonistes de la voie de l'AR mènent à une absence des marqueurs normalement exprimés dans le foie et le pancréas. A l'inverse, un excès d'AR mène à une expansion de ces marqueurs au niveau de l'endoderme antérieur (Stafford and Prince, 2002). De plus, l'AR promeut la différenciation de cellules progénitrices en cellules béta dans le pancréas en activant des gènes de différenciation (Ostrom et al., 2008). Cependant, les organes dérivés de l'endoderme postérieur comme l'intestin ne semblent pas requérir l'AR pour leur développement ce qui suggère qu'au sein de l'endoderme, il existe des zones répondant à l'AR et d'autres qui ne le sont pas. Ces zones de réponse à l'AR le long de l'axe A/P de l'endoderme sont définies par l'expression dynamique spatio-temporelle des enzymes de synthèse (RALDH) et de dégradation (CYP26) de l'AR (Blentic et al., 2003; Reijntjes et al., 2005). De plus, l'expression localisée des différents sous-types de RAR dans l'endoderme parachève l'établissement des zones d'activités de l'AR le long de l'axe A/P (Mollard et al., 2000). Aussi, l'expression différentielle des RAR le long de l'axe D/V de l'endoderme pourrait jouer un rôle dans la régionalisation de l'action de l'AR au sein de l'endoderme (Pan et al., 2007).

III. EVO-DEVO : COMMENT L'AR PEUT-IL ENGENDRER DES INNOVATIONS PHENOTYPIQUES ?

1. L'EVO-DEVO : L'EVOLUTION PAR LE DEVELOPPEMENT

1.1. Une tentative de définition

La biologie de l'évolution cherche à comprendre comment les organismes ont évolué, à disséquer les patrons et les processus qui ont permis la mise en place de l'incroyable biodiversité que l'on constate et à savoir comment de nouvelles formes sont apparues. Elle cherche notamment si l'origine de ces modifications réside dans les mécanismes développementaux contrôlant la taille et la forme de l'organisme et cherche à comprendre comment les contraintes développementales influent sur les mécanismes évolutifs. De son côté, la biologie du développement a pour but d'étudier comment des altérations dans l'expression et la fonction de certains gènes mènent à des changements phénotypiques. Ainsi, il existe une interconnexion évidente entre ces deux domaines de la biologie : l'évo-dévo. L'évo-dévo, qui correspond donc à la Biologie de l'EVOlution et du DEVelOppement a pour but de comprendre les mécanismes développementaux qui sont à l'origine des changements évolutifs observés et de proposer des scénarios fonctionnels de l'évolution des morphologies (Figure 26). Finalement, en comparant les mécanismes développementaux de différentes espèces, l'évo-dévo décrypte les différences à l'origine de modifications morphologiques pour établir un scénario évolutif. L'origine de cette discipline remonte à la fin du XIX^{ème} siècle avec les

Figure 26: Evo-Devo: des gènes du développement à l'évolution. Quelles modifications génétiques sont à l'origine de la diversification des phénotypes?

premières études d'embryologie. En fait elle a été récemment remis en lumière après la découverte des gènes homéotiques et de leur unicité fonctionnelle chez les métazoaires (Duboule and Dolle, 1989; Graham et al., 1989) et plus récemment par la disponibilité de plus en plus de génomes d'espèces non modèles (Carroll, 2005).

1.2. Où et comment agit l'évolution ?

i. Mutations

Les mutations génétiques sont indéniablement essentielles pour la modification des gènes du développement et donc du plan d'organisation embryonnaire. La fonction propre de chaque gène dépend à la fois de la fonction de la protéine codée, mais également des conditions sous lesquelles la protéine est exprimée. Ainsi, les mutations peuvent intervenir dans des régions régulatrices (régions *cis*) pour modifier l'expression de gènes du développement, ou bien peuvent cibler les séquences codantes des protéines et ainsi changer leur fonction.

Les mutations dans les séquences codantes des gènes sont supposées être moins fréquentes et plus facilement délétères car la fonction propre d'une protéine soumet sa séquence codante à une forte pression de sélection. Cependant, quelques exemples de mutations dans la partie codante de gènes du développement ont été mis en évidence comme étant à l'origine d'innovations morphologiques adaptatives. C'est par exemple le cas dans l'acquisition de l'albinisme chez différentes populations de poissons cavernicoles après mutation du gène *Oca2* (Protas et al., 2006). Des mutations au sein de ce gène sont également responsables de différences de pigmentation de la peau observées chez l'homme (Lamason et al., 2005). Il est proposé que des mutations dans les séquences codantes ne sont possibles que si le gène ne présente qu'une faible pléiotropie. En effet, des mutations dans des gènes pleiotropes vont avoir de multiples conséquences et seront ainsi fréquemment délétères (Carroll et al., 2004). Cependant, certaines mutations dans des facteurs pléiotropes comme le récepteur de l'ectodysplasine (EDAR) ont été décrites et semblent avoir été sélectionnées dans certaines populations d'Asie de l'est (Sadier et al., 2013).

A l'inverse, King et Wilson déclaraient en 1975 que « *le faible niveau de divergence des séquences codantes des gènes entre l'homme et le chimpanzé, ne pouvait à lui seul,*

expliquer les différences phénotypiques importantes entres ces deux espèces » (King and Wilson, 1975). Depuis, il a été généralement admis que les modifications au niveau des séquences *cis*-régulatrices jouent un rôle principal dans la genèse de nouveaux traits phénotypiques (Wray, 2007). En effet, de nombreuses validations expérimentales ont montré que des mutations dans les séquences *cis*-régulatrices des gènes peuvent induire des modifications développementales et des changements morphologiques adaptatifs. Cela a par exemple été montré pour des gènes impliqués dans la pigmentation des ailes de drosophile (Gompel et al., 2005) ou encore pour le gène *pitx1* dont les changements d'expression chez différentes espèces d'épinoches ont conduit à la perte de leur armure osseuse (Chan et al., 2010).

Mes travaux de thèse ont mis en évidence que des substitutions dans les séquences codantes des gènes peuvent permettre une régulation plus fine de l'activité des protéines via l'acquisition de résidus ciblés par des modifications post-traductionnelles comme les phosphorylations. Mes résultats suggèrent ainsi que des mutations dans des gènes exprimés dans de nombreux organes et ayant un rôle pléiotrope peuvent ne pas avoir d'effets nécessairement délétères. (Chapitre 3, page 139)

ii. Duplication de gènes

Il a été suggéré que l'apparition de nouvelles morphologies et de nouveaux phénotypes pouvait être corrélée à l'augmentation de la complexité génomique induite par des évènements de duplication de génome (WGD)(Ohno, 1970). En particulier, deux évènements de duplication de génome au cours de l'évolution ont joué un rôle majeur dans la diversification du plan d'organisation des différentes espèces de vertébrés (Dehal and Boore, 2005). Cette théorie a été également été validée expérimentalement chez les poissons téléostéens dont le génome a subit une duplication supplémentaire (FSGD) par rapport aux autres vertébrés (Jaillon et al., 2004) (Figure 19, page 27).

Selon l'hypothèse de Ohno, l'augmentation du nombre de gènes suite aux évènements de WGD permet l'acquisition de nouvelles fonctions pour les gènes dupliqués (gènes paralogues) et est ainsi à l'origine de la diversification des phénotypes. Selon le modèle dit de Duplication, Dégénération, Complémentation (DDC) (Force et al., 1999), il est généralement admis que la duplication d'un gène ancestral en plusieurs gènes paralogues peut mener à trois scénarios différents (Figure 27):

- Pseudogénisation: l'un des gènes paralogues est perdu, perd sa fonction ou n'est plus exprimé et se transforme donc en pseudogène qui dégénère et n'est ainsi plus reconnaissable.
- Sous-fonctionnalisation : les gènes paralogues acquièrent chacun une sous-fonction du gène ancestral. Ainsi, ils assurent à eux deux, ensemble, la fonction d'origine du gène avant duplication.
- Néofonctionnalisation : Un des gènes paralogues conserve la fonction du gène ancestral alors que le second acquière de nouvelles fonctions. C'est en particulier ce type d'évènement qui est le plus recherché en Evo-Devo car il va dans le sens d'une corrélation entre duplication et innovation.

A l'échelle du génome entier, un évènement de WGD va doubler la quasi-totalité des gènes du développement, leur laissant ainsi champ libre pour acquérir de nouvelles fonctions et permettre ainsi l'évolution du plan d'organisation de l'embryon. Concernant les RAR, il a été suggéré que l'acquisition de plusieurs paralogues à l'issue d'évènements de WGD a mené à la néofonctionnalisation des sous-types alpha et gamma des RAR de mammifères (section B-I-2., page 27).

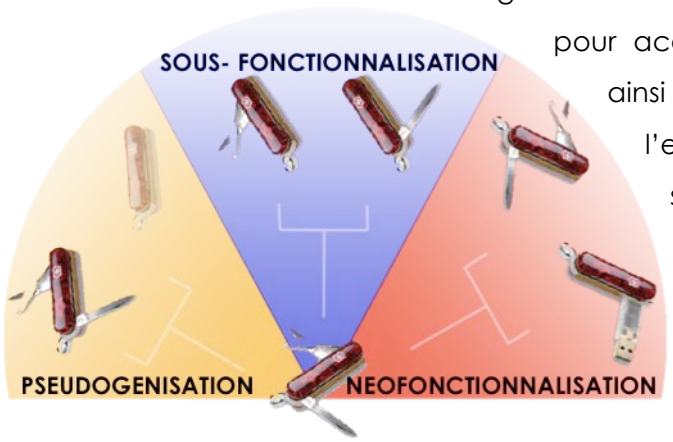


Figure 27: Différents scénarios après une duplication de gènes. Lors d'une duplication de gène, la fonction du gène ancestral peut être conservé chez un seul des paralogues menant à la perte du second: pseudogénisation. La fonction ancestrale peut être partagé entre les deux nouveaux paralogues: sous-fonctionnalisation. Un paralogue peut conserver la fonction ancestrale alors que l'autre acquiert de nouvelles fonctions: neofonctionnalisation.

1.3. Evo-Devo de la voie de l'AR

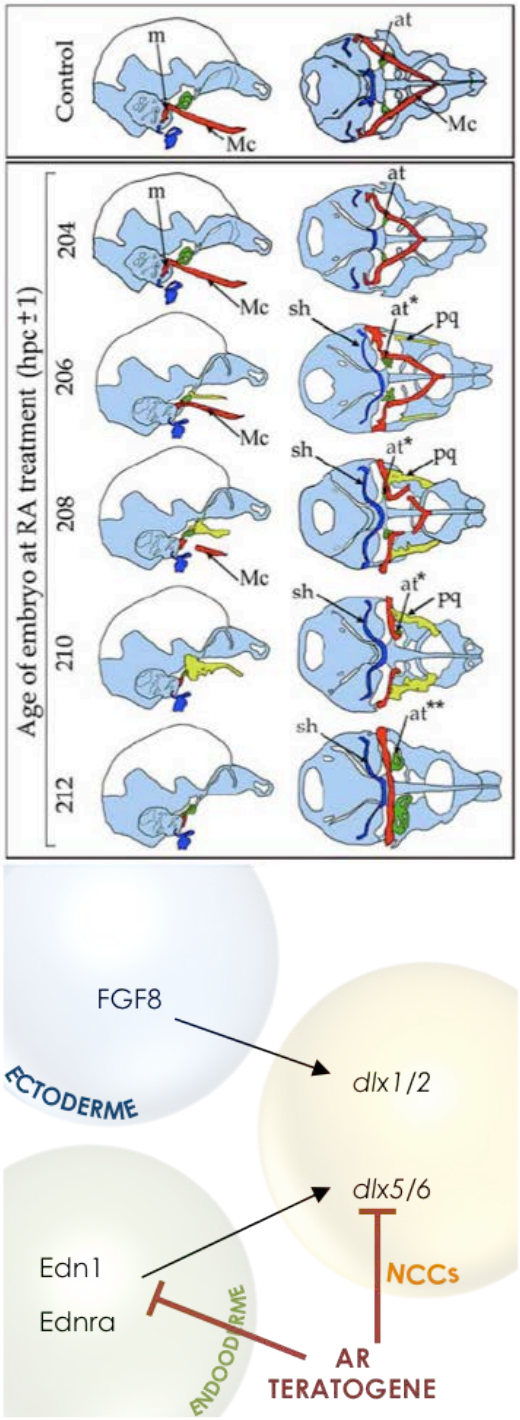
La voie de l'AR ayant de multiples rôles développementaux, il est tentant de proposer que des modifications de l'activité et/ou de l'expression de certains membres de cette voie puissent être à l'origine de changements morphologiques observés chez les vertébrés. En effet, la forte pléiotropie de la voie de l'AR offre une palette d'action importante pour l'évolution. Cependant, les effets tératogènes de l'AR laissent supposer que de fortes contraintes développementales vont s'opposer à la modification de la voie par des mutations (Carroll, 2008). Quelques études présentées ci-dessous ont mis en

évidence un rôle de la voie de l'AR dans la genèse de nouvelles morphologies. Mes travaux de thèse ont également participé à montrer que des modifications subtiles du métabolisme local de l'AR *in vivo* pouvaient être à l'origine de la diversité de microstructures comme les dents pharyngiennes de poissons (chapitre 3, page 163)

2. IMPLICATION DE LA VOIE DE L'AR DANS L'EVOLUTION DES STRUCTURES CRANIO-FACIALES

2.1. Effets tératogènes craniaux-faciaux induits par l'AR

L'AR est un puissant tératogène lorsqu'il est administré durant le développement embryonnaire des vertébrés et de certains invertébrés (Escriva et al., 2002; Soprano and Soprano, 1995). Il existe un continuum d'anomalies développementales induites par l'AR selon la concentration, la durée et le stade développemental de l'embryon exposé et l'ensemble des ces malformations est regroupé sous le terme d'embryopathie dépendante de l'AR (RAE). Une des malformations caractéristiques induites par l'AR, est un défaut dans la morphogénèse des structures craniofaciales dont la mâchoire et l'oreille interne (Figure 28) (Lammer et al., 1985; Mallo, 1997). Chez l'embryon de souris, les défauts craniofaciaux sont caractérisés par une hypoplasie et une fusion des deux premiers arcs pharyngiens qui sont à l'origine de nombreuses structures craniofaciales affectées par l'effet tératogène de l'AR. Les arcs pharyngiens sont rapidement colonisés par les NCC craniales qui vont participer à leur morphogénèse ainsi qu'au développement des structures dérivées des arcs pharyngiens (Santagati and Rijli, 2003). Il a été montré que les effets tératogènes de l'AR n'impliquent pas un effet de l'AR sur la survie ou la migration des NCC chez la souris. Plutôt, l'AR perturbe les voies de signalisation en provenance de l'épithélium et de l'endoderme pharyngien et qui déterminent la destinée des NCC via l'expression des gènes *dlx* (Figure 28). Les gènes *dlx* codent pour des facteurs de transcriptions à homéodomaines qui, à l'image des gènes *hox*, sont essentiels dans la spécification de l'identité des structures craniofaciales (Acampora et al., 1999). Ainsi, il est admis que les gènes *dlx* ont un rôle clé dans le développement et l'évolution de la mâchoire chez les vertébrés (Depew et al., 2005; Takechi et al., 2013). De manière intéressante, les mêmes anomalies que celles induites par les effets tératogènes de l'AR sur la mâchoire sont retrouvées chez des mutants de souris pour les gènes *dlx5* et *dlx6* (Beverdam et al., 2002). En cohérence avec ces observations, il a été montré que l'AR perturbe l'activation des gènes *dlx* dans les NCC craniales



normalement régulée par un mécanisme dépendant de l'endothelin-1 (Edn) et de son récepteur (Ednra) dans l'endoderme pharyngien (Figure 28) (Charite et al., 2001; Ruest et al., 2004). Il est intéressant de noter que les enzymes CYP26A1 et CYP26C1 sont exprimées dans le premier arc pharyngien lors sa colonisation par les NCC (Sakai et al., 2001; Tahayato et al., 2003) et que cela protégerait ces zones des effets tératogènes de l'AR.

2.2. Le rôle de l'AR dans l'évolution des structures cranio-faciales.

De manière surprenante, certaines morphologies craniofaciales observées après exposition à l'AR ou mutation des gènes de la voie, ressemblent à des structures retrouvées chez d'autres espèces et en particulier à des structures primitives de mâchoires au sein des gnathostomes (atavisme). Ainsi, les voies de signalisation nécessaires pour la mise en place de la mâchoire des mammifères pourraient avoir été sujettes à des modulations au cours de l'évolution des vertébrés. En particulier, il est envisageable que l'acquisition d'une protection contre la voie de l'AR par l'expression des enzymes CYP26 au niveau de la région du premier arc pharyngien, ait contribué à

Figure 28: Effets tératogènes de l'AR sur les structures craniofaciales. Figure de Vieux-Rochas et al., 2010 schématisant les anormalités craniofaciales chez des embryons de souris à 14 dpc après exposition à l'AR à partir de différents temps (indiqués à gauche).Le cartilage de Meckel est indiqué en rouge (Mc), le palatoquadrate en jaune (pq), l'aile temporale de l'os sphénoïde en vert (at) et le cartilage stylo-hyoïdien en bleu foncé (sh). Le premier défaut induit est une déformation proximale du Mc et une perte rapide du marteau de l'oreille interne (m). Les duplications de structures sont indiquées par un astérisque.
Interaction moléculaire entre l'ectoderme, l'endoderme et les NCC craniales: l'expression correcte des gènes *dlx* dans les NCC est régulée par FGF8 depuis l'ectoderme et par l'action de l'endotheline-1 (Edn1) sur son récepteur (Ednra) depuis l'endoderme pharyngien. Les effets tératogènes de l'AR inhibent l'expression des gènes *dlx5/6* induite par Edn1/Ednra.

l'élaboration de la mâchoire des gnathostomes modernes (Vieux-Rochas et al., 2007). Une étude récente a montré que chez le xénope, il existe une période très courte de 6 heures pendant laquelle une exposition à l'AR induit des malformations craniofaciales sans affecter le reste de l'embryon (Vieux-Rochas et al., 2010). Ces fenêtres d'action sont des périodes clés durant lesquelles une modulation de la voie de l'AR pourrait permettre l'acquisition de nouvelles caractéristiques au niveau de la mâchoire, tout en contournant les fortes contraintes développementales liées à l'action de l'AR sur le développement. De plus, une exposition de seulement une minute à l'AR suffit à induire des malformations craniofaciales irréversibles, suggérant que des modifications très fines dans les voies de régulation sont suffisantes pour l'apparition de structures modifiées (Vieux-Rochas et al., 2010). Ces effets très subtils pourraient être à l'origine de légères différences morphologiques craniofaciales intra-spécifiques, mais aussi plus largement à la base de la diversité interspécifique acquise au cours de l'évolution.

Enfin cette discussion du rôle éventuel de l'AR dans l'évolution des caractéristiques de la mâchoire est renforcée par des données plus anciennes issues de mutants de souris dans la voie de l'AR. En effet, la double invalidation (KO) de RARγ et RARα chez la souris induit la formation de structures surnuméraires au niveau du crâne des embryons (Lohnes et al., 1994). De manière intéressante, ces éléments correspondent à des structures atavistiques retrouvées chez des fossiles de reptiles mammaliens (thérapsides). Ainsi, bien que l'interprétation de ces mutations atavistes soit très délicate, cela suggère qu'au cours de l'évolution, et plus précisément lors de la transition reptile-mammifère (divergence synapside/diapside), des mécanismes dépendants de la voie de l'AR pourraient être à l'origine de certaines modifications des structures crâniennes de reptiles ayant généré des caractéristiques propres aux crânes de mammifères,

3. VARIATION INTERSPECIFIQUE DE LA MORPHOLOGIE DE L'INTESTIN DE XENOPE INDUIT PAR L'AR

La morphologie et la physiologie de l'appareil digestif des animaux sont souvent corrélées à la niche écologique occupée par l'animal et la diversification de ces organes au cours de l'évolution va de pair avec l'alimentation disponible (Grant and Grant, 2011). En particulier, le régime alimentaire est influencé par la configuration et la longueur du tube digestif (Stevens and Hume, 2004). Chez les larves d'anoures (amphibiens sans queue), on

retrouve des changements morphologiques au niveau de l'appareil digestif en fonction du régime alimentaire larvaire. La plupart des larves d'anoures sont herbivores comme chez le genre *Chacophrys* mais on retrouve quelques espèces à larves carnivores comme *Lepidobatrachus laevi* s'alimentant de petites proies ou encore de grosses proies entières comme *Ceratophrys cranwellii*. Chez les carnivores, l'estomac est souvent d'une plus grande contenance et possède une activité protéolytique importante et un intestin relativement court alors que chez les herbivores, l'intestin est beaucoup plus long et l'estomac ne détient pas d'activité enzymatique complexe. Ainsi, ces espèces représentent un modèle d'étude intéressant pour comprendre les mécanismes menant à la formation d'appareils digestifs aux morphologies différentes, adaptés à l'alimentation carnivore des larves.

Récemment, un crible chimique par traitement pharmacologique a été réalisé sur des larves de xénope (*Xenopus*) pour rechercher des molécules capables de modifier la morphologie primitive de l'intestin de xénope en une morphologie plus complexe, proche de celle retrouvée chez les larves d'anoures carnivores (Bloom et al., 2013). Parmi celles-ci, le DEAB (4-Diethylaminobenzaldehyde), un inhibiteur des enzymes RALDH qui synthétisent l'AR et la molécule Ro-41-5253, un antagoniste soluble de RARα (Keidel et al., 1994). En effet, le traitement de larves de xénope avec l'une de ces molécules qui altèrent la voie de l'AR induit un changement de morphologie de l'intestin de manière dose-dépendante, rappelant celle observée chez des larves d'anoures carnivores (Figure 29) (Bloom et al.,

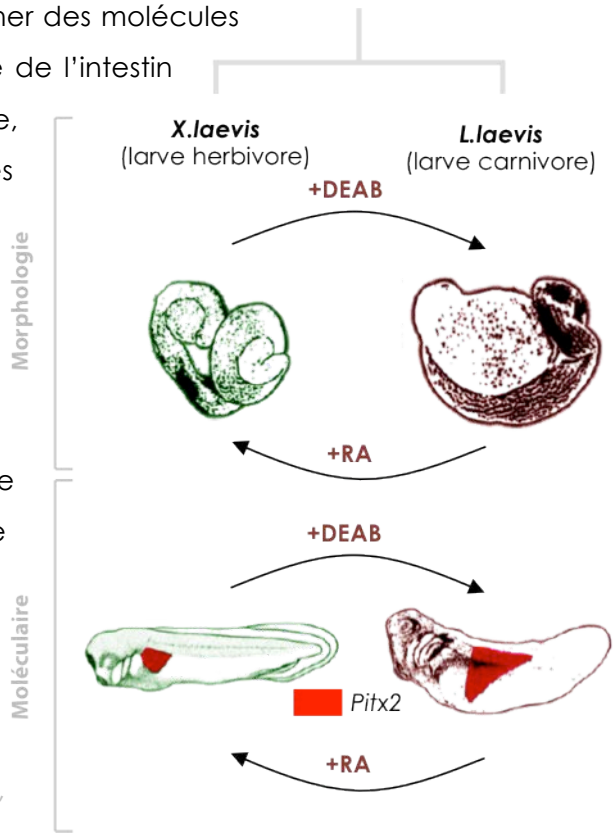


Figure 29–Variation interspécifique de la morphologie de l'intestin de larves d'anoures induite par l'AR. Figure adaptée de Bloom et al., 2013 montrant les différences morphologiques de l'intestin et de l'expression du marqueur Pitx2 chez la larve entre deux espèces d'anoures. La larve herbivore de *Xenopus laevis* possède un intestin petit et le domaine d'expression de *Pitx2* est restreint. A l'inverse, l'intestin de la larve carnivore de *Lepidobatrachus laevis* est plus gros et l'expression de *Pitx2* est plus étendue antérieurement. Un traitement à l'AR de larve carnivore récapitule une morphologie d'intestin plus simple et réduit le domaine d'expression de *Pitx2*. A l'inverse, une inhibition de la voie de l'AR par traitement au DEAB chez des larves herbivores induit une expansion du territoire d'expression de *Pitx2*, associée au développement d'un intestin plus gros et complexe ressemblant à celui de larves carnivores.

2013). Des changements similaires sont observés après traitement de larves de *Chacophrys*, une espèce herbivore à la morphologie intestinale simple. A l'inverse, une exposition à l'AR de larves d'anoures carnivores (*Lepidobatrachus*) semble restaurer une morphologie intestinale plus simple ressemblant aux structures primitives des espèces herbivores. Au niveau moléculaire, l'expression du marqueur *pitx2*, un facteur de transcription impliqué dans la mise en place de l'asymétrie G/D des viscères, est différente entre les espèces à larves herbivores et carnivores. En effet, l'expression de *pitx2* est plus étendue chez les larves carnivores ce qui pourrait expliquer la formation de structures digestives de plus grande taille en déterminant un plus grand territoire présomptif. De manière intéressante, une altération de la voie de l'AR chez des larves herbivores étend le territoire d'expression de *pitx2* comme observé chez les larves carnivores et vice-versa (Bloom et al., 2013) (Figure 29). Ainsi, les auteurs concluent qu'une modification de l'expression de gènes impliqués dans l'asymétrie G/D et régulés par l'AR (comme *pitx2*) pourrait avoir participé à l'émergence de nouvelles morphologies intestinales interspécifiques chez les anoures. Celles-ci auraient été sélectionnées positivement au cours de l'évolution car étant plus adaptées à un certain régime alimentaire. Reste à déterminer comment une telle modulation de la voie de l'AR a pu être mise en place sans altérer le développement global d'autres structures embryonnaires dépendantes de l'AR.

4. LA PERTE DES DENTS ORALES CHEZ LES CYPRINIFORMES

4.1. Les dents de poisson comme modèle d'étude en évo-dévo.

Les dents correspondent à des structures spécifiques des vertébrés et leur mode de développement semble conservé au sein du groupe (Jernvall and Salazar-Ciudad, 2007; Smith et al., 2009). De part leur grande diversité de nombre, de forme et de localisation, l'étude des mécanismes menant au développement d'un certain nombre de dents à différents endroits dans la cavité buccale et/ou le pharynx est d'un intérêt tout particulier pour des études évo-dévo. Plus de la moitié des espèces des vertébrés sont des acténoptérygiens c'est à dire des poissons à nageoires rayonnées, le groupe de vertébré le plus diversifié avec plus de 27 000 espèces recensées (Nelson, 2006). Ils présentent une incroyable diversité morphologique de forme et de couleur et leur denture est également très diversifiée. En effet, selon les espèces on peut retrouver des dents sur les mâchoires inférieures et/ou supérieures, sur le palais, dans différentes régions de la bouche et tout le long du pharynx (Pasco-Viel et al., 2010; Stock, 2007) (Figure 30). A titre d'exemple, entre les

trois principaux modèles de poisson étudiés pour le développement des dents, à savoir le poisson zèbre (*Danio rerio*), le medaka (*Oryzias latipes*) et le tétra (*Astyanax mexicanus*), on retrouve des différences dans la forme et la localisation des dents. Toutefois, une caractéristique de la dentition du poisson zèbre est qu'il ne développe des dents que sur le dernier arc pharyngien et cette réduction de la denture est un caractère commun à tous les cypriniformes (Stock et al., 2006) (Figure 30). Il est intéressant de constater qu'une telle réduction de la denture n'a jamais été inversée au cours de l'évolution de ce groupe malgré une diversification très importante des différentes espèces et de nombreux régimes alimentaires différents. A noter que des structures osseuses ressemblant à des dents ont été décrites au niveau de la mâchoire chez un cypriniforme miniature *Danionella dracula* suggérant que de forte contraintes développementales s'opposent à la restauration de la dentition orale chez les cypriniformes (Britz et al., 2009).

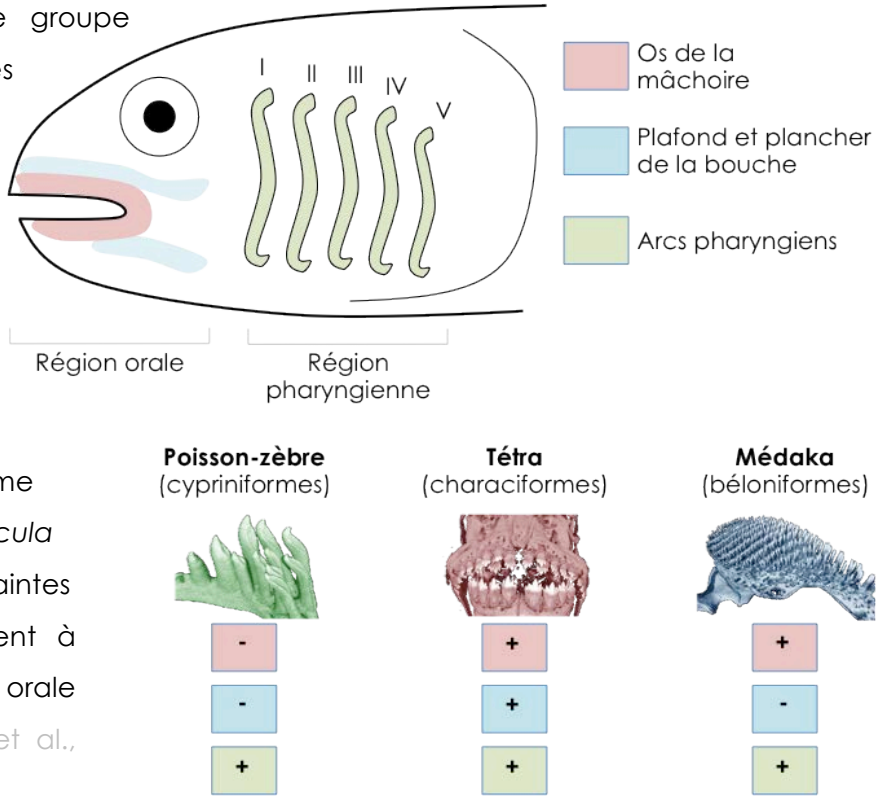


Figure 30: Localisation des dents orales et pharyngiennes chez les poissons. Vue latérale schématisée illustrant les régions dentées chez les poissons (orales et pharyngienne). Pour chacune des trois espèce modèles (poisson-zèbre, tétra, médaka), la présence ou l'absence de dents dans les différentes régions est indiquée par un + ou un -. Une image de la dentition par microtomographie est présentée pour chaque espèce et illustre la diversité de forme et de localisation des dents. A noter que le poisson-zèbre ne possède des dents pharyngiennes que sur la partie ventrale du 5ème arc pharyngien (V) alors que le tétra et le médaka ont des dents sur d'autres arcs antérieurs ainsi qu'en position dorsale sur l'arc. *(images de microCT: Emmanuel Pasco-Viel)*.

4.2. Scénario impliquant la voie de l'AR pour la perte des dents orales chez les cypriniformes.

Quels sont les mécanismes développementaux à l'origine de la réduction de la denture chez les cypriniformes ? Comment une telle réduction n'a-t-elle jamais été inversée ? Une étude comparative du développement des dents entre le poisson-zèbre et d'autres espèces possédant des dents dans la cavité buccale (médaka, tétra) permettrait ainsi de comprendre le scénario évolutif de la réduction de la denture des cypriniformes. Une étude du laboratoire de Vincent Laudet a montré que la voie de l'AR était nécessaire pour le développement des dents pharyngiennes chez le poisson-zèbre (Gibert et al., 2010). En effet, l'inhibition de la synthèse endogène d'AR ou le traitement des embryons à des antagonistes des RAR empêche l'expression des marqueurs dentaires comme *dlx2b* ou *pitx2a* et inhibe ainsi le développement de la première paire de dents. Au niveau du pharynx, l'enzyme RALDH2 est exprimée dans le pharynx postérieur ventral entourant le cinquième arc branchial. Ainsi, il est proposé que RALDH2 synthétise localement l'AR qui va servir de signal pour l'induction des bourgeons dentaires chez le poisson-zèbre via l'expression de marqueurs spécifiques comme *pitx2a* ou *dlx2a*. Cependant, de manière surprenante, l'induction des dents orales et pharyngiennes chez le médaka et le tétra n'est pas dépendante de la voie de l'AR (Gibert et al., 2010). Ces données sont soutenues par l'observation que chez le tétra, l'enzyme RALDH2 n'est pas exprimée dans la région du pharynx où sont induites les dents (Figure 31).

Ainsi, l'ensemble de ces résultats suggère que chez les cypriniformes, le développement des dents a acquis une dépendance à la voie de l'AR, qui peut être corrélée au décalage de l'expression de RALDH2 au niveau du pharynx postérieur chez ces espèces. Etant donné que le développement des dents pharyngiennes et orales repose sur des mécanismes moléculaires similaires, il est vraisemblable que la dépendance à l'AR ait également touché le développement des dents orales. Or, les enzymes RADLH2 et RALDH3 n'étant pas exprimées dans la région orale chez le poisson zèbre (White et al., 2007) l'AR n'y est donc pas synthétisé empêchant ainsi le développement de dents orales (Figure 31).

Il est peu probable que la perte des dents orales chez les cypriniformes soit due à des mutations touchant la fonction et la régulation des gènes impliqués dans le développement des dents comme les gènes *dlx*. En effet, de telles mutations auraient certainement des effets importants sur le développement d'autres structures, comme les arcs pharyngiens, et mèneraient à des défauts délétères. Ainsi, ce scénario basé sur l'acquisition d'une

dépendance à une voie de signalisation est plus parcimonieux et contourne les contraintes développementales. Il est probable que la dépendance à l'AR s'est traduite par l'acquisition ou la modification d'éléments *cis*-régulateurs dans le promoteur de gènes contrôlant l'induction et le développement des dents comme *pitx2a* (Figure 31).

Toutefois, il est probable que d'autres voies de signalisation régulant le développement des dents, en particulier la voie FGF (Jackman et al., 2004) soit impliquées dans la perte des dents orales chez les cypriniformes. En effet, l'inhibition de la voie des FGF chez le tétra (qui possède des dents orales) abolit l'expression de marqueurs de dents orales comme *dlx2a* ou *dlx2b*, mais n'a pas d'effet sur l'expression d'autres marqueurs de dents pharyngiennes comme *pitx2* (Stock et al., 2006). Ainsi, la voie des FGF est aussi putativement impliquée dans la perte de ce caractère et le fait que plusieurs voies de signalisation puissent avoir été touchées pourrait expliquer la non-réversibilité de cette perte.

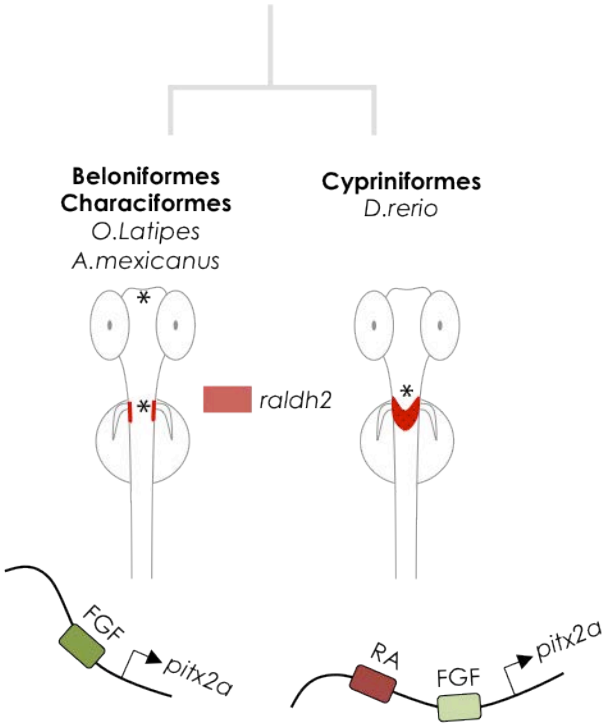


Figure 31: Scénario évolutif de la perte des dents chez les cypriniformes. L'expression du marqueur dentaire *pitx2a* est régulée par la voie des FGF. Ainsi, chez les beloniformes et characiformes, l'induction et la formation des dents orales et pharyngiennes (indiquées par une astérisque) est régulée par la voie des FGF. D'autre part, spécifiquement chez les cypriniformes, la régulation de *pitx2a* est aussi dépendante de la voie de l'AR. Le champ d'expression de l'enzyme RALDH2 (en rouge) est plus important dans le pharynx postérieur des cypriniformes et permet la synthèse locale d'AR nécessaire pour l'induction et le développement des dents. Cependant, l'AR n'est pas synthétisé dans la zone d'induction des dents orales, d'où l'absence de ce type de dentition chez les cypriniformes. C'est donc l'acquisition d'une dépendance à l'AR qui serait à l'origine de la perte des dents orales chez les cypriniformes.

Durant mon travail de thèse, je me suis intéressé au rôle de l'AR dans la diversité de denture des poissons téléostéens ainsi que dans la diversité de forme et du nombre de dents chez les cypriniformes. Mes travaux ont ainsi permis d'établir un rôle clé de l'AR dans l'élaboration d'une telle diversité de denture chez les poissons (chapitre 3, page 163).

C. LES RAR COORDONNENT LE SIGNAL DE L'AR AU COURS DU DEVELOPPEMENT

Cette partie introductive correspond à une revue publiée en 2011 dans un numéro spécial du journal *Molecular and Cellular Endocrinology* intitulé « *Nuclear Receptor Structure : Dynamics and Function* ». Elle a pour but d'exposer le rôle des RAR dans la transduction du signal de l'AR au cours du développement dans un contexte *in vivo*. Plutôt que de cataloguer les nombreux effets de l'AR décrits jusqu'alors, principalement chez la souris, nous avons voulu corréler les aspects moléculaires de la structure des RAR à leur fonction *in vivo* au cours du développement. Cette revue illustre le rôle central des RAR comme coordinateurs de la réponse de l'AR, mais nécessitant l'intégrité de nombreux partenaires et d'autres voies de signalisation pour transduire harmonieusement le signal de l'AR au cours du développement.

Molecular and Cellular Endocrinology 348 (2012) 348–360

Contents lists available at SciVerse ScienceDirect

Molecular and Cellular Endocrinology

journal homepage: www.elsevier.com/locate/mce

Review

Nuclear retinoic acid receptors: Conductors of the retinoic acid symphony during development

Eric Samarut, Cécile Rochette-Egly *

IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire), INSERM, U596; CNRS, UMR7104; Université de Strasbourg, 1 rue Laurent Fries, BP 10142, 67404 Illkirch Cedex, France

ARTICLE INFO

Article history:
Received 7 January 2011
Received in revised form 21 March 2011
Accepted 31 March 2011
Available online 8 April 2011

Keywords:
RA
RAR
Phosphorylation
Transcription
Development
Differentiation

ABSTRACT

The vitamin A derivative, retinoic acid (RA), is essential for embryonic development through the activation of cognate nuclear receptors, RARs, which work as ligand dependent regulators of transcription. *In vitro* studies revealed how RARs control gene expression at the molecular level and now it appears that it is fine-tuned by a phosphorylation code. In addition, several genetic approaches provided valuable insights on the functions of RARs during development and on the influence of other actors such as the enzymes involved in RA synthesis and degradation and other signaling pathways. It appears that RARs are the conductors of the RA signaling symphony through controlling the dynamics and the coordination of the different players and development steps.

Contents

1. Introduction

Vitamin A (retinol) is essential for embryonic development, organogenesis and homeostasis of most body tissues as well as for vision, immune functions and reproduction. Such an important

* Corresponding author. Tel.: +33 3 88 65 34 59; fax: +33 3 88 65 32 01.
E-mail addresses: esamarut@igbmc.fr (E. Samarut), cegly@igbmc.fr (C. Rochette-Egly).

0303-7207/$ - see front matter © 2011 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.mce.2011.03.025

role has been revealed by both experimental approaches and clinical observations. Indeed deficiency of vitamin A leads to neonatal growth retardation and a large array of congenital malformations affecting the ocular, cardiac, respiratory and urogenital systems collectively referred to as the fetal vitamin A deficiency (VAD) syndrome (Clagett-Dame and DeLuca, 2002; Zile, 2001).

The only source of vitamin A in the organism is diet-derived and retinol has to be converted to active compounds such as retinoic acid (RA). Then in cells and tissues, RA levels must be controlled through a balance between synthesis and catabolism (Fig. 1A). Now it is admitted that RA supports the effects of vitamin A. This is based on the ability of RA to replace vitamin A during embryogenesis and to reverse some developmental defects in VAD animals.

Newly synthesized RA binds to cellular RA binding proteins (CRABP-I and CRABP-II) and enters the nucleus (Budhu and Noy, 2002; Delva et al., 1999) where it controls gene expression through the activation of specific nuclear receptors, RARs, which consist of three subtypes, α (NR1B1), β (NR1B2) and γ (NR1B3) encoded by separate genes (for review see Chambon, 1996; Germain et al., 2006a,c). In mammals, for each subtype, there are several isoforms, which are generated by differential promoter usage and alternative splicing and differ only in their N-terminal regions. RARs function as ligand-dependent transcriptional regulators, heterodimerized with retinoid X receptors (RXRs), which also consist of three types, α (NR2B1), β (NR2B2) and γ (NR2B3) (Germain et al., 2006b). As such, they regulate the expression of subsets of target genes involved in cellular differentiation, proliferation and apoptosis, and mediate the pleiotropic effects of RA.

In this review we will recapitulate how RARs conduct RA signaling during development, focusing on molecular aspects and on three main questions: (i) when and where are RARs needed during development (ii) when and where is RA available (iii) do RARs integrate other developmental networks and then when and where.

2. Nuclear retinoic acid receptors are the functional effectors of RA signaling

2.1. RARs structure and canonical RAR-mediated regulation of transcription

The basics of RARs' structure and function have been recapitulated in several recent reviews (Bastien and Rochette-Egly, 2004; Chambon, 1996; Laudet and Gronemeyer, 2001; Lefebvre et al., 2005; Rochette-Egly and Germain, 2009). Briefly, RARs exhibit a well-defined domain organization and structure consisting mainly of a central DNA-binding domain (DBD) and a C-terminal ligand-binding domain (LBD) (Fig. 1B). Crystallographic and nuclear magnetic resonance analysis of the structures of these domains and the characterization of the associated multiprotein complexes provided a wealth of information on how RARs regulate transcription. However, recent studies highlighted the importance of the N-terminal domain (NTD), which, in contrast to the DBD and the LBD, is naturally disordered. Nevertheless, it contains a conserved proline-rich motif, which forms helices that can bind proteins with SH3 or WW domains (Lalevee et al., 2010). It has been suggested that binding of such coregulatory proteins might have profound impacts on the structural properties of the NTD and/or on the dynamics of the adjacent domains.

The LBD is formed by 12 conserved alpha helices and a beta-turn, which are folded into a three-layered, and parallel helical sandwich (Fig. 1B) (Renaud and Moras, 2000; Renaud et al., 1995). It is functionally complex as it contains the ligand-binding pocket (LBP), the main dimerization domain and a hydrophobic cleft involved in coregulators binding. Analysis of the crystal structures of the unliganded and ligand-bound LBDs highlighted that RA binding induces conformational changes, the most striking one being the repositionment of helix 12 (le Maire et al., 2010; Renaud et al., 1995).

The core of the DBD, which confers sequence specific DNA recognition, is composed of two zinc-nucleated modules and two α-helices, which cross at right angles and fold into a single globular domain (Zechel et al., 1994a,b). The DBD also includes a dimerization surface that allows RARs bind to DNA as asymmetric, oriented heterodimers with RXRs. Note that the DBD of RXR depicts a third α-helix that facilitates DNA binding and homodimerization (Lee et al., 1993).

RXR/RAR heterodimers bind to specific DNA sequences or RA response elements (RAREs), composed typically of two direct repeats of a core hexameric motif, (A/G)G [G/T] TCA and located in the regulatory sequences of target genes (Leid et al., 1992; Mangelsdorf and Evans, 1995). The classical RARE is a 5 bp-spaced direct repeat (referred to as DR5). However, the heterodimers also bind to direct repeats separated by 1 bp (DR1) or 2 bp (DR2). RAREs have been identified in the promoters of a large number of RA-target genes implicated in a wide variety of functions including RA metabolism and cellular differentiation (Balmer and Blomhoff, 2005; Bour et al., 2006). For instance, the classical DR5 elements are found in the promoters of the *rarβ2* gene itself, the *cyp26A1* gene, the *foxA1* gene and several *Homeobox* (*hox*) genes. Recently the development of high throughput technologies such as chromatin immunoprecipitation (ChIP) coupled with array hybridization or flow cell sequencing expanded the repertoire of potential high affinity response elements (Delacroix et al., 2010; Hua et al., 2009). However it also revealed degenerative and non-identifiable DRs, as well as DRs overlapping with binding sites for other transcription factors, increasing the complexity of RAR-mediated gene transcription.

The transcriptional regulation of RA-target genes relies on RARs binding to DNA and on ligand-induced conformational changes in the LBD that direct the association/dissociation of several coregulator complexes (Fig. 1C). According to the canonical model (Dilworth and Chambon, 2001), in the absence of ligand and in a context of chromatin where the nucleosomes do not impede binding to RAREs, RAREs are occupied by RAR/RXR heterodimers, associated with large protein complexes endowed with enzymatic activities, which maintain chromatin in a condensed and repressed state. According to recent studies, repression in the absence of RA may also involve Topoisomerase II (McNamara et al., 2008) and polycomb group proteins (Gillespie and Gudas, 2007a,b). Upon ligand binding, RARs undergo conformational changes that provoke corepressors release and initiate an ordered and coordinated recruitment of coactivators within large complexes with different enzymatic activities (histone acetyl and methyl transferases and DNA-dependent ATPases) (Perissi et al., 2010; Perissi and Rosenfeld, 2005), which will alter the chromatin structure surrounding the promoter of the target genes and pave the way for the recruitment of the transcription machinery including the multisubunit mediator complex, RNA polymerase II and the general transcription factors.

However, recent studies highlighted that, depending on the target gene promoter, RARs can employ different programs for gene activation. Indeed, the mediator and RNA PolII can already occupy with RARα the promoters of some endogenous genes even in the absence of RA (Flajollet et al., 2006; Pavri et al., 2005). In that context, gene-specific factors like PARP-1 are determinant for switching the Mediator to its active form. According to other studies, only a small fraction of RAREs are occupied by RXR/RAR heterodimers in the absence of RA and RARs are recruited to RAREs upon ligand binding (Biddie et al., 2010; Bruck et al.,

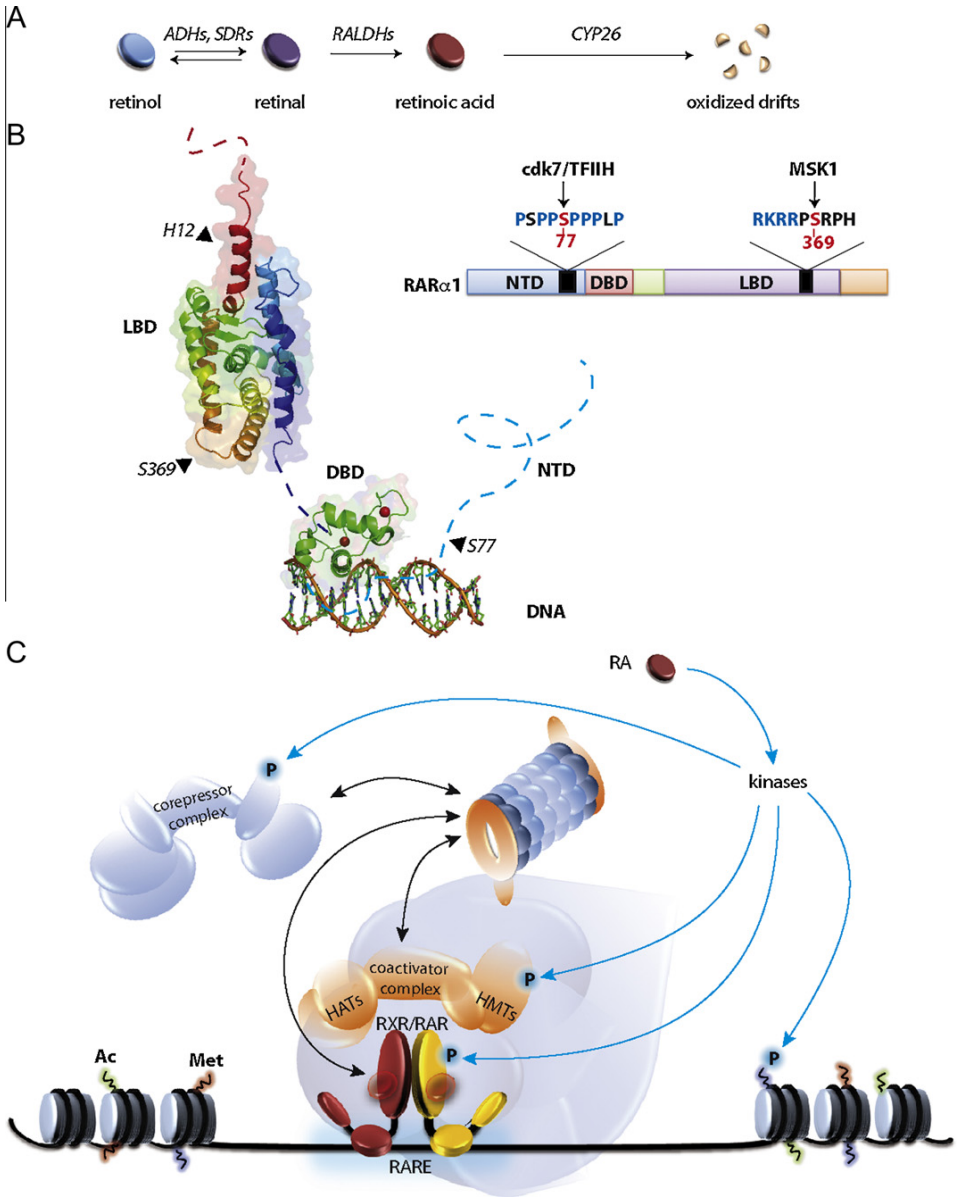


Fig. 1. Retinoic acid synthesis and signaling through RARs. (A) Vitamin A (retinol) is reversibly oxidized into retinal by alcohol dehydrogenase and short-chain dehydrogenases/reductases enzymes (ADHs and SDRs). Retinaldehyde dehydrogenases (RALDHs) irreversibly oxidize retinal into retinoic acid, which is then catabolized into oxidized drifts through the CYP26 subfamily of enzymes. (B) RARs depict a modular organization with an unstructured N-terminal domain (NTD), and two well structured domains: a central DNA binding domain (DBD) and a C-terminal ligand-binding domain (LBD). The phosphorylation sites located in the NTD and the LBD of RARα are also shown. (C) In response to RA, RAR/RXR heterodimers are recruited to response elements (RARE) located in the promoters of target genes. The conformational changes induced by RA binding provoke the dissociation of corepressor complexes and the binding of coactivator complexes that modify and remodel chromatin in order to pave the way for the recruitment of the transcriptional machinery. In parallel, RA activates kinase pathways that target (1) histones (2) corepressors and coactivators and (3) RARs themselves. Histones phosphorylation participates in chromatin decompaction. Phosphorylation of coregulators controls the dynamics of their interactions with RARs. It also controls their ubiquitination and degradation by the proteasome and thereby the dynamic and coordinated recruitment of the different transcriptional actors. RARs phosphorylation controls the association/dissociation of coregulators and DNA binding. It also signals the end of the RA signal through their degradation by the ubiquitin proteasome system.

2009; Martens et al., 2010). Finally, recent chromatin conformation capture technologies revealed that RARs bound at different enhancer elements of a gene can form loops (Bruck et al., 2009). Now one emerging view would be that RARs, as other transcription factors, would create long range chromatin loops, bridging genomic loci located even on different chromosomes (Nunez et al., 2009; Schoenfelder et al., 2010), thus creating hot-spots of transcription.

2.2. RA synthesis and degradation

As mentioned above, the transcriptional activity of RARs relies on the binding of the ligand, i.e. retinoic acid either all-trans or 9-cis RA. However, the natural *in vivo* form is essentially all-trans-RA and no significant levels of 9-cis RA can be detected (Wolf, 2006). As RXRs, the heterodimeric partners of RARs, were shown to bind 9-cis RA *in vitro*, this makes controversial the

in vivo relevance of retinoids binding to RXRs (Calleja et al., 2006; Mic et al., 2003).

Retinoic acid is a derivative of vitamin A that cannot be synthesized *de novo*. Therefore RA supply in the organism is exclusively diet-derived and requires several conversions by key enzymes that represent an important part of the large RA signaling family components (Duester, 2008; Maden, 2007; Theodosiou et al., 2010). Briefly, RA is synthesized through retinol oxidation to retinal by alcohol dehydrogenase and short-chain dehydrogenases/reductases enzymes (ADHs and SDRs) (Fig. 1A). Retinal is then irreversibly oxidized into RA through the action of retinaldehyde dehydrogenases (RALDHs). Three RALDHs (RALDH1, RALDH2 and RALDH3) have been identified and exhibit distinct tissue expression patterns during mouse development, which collectively correlate with the dynamics of RA production. Note that according to some reports, RA synthesis during embryonic patterning may also occur independently of RALDHs and involve CYP1B1 or CYP1A1, members of the cytochrome p450 family of mono-oxygenases (Chambers et al., 2007; Chen et al., 2000).

On the other hand, RA is catabolized into more polar oxidized metabolites such as 4-hydroxy or 4-oxo RA through enzymes of the cytochrome P450 family (Fig. 1A). The mammalian genomes contain three *cyp26* genes, *cyp26A1*, *cyp26B1* and *cyp26C1* and orthologs have been cloned in other vertebrates like zebrafish (White and Schilling, 2008). They depict specific expression patterns, suggesting individual and combinatorial roles for each one in the control of RA levels.

Interestingly, both RA metabolic cascade and RAR-mediated RA signaling are conserved during evolution (Campo-Paysaa et al., 2008; Escriva et al., 2006; Marletaz et al., 2006; Samarut et al., 2011; Theodosiou et al., 2010), emphasizing the robustness and the significance of this pathway in many biological processes including embryonic development.

2.3. New pictures of RAR signaling: kinase cascade activation

Today it is becoming increasingly evident that, in addition to the classical genomic effects, liganded RARs also induce non-genomic responses such as the rapid and transient activation of several kinase cascades (Fig. 1C). These non-genomic effects are mediated by a subpopulation of RARs anchored at the cytoplasmic side of the cell membrane (Masia et al., 2007; Piskunov and Rochette-Egly, in press). Indeed, in several cell types, RA-bound RARs activate p38MAPK, which in turn, translocates into the nucleus and activates a downstream mitogen and stress-activated kinase, MSK1 (Bruck et al., 2009). However, this process appears to be cell-type specific as RARs rather activate Erks in neuronal and Sertoli cells (Chen and Napoli, 2008; Gupta et al., 2008; Masia et al., 2007; Pan et al., 2005). Nevertheless this suggests a new paradigm by which RARs integrate membrane/cytoplasm events orchestrating several coordinated cascades of phosphorylations of nuclear proteins, in order to fine-tune transcription (Fig. 1C).

In line with this, it emerged that in the nucleus, RA-activated MSK1 phosphorylates rapidly RARα at a serine residue (S369) located in loop L9-10 within the LBD (Bruck et al., 2009) (Fig. 1B). Phosphorylation of this serine increases the dynamics/flexibility of the nearby loop L8-9 (Samarut et al., 2011), which corresponds to the docking site of cyclin H, that forms with CDK7 and MAT1 the CAK subcomplex of the general transcription factor TFIIH. Consequently, RARα becomes able to bind cyclin H, allowing the phosphorylation of S77 located in the proline-rich domain of the NTD by the CDK7 kinase (Bour et al., 2005; Gaillard et al., 2006) (Fig. 1B). Finally, phosphorylation of the NTD drives the recruitment of RARα to promoters. Note that in the case of the RARγ subtype, phosphorylation of the NTD also induces the dissociation of vinexinβ, an actin-binding protein with SH3 domains, which represses RARγ activity (Lalevee et al., 2010). Thus RA-induced phosphorylations appear to play an important role in the dynamics of RAR-mediated transcription.

2.4. Multi-modulation of RAR-target genes through phosphorylation

Now it emerges that RARs act through the combination of genomic and non-genomic effects, introducing the notion of transcriptional dynamics so that the right proteins are present with the right activity at the right place and at the right time.

Indeed, at the very beginning of the RA switch, RA-activated MSK1 is required for making RAREs accessible to RAR/RXR heterodimers through binding to chromatin and phosphorylation of histones H3 (H3S10), an event that promotes other histones modifications and chromatin remodeling, according to the histone code (Bruck et al., 2009; Vermeulen et al., 2009). MSK1 is also required for DNA recruitment of RARα (Bruck et al., 2009), through RARα phosphorylation (see Section 2.1).

Whether MSK1 also controls the association/dissociation of coregulator complexes is still unclear but it has been shown that the upstream kinase p38MAPK phosphorylates corepressors (Perissi et al., 2008; Rosenfeld et al., 2006), as well as the coactivator SRC-3 (Gianni et al., 2006). Then phosphorylation promotes their ubiquitination according to a code very similar to the histone code and finally their degradation by the 26S proteasome (Gianni et al., 2006) (Fig. 1C). It has been proposed that phosphorylation, ubiquitination and degradation cooperate to control the corepressors/coactivators exchanges in line with the idea that transcription is a highly dynamic process. Lastly, at the end of the RA response, RARs are also ubiquitinated and degraded (Gianni et al., 2002; Kopf et al., 2000), a radical process for turning off the transcriptional signal (Bour et al., 2007; Rochette-Egly and Germain, 2009).

3. Dynamic expression of the RA signaling players provides a robust and coordinated RA response

3.1. RARs are the conductors transducing RA signaling

Numerous studies described the importance of RA signaling during early development and organogenesis (Clagett-Dame and DeLuca, 2002; Maden, 2007; Mark et al., 2006; Zile, 2001) with mouse being one of the most studied vertebrate model. These studies demonstrated that RA is crucial for the correct patterning of early structures in the embryo such as neural tube development and closure, spinal chord formation, rhombomeres patterning and specification, somitogenesis, limb bud morphogenesis and heart development.

How RA could exert such pleiotropic effects was a long-standing question until the discovery of RARs. Subsequently, genetic approaches in the animal corroborated that RARs are indeed the major conductors of RA signaling during embryonic development and decoded how they function at the molecular level (Mark et al., 2006, 2009). Mutant mice lacking RARα, RARβ or RARγ were first generated. Unexpectedly these single KO mice were viable and displayed a limited number of defects that did not mirror all those seen in the fetal and postnatal VAD syndromes (Table 1). In fact, complete VAD phenotypes were recapitulated only in double KO mutant mice (Table 1), suggesting that RARs are largely functionally redundant. The exhaustive analysis of the different double KO combinations allowed drawing a list assigning specific phenotypes and thus specific functions for RARs during development. Note that RAR double mutants also exhibited congenital abnormalities that were not

Table 1
Single RAR, double RAR and RXR/RAR KO phenotypes.

Simple KO abnormalities		Double KO fetal VAD-associated syndromes		
PN	growth retardation male sterility	RARα	+RARβ	respiratory system defects heart outflow tract defects kidney hypoplasia urogenital system defects
CD	webbed digit cervical vertebrae malformations altered alveolar formation		+RXRα	respiratory system defects heart outflow tract defects urogenital tract defects
PN	growth retardation	RARβ	+RARγ	heart outflow tract defects ureteral defects ocular defects
CD	cervical vertebrae malformations hyperplasia of the primary vitreous body (fetal VAD) altered alveolar formation behavioural defects		+RXRα	ocular defects heart outflow tract defects
PN	growth deficiency male sterility squameous metaplasia of several epithelia	RARγ	+RARα	hypoplasia of the ventricular myocardium heart outflow tract defects kidney hypoplasia urogenita tract defects ocular defects
CD	webbed digit cervical vertebrae malformations malformed laryngeal cartilage agenesis of the Harderian glands agenesis of the metoptic pillar of the skull altered alveolar formation abnormal differentiation of granular keratinocytes		+RXRα	hypoplasia of the sub-maxillary gland ocular defects heart outflow tract defects skeletal defects

Single KO of one RAR subtype in mouse (α, β, γ) is associated with congenital defects (CD) or post-natal (PN) abnormalities (left panel). Only double KO (RAR/RAR or RAR/RXRα) is associated with fetal VAD syndromes (right panel). These data are mainly adapted from the reviews by Mark et al. (2006, 2009). Refer to them for further details

described in the VAD syndrome such as skeletal defects, most probably due to the difficulty to achieve by dietary deprivation a state of profound VAD compatible with pregnancy.

Most interestingly, a marked synergy was observed between RARs and RXRs mutations (Kastner et al., 1997; Mark et al., 2009), indicating that there is much less redundancy between RARs in a RXRα mutant background. These data suggested that RXR/RAR heterodimers are the functional units transducing RA signal for a large number of RA-dependent processes. This is in full agreement with the *in vitro* studies showing that RXR/RAR heterodimers bind more tightly and specifically than RARs on their own to RAREs. Several malformations were seen in several types of RXRα/RAR double mutants while other given defects were seen in only one type of RXRα/RAR mutant combination (Table 1). As an example, respiratory and urogenital defects were observed specifically in the RXRα/RARα mutants while skeletal defects were seen in the RXRα/RARγ mutants (Mark et al., 2006, 2009). This multiplicity is not easy to interpret but suggests that the developmental processes are complex, involving the differential and sequential activation of different RARs.

Similar strategies were performed with zebrafish, which is another established vertebrate model widely used to study embryonic development, and which express two *RARα* (A and B) and two *RARγ* (A and B) genes (Bertrand et al., 2007; Hale et al., 2006; Linville et al., 2009). Using antisense morpholino oligonucleotides, it has been shown that the four RARs are required for hindbrain patterning, while RARα-B and RARγ-A are uniquely required in pectoral fin budding and pharyngeal arches development, respectively (Linville et al., 2009).

In conclusion, all these *in vivo* strategies revealed the powerful and complex potential of RARs as conductors of RA signaling during specific developmental processes.

Note however that some recent *in vitro* studies suggested that RA signaling could be mediated by other receptors such as RORβ (NR1F2) (Stehlin-Gaon et al., 2003), COUP-TFII (NR2F2) (Kruse et al., 2008), TR2/4 (Zhou et al., 2011) or PPARβ/δ (NR1C2) (Schug et al., 2007). Nevertheless the *in vivo* relevance of such observations during development remains to be determined.

3.2. RAR expression is subtly orchestrated during development

Given that a particular RXRα/RAR heterodimer plays a predominant role in particular developmental events, one can suggest that it would reflect the predominant presence of the corresponding RAR and RXR partners in the target tissue. Therefore the expression pattern of RARs, is expected to determine where and when RA signaling is important during development.

In situ hybridization experiments performed with mouse and zebrafish embryos showed that the expression of each RAR subtype and even of specific isoforms of a given RAR is regionally and temporally regulated (Dolle, 2009; Mollard et al., 2000) (Fig. 2). Indeed, at early developmental stages, RAR genes are already expressed in mouse and zebrafish as well as in most vertebrate species but their expression is overlapping and diffuse. The first signs of localized expression are seen at the gastrulation stage. Then the different RARs adopt specific patterns of tissue expression that change from the early to the late stages of embryonic development in line with their functions (Fig. 2). As an example, in mouse embryos, RARγ concentrates throughout the limb bud mesenchyme before being localized to precartilaginous cell condensations in line with its role in limb development (Fig. 2A). Similarly, in zebrafish (Fig. 2B), RARα-B distributes throughout the spinal cord and in the lateral plate mesoderm in line with its role in pectoral fin budding (Bertrand et al., 2007; Hale et al.,

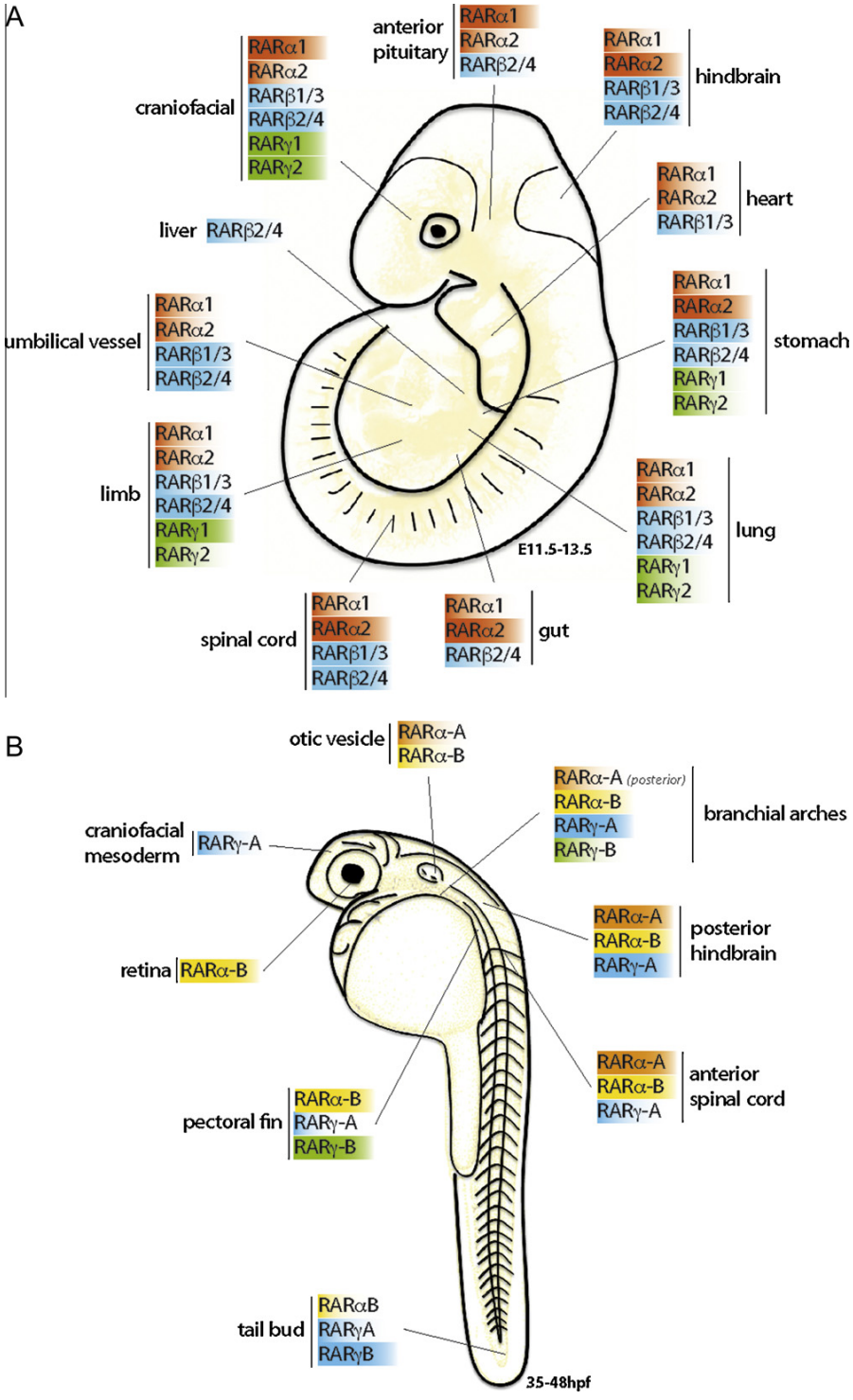


Fig. 2. Spatial RARs expression in mouse and zebrafish embryos. Schematic representation of the main expression patterns of the different RAR subtypes in mouse embryo (α, red; β, blue; γ, green) at stages E11.5-13.5 (A) and in zebrafish embryo (α-A, orange; α-B, yellow; γ-A, blue; γ-B, green) between 35 and 48hpf (B). Higher color intensities correspond to the highest expression level.
These schemes have been built from the data in (Dolle, 2009; Mollard et al., 2000) in the case of mouse embryos and in (Bertrand et al., 2007; Hale et al., 2006; Linville et al., 2009) for zebrafish. Refer to these references for further details.

2006; Linville et al., 2009). In contrast, RARγ-A expression restricts to cranial mesoderm where it plays a role in pharyngeal arches development (Linville et al., 2009).

3.3. Other players in the RA symphony: RAR coregulators

At the molecular level, corepressors and coactivators are involved in the fine-tuning of RARs transcriptional activity (see above) and thus should be other important actors of the RA symphony during development. In contrast to RARs, most corepressors are widely and ubiquitously expressed, suggesting that the specificity of RA signaling would be conferred more by the RARs themselves than by their coregulators (Bertrand et al., 2007). Nevertheless, in mouse embryos, knockout of the co-repressors N-CoR (Jepsen et al., 2000) and SMRT (Jepsen et al., 2007) led to altered patterns of RA-target genes transcription and of development. Similarly, in zebrafish, depletion of N-CoR disrupted early hindbrain patterning and phenocopied an increase in RA signaling (Xu et al., 2009). The conclusion was that corepressors are required for correct RAR-mediated RA signaling in diverse aspects of development by keeping off certain differentiation programs at certain times (Koide et al., 2001).

Interestingly, the p160 family coactivator NCoA3 (SRC-3/AIB1/ACTR) shows spatially restricted expression patterns, which are reminiscent of that of RARs and RXRs (Bertrand et al., 2007; Chen et al., 2010), suggesting possible crosstalks between RARs and this coactivator (Fig. 3). However, the link with RA signaling is still unclear.

Recently, RERE (a protein of the atrophin family) has been identified as a novel coregulator of RARs that is recruited to the core RARE promoter of RA-target genes in response to RA (Vilhais-Neto et al., 2010). Most interestingly, *rere* is ubiquitously expressed in the presomitic mesoderm and the analysis of Rere mutant mice demonstrated that this coregulator controls somite bilateral symmetry (Vilhais-Neto et al., 2010). Thus, the coregulator complexes associated to RARs are required to promote correct somite formation through the correct activation of the transcriptional activity of RA target genes.

3.4. Importance of the RA "sources and sinks"

Given that ligand binding controls the transcriptional activity of RARs, it is evident that during development, RA signaling should also require the coordinated production of RA through a balance between RALDH and CYP26 enzymes, which catalyze RA synthesis and degradation, respectively.

RALDH2 is the first RALDH to be expressed during embryogenesis (Niederreither et al., 1997, 1999; Theodosiou et al., 2010). Its expression starts during gastrulation in the posterior embryonic mesoderm and correlates with the stage at which RA signaling

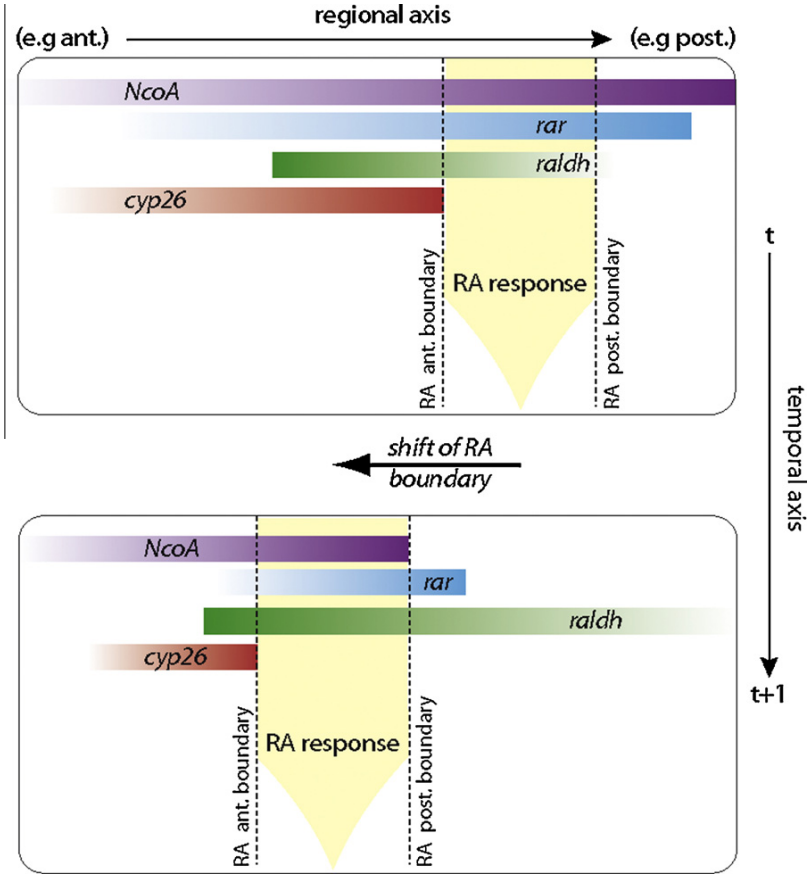


Fig. 3. Dynamic expression of RA actors delimit RA signaling boundaries. Schematic figure illustrating how the dynamic and coordinated expression of the different RA signaling players controls correct development. RA signaling players (i.e. coactivators (NCoA), nuclear receptors (rar), metabolic enzymes (raldh, cyp26)) depict a specific regional expression pattern (e.g. along an antero-posterior axis) at a t time (upper panel), which shapes RA boundaries and defines a correct RA responsive territory (yellow arrow). The expression pattern of the different players changes with time (from *t* to *t* + 1, lower panel), thus shifting RA boundaries and therefore RA responsive territories. In conclusion, the expression of all actors of RA signaling has to be coordinated in order to reach a correct, dynamic and harmonious development.

starts. From gastrulation to mid-gestation, the *raldh2* gene is expressed in a dynamic and spatially restricted manner in many mesodermal derivatives including somites, branchial arch mesenchyme and the lateral plate mesenchyme of the limb-forming region, highlighting the importance of RA signaling in the development fate of these cells. Accordingly, targeted inactivation of *raldh2* in mouse or in zebrafish (*neckless* (*nls*) mutants) resulted in hindbrain, neural crest, somites and limb defects (Begemann et al., 2001; Niederreither and Dolle, 2006; Niederreither et al., 1999). Moreover, in mouse, loss of function of *raldh2* mimics the most severe phenotypes associated with VAD.

It is worthy to note that the CYP26 enzymes that are involved in RA degradation (and thus supposed to inhibit the transcriptional activity of the surrounding RARs) are expressed in regions adjacent (and non overlapping) to those producing RA (Niederreither and Dolle, 2006; Ribes et al., 2007). Accordingly, during gastrulation, CYP26A1 and RALDH2 are expressed in the anterior and posterior regions of mouse embryos, respectively. Such opposite locations suggested the existence of a "source-and-sink" mechanism for regulation of RA signaling. However, the dynamic expression of the different CYP26 enzymes especially during hindbrain patterning (Glover et al., 2006), suggested that they would rather define sequential sharp boundaries of RA degradation. Such boundaries would regulate the formation of waves or pulses of RA so that right levels of RA are available at the right place and at the right time in order to provide a fine-shaped RA responsiveness profile during development (Hernandez et al., 2007; Sirbu et al., 2005; White et al., 2007; White and Schilling, 2008) (Fig. 3).

Remarkably, disruption of the *cyp26A1* gene in mouse (Abu-Abed et al., 2001; White and Schilling, 2008) and zebrafish embryos (Hernandez et al., 2007; Pennimpede et al., 2010, 2006) as well as inactivation of *cyp26A1* in the *giraffe* (*gir*) zebrafish mutants (Emoto et al., 2005), generate multiple patterning defects that are opposite to the defects observed in the *raldh2* mutants and that are reminiscent of the teratogenic effects of RA. Therefore, it has been proposed that *cyp26* expression would bring a robust adaptive feature in order to protect the developing embryos against unappropriate fluctuations in RA levels (White and Schilling, 2008). This is at the basis of the long-standing observation that depletion of endogenous RA in *raldh2* null embryos can be rescued by uniform exposure to exogenous RA (Niederreither et al., 1999). Thus RA responsive territories would not be shaped by a simple gradient of RA diffusing over long distances but by dynamic stepwise confinements of RA through the degrading CYP26 enzymes. In addition, given that *raldh2* and *cyp26A1* are down and up-regulated respectively by RA (Loudig et al., 2005; Niederreither et al., 1997), it has been proposed that complex feedback interactions exist between the different actors

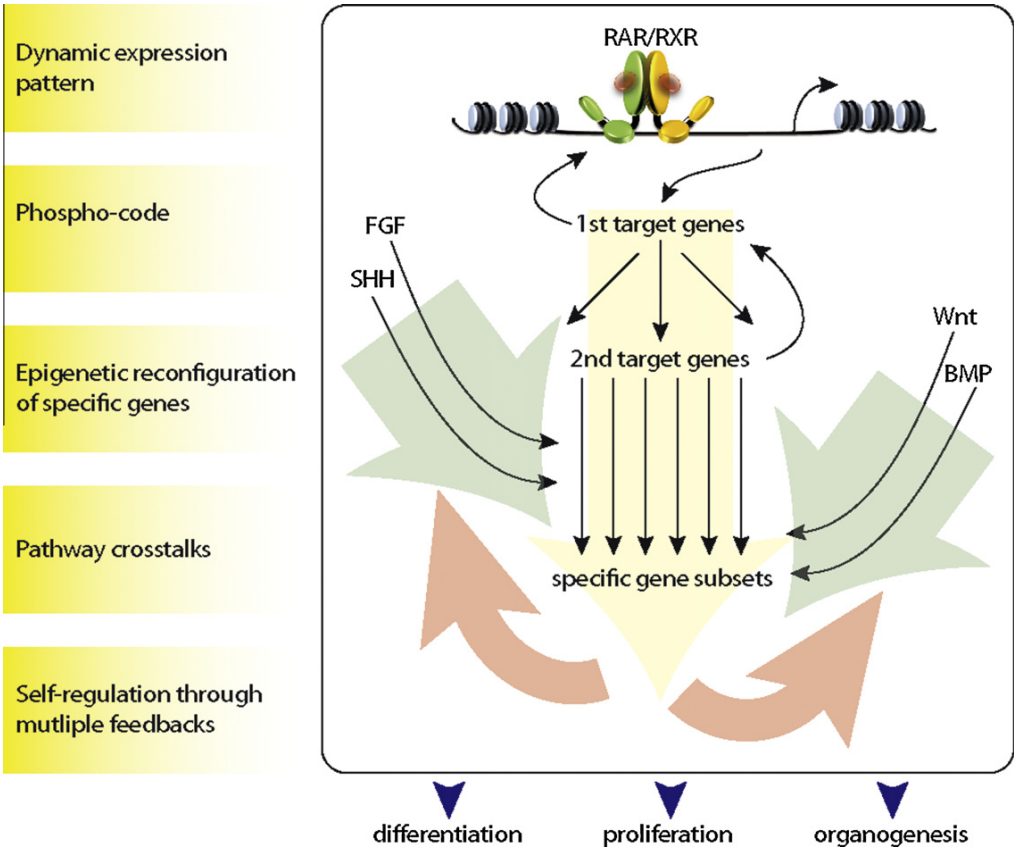


Fig. 4. Molecular aspects of RAR-mediated RA signaling: complex and fine-tuned regulation. Liganded RXR/RAR heterodimers rapidly activate direct target genes (e.g. *hox*, *cyp26a1*, etc.) through binding to response elements located in their promoters. In turn, the products of these primary genes activate their own target genes (i.e. 2nd target genes) resulting into a cascade of genes activation and expression leading to specific phenotypes. This linear activation cascade (yellow arrow) integrates multiple other signaling pathways such as the FGF, Shh, Wnt and BMP pathways which crosstalk with RA signaling. It is also controlled by phosphorylations, histone "language" and several feedbacks that increase the complexity of the RA signaling network. As a consequence, we observe an amplified gene response far from the canonical model of the direct and simple activation of target genes.

of the RA signaling pathway, providing an extremely robust system (Fig. 4).

In conclusion, dynamic and coordinated spatial and temporal regulations of RA metabolism also fine-tune the transcriptional activity of RARs through providing them the right amount of ligand at the right time and at the right place in order to reach harmonious effects of RA during development (Romand et al., 2004) (Fig. 3).

4. RARs conduct RA signaling in an integrative and coordinated manner

4.1. Crosstalks with other signaling pathways fine-shape RA boundaries

Embryonic development is intrinsically dynamic as illustrated by hindbrain patterning and somitogenesis. This relies on the dynamics of RALDH2 and CYP26 enzymes, which control coordinated extensions and retreats of the RA signal during the course of development (see Section 3.4). However, embryonic patterning depends not only on RA signaling but also on several other pathways such as the fibroblast growth factor (FGF), Wnt, sonic hedgehog (SHH), Notch, BMP and others, which act according to combinatorial gradient systems or as oscillators (Aulehla and Pourquie, 2008; Dubrulle and Pourquie, 2004b; Goldbeter and Pourquie, 2008; Ma and Raible, 2009). The interesting point is that all these pathways interfere with each other and also with RA, thereby increasing the complexity of RA signaling during development (Engberg et al., 2010; Olivera-Martinez and Storey, 2007).

As an example, FGFs, which are extracellular signaling molecules that bind transmembrane receptors with tyrosine kinase activity (Itoh, 2007), crosstalk with RA for the coordination of somitogenesis (Aulehla and Pourquie, 2010; Dubrulle and Pourquie, 2004a; Rhodes and Lohnes, 2006; Vermot et al., 2005; Vermot and Pourquie, 2005), hindbrain patterning (White et al., 2007) and early neural patterning (Kudoh et al., 2002; Shiotsugu et al., 2004). Remarkably, FGF and RA signalings appear to antagonize each other, suggesting that embryonic patterning would be controlled by series of mutually interactive feedback loops between FGF and RARs (Duester, 2008; Zhao and Duester, 2009).

However, a molecular understanding of how RA and FGF oppose each other is still controversial. In the context of bilateral symmetry of somite formation, which relies on a clock and wavefront mechanism with rhythmic and oscillating expression of genes including *fgf8* along the presomitic mesoderm, it has been proposed that RA would synchronize the segmentation clock oscillations through repressing *fgf8* expression (Sirbu and Duester, 2006; Vermot et al., 2005; Vermot and Pourquie, 2005). In contrast, during hindbrain development, FGF suppresses expression of *cyp26A1* (Kudoh et al., 2002; White et al., 2007), while RA has the opposite effect (Loudig et al., 2005), suggesting that CYP26 enzymes would stand as linkers between the two pathways. It has been proposed that these opposing forces would create an extremely robust system to control RA availability and therefore the correct temporal and regional expression of specific developmental genes.

Finally, given that FGF signaling is mediated by several downstream effectors including PI3K, Akt, ERKs and p38MAPK and that the graded distribution of the FGF proteins has been correlated with a similar gradient of MAPKs and Akt (Aulehla and Pourquie, 2010; Benazeraf et al., 2010), one cannot exclude that FGF controls RA signaling through the phosphorylation of several actors of the RA signaling pathway, including RARs themselves and their coregulators. Accordingly, it has been shown that phosphorylation by Akt inhibits RARα binding to DNA and SMRT dissociation (Lefebvre et al., 2006; Srinivas et al., 2006).

In conclusion whatever the mechanism is, both the RA and FGF pathways include many feedback and feed-forward loops. This forms a robust integrated system that transduces the right signal at the right time in a harmonized fashion.

4.2. RARs molecular targets: from direct target genes to coordinated gene cascades

During development, cell differentiation is one of the most crucial steps and requires on one hand the down-regulation of gene subsets involved in keeping the cells undifferentiated and in controlling their self-renewal features and on the other hand the activation of new gene combinations driving lineage-specific differentiation. The role of repression and/or activation of definite genes during differentiation has been exemplified in the context of embryocarcinoma cells (Bour et al., 2006) and embryonic stem (ES) cells (Gudas and Wagner, 2011), which are pluripotent cells derived from the inner cell mass of the blastocyst. The interesting feature of ES cells is that they are capable of differentiating into a very large variety of cell types *in vitro* and contribute to the generation of the notochord, neural tube and somites in mouse embryos (Wilson et al., 2009).

The earliest action of RA is to promote the repression of a gene regulatory network centered on *oct4* and *nanog*, which are essential for the establishment and the maintenance of the pluripotential state (Chambers and Tomlinson, 2009; Chan et al., 2011; Kalmar et al., 2009). The expression of these genes is regulated by a combination of several transcription factors and nuclear receptors (Fuhrmann et al., 2001; Gu et al., 2005; Mullen et al., 2007) but how they are repressed by RA signaling is still unclear. However it must be noted that their promoters depict RAR-binding sequences associated with repression (Barnea and Bergman, 2000; Fuhrmann et al., 2001).

Then the new high-throughput technologies revealed that in mouse embryocarcinoma F9 cells (Bour et al., 2006) and ES cells (Delacroix et al., 2010), RA rapidly activates the expression of more than 300 genes that have been correlated to differentiation. However, only some of these genes are direct RA targets with RAREs located in their promoters. These RARE-driven genes include among others, the *cyp26A1*, *rarβ2*, *foxA1/hnf3a*, and several *hox* genes such as the *hoxa1*, *hoxb1*, *hoxc4* and *hoxd4* genes. Depending on the differentiation step and on the target gene, the activation of these genes requires different RAR subtypes (Kashyap et al., 2011; Rochette-Egly and Chambon, 2001; Zechel, 2005). These direct target genes are the initiators of a network of cascades of activation of several other downstream genes that will finally lead to the expression of components characteristic of the differentiation state and involved in tissue patterning and organogenesis (Fig. 4).

Remarkably, as mentioned in the context of embryonic development, to regulate gene expression and promote ES cells differentiation, RA also crosstalks with several other pathways such as FGF signaling (Engberg et al., 2010; Lu et al., 2009; Stavridis et al., 2010) (Fig. 4). According to recent reports, in ES cells, this crosstalk involves a dual mechanism with first an increase and then a decrease in FGF signaling (Stavridis et al., 2010). The rapid induction by RA of *fgf8* and downstream Erk activity has been correlated to the loss of self-renewal. In contrast, the long-term repression of *fgf4* drives ES cells neuronal differentiation, corroborating that attenuation of FGF signaling by RA is a conserved fundamental mechanism driving differentiation towards somatic cell fates. How RA signaling down-regulates FGF signaling is still unclear, but several complex RAREs have been identified in the *fgf8* gene promoter (Brondani et al., 2002; Zhao et al., 2009). In contrast, *fgf4* repression would be indirect and would rather reflect the upstream repression of *oct4* (Gu et al., 2005).

Finally, one must not forget that, in several cell types including multipotent embryocarcinoma cells (Rochette-Egly and Germain, 2009), RA activates rapidly MAPKs through non-genomic mechanisms, which have been correlated to the repression of *oct4* (Gupta et al., 2008) and to the activation of several RA-target genes involved in RA metabolism such as *cyp26A1* (Bruck et al., 2009) or in cell differentiation (Bour et al., 2006; Rochette-Egly and Chambon, 2001; Taneja et al., 1997). However, in the context of ES cells, whether MAPKs are also activated through this process or through the genomic activation of *fgf8* remains to be determined (Stavridis et al., 2007, 2010).

Taken together these data highlight the immense complexity of the signals involved in gene regulation during cell differentiation and embryonic development. One of the best examples is that of *hox* genes which are organized in four clusters (*hoxa*, *hoxb*, *hoxc* and *hoxd*). A number of genetic studies revealed that several RA-regulated waves of transcriptional activation or repression of *hox* genes occur in a precise spatio-temporal pattern along the antero-posterior axis during neural development. The interesting point is that while some *hox* genes are regulated through RARs binding to RAREs, others are controlled by the interconnected FGF and Wnt signaling pathways. In addition, critical cross-regulations between the different *hox* genes have been described. Thus, particular combinations of *hox* gene products referred to as the Hox code must be strictly controlled both spatially and temporally for correct embryonic patterning (Glover et al., 2006; Rhodes and Lohnes, 2006).

4.3. Epigenetic and chromatin modifications: a dynamic "language" to modulate RAR-target gene expression

During cell differentiation and development, correct gene expression requires a proper chromatin structure and compaction state, which relies on a subtle code of histone post-translational modifications, including acetylation, methylation and phosphorylation (Bloushtain-Qimron et al., 2009; Lee et al., 2010; Smith and Shilatifard, 2010). In general, trimethylation of lysines 27 and 9 on histone H3 (H3K27,9) that involves binding of the polycomb repressive complex 2 (PRC2) is a repressive modification. In contrast, trimethylation and acetylation at lysines 4 and 9 respectively, are permissive.

In undifferentiated ES cells, the genes involved in development and cell fate decisions, including the *hoxa1* and *hoxb1* genes are trimethylated at H3K27 (Kashyap et al., 2011) but maintained in a poised state so that they may be rapidly transcribed when triggered by developmental cues. Accordingly, during RA-induced differentiation, the H3K27 trimethylation mark decreases (Lee et al., 2007). However, this loss is not sufficient and needs to be associated to a rapid and broad epigenetic reorganization of the *hoxa* and *hoxb* clusters with H3K4 methylation and H3 acetylation, through the recruitment to response elements of the RARγ subtype (Kashyap et al., 2011) with the associated large complexes endowed with histone modifying and remodeling activities. Whether these steps are also modulated by the phosphorylation of histones and RARs, as in other cell types (Bruck et al., 2009) will require further investigations. Nevertheless, it is tempting to speculate that such an epigenomic reconfiguration would favorize the assembly or recruitment of several other factories in order to facilitate coordinated intra-cluster regulation. Again these events are driven by RARs, underlining their central role in RA-signaling.

5. Conclusions and perspectives

The discovery of RARs coupled to recent investigations into their roles, led to substantial new insights into how RA signaling participates to embryonic development. Now there is compelling evidence that RARs are the conductors of the RA symphony. Indeed they conduct spatio-temporally the different steps of development through controlling the expression of the different developmental genes. In this role they are helped by several players, (i) the enzymes involved in RA synthesis and degradation that control the right amount of ligand at the right time and at the right place, (ii) other signaling pathways that interfere with RA synthesis and the genes activation cascades. The interconnections between these different players have been deciphered in specific tissues and patterning processes as well as in embryonic stem cells but the complexity of the networks is still increasing. Thus, as in an orchestra, all the RA signaling players have a specific role and are coordinated by a conductor, RAR, which is the main protagonist coordinating the others, in order to reach a robust and harmonious development. Now the challenge would be to visualize in real-time the different players of the RA symphony during embryonic development.

Acknowledgements

We specially thank Ziad Al Tanoury and Gabriella Pankotai-Bodó for critical reading of this manuscript. Funds from CNRS, INSERM, the Agence Nationale pour la Recherche (ANR-05-BLAN-0390-02 and ANR-09-BLAN-0297-01), the Foundation pour la Recherche Medicale (FRM, DEQ20090515423), the Institut National du Cancer (INCa-PLO7-96099 and PL09-194) and the Association pour la Recherche sur le Cancer (ARC 3169) supported this work. The Ministère de l'Enseignement Supérieur et de la Recherche supported ES.

References

Abu-Abed, S., Dolle, P., Metzger, D., Beckett, B., Chambon, P., Petkovich, M., 2001. The retinoic acid-metabolizing enzyme, CYP26A1, is essential for normal hindbrain patterning, vertebral identity, and development of posterior structures. Genes Dev. 15, 226–240.

Aulehla, A., Pourquie, O., 2008. Oscillating signaling pathways during embryonic development. Curr. Opin. Cell Biol. 20, 632–637.

Aulehla, A., Pourquie, O., 2010. Signaling gradients during paraxial mesoderm development. Cold Spring Harb. Perspect. Biol. 2, a000869.

Balmer, J.E., Blomhoff, R., 2005. A robust characterization of retinoic acid response elements based on a comparison of sites in three species. J. Steroid Biochem. Mol. Biol. 96, 347–354.

Barnea, E., Bergman, Y., 2000. Synergy of SF1 and RAR in activation of Oct-3/4 promoter. J. Biol. Chem. 275, 6608–6619.

Bastien, J., Rochette-Egly, C., 2004. Nuclear retinoid receptors and the transcription of retinoid-target genes. Gene 328, 1–16.

Begemann, G., Schilling, T.F., Rauch, G.J., Geisler, R., Ingham, P.W., 2001. The zebrafish neckless mutation reveals a requirement for raldh2 in mesodermal signals that pattern the hindbrain. Development 128, 3081–3094.

Benazeraf, B., Francois, P., Baker, R.E., Denans, N., Little, C.D., Pourquie, O., 2010. A random cell motility gradient downstream of FGF controls elongation of an amniote embryo. Nature 466, 248–252.

Bertrand, S., Thisse, B., Tavares, R., Sachs, L., Chaumot, A., Bardet, P.L., Escriva, H., Duffraisse, M., Marchand, O., Safi, R., Thisse, C., Laudet, V., 2007. Unexpected novel relational links uncovered by extensive developmental profiling of nuclear receptor expression. PLoS Genet. 3, e188.

Biddie, S.C., John, S., Hager, G.L., 2010. Genome-wide mechanisms of nuclear receptor action. Trends Endocrinol. Metab. 21, 3–9.

Bloushtain-Qimron, N., Yao, J., Shipitsin, M., Maruyama, R., Polyak, K., 2009. Epigenetic patterns of embryonic and adult stem cells. Cell Cycle 8, 809–817.

Bour, G., Gaillard, E., Bruck, N., Lalevee, S., Plassat, J.L., Busso, D., Samama, J.P., Rochette-Egly, C., 2005. Cyclin H binding to the RAR{alpha} activation function (AF)-2 domain directs phosphorylation of the AF-1 domain by cyclin-dependent kinase 7. Proc. Natl. Acad. Sci. U.S.A. 102, 16608–16613.

Bour, G., Lalevee, S., Rochette-Egly, C., 2007. Protein kinases and the proteasome join in the combinatorial control of transcription by nuclear retinoic acid receptors. Trends Cell Biol. 17, 302–309.

Bour, G., Taneja, R., Rochette-Egly, C., 2006. Mouse embryocarcinoma F9 cells and retinoic acid. A model to study the molecular mechanisms of endodermal differentiation. In: Taneja, R. (Ed.), Nuclear Receptors in Development, vol. 16. Elsevier Press Inc, pp. 211–253.

Brondani, V., Klimkait, T., Egly, J.M., Hamy, F., 2002. Promoter of FGF8 reveals a unique regulation by unliganded RARalpha. J. Mol. Biol. 319, 715–728.

Bruck, N., Vitoux, D., Ferry, C., Duong, V., Bauer, A., de The, H., Rochette-Egly, C., 2009. A coordinated phosphorylation cascade initiated by p38MAPK/MSK1 directs RARalpha to target promoters. EMBO J. 28, 34–47.

Budhu, A.S., Noy, N., 2002. Direct channeling of retinoic acid between cellular retinoic acid-binding protein II and retinoic acid receptor sensitizes mammary carcinoma cells to retinoic acid-induced growth arrest. Mol. Cell. Biol. 22, 2632–2641.

Calleja, C., Messaddeq, N., Chapellier, B., Yang, H., Krezel, W., Li, M., Metzger, D., Mascrez, B., Ohta, K., Kagechika, H., Endo, Y., Mark, M., Ghyselinck, N.B., Chambon, P., 2006. Genetic and pharmacological evidence that a retinoic acid cannot be the RXR-activating ligand in mouse epidermis keratinocytes. Genes Dev. 20, 1525–1538.

Campo-Paysaa, F., Marletaz, F., Laudet, V., Schubert, M., 2008. Retinoic acid signaling in development: tissue-specific functions and evolutionary origins. Genesis 46, 640–656.

Chambers, D., Wilson, L., Maden, M., Lumsden, A., 2007. RALDH-independent generation of retinoic acid during vertebrate embryogenesis by CYP1B1. Development 134, 1369–1383.

Chambers, I., Tomlinson, S.R., 2009. The transcriptional foundation of pluripotency. Development 136, 2311–2322.

Chambon, P., 1996. A decade of molecular biology ot retinoic acid receptors. FASEB J. 10, 940–954.

Chan, Y.S., Yang, L., Ng, H.H., 2011. Transcriptional regulatory networks in embryonic stem cells. Prog. Drug Res. 67, 239–252.

Chen, H., Howald, W.N., Juchau, M.R., 2000. Biosynthesis of all-trans-retinoic acid from all-trans-retinol: catalysis of all-trans-retinol oxidation by human P-450 cytochromes. Drug Metab. Dispos. 28, 315–322.

Chen, N., Napoli, J.L., 2008. All-trans-retinoic acid stimulates translation and induces spine formation in hippocampal neurons through a membrane-associated RARalpha. FASEB J. 22, 236–245.

Chen, X., Liu, Z., Xu, J., 2010. The cooperative function of nuclear receptor coactivator 1 (NCOA1) and NCOA3 in placental development and embryo survival. Mol. Endocrinol. 24, 1917–1934.

Clagett-Dame, M., DeLuca, H.F., 2002. The role of vitamin A in mammalian reproduction and embryonic development. Annu. Rev. Nutr. 22, 347–381.

Delacroix, L., Moutier, E., Altobelli, G., Legras, S., Poch, O., Choukrallah, M.A., Bertin, I., Jost, B., Davidson, I., 2010. Cell-specific interaction of retinoic acid receptors with target genes in mouse embryonic fibroblasts and embryonic stem cells. Mol. Cell. Biol. 30, 231–244.

Delva, L., Bastie, J.N., Rochette-Egly, C., Kraiba, R., Balitrand, N., Despouy, G., Chambon, P., Chomienne, C., 1999. Physical and functional interactions between cellular retinoic acid binding protein II and the retinoic acid-dependent nuclear complex. Mol. Cell. Biol. 19, 7158–7167.

Dilworth, F.J., Chambon, P., 2001. Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription. Oncogene 20, 3047–3054.

Dolle, P., 2009. Developmental expression of retinoic acid receptors (RARs). Nucl. Recept. Signal. 7, e006.

Dubrulle, J., Pourquie, O., 2004a. Coupling segmentation to axis formation. Development 131, 5783–5793.

Dubrulle, J., Pourquie, O., 2004b. fgf8 mRNA decay establishes a gradient that couples axial elongation to patterning in the vertebrate embryo. Nature 427, 419–422.

Duester, G., 2008. Retinoic acid synthesis and signaling during early organogenesis. Cell 134, 921–931.

Emoto, Y., Wada, H., Okamoto, H., Kudo, A., Imai, Y., 2005. Retinoic acid-metabolizing enzyme Cyp26a1 is essential for determining territories of hindbrain and spinal cord in zebrafish. Dev. Biol. 278, 415–427.

Engberg, N., Kahn, M., Petersen, D.R., Hansson, M., Serup, P., 2010. Retinoic acid synthesis promotes development of neural progenitors from mouse embryonic stem cells by suppressing endogenous, Wnt-dependent nodal signaling. Stem Cells 28, 1498–1509.

Escriva, H., Bertrand, S., Germain, P., Robinson-Rechavi, M., Umbhauer, M., Cartry, J., Duffraisse, M., Holland, L., Gronemeyer, H., Laudet, V., 2006. Neofunctionalization in vertebrates: the example of retinoic acid receptors. PLoS Genet. 2, e102.

Flajollet, S., Lefebvre, B., Rachez, C., Lefebvre, P., 2006. Distinct roles of the steroid receptor coactivator 1 and of MED1 in retinoid-induced transcription and cellular differentiation. J. Biol. Chem. 281, 20338–20348.

Fuhrmann, G., Chung, A.C., Jackson, K.J., Hummelke, G., Baniahmad, A., Sutter, J., Sylvester, I., Scholer, H.R., Cooney, A.J., 2001. Mouse germline restriction of Oct4 expression by germ cell nuclear factor. Dev. Cell 1, 377–387.

Gaillard, E., Bruck, N., Brelivet, Y., Bour, G., Lalevee, S., Bauer, A., Poch, O., Moras, D., Rochette-Egly, C., 2006. Phosphorylation by PKA potentiates retinoic acid receptor alpha activity by means of increasing interaction with and phosphorylation by cyclin H/cdk7. Proc. Natl. Acad. Sci. U.S.A. 103, 9548–9553.

Germain, P., Chambon, P., Eichele, G., Evans, R.M., Lazar, M.A., Leid, M., De Lera, A.R., Lotan, R., Mangelsdorf, D.J., Gronemeyer, H., 2006a. International Union of Pharmacology. LX. Retinoic acid receptors. Pharmacol. Rev. 58, 712–725.

Germain, P., Chambon, P., Eichele, G., Evans, R.M., Lazar, M.A., Leid, M., De Lera, A.R., Lotan, R., Mangelsdorf, D.J., Gronemeyer, H., 2006b. International Union of Pharmacology. LXIII. Retinoid X receptors. Pharmacol. Rev. 58, 760–772.

Germain, P., Staels, B., Dacquet, C., Spedding, M., Laudet, V., 2006c. Overview of nomenclature of nuclear receptors. Pharmacol. Rev. 58, 685–704.

Gianni, M., Bauer, A., Garattini, E., Chambon, P., Rochette-Egly, C., 2002. Phosphorylation by p38MAPK and recruitment of SUG-1 are required for RA-indced RARγ degradation and transactivation. EMBO J. 21, 3760–3769.

Gianni, M., Parrella, E., Raska, I., Gaillard, E., Nigro, E.A., Gaudon, C., Garattini, E., Rochette-Egly, C., 2006. P38MAPK-dependent phosphorylation and degradation of SRC-3/AIB1 and RARalpha-mediated transcription. EMBO J. 25, 739–751.

Gillespie, R.F., Gudas, L.J., 2007a. Retinoic acid receptor isotype specificity in F9 teratocarcinoma stem cells results from the differential recruitment of coregulators to retinoic response elements. J. Biol. Chem. 282, 33421–33434.

Gillespie, R.F., Gudas, L.J., 2007b. Retinoid regulated association of transcriptional co-regulators and the polycomb group protein SUZ12 with the retinoic acid response elements of Hoxa1, RARbeta(2), and Cyp26A1 in F9 embryonal carcinoma cells. J. Mol. Biol. 372, 298–316.

Glover, J., Renaud, J., Lampe, X., Rijli, F., 2006. Hindbrain development and retinoids. In: Taneja, R. (Ed.), Nuclear Receptors in Development, vol. 16. Elsevier, pp. 145–180.

Goldbeter, A., Pourquie, O., 2008. Modeling the segmentation clock as a network of coupled oscillations in the Notch, Wnt and FGF signaling pathways. J. Theor. Biol. 252, 574–585.

Gu, P., LeMenuet, D., Chung, A.C., Mancini, M., Wheeler, D.A., Cooney, A.J., 2005. Orphan nuclear receptor GCNF is required for the repression of pluripotency genes during retinoic acid-induced embryonic stem cell differentiation. Mol. Cell. Biol. 25, 8507–8519.

Gudas, L.J., Wagner, J.A., 2011. Retinoids regulate stem cell differentiation. J. Cell. Physiol. 226, 322–330.

Gupta, P., Ho, P.C., Huq, M.M., Ha, S.G., Park, S.W., Khan, A.A., Tsai, N.P., Wei, L.N., 2008. Retinoic acid-stimulated sequential phosphorylation, PML recruitment, and SUMOylation of nuclear receptor TR2 to suppress Oct4 expression. Proc. Natl. Acad. Sci. U.S.A. 105, 11424–11429.

Hale, L.A., Tallafuss, A., Yan, Y.L., Dudley, L., Eisen, J.S., Postlethwait, J.H., 2006. Characterization of the retinoic acid receptor genes raraa, rarab and rarg during zebrafish development. Gene. Expr. Patterns 6, 546–555.

Hernandez, R.E., Putzke, A.P., Myers, J.P., Margaretha, L., Moens, C.B., 2007. Cyp26 enzymes generate the retinoic acid response pattern necessary for hindbrain development. Development 134, 177–187.

Hua, S., Kittler, R., White, K.P., 2009. Genomic antagonism between retinoic acid and estrogen signaling in breast cancer. Cell 137, 1259–1271.

Itoh, N., 2007. The Fgf families in humans, mice, and zebrafish: their evolutional processes and roles in development, metabolism, and disease. Biol. Pharm. Bull. 30, 1819–1825.

Jepsen, K., Hermanson, O., Onami, T.M., Gleiberman, A.S., Lunyak, V., McEvilly, R.J., Kurokawa, R., Kumar, V., Liu, F., Seto, E., Hedrick, S.M., Mandel, G., Glass, C.K., Rose, D.W., Rosenfeld, M.G., 2000. Combinatorial roles of the nuclear receptor corepressor in transcription and development. Cell 102, 753–763.

Jepsen, K., Solum, D., Zhou, T., McEvilly, R.J., Kim, H.J., Glass, C.K., Hermanson, O., Rosenfeld, M.G., 2007. SMRT-mediated repression of an H3K27 demethylase in progression from neural stem cell to neuron. Nature 450, 415–419.

Kalmar, T., Lim, C., Hayward, P., Munoz-Descalzo, S., Nichols, J., Garcia-Ojalvo, J., Martinez Arias, A., 2009. Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. PLoS Biol. 7, e1000149.

Kashyap, V., Gudas, L.J., Brenet, F., Funk, P., Viale, A., Scandura, J.M., 2011. Epigenomic reorganization of the clustered hox genes in embryonic stem cells induced by retinoic acid. J. Biol. Chem. 286, 3250–3260.

Kastner, P., Mark, M., Ghyselinck, N., Krezel, W., Dupe, V., Grondona, J.M., Chambon, P., 1997. Genetic evidence that the retinoid signal is transduced by heterodimeric RXR/RAR functional units during mouse development. Development 124, 313–326.

Koide, T., Downes, M., Chandraratna, R.A., Blumberg, B., Umesono, K., 2001. Active repression of RAR signaling is required for head formation. Genes Dev. 15, 2111–2121.

Kopf, E., Plassat, J.L., Vivat, V., de The, H., Chambon, P., Rochette-Egly, C., 2000. Dimerization with retinoid X receptors and phosphorylation modulate the retinoic acid-induced degradation of retinoic acid receptors alpha and gamma through the ubiquitin-proteasome pathway. J. Biol. Chem. 275, 33280–33288.

Kruse, S.W., Suino-Powell, K., Zhou, X.E., Kretschman, J.E., Reynolds, R., Vonrhein, C., Xu, Y., Wang, L., Tsai, S.Y., Tsai, M.J., Xu, H.E., 2008. Identification of COUP-TFII orphan nuclear receptor as a retinoic acid-activated receptor. PLoS Biol. 6, e227.

Kudoh, T., Wilson, S.W., Dawid, I.B., 2002. Distinct roles for Fgf, Wnt and retinoic acid in posteriorizing the neural ectoderm. Development 129, 4335–4346.

Lalevee, S., Bour, G., Quinternet, M., Samarut, E., Kessler, P., Vitorino, M., Bruck, N., Delsuc, M.A., Vonesch, J.L., Kieffer, B., Rochette-Egly, C., 2010. Vinexin{beta}, an atypical "sensor" of retinoic acid receptor {gamma} signaling: union and sequestration, separation, and phosphorylation. FASEB J. 24, 4523–4534.

Laudet, V., Gronemeyer, H., 2001. Nuclear Receptor Factsbook. Acedemic Press, London.

le Maire, A., Teyssier, C., Erb, C., Grimaldi, M., Alvarez, S., de Lera, A.R., Balaguer, P., Gronemeyer, H., Royer, C.A., Germain, P., Bourguet, W., 2010. A unique secondary-structure switch controls constitutive gene repression by retinoic acid receptor. Nat. Struct. Mol. Biol. 17, 801–807.

Lee, E.R., Murdoch, F.E., Fritsch, M.K., 2007. High histone acetylation and decreased polycomb repressive complex 2 member levels regulate gene specific transcriptional changes during early embryonic stem cell differentiation induced by retinoic acid. Stem Cells 25, 2191–2199.

Lee, J.S., Smith, E., Shilatifard, A., 2010. The language of histone crosstalk. Cell 142, 682–685.

Lee, M.S., Kliewer, S.A., Provencal, J., Wright, P.E., Evans, R.M., 1993. Structure of the retinoid X receptor alpha DNA binding domain: a helix required for homodimeric DNA binding. Science 260, 1117–1121.
Lefebvre, B., Brand, C., Flajollet, S., Lefebvre, P., 2006. Down-regulation of the tumor suppressor gene retinoic acid receptor beta2 through the phosphoinositide 3-kinase/Akt signaling pathway. Mol. Endocrinol. 20, 2109–2121.
Lefebvre, P., Martin, P.J., Flajollet, S., Dedieu, S., Billaut, X., Lefebvre, B., 2005. Transcriptional activities of retinoic acid receptors. Vitam. Horm. 70, 199–264.
Leid, M., Kastner, P., Chambon, P., 1992. Multiplicity generates diversity in the retinoic acid signalling pathways. Trends Biochem. Sci. 17, 427–433.
Linville, A., Radtke, K., Waxman, J.S., Yelon, D., Schilling, T.F., 2009. Combinatorial roles for zebrafish retinoic acid receptors in the hindbrain, limbs and pharyngeal arches. Dev. Biol. 325, 60–70.
Loudig, O., Maclean, G.A., Dore, N.L., Luu, L., Petkovich, M., 2005. Transcriptional co-operativity between distant retinoic acid response elements in regulation of Cyp26A1 inducibility. Biochem. J. 392, 241–248.
Lu, J., Tan, L., Li, P., Gao, H., Fang, B., Ye, S., Geng, Z., Zheng, P., Song, H., 2009. All-trans retinoic acid promotes neural lineage entry by pluripotent embryonic stem cells via multiple pathways. BMC Cell Biol. 10, 57.
Ma, E.Y., Raible, D.W., 2009. Signaling pathways regulating zebrafish lateral line development. Curr. Biol. 19, R381–R386.
Maden, M., 2007. Retinoic acid in the development, regeneration and maintenance of the nervous system. Nat. Rev. Neurosci. 8, 755–765.
Mangelsdorf, D.J., Evans, R.M., 1995. The RXR heterodimers and orphan receptors. Cell 83, 841–850.
Mark, M., Ghyselinck, N.B., Chambon, P., 2006. Function of retinoid nuclear receptors: lessons from genetic and pharmacological dissections of the retinoic acid signaling pathway during mouse embryogenesis. Annu. Rev. Pharmacol. Toxicol. 46, 451–480.
Mark, M., Ghyselinck, N.B., Chambon, P., 2009. Function of retinoic acid receptors during embryonic development. Nucl. Recept. Signal. 7, e002.
Marletaz, F., Holland, L.Z., Laudet, V., Schubert, M., 2006. Retinoic acid signaling and the evolution of chordates. Int. J. Biol. Sci. 2, 38–47.
Martens, J.H., Rao, N.A., Stunnenberg, H.G., 2010. Genome-wide interplay of nuclear receptors with the epigenome. Biochim. Biophys. Acta. doi:10.1016/j.bbadis.2010.10.005.
Masia, S., Alvarez, S., de Lera, A.R., Barettino, D., 2007. Rapid, nongenomic actions of retinoic acid on phosphatidylinositol-3-kinase signaling pathway mediated by the retinoic acid receptor. Mol. Endocrinol. 21, 2391–2402.
McNamara, S., Wang, H., Hanna, N., Miller Jr., W.H., 2008. Topoisomerase IIbeta negatively modulates retinoic acid receptor alpha function: a novel mechanism of retinoic acid resistance. Mol. Cell. Biol. 28, 2066–2077.
Mic, F.A., Molotkov, A., Benbrook, D.M., Duester, G., 2003. Retinoid activation of retinoic acid receptor but not retinoid X receptor is sufficient to rescue lethal defect in retinoic acid synthesis. Proc. Natl. Acad. Sci. U.S.A. 100, 7135–7140.
Mollard, R., Viville, S., Ward, S.J., Decimo, D., Chambon, P., Dolle, P., 2000. Tissue-specific expression of retinoic acid receptor isoform transcripts in the mouse embryo. Mech. Dev. 94, 223–232.
Mullen, E.M., Gu, P., Cooney, A.J., 2007. Nuclear receptors in regulation of mouse ES cell pluripotency and differentiation. PPAR Res. 2007, 61563.
Niederreither, K., Dolle, P., 2006. Molecular mediators od retinoic acid signaling during development. In: Taneja, R. (Ed.), Nuclear Receptors in Development, vol. 16. Elsevier, pp. 105–143.
Niederreither, K., McCaffery, P., Drager, U.C., Chambon, P., Dolle, P., 1997. Restricted expression and retinoic acid-induced downregulation of the retinaldehyde dehydrogenase type 2 (RALDH-2) gene during mouse development. Mech. Dev. 62, 67–78.
Niederreither, K., Subbarayan, V., Dolle, P., Chambon, P., 1999. Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. Nat. Genet. 21, 444–448.
Nunez, E., Fu, X.D., Rosenfeld, M.G., 2009. Nuclear organization in the 3D space of the nucleus – cause or consequence? Curr. Opin. Genet. Dev. 19, 424–436.
Olivera-Martinez, I., Storey, K.G., 2007. Wnt signals provide a timing mechanism for the FGF-retinoid differentiation switch during vertebrate body axis extension. Development 134, 2125–2135.
Pan, J., Kao, Y.L., Joshi, S., Jeetendran, S., Dipette, D., Singh, U.S., 2005. Activation of Rac1 by phosphatidylinositol 3-kinase in vivo: role in activation of mitogen-activated protein kinase (MAPK) pathways and retinoic acid-induced neuronal differentiation of SH-SY5Y cells. J. Neurochem. 93, 571–583.
Pavri, R., Lewis, B., Kim, T.K., Dilworth, F.J., Erdjument-Bromage, H., Tempst, P., de Murcia, G., Evans, R., Chambon, P., Reinberg, D., 2005. PARP-1 determines specificity in a retinoid signaling pathway via direct modulation of mediator. Mol. Cell 18, 83–96.
Pennimpede, T., Cameron, D.A., MacLean, G.A., Li, H., Abu-Abed, S., Petkovich, M., 2010. The role of CYP26 enzymes in defining appropriate retinoic acid exposure during embryogenesis. Birth Defects Res. A Clin. Mol. Teratol. 88, 883–894.
Pennimpede, T., Cameron, D.A., Petkovich, M., 2006. Regulation of murine embryonic patterning and morphogenesis by retinoic acid signaling. In: Taneja, R. (Ed.), Nuclear Receptors in Development, vol. 16. Elsevier, pp. 65–104.
Perissi, V., Jepsen, K., Glass, C.K., Rosenfeld, M.G., 2010. Deconstructing repression: evolving models of co-repressor action. Nat. Rev. Genet. 11, 109–123.
Perissi, V., Rosenfeld, M.G., 2005. Controlling nuclear receptors: the circular logic of cofactor cycles. Nat. Rev. Mol. Cell Biol. 6, 542–554.
Perissi, V., Scafoglio, C., Zhang, J., Ohgi, K.A., Rose, D.W., Glass, C.K., Rosenfeld, M.G., 2008. TBL1 and TBLR1 phosphorylation on regulated gene promoters overcomes dual CtBP and NCoR/SMRT transcriptional repression checkpoints. Mol. Cell 29, 755–766.
Piskunov, A., Rochette-Egly, C. MSK1 and nuclear receptor signaling. In: Arthur, S., Vermeulen, L. (Eds.), MSKs. Landes Biosciences, Austin, in press.
Renaud, J.P., Moras, D., 2000. Structural studies on nuclear receptors. Cell. Mol. Life Sci. 57, 1748–1769.
Renaud, J.P., Rochel, N., Ruff, M., Vivat, V., Chambon, P., Gronemeyer, H., Moras, D., 1995. Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid. Nature 378, 681–689.
Rhodes, c., Lohnes, D., 2006. Retinoid receptors in vertebral patterning. In: Taneja, R. (Ed.), Nuclear Receptors in Development, vol. 16. Elsevier, pp. 181–210.
Ribes, V., Fraulob, V., Petkovich, M., Dolle, P., 2007. The oxidizing enzyme CYP26a1 tightly regulates the availability of retinoic acid in the gastrulating mouse embryo to ensure proper head development and vasculogenesis. Dev. Dyn. 236, 644–653.
Rochette-Egly, C., Chambon, P., 2001. F9 embryocarcinoma cells: a cell autonomous model to study the functional selectivity of RARs and RXRs in retinoid signaling. Histol. Histopathol. 16, 909–922.
Rochette-Egly, C., Germain, P., 2009. Dynamic and combinatorial control of gene expression by nuclear retinoic acid receptors. Nucl. Recept. Signal. 7, e005.
Romand, R., Niederreither, K., Abu-Abed, S., Petkovich, M., Fraulob, V., Hashino, E., Dolle, P., 2004. Complementary expression patterns of retinoid acid-synthesizing and -metabolizing enzymes in pre-natal mouse inner ear structures. Gene Expr. Patterns 4, 123–133.
Rosenfeld, M.G., Lunyak, V.V., Glass, C.K., 2006. Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. Genes Dev. 20, 1405–1428.
Samarut, E., Amal, I., Markov, G., Stote, R., Dejaegere, A., Laudet, V., Rochette-Egly, C., 2011. Evolution of nuclear retinoic acid receptors alpha (RARα) phosphorylation sites. Serine gain provides fine-tuned regulation. Mol. Biol. Evol., doi:10.1093/molbev/msrO35.
Schoenfelder, S., Clay, I., Fraser, P., 2010. The transcriptional interactome: gene expression in 3D. Curr. Opin. Genet. Dev. 20, 127–133.
Schug, T.T., Berry, D.C., Shaw, N.S., Travis, S.N., Noy, N., 2007. Opposing effects of retinoic acid on cell growth result from alternate activation of two different nuclear receptors. Cell 129, 723–733.
Shiotsugu, J., Katsuyama, Y., Arima, K., Baxter, A., Koide, T., Song, J., Chandraratna, R.A., Blumberg, B., 2004. Multiple points of interaction between retinoic acid and FGF signaling during embryonic axis formation. Development 131, 2653–2667.
Sirbu, I.O., Duester, G., 2006. Retinoic-acid signalling in node ectoderm and posterior neural plate directs left-right patterning of somitic mesoderm. Nat. Cell Biol. 8, 271–277.
Sirbu, I.O., Gresh, L., Barra, J., Duester, G., 2005. Shifting boundaries of retinoic acid activity control hindbrain segmental gene expression. Development 132, 2611–2622.
Smith, E., Shilatifard, A., 2010. The chromatin signaling pathway: diverse mechanisms of recruitment of histone-modifying enzymes and varied biological outcomes. Mol. Cell 40, 689–701.
Srinivas, H., Xia, D., Moore, N.L., Uray, I.P., Kim, H., Ma, L., Weigel, N.L., Brown, P.H., Kurie, J.M., 2006. Akt phosphorylates and suppresses the transactivation of retinoic acid receptor alpha. Biochem. J. 395, 653–662.
Stavridis, M.P., Collins, B.J., Storey, K.G., 2010. Retinoic acid orchestrates fibroblast growth factor signalling to drive embryonic stem cell differentiation. Development 137, 881–890.
Stavridis, M.P., Lunn, J.S., Collins, B.J., Storey, K.G., 2007. A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. Development 134, 2889–2894.
Stehlin-Gaon, C., Willmann, D., Zeyer, D., Sanglier, S., Van Dorsselaer, A., Renaud, J.P., Moras, D., Schule, R., 2003. All-trans retinoic acid is a ligand for the orphan nuclear receptor ROR beta. Nat. Struct. Biol. 10, 820–825.
Taneja, R., Rochette-Egly, C., Plassat, J.L., Penna, L., Gaub, M.P., Chambon, P., 1997. Phosphorylation of activation functions AF-1 and AF-2 of RAR alpha and RAR gamma is indispensable for differentiation of F9 cells upon retinoic acid and cAMP treatment. EMBO J. 16, 6452–6465.
Theodosiou, M., Laudet, V., Schubert, M., 2010. From carrot to clinic: an overview of the retinoic acid signaling pathway. Cell. Mol. Life Sci. 67, 1423–1445.
Vermeulen, L., Berghe, W.V., Beck, I.M., De Bosscher, K., Haegeman, G., 2009. The versatile role of MSKs in transcriptional regulation. Trends Biochem. Sci. 34, 311–318.
Vermot, J., Gallego Llamas, J., Fraulob, V., Niederreither, K., Chambon, P., Dolle, P., 2005. Retinoic acid controls the bilateral symmetry of somite formation in the mouse embryo. Science 308, 563–566.
Vermot, J., Pourquie, O., 2005. Retinoic acid coordinates somitogenesis and left-right patterning in vertebrate embryos. Nature 435, 215–220.
Vilhais-Neto, G.C., Maruhashi, M., Smith, K.T., Vasseur-Cognet, M., Peterson, A.S., Workman, J.L., Pourquie, O., 2010. Rere controls retinoic acid signalling and somite bilateral symmetry. Nature 463, 953–957.
White, R.J., Nie, Q., Lander, A.D., Schilling, T.F., 2007. Complex regulation of cyp26a1 creates a robust retinoic acid gradient in the zebrafish embryo. PLoS Biol. 5, e304.
White, R.J., Schilling, T.F., 2008. How degrading: Cyp26s in hindbrain development. Dev. Dyn. 237, 2775–2790.
Wilson, V., Olivera-Martinez, I., Storey, K.G., 2009. Stem cells, signals and vertebrate body axis extension. Development 136, 1591–1604.

Wolf, G., 2006. Is 9-cis-retinoic acid the endogenous ligand for the retinoic acid-X receptor? Nutr. Rev. 64, 532–538.
Xu, F., Li, K., Tian, M., Hu, P., Song, W., Chen, J., Gao, X., Zhao, Q., 2009. N-CoR is required for patterning the anterior-posterior axis of zebrafish hindbrain by actively repressing retinoid signaling. Mech. Dev. 126, 771–780.
Zechel, C., 2005. Requirement of retinoic acid receptor isotypes alpha, beta, and gamma during the initial steps of neural differentiation of PCC7 cells. Mol. Endocrinol. 19, 1629–1645.
Zechel, C., Shen, X.Q., Chambon, P., Gronemeyer, H., 1994a. Dimerization interfaces formed between the DNA binding domains determine the cooperative binding of RXR/RAR and RXR/TR heterodimers to DR5 and DR4 elements. EMBO J. 13, 1414–1424.
Zechel, C., Shen, X.Q., Chen, J.Y., Chen, Z.P., Chambon, P., Gronemeyer, H., 1994b. The dimerization interfaces formed between the DNA binding domains of RXR, RAR and TR determine the binding specificity and polarity of the full-length receptors to direct repeats. EMBO J. 13, 1425–1433.
Zhao, X., Duester, G., 2009. Effect of retinoic acid signaling on Wnt/beta-catenin and FGF signaling during body axis extension. Gene Expr. Patterns 9, 430–435.
Zhao, X., Sirbu, I.O., Mic, F.A., Molotkova, N., Molotkov, A., Kumar, S., Duester, G., 2009. Retinoic acid promotes limb induction through effects on body axis extension but is unnecessary for limb patterning. Curr. Biol. 19, 1050–1057.
Zhou, X.E., Suino-Powell, K.M., Xu, Y., Chan, C.W., Tanabe, O., Kruse, S.W., Reynolds, R., Engel, J.D., Xu, H.E., 2011. The orphan nuclear receptor TR4 is a vitamin A-activated nuclear receptor. J. Biol. Chem. 286, 2877–2885.
Zile, M.H., 2001. Function of vitamin A in vertebrate embryonic development. J. Nutr. 131, 705–708.

D. ZEBRAR : LE ROLE DE L'AR AU COURS DU DEVELOPPEMENT DU POISSON-ZEBRE

Cette revue a été soumise pour publication au journal *Biochimica et Biophysica Acta* dans le cadre d'un numéro spécial intitulé *Nuclear receptors in animal development* qui sera publié en 2014. L'objectif de cette revue est de mettre en avant les avantages du poisson-zèbre en tant que modèle de biologie du développement pour l'étude *in vivo* du rôle de RA et de ses récepteurs. De plus, elle illustre la pléiotropie de la voie de l'AR dans le développement en se basant sur des études utilisant le poisson-zèbre. De ce fait, c'est la première revue qui récapitule les principaux rôles de l'AR au cours de l'embryogénèse du poisson-zèbre et devrait être ainsi d'un fort intérêt pour la communauté.

ZebRA: an overview of retinoic acid effects during zebrafish development

Eric Samarut[1], Daniel Fraher[2], Vincent Laudet[1] and Yann Gibert[2]*

1: Institut de Génomique Fonctionnelle de Lyon; Ecole Normale Supérieure de Lyon; CNRS UMR 5242 ; Université Lyon 1; Université de Lyon; 46 allée d'Italie, 69364 Lyon Cedex 07, France;

2: Metabolic Genetic Diseases Laboratory, Metabolic Research Unit, Deakin School of Medicine, 75 Pigdons Road, Waurn Ponds VIC 3217, Australia

*: Corresponding author: E-mail: y.gibert@deakin.adu.au

Keywords: Retinoic acid, RAR, zebrafish, organogenesis, patterning

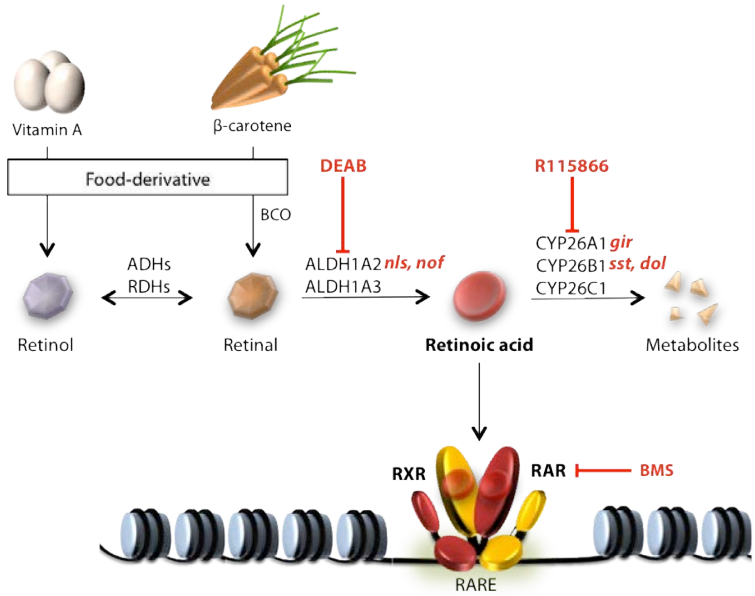
ABSTRACT

Retinoic Acid (RA), the main active vitamin A derivative is undeniably crucial for embryo development regulating cellular processes, embryo patterning and organogenesis. Many studies performed in mammalian or avian models undertook successfully the investigation of the role played by RA during embryogenesis. However, since the early 1980's, zebrafish (*Danio rerio*) has emerged as powerful developmental model. Unlike other vertebrate models, zebrafish embryogenesis is external allowing the observation of the see-through embryo from the earliest steps, but also providing an easily accessible system for pharmacological treatment or genetic approaches. Therefore, zebrafish research largely participates in deciphering the role of RA during development and this review aims at illustrating different concepts of RA signaling based on the research performed on zebrafish. Indeed, RA action relies on many cross-talks with different other signaling pathways and requires a coordinated, dynamic and fine-regulation of its level and activity in both temporal and spatial dimensions. This review also highlights major advances that have been discovered using zebrafish such as the observation for the first time of the RA gradient *in vivo*. Finally, it underlines that zebrafish is a convenient and powerful model for functional genomics.

INTRODUCTION

Retinoic acid (RA) is the active derivate of retinol, the form of vitamin A ingested from animal food. Other forms of vitamin A can be found in carotenes (alpha, beta, gamma-carotene and the xanthophyll beta-cryptoxanthin) which act as provitamin A and can be cleaved by 15-15'-dioxygenase (BCO) and therefore transformed into retinol (Figure 1). Animals, including the zebrafish, are dependant of their nutrition as source of vitamin A as it cannot be synthesized *de novo*. Importance of Vitamin A during embryonic development comes from work in avian and rodent where it has been determine that Vitamin A deficiency (VAD) syndrome lead to a wide range of abnormalities to the developing foetus. In human, VAD is still common in developing countries and manifest as xerophthalmia, keratomalacia or even complete blindness (WHO). Due to its impact on human health, numerous studies have tried to elucidate the molecular basis of Vitamin A. It was only in 1987 that two independent research group identified a receptor protein that was able to bind all-trans-retinoic acid [1, 2]. This receptor protein was then name retinoic acid receptor (RAR) and found to belong to a gene family with three members in human: RARα (NR1B1), RARβ (NR1B2) and RARγ (NR1B3).

In vertebrates, RA is produced in the body by two successive steps (Figure 1). All-trans-retinoic acid is first converted to retinladehyde by oxidative enzymes. Two main classes of enzymes are able to perform this step: alcohol dehydrogenases (ADHs) and retinol dehydrogenases (RDHs). The second limiting step in RA synthesis is the oxidation of retinaldehyde to RA which is carried out by three aldehyde dehydrogenases 1 family member A enzymes in human: Aldh1a1, Aldh1a2 and Aldh1a3 (also known as retinaldehyde dehydrogenases: Raldh1;Raldh2 and Raldh3). In zebrafish however *aldh1a1* is not found in genomic sequences [3]. As previously stated once synthesized, RA binds to and activates RARs. In absence of their natural ligand, RARs act as active repressor of gene expression [4, 5]. When RA binds to its receptor, this induces a conformational change releasing t he co-repressors and recruiting the co-activators allowing transcription of target genes. Several genes have been reported to be RA inducible and contain RA responsive elements (RAREs). Hox gene remains to date the most studied RA induced genes. Although human and mammalian genome has three RAR genes, the zebrafish genome due to the fish specific genome duplication and subsequent secondary loss contains fours RAR genes: RARα-A; RARα-B; RARγ-A and RARγ-B while RARβ has been specifically lost in the zebrafish genome [6].



RA levels in the organism are tightly regulated especially during embryogenesis with cells that require RA for their normal development and function while others must not be in contact with RA. This level of control is carried out by the cytochrome P450 26 subfamily: cyp26a1, cyp26b1 and cyp26c1. The number of cyp26 enzymes is similar in the zebrafish and in mammals. These enzymes

Figure 1: Pharmacological and genetic zebrafish toolkit to study Retinoic Acid. The main source of retinol come from dietary vitamin A intake. It is reversibly converted in retinal through ADH and RDH enzymes. Retinal can also be formed after cleavage of β-carotene by BCO enzymes. Retinoic Acid is synthetized from oxidation of retinal by ALDH1A enzymes (two enzymes in zebrafish). These enzymes can be chemically inhibited by diethylaminobenzaldehyde (DEAB) treatment. Two mutant lines for aldh1a2 have been characterized: *no-fin* (*nof*) and *neckless* (*nls*). Retinoic acid bind to its cognate nuclear receptor (RAR) to activate target gene expression. RAR activation by RA can be inhibited by subtype-specific antagonists or inverse agonists (BMS). Retinoic Acid is metabolised into more polar metabolites by CYP26 enzymes (three in zebrafish). The *cyp26a1* gene is disrupted in the *giraffe* (*gir*) mutant, and two mutants have been generated with a non-sens mutation in *cyp26b1* gene: *stocksteif* (*sst*) and *dolphin* (*dol*).

catalyse reactions that convert RA into 4-hydroxy-RA that is in turn oxidized to 4-oxo-RA (Figure 1). Therefore cyp26 enzymes function is to prevent activation of RARs by RA in specific cells. The complementary expression pattern of RA synthesizing enzymes (aldh1a) and RA degrading enzymes (cyp26) allow normal distribution of this morphogen and therefore normal development of embryos and foetuses.

Thanks to its external development, zebrafish is a convenient model to assay pharmacological treatment. By simply treating the medium in which embryos are developing with chemicals altering RA metabolism, researchers have a unique occasion to investigate the role of RA pathway during embryogenesis. Different compounds are commercially available to specifically inhibit (i) ALDH1A-mediated endogenous RA synthesis with diethylaminobenzaldehyde (DEAB), (ii) endogenous CYP26-mediated RA metabolism with R115866 inhibitor, or even (iii) RAR activation by RA thanks to specific RAR-antagonist or inverse agonist (e.g BMS493) (Figure 1). Moreover, zebrafish is a very suitable model for genetic approach and different mutant lines have already been generated with mutation in *aldh1a2* gene (*neckless*, *no-fin*), in *cyp26a1* (giraffe) or in *cyp26b1* (*stocksteif*, *dolphin*) [7-11]. As a result, both genetic and chemical tools are available for studying RA effects on the see-through and fast-developing zebrafish embryo but also in adult fish.

EARLY EMBRYOGENESIS

Brain patterning

Numerous studies have focused on the implication of RA signaling during hindbrain patterning. Hindbrain patterning is characterized by the formation of equal segments called rhombomeres (r). The number of rhombomeres varies depending on the authors between 7 and 8. Each rhombomere acquires a specific identity depending on its position along the antero-posteiror axis. Identification of RA as a major signal in rhombomere development came from analysis of VAD syndrome in avian where a lack of RA leads to the absence of rhombomere 4-8 into an enlarged rhombomere 3 [12]. In zebrafish the first data on hindbrain patterning and RA signaling came from the analysis of the *neckless* (*nls*) mutant deficient in *alhh1a2* enzyme [7], later confirmed by the identification of a second *aldh1a2* mutant named *no fin* (*nof*) [8]. However in these zebrafish mutants the severity of posterior rhombemere loss was less that in the VAD quail syndrome. Therefore it has been postulated that other source(s) of RA synthesis were present in the zebrafish or that maternal supply of RA might help early hindbrain patterning. Use of pharmacological compounds that prevent RA synthesis such as 4-Diethylaminobenzaldehyde (DEAB) or the pan-RAR inverse agonist such as BMS 493 after the maternal to zygotic transition

generated embryos with a much reduced hindbrain with the posterior limit of the hindbrain being found at the r4/r5 boundary [13]. Therefore during zebrafish embryonic development, at least one other source of RA independent of *aldh1a2* is implicated in hindbrain patterning. This other source of RA production is yet to be identified. The most obvious secondary source of RA production will lies in *aldh1a3*, however morpholino knock-down of *aldh1a3* does not seem to have any hindbrain patterning abnormality [14]. Therefore the zebrafish may possess another enzyme with retinaldehyde dehydrogenase activity or has another way to produce RA in an *aldh1a* independent manner.

Although the anterior part of the brain is qualified as being a RA free zone in zebrafish, *Aldh1a2* homozygous mutant in mouse display forebrain deficiencies such as defective growth and morphogenesis of the optic tectum [15, 16] while analysis of the *Aldh1a2*/*Aldh1a3* double homozygous mutant challenges the view of a role played by RA in forebrain formation [17]. A definitive role of RA signalling during forebrain patterning came from recent studies in zebrafish where Gongal and colleagues demonstrated that *aldh1a2* expression is controlled and maintained by the H6-homeobox (HMX) 4 [18]. Surprisingly, exogenous RA can rescue the forebrain phenotype in Hmx4 morphants while DEAB treatments do not affect forebrain patterning. Therefore the authors proposed a model in which Hmx4 regulates RA production which in turns control sonic hedgehog (shh) signalling for a normal forebrain patterning in zebrafish.

RA Gradient

RA is a morphogen molecule that diffuses from the posterior to the anterior of the embryo through a concentration gradient and specifies positional information within paraxial mesoderm and for rhombomere patterning in the hindbrain [13, 19, 20]. In the past few years, studies using zebrafish as a model provided new insights in how this RA gradient is dynamically regulated during early embryogenesis, and is RA accurately acting through a canonical morphogen mechanism. Many studies performed in mammals, birds and fish pointed out the posteriorizing effect of RA. Indeed, when applied exogenously, RA promotes the formation of posterior structures at the expanse of anterior features (e.g eyes, forebrain) [8] and conversely, depleting RA synthesis prevent the formation of posterior rhombomeres in the hindbrain [19]. This posteriorizing effect of RA is consistent with the expression of ALDH1A2 enzyme in the posterior mesoderm where RA is synthetized, therefore providing the "source" of the gradient. However, the fact that RA deficiency can be rescued by a homogeneous exposure to exogenous RA [21] means that RA gradient is not simply generated through diffusion from a synthetizing zone. In fact, many studies on zebrafish embryos revealed that RA availability is prevented in the anterior

region of the embryo through the local activity of CYP26 enzymes that degrade RA and therefore can be considered as a "sink" [21-23]. As a result, RA gradient is piloted through the dynamic expression of both synthetizing (i.e *aldh*) and degrading (i.e *cyp26*) enzymes. *Cyp26a1* is the major RA-metabolizing enzyme that is expressed during gastrulation and RA positively regulates its expression. However, this single feedback regulatory loop cannot explain why RA is not degraded posteriorly where it should promote its own degradation. Indeed, it has been shown that *cyp26a1* expression is not only regulated by RA but also by FGF pathway that indirectly inhibits the expression of *cyp26a1* by RA [23] (Figure 2). As a result, as confirmed by computational simulations, posterior FGF signaling inhibits the activation of *cyp26a1* by RA therefore allowing a robust and adaptable RA gradient from posterior to anterior [23].

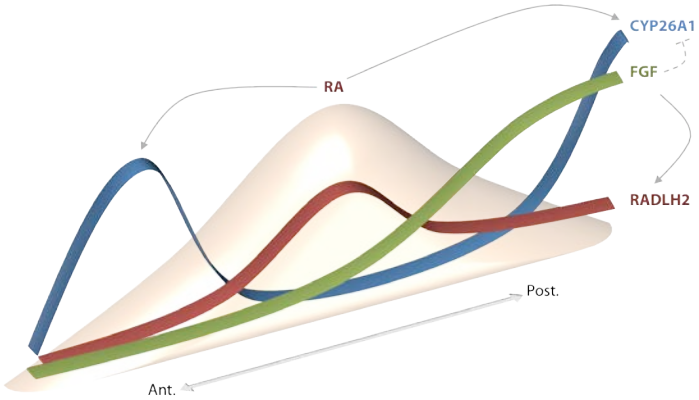


Figure 2: Molecular maintenance of a two-tailed RA gradient in the zebrafish embryo. Strong expression of *cyp26a1* in extreme anterior and posterior regions of the embryo is retro-activated by RA (arrows) and prevent its availability by degrading it. The posterior expression of FGF signaling activates *raldh2* expression for RA synthesis while inhibiting *cyp26a1* activation by RA (dotted arrow). As a result, this molecular network lead to the synthesis of RA following a two-tailed gradient with the highest RA level in the trunk region.

In order to confirm that RA is a morphogen and that it is distributed in a gradient fashion, transgenic zebrafish lines expressing YFP under the control of retinoic acid response elements (RAREs) were generated and allowed the visualizing of RA transcriptional activity *in vivo* [23, 24]. YFP expression in non-treated embryos was graded along the A/P axis, but more interestingly, YFP expression can be rescued, after inhibiting the endogenous synthesis of RA by *aldh1a2*, with RA-soaked beads. This rescue was observed in a long-range distance from the implanted beads and in a dose-dependent fashion. Furthermore, this artificial RA gradient generated from the beads is sufficient to induce different *hox* genes that are necessary for hindbrain patterning, differentially sensitive to specific RA concentration thresholds along the gradient. These data confirm that RA acts as a morphogen and acts directly on the expression of genes that define A/P patterning [23]. However, transgenic RARE-YFP line is only informative on the transcriptional activity of RA signaling but not on the rigorous RA level. However, the endogenous RA concentration gradient was recently observed *in vivo* in zebrafish thanks to the elegant design of Genetically Encoded Probes for RA (GEPRAs) [25]. Briefly, these GEPRAs are the fusion of RAR LBDs with different fluorescent proteins. The binding of RA drives drastic conformational changes

within the LBD leading to a detectable FRET (fluorescence resonance energy transfer) and allowing a live and quantitative observation of the endogenous RA level during early zebrafish embryogenesis. This technique revealed the presence of a linear "two-tailed" RA gradient, with the highest RA level found in the mid-trunk and the lowest in the head and the tail, which fits with the "source and sink" model shaped by both *aldh1a2* and *cyp26* expressions (Figure 2). Moreover, it explain why RA-deficient embryos treated with a uniform exposure to exogenous RA depict a restored RA gradient, therefore proving that the establishment of RA gradient relies on local degradation of RA. Lastly, it allowed confirming that FGF signaling (i.e *fgf8*) is required for the maintenance of RA gradient during somitogenesis through the maintenance of *aldh1a2* expression.

Altogether, these zebrafish data confirmed that RA can be considered as a true morphogen by providing long-range distance positional information through a gradient distribution along the A/P axis that mainly rely on local degradation of diffusing RA from its synthetizing source. Furthermore, this RA gradient as to be considered in tight relationship with FGF signaling that dynamically regulates it during embryogenesis through complex cross-talks therefore providing robustness and adaptable features to the gradient.

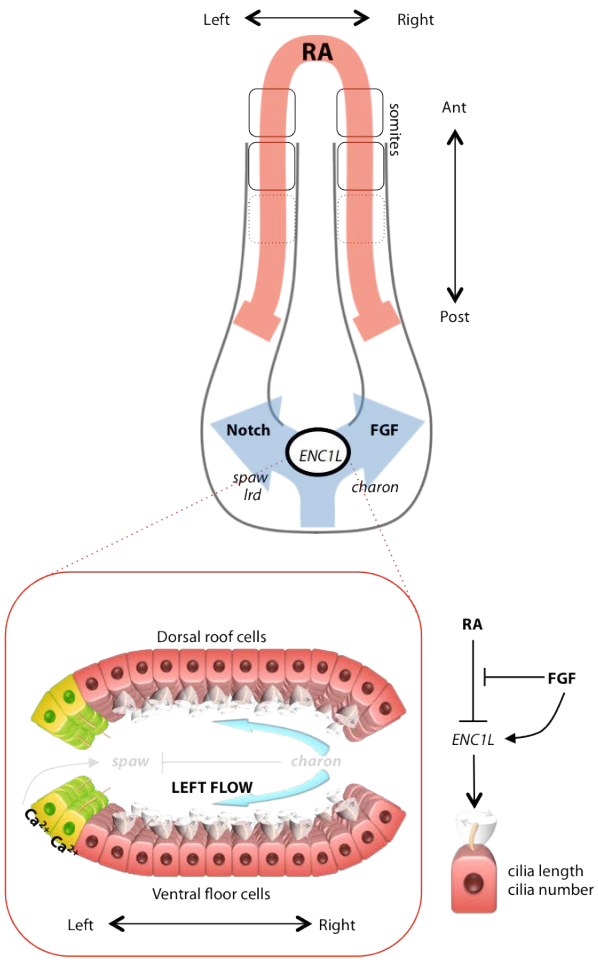
Left right asymmetry

The vertebrate body plan display a bilateral symmetry. In 2005 two papers published almost simultaneously reported that RA controls the bilateral symmetry of somite formation in mouse and chicken [26, 27]. In mouse for example *Raldh2* homozygous mutant often exhibit fewer somites on one side of the embryo (generally the right) [26]. In zebrafish a similar phenomenon has been demonstrated, using the *nls* mutant and *aldh1a2* morphants, Kawakami and colleagues demonstrated that lack of RA signaling impaired somite symmetry during zebrafish embryogenesis with a trend to increase somite number on the left side of the embryo [28]. The authors were able to link the somite asymmetry formation to a desynchronization of the molecular clock with uneven expression of *deltaC*, *her1* and *her7* in *aldh1a2* morphants. Moreover, this study showed that a lack of RA signaling expands anteriorly the expression of the FGF ligand *fgf8*, on the right side of the embryo. Therefore the following model was proposed in left-right asymmetry during somitogenesis: RA signaling produces a left-right buffer of the laterality information flow [28] (Figure 3).

During zebrafish development, RA signalling plays also a fundamental role in the control of visceral and heart laterality. RA controls laterality decision at two stages during zebrafish embryogenesis. Before the 2-somites stage (around 11 hpf) inhibition of RA signalling treatments with the pan-RAR receptor antagonist BMS453 or using the RA synthesis inhibitor DEAB,

randomized the left-right positioning of visceral organs [29]. This lack of correct left right asymmetry of visceral organ was a consequence of bilateral expression of the asymmetric genes *southpaw* and *lefty2* in the lateral plate mesoderm [29]. Analysis of the RAR revealed that knockdown of RARγ-B results in bilateral livers without impacting liver size [30]. This study therefore revealed that the primitive endoderm is able to form liver on both the left and right side of the embryo. This study also reveals that the bilaterality of liver observed in RARγ-B morphants can be rescued by overexpressing *bmp2b* indicating that RA signalling acts upstream of BMP signalling during liver positioning [30].

Depletion or RA signalling after the 2-somite stage has no effect of the laterality of visceral organ (*southpaw* and *lefty2* are expressed normally) but perturbs the expression of *bmp4* in the cardiac field resulting in a randomized cardiac looping of the heart. Analysis of Tg(cmlc2:GFP) embryos, which express GFP in both the atrium and the ventricle, depleted of RA signalling lead to a majority of embryos with heart displaying a L-Looped heart or a non-looped heart at the midline [29].



The question remains to understand how RA signalling can influence left-right asymmetry of visceral organs during zebrafish development. The answer lies within the study of the Kupffer's vesicle (KV), the assumed equivalent of the murine embryonic node implicated in left-right asymmetry. Depletion of RA signalling during early zebrafish development does not affect KV formation but increases cilia length as marked by alpha-Acetylated tubulin antibody staining [29]. As a result, cilia in the Kupffer's vesicle lose their well-organized counter-clockwise directional flow to a non-directional flow in RA

Figure 3: Multiple roles of RA in controlling asymmetric flow within the Kuppfer's vesicle (KV) and buffering assymetric signal for bilateral somite development. Within the KV, ENC1L expression is expressed and ensure correct cilia length and number in KV cells. Its expression is regulated by both RA and FGF signalling thus RA is required for correct flow within the KV. This flow is responsible of an increase in Ca^{2+} influx in left cells of the KV (green cells) and therefore promote the unilateral activation of genes (e.g *spaw*) involved in L/R asymmetry. Furthermore, anterior RA signaling from the forming somites acts as a buffer against assymetric signals (e.g Notch, FGF) from the posterior, and allow synchronized bilateral somite formation.

compromised embryos. Using an ectoderm-Neural Cortex1-like (*enc1l*) morpholino, Qian and collaborators showed that this gene was a direct downstream effector of RA signalling pathway in modulating KV ciliogenesis in zebrafish [31] (Figure 3). The length and number of cilia in cells within the KV is important to maintain a directional flow, equivalent to the nodal flow described in mammals [32, 33], and leading to the unilateral activation of genes required for asymmetrical patterning [34].

ORGANOGENESIS

Heart development

Thanks to its external development, see-through zebrafish embryos allow easy and non-invasive heart and blood cell observation during embryogenesis. Moreover, embryogenesis is not dependent of blood circulation since the embryo can receive enough oxygen from the external medium by passive diffusion. As a result and unlike mammalian models, many zebrafish mutants with strong cardiovascular defects can still survive until the end of organogenesis, therefore providing unprecedented investigations on cardiovascular developmental.

The heart is the first organ to develop and to function and its development requires coordinated and combinatorial mechanisms that rely on many developmental pathways. In zebrafish, the embryonic heart originates from cardiogenic precursors that are already determined in the late blastula and which involute in the early gastrula to then migrate to the lateral plate mesoderm (LPM) near the midbrain-hindbrain boundary [35]. RA signaling is firstly involved in the mesoderm commitment to cardiac lineage and early RA exposure of zebrafish gastrula embryos lead to early loss of all cardiac precursors [36, 37]. The bilateral populations of cardiogenic precursors from LPM are already composed of myocardial precursors (that will form the outer myocardial layer of the heart) and endocardial precursors (that will form the inner endocardial layer), both expressing specific markers. At around 19 hpf, they all migrate towards the embryo midline where they fuse to form a linear heart tube. While migrating, myocardial precursors terminate their differentiation by expressing specific genes conferring them contractile activity. The linear heart tube then undergo many morphogenetic changes that will shape it as a multi-chambered organ, the main observables structures being the ventricule and the atrium along the A/P axis [38]. The formation of cardiac chambers and of its A/P patterning requires RA signaling. Indeed, exogenous RA treatment of zebrafish embryo leads to a posteriorized heart by inhibiting the developmental of ventricular cells [37]. This is consistent with the expression of *aldh1a2* in the preatrial region of the murine heart tube [39] that suggests that endogenous RA signaling favors atrial development at the expense of ventricular development, during heart chamber formation. However, RA signaling also plays a major role before chamber formation in

the commitment of cardiogenic precursors into atrial *versus* ventricular differentiation. In the *neckless* zebrafish mutant that lacks endogenous synthesis of RA through ALDH1A2 activity, an excess of cardiomyocytes is observed [40]. Consistently, by treating zebrafish embryos with a pan-RAR antagonist (BMS 189453) at the gastrula stage, both ventricular and atrial precursor populations are more numerous. When applied at early somite-stage, BMS 189453 treatment only increases the number of atrial cells thus suggesting a role of RA signaling in restricting atrial cell number at this stage. Interestingly, this RA-dependent restriction of atrial cell fate is mediated through a non-autonomous action of *hoxb5b*. In fact, this *hox* gene is expressed in the nearby pectoral fin (forelimb) field at the early-somite stage and its expression in induced by RA. Morpholino knockdown of *hoxb5b* phenocopies the lack of RA signaling induced by BMS 189453 treatment and this phenotype can be rescued by ectopic sources of *hoxb5b* delivered by transplanted cells expressing an hyperactive *vp16-hoxb5b* [41]. More recently, FGF signaling (*fgf8*) has been confirmed as acting downstream to RA signaling in this forelimb and heart coordinated development [42] which is consistent with some studies in mice. In fact, increasing FGF signaling lead to an increased number of cardiac cells and conversely, down-regulation of *fgf8a* (that is expressed in cardiac progenitors) by morpholino injection, prevent the phenotype induced by RA deficiency. *Fgf8* is also involved in the development of cardiac precursors and is required for the proper expression of the heart precursor marker within the LPM [43].

Later in development, RA exposure rapidly induces cardio-toxicity in the zebrafish larvae such as edema or heart elongation [37]. In fact, RA-treated embryos at 72 hpf develop cardiac malformations less than 24 hours post-treatment. The precise mechanisms induced by RA that lead to such toxicity is not well understood but a micro-array assay pointed out NR2F5 (a member of the COUP-TF family) as being rapidly activated by RA in the heart. Interestingly, NR2F5 overexpression phenocopies RA-induced cardiac phenotype and conversely, morpholino injection against NR2F5 prevents RA-induced cardio-toxicity, therefore suggesting a role for NR2F5 in mediating RA-teratogenic effect on the heart [44].

Hematopoiesis

Studies mainly performed in mice, pointed out that RA signaling regulate developmental hematopoiesis, and that it tightly balances the determination of precursor cells to a specific hematopoietic lineage, but generalized conclusions about RA effects were not easily drawn. Complementary studies using zebrafish as a model helped to have a better view of RA effects on the early phases of hematopoiesis [45]. Briefly, like in all vertebrates, hematopoiesis is established in two main waves in zebrafish during embryogenesis [46]. The first wave called "primitive" hematopoiesis mainly generates red blood cell population and few myeloid cells from cells

located in the anterior and posterior LPM respectively, where they express specific markers such as *scl* or genes from the *gata* family [47]. The second wave called "definitive" hematopoiesis starts at around 36 hpf and leads to the production of all the different blood cell lineages for lifetime, from a pool of hematopoietic stem cells located in the kidney marrow. Interestingly, RA signaling has been reported as playing a major role in the early primitive hematopoiesis. In fact, zebrafish embryos treated with exogenous RA show a strong dose-dependent reduction of *gata1*-expressing blood cells at the 5-somites stage [45]. Moreover, this reduced expression is rescued by the injection of *scl* mRNA, therefore proving that the RA effects observed are specific to the hematopoietic cells and also suggesting that RA is acting upstream to *scl* for correct primitive erythropoiesis. Conversely, inhibition of RA signaling by DEAB treatment leads to an increased expression of *gata1*. However, these effects of RA are only observed after early treatment of zebrafish embryos at the gastrula stage suggesting that RA signaling is regulating primitive hematopoiesis during the early step of mesodermal patterning. Recently, RA signaling has been reported as also regulating restrictively zebrafish primitive myelopoiesis [48].

Neural crest, pharyngeal arches and tooth

Neural crest cells (NCCs) are transient migratory cells that emerge from the dorsal part of the neural tube (neural folds) during its closure (Figure 4). Of note is that in teleosts, the formation of the neural tube is processed through a mechanism called secondary neurulation that corresponds to a thickening of the ectoderm followed by a lumen cavitation [49]. Basically, at the molecular level, the induction model commonly admitted is that ectoderm can give rise to both epidermis and neural tissue following a BMP gradient. While BMP signaling promotes induction of epidermis, BMP antagonists (e.g *noggin*, *chordin*, *follistatin*) from the organizer region, potentiate the determination of neural tissue. Following this model, an intermediate level of BMP give rise to an intermediate fate therefore forming neural tube borders, roughly corresponding to the neural folds [50, 51] (Figure 4). Secondary signaling emerging from the epidermis and/or the lateral mesoderm are required to posteriorize the anterior neural plate for the induction of neural crests. The main signaling pathways involved are FGF, Wnt and BMP [50, 52, 53] but RA signaling has also been reported in *Xenopus* as being required for the induction of neural crests that express specific markers such as *Slug* [54]. Although paraxial mesoderm has been shown to be required for NC induction in *Xenopus* [55], its requirement in zebrafish is less obvious since absence of involuted mesoderm in zebrafish embryos in which *Nodal* signaling is impaired, does not impair NC induction [56]. Once induced, NCCs delaminate from the dorsal neural tube to migrate within the entire embryo through a complex journey driven by many repulsive signals [52, 57]. Trunk NCCs mainly generate pigment cells or ventral derivatives such as

enteric neurons [52] whereas craniofacial NCCs from the anterior brain-forming-region migrate into the fronto-nasal process and those from more posterior rhombomeres colonize pharyngeal arches [57]. Driven by a complex combination of environmental signals from their neighboring structures, cranial NCCs migrate following specific migratory pathways or streams depending of their origin along the A/P axis. These NCCs will mainly generate cranio-facial skeletal structures such as bone and cartilage [58]. Using the Tg[foxd3::GFP] zebrafish transgenic line that express GFP in NCCs that migrated into the pharyngeal arches (PAs), Kopinke and colleagues showed that RA signaling play a role in the correct migration of these NCCs. Indeed, blockade of RA signaling by DEAB treatment during gastrulation results in the lack of NCCs in the three posterior PAs probably due to a problem in their specification associated with posterior rhombomere patterning defect induced by RA [59]. Consistently, this impairment of posterior PA formation was also observed in *raldh2*$^{-/-}$ zebrafish mutant [8]. Furthermore, short-time exposure to exogenous RA from 24-30 hpf lead to disorganization of late migratory NCCs and post-migratory cranial NCCs in the posterior pharynx associated with abnormal gene expression (i.e *crestin* or *dlx2a*) [60]. As expected and as previously observed [61], the cartilage structures derived from these NCCs within the posterior PAs are absent or

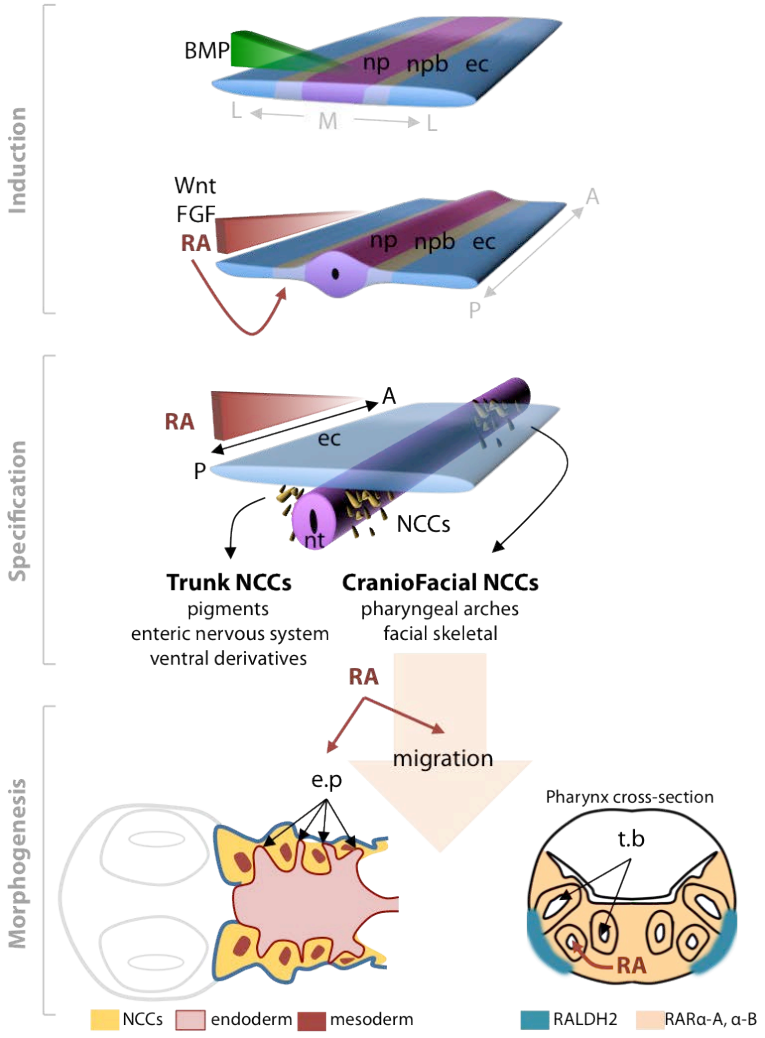


Figure 4: Retinoic Acid involvement in NCC induction, specification and morphogenesis of derived structures. BMP signaling primarily pattern the neuroectoderm in a lateromedial gradient (np: neural plate; npb: neural plate borders; ec:ectoderm). RA, FGF and Wnt signalings act as secondary signals to finally induce NCCs from neural plate borders. The dorso-ventral RA gradient is also involved in the specification of the different types of NCCs along the A/P axis and therefore in the patterning of the derived structures. In particular, RA is involved in the correct migration of cranial NCCs into pharyngeal arches (PA). RA is lastly required during PA formation for endodermal pouches (e.p) morphogenesis. Later, RA is synthetized from ventral posterior pharynx and act through RARα to induce tooth buds (t.b).

malformed whereas cartilage structures derives from the anterior PAs are developing almost normally [59]. Linville and colleagues confirmed this phenotype (i.e the lack of *dlx2* expression in the posterior PAs) and they showed that this RA effect is mediated through RARγ-A subtype since specific morpholino knockdown of this receptor subtype is sufficient to prevent the formation of posterior PAs [62]. Interestingly, if RA synthesis inhibition is achieved after gastrulation, NCCs do migrate within the five PAs but the correct separation between the most posterior PAs is not accomplished. This is explained by the fact that RA also plays a role in the formation and morphogenesis of the endodermal pouches that normally separate PAs [59, 63] (Figure 4). All these data confirmed that RA, by acting within different tight time-windows, is required all along NCCs journey, from their specification to their migration and for the correct morphogenesis and segmentation of their derived structures such as PAs.

Cranial NCCs are also giving rise to the dental mesenchyme required for tooth development [64] and recent studies using zebrafish showed the involvement of RA signaling in this process. Zebrafish, like all cypriniformes, lack oral teeth and only depict pharyngeal dentition. Impairing RA synthesis by DEAB treatment during the timeframe of induction of the first tooth-pair (named 4V1) result in the lack of tooth induction [65]. By using RAR-subtype specific antagonists, we were able to show that RA mediates its actions on tooth induction through the alpha subtypes of RARs (i.e RARα-A and RARα-B). Furthermore, *aldh1a2* expression pattern observed by in situ hybridization revealed a local and transient expression in the posterior ventral pharynx, that provide local source of RA that can diffuse toward the tooth bud mesenchyme [65]. Recently, we showed that *cyp26b1*, a RA-degrading enzyme, is expressed from 48 hpf in the tooth-forming region and play the role of a sink to deplete RA from this region (unpublished). As a result, the temporal expression of both synthetizing (i.e *aldh1a2*) and degrading (i.e *cyp26b1*) enzymes drive a precise temporal and regional window of action for RA signaling to regulate tooth development. Interestingly, we showed that small changes in the temporal expression of *cyp26b1* are correlated with different dental phenotypes within wild cypriniformes species (unpublished). Besides, exposure of zebrafish embryos to exogenous RA leads to different dental phenotypes that resemble dentition of other teleosts, such as anterior expansion of dentition [60] or change in tooth shape and number depending of the time and amount of exogenous RA dispensed (unpublished). Altogether, these results suggest that RA signaling is a good molecular candidate that is thought to have generated a huge diversity of dentition within cypriniformes and more widely in vertebrates.

Bone

As with many other organ systems, bone tissue and its development are highly conserved between zebrafish and mammals, positioning the zebrafish as a reliable model organism to study bone. As is seen in mammals, zebrafish create bone by both endochondral ossification, in which bone progenitor cells condense and differentiate into chondrocytes before differentiating into bone cells, and intramembranous ossification, where the progenitors condense and differentiate directly into bone cells [66-68]. Consistent with mammals, zebrafish also have osteoblasts (bone-building cells) and osteoclasts (bone-resorbing cells).[69, 70] In zebrafish, osteoblasts begin differentiating at 36 hpf, initially in the craniofacial region where the first bones will develop, including the opercle and the cleithrum.[71]. As these osteoblasts mature, they express many of the same genes that are seen in mammals including *runx2* and *osterix* during bone development, and *osteocalcin* and *osteonectin* as fully mature cells.[71]. As in mammals, zebrafish osteoclasts are hematopoietically derived, firstly present in zebrafish around 10-12 dpf, and are detectable by tartrate-resistant acid phosphatase staining (TRAP) [70]. Similar to other vertebrates, zebrafish osteoclasts are responsible for degrading bone matrix [70]. Taken together, these similarities show that the zebrafish is an adequate model for studying bone development.

Along with the critical roles of RA in early patterning and organogenesis of the vertebrate embryo, it is also involved in bone development. Early research identified that RA treatment caused changes in the size and shape of bones in embryonic mice and could induce osteoblast differentiation in murine cell lines [72, 73]. However, subsequent research has provided conflicting results of the effects that RA has on bone development. Some publications have claimed that RA caused an increase in osteoblast differentiation and bone development while others have stated that it caused a decrease in differentiation [74-77]. Much of this research was performed in cell lines or in mammalian models, hindering observations at stages of early bone development in an *in vivo* context. The zebrafish *stocksteif* and *dolphin* mutants both have a non-sense mutation in *cyp26b1*, causing increased endogenous RA levels. These mutants present phenotypes of hyperossification that are visible with alizarin red staining by 8 dpf [9, 11]. Fittingly, wild-type fish pharmacologically treated with RA phenocopy both of these mutations. Interestingly, it was shown that the increase in bone mineralization of the *stocksteif* mutant was not caused by an increase in osteoblast number or by mis-localization of osteoblasts, but was rather proposed to be caused by increased osteoblast activity [9]. Also, *in situ* hybridization showed that expression of *osterix* and *col10a1*, two markers of mature osteoblasts in zebrafish [71] was unaffected in the *stocksteif* mutant compared to siblings at 4 dpf. Further research using exogenous RA and DEAB treatment has looked at acute effects that RA signalling has on bone development [78]. It was shown that, in contrast to the *stosckseif* mutant, RA treatment at 48-52

hpf decreased the expression of *osterix* along with *runx2b*, an earlier marker of osteoblasts, while DEAB treatment at this time interval increased the expression of these genes. RA treatment from 48-54 hpf was also shown to decrease bone mineralization at 144 hpf. While this data may seem to conflict with other results, it was shown that RA treatment of zebrafish at 84-96 hpf, when mature osteoblasts are present, increased bone mineralization. However, it remains unclear why the *stocksteif* mutant shows normal development of osteoblasts while acute RA treatment hinders osteoblast maturation. Interestingly, these experiments reveal a possible dual role for RA in bone development. RA may play a role inhibiting osteoblast differentiation, as shown by the exogenous RA treatments early in development. Later in development, after osteoblasts have matured, RA may have an opposite role in promoting osteoblast activity to enhance bone mineralization. This proposed dual role of RA in bone development may also explain why there have been conflicting results from *in vitro* experiments.

Lipid

It has been known for over thirty years that RA signalling has an effect on fat development. It was initially shown that RA inhibited the differentiation of mammalian cells into adipocytes [79, 80]. However, subsequent *in vitro* studies have produced conflicting results, some showing a pro-adipogenic effect of RA treatment [81-83] while others claiming an inhibitory role for RA in adipocyte differentiation [84, 85]. The contradictory effects seen in these *in vitro* studies may have arisen from differences in RA concentration or from the differences in the relative timing that the treatment was given during differentiation. Mammalian studies have primarily shown that RA acts to reduce adiposity and overall body weight and can increase energy expenditure in mature animals [86-88]. It has been proposed that the reduced adiposity seen in RA treated rodents is caused by the combination of two aspects of RA signalling. Firstly, RA has been shown to increase expression of *pref-1*, *SOX9*, and *KLF2*, which are expressed in preadipocytes and may inhibit the cells from further differentiating into mature adipocytes [87]. Secondly, in already mature adipocytes, RA may signal through RARs and PPARβ and PPARδ to increase energy expenditure [87]. While these studies reveal a role for RA in reducing adiposity in rodent models, the effects of RA on adiposity in humans is less clear. There is an association between low dietary vitamin A intake and high adiposity [89]. However, humans receiving RA as a treatment for cancer have experienced increased weight gain [90]. In fact, the effect of RA on adiposity in mammals is further complicated by some rodent studies that have shown that RA and increased vitamin A intake can cause higher adiposity [91, 92].

Over the past few years, the use of zebrafish for lipid research has grown in popularity. Chemicals such as Oil-Red-O (ORO) and Nile Red have been used to stain for neutral lipids and

to fluorescently label adipocytes respectively [93]. Fluorescent lipid analogs have also been developed to study the metabolic functions of zebrafish once feeding has begun [94]. These fluorescent analogs can be used to trace lipid intake, processing, and incorporation into tissue. These visualization techniques have helped to establish the homology between zebrafish and mammalian adipocyte development and lipid metabolism [93, 95, 96]. Because of these recent developments, the zebrafish has recently become a model used to analyze the effects of RA on adipose development. In the *neckless* (*aldh1a2*) mutant, which has a mutation in the RA-synthezing enzyme retinaldehyde dehydrogenase type 2, causing a reduction of endogenous RA, there was decreased lipid deposition labelled with ORO [78]. Consistent with the mutant phenotype, DEAB treatment, inhibiting RA synthesis, also showed a reduction in ORO labelling. Not only did both the mutant and DEAB treated embryos show a decrease in ORO, but they also showed decreased expression of *C/ebpα*, an adipocyte lineage marker. Fittingly, it was also shown that exogenous RA treatment increased the expression of *C/ebpα*. These results show that RA has a pro-adipogenic role during early development in zebrafish, which is contradictory to the effects primarily seen in the adult mammal studies discussed previously. These differing effects could be due to differences in lipid metabolism during development and adulthood. Further investigation is required to elucidate a specific role for RA in lipid metabolism early and late in zebrafish life.

CONCLUSION

In this review, we aimed at illustrating the power of zebrafish as a developmental model. It is not only used for standard forward and reverse genetics, but is also easily accessible for pharmacological treatments. Moreover it is also a suitable model for the generation of transgenic lines integrating the latest genetic engineering advances. Its use for RA research is now largely widespread and it allowed shedding light on new knowledge of the role of RA mainly in early patterning and also in organogenesis. Its fast external development and its transparency are major advantages to investigate in details developing structures.

REFERENCES

[1] M. Petkovich, N.J. Brand, A. Krust, P. Chambon, A human retinoic acid receptor which belongs to the family of nuclear receptors, Nature, 330 (1987) 444-450.
[2] V. Giguere, E.S. Ong, P. Segui, R.M. Evans, Identification of a receptor for the morphogen retinoic acid, Nature, 330 (1987) 624-629.
[3] S. Pittlik, S. Domingues, A. Meyer, G. Begemann, Expression of zebrafish aldh1a3 (raldh3) and absence of aldh1a1 in teleosts, Gene expression patterns : GEP, 8 (2008) 141-147.
[4] T. Koide, M. Downes, R.A. Chandraratna, B. Blumberg, K. Umesono, Active repression of RAR signaling is required for head formation, Genes & development, 15 (2001) 2111-2121.
[5] A.D. Weston, B. Blumberg, T.M. Underhill, Active repression by unliganded retinoid receptors in development: less is sometimes more, The Journal of cell biology, 161 (2003) 223-228.

[6] S. Bertrand, B. Thisse, R. Tavares, L. Sachs, A. Chaumot, P.L. Bardet, H. Escriva, M. Duffraisse, O. Marchand, R. Safi, C. Thisse, V. Laudet, Unexpected novel relational links uncovered by extensive developmental profiling of nuclear receptor expression, PLoS genetics, 3 (2007) e188.
[7] G. Begemann, T.F. Schilling, G.J. Rauch, R. Geisler, P.W. Ingham, The zebrafish neckless mutation reveals a requirement for raldh2 in mesodermal signals that pattern the hindbrain, Development, 128 (2001) 3081-3094.
[8] H. Grandel, K. Lun, G.J. Rauch, M. Rhinn, T. Piotrowski, C. Houart, P. Sordino, A.M. Kuchler, S. Schulte-Merker, R. Geisler, N. Holder, S.W. Wilson, M. Brand, Retinoic acid signalling in the zebrafish embryo is necessary during pre-segmentation stages to pattern the anterior-posterior axis of the CNS and to induce a pectoral fin bud, Development, 129 (2002) 2851-2865.
[9] K.M. Spoorendonk, J. Peterson-Maduro, J. Renn, T. Trowe, S. Kranenbarg, C. Winkler, S. Schulte-Merker, Retinoic acid and Cyp26b1 are critical regulators of osteogenesis in the axial skeleton, Development, 135 (2008) 3765-3774.
[10] Y. Emoto, H. Wada, H. Okamoto, A. Kudo, Y. Imai, Retinoic acid-metabolizing enzyme Cyp26a1 is essential for determining territories of hindbrain and spinal cord in zebrafish, Developmental biology, 278 (2005) 415-427.
[11] K. Laue, M. Jänicke, N. Plaster, C. Sonntag, M. Hammerschmidt, Restriction of retinoic acid activity by Cyp26b1 is required for proper timing and patterning of osteogenesis during zebrafish development, Development, 135 (2008) 3775-3787.
[12] E. Gale, M. Zile, M. Maden, Hindbrain respecification in the retinoid-deficient quail, Mechanisms of development, 89 (1999) 43-54.
[13] G. Begemann, M. Marx, K. Mebus, A. Meyer, M. Bastmeyer, Beyond the neckless phenotype: influence of reduced retinoic acid signaling on motor neuron development in the zebrafish hindbrain, Developmental biology, 271 (2004) 119-129.
[14] M. Yahyavi, H. Abouzeid, G. Gawdat, A.S. de Preux, T. Xiao, T. Bardakjian, A. Schneider, A. Choi, E. Jorgenson, H. Baier, M. El Sada, D.F. Schorderet, A.M. Slavotinek, ALDH1A3 loss of function causes bilateral anophthalmia/microphthalmia and hypoplasia of the optic nerve and optic chiasm, Human molecular genetics, 22 (2013) 3250-3258.
[15] V. Ribes, Z. Wang, P. Dolle, K. Niederreither, Retinaldehyde dehydrogenase 2 (RALDH2)-mediated retinoic acid synthesis regulates early mouse embryonic forebrain development by controlling FGF and sonic hedgehog signaling, Development, 133 (2006) 351-361.
[16] F.A. Mic, A. Molotkov, N. Molotkova, G. Duester, Raldh2 expression in optic vesicle generates a retinoic acid signal needed for invagination of retina during optic cup formation, Dev Dyn, 231 (2004) 270-277.
[17] N. Molotkova, A. Molotkov, G. Duester, Role of retinoic acid during forebrain development begins late when Raldh3 generates retinoic acid in the ventral subventricular zone, Developmental biology, 303 (2007) 601-610.
[18] P.A. Gongal, L.D. March, V.L. Holly, L.M. Pillay, K.M. Berry-Wynne, H. Kagechika, A.J. Waskiewicz, Hmx4 regulates Sonic hedgehog signaling through control of retinoic acid synthesis during forebrain patterning, Developmental biology, 355 (2011) 55-64.
[19] L. Maves, C.B. Kimmel, Dynamic and sequential patterning of the zebrafish posterior hindbrain by retinoic acid, Developmental biology, 285 (2005) 593-605.
[20] A. Aulehla, O. Pourquie, Signaling gradients during paraxial mesoderm development, Cold Spring Harb Perspect Biol, 2 (2010) a000869.
[21] R.E. Hernandez, A.P. Putzke, J.P. Myers, L. Margaretha, C.B. Moens, Cyp26 enzymes generate the retinoic acid response pattern necessary for hindbrain development, Development, 134 (2007) 177-187.
[22] R.J. White, T.F. Schilling, How degrading: Cyp26s in hindbrain development, Dev Dyn, 237 (2008) 2775-2790.
[23] R.J. White, Q. Nie, A.D. Lander, T.F. Schilling, Complex regulation of cyp26a1 creates a robust retinoic acid gradient in the zebrafish embryo, PLoS biology, 5 (2007) e304.
[24] A. Perz-Edwards, N.L. Hardison, E. Linney, Retinoic acid-mediated gene expression in transgenic reporter zebrafish, Developmental biology, 229 (2001) 89-101.
[25] S. Shimozono, T. Iimura, T. Kitaguchi, S. Higashijima, A. Miyawaki, Visualization of an endogenous retinoic acid gradient across embryonic development, Nature, 496 (2013) 363-366.
[26] J. Vermot, J. Gallego Llamas, V. Fraulob, K. Niederreither, P. Chambon, P. Dolle, Retinoic acid controls the bilateral symmetry of somite formation in the mouse embryo, Science, 308 (2005) 563-566.
[27] J. Vermot, O. Pourquie, Retinoic acid coordinates somitogenesis and left-right patterning in vertebrate embryos, Nature, 435 (2005) 215-220.

[28] Y. Kawakami, A. Raya, R.M. Raya, C. Rodriguez-Esteban, J.C. Izpisua Belmonte, Retinoic acid signalling links left-right asymmetric patterning and bilaterally symmetric somitogenesis in the zebrafish embryo, Nature, 435 (2005) 165-171.
[29] S. Huang, J. Ma, X. Liu, Y. Zhang, L. Luo, Retinoic acid signaling sequentially controls visceral and heart laterality in zebrafish, The Journal of biological chemistry, 286 (2011) 28533-28543.
[30] M.K. Garnaas, C.C. Cutting, A. Meyers, P.B. Kelsey, Jr., J.M. Harris, T.E. North, W. Goessling, Rargb regulates organ laterality in a zebrafish model of right atrial isomerism, Developmental biology, 372 (2012) 178-189.
[31] M. Qian, S. Yao, L. Jing, J. He, C. Xiao, T. Zhang, W. Meng, H. Zhu, H. Xu, X. Mo, ENC1-like integrates the retinoic acid/FGF signaling pathways to modulate ciliogenesis of Kupffer's Vesicle during zebrafish embryonic development, Developmental biology, 374 (2013) 85-95.
[32] N. Hirokawa, Y. Tanaka, Y. Okada, Left-right determination: involvement of molecular motor KIF3, cilia, and nodal flow, Cold Spring Harb Perspect Biol, 1 (2009) a000802.
[33] J.J. Essner, K.J. Vogan, M.K. Wagner, C.J. Tabin, H.J. Yost, M. Brueckner, Conserved function for embryonic nodal cilia, Nature, 418 (2002) 37-38.
[34] T. Matsui, Y. Bessho, Left-right asymmetry in zebrafish, Cell Mol Life Sci, 69 (2012) 3069-3077.
[35] D.Y. Stainier, Zebrafish genetics and vertebrate heart formation, Nature reviews. Genetics, 2 (2001) 39-48.
[36] K.E. Yutzey, D. Bader, Diversification of cardiomyogenic cell lineages during early heart development, Circulation research, 77 (1995) 216-219.
[37] D.Y. Stainier, M.C. Fishman, Patterning the zebrafish heart tube: acquisition of anteroposterior polarity, Developmental biology, 153 (1992) 91-101.
[38] D. Yelon, D.Y. Stainier, Patterning during organogenesis: genetic analysis of cardiac chamber formation, Seminars in cell & developmental biology, 10 (1999) 93-98.
[39] J.B. Moss, J. Xavier-Neto, M.D. Shapiro, S.M. Nayeem, P. McCaffery, U.C. Drager, N. Rosenthal, Dynamic patterns of retinoic acid synthesis and response in the developing mammalian heart, Developmental biology, 199 (1998) 55-71.
[40] B.R. Keegan, J.L. Feldman, G. Begemann, P.W. Ingham, D. Yelon, Retinoic acid signaling restricts the cardiac progenitor pool, Science, 307 (2005) 247-249.
[41] J.S. Waxman, B.R. Keegan, R.W. Roberts, K.D. Poss, D. Yelon, Hoxb5b acts downstream of retinoic acid signaling in the forelimb field to restrict heart field potential in zebrafish, Developmental cell, 15 (2008) 923-934.
[42] M.R.J. Sorrell, J.S. Waxman, Restraint of Fgf8 signaling by retinoic acid signaling is required for proper heart and forelimb formation, Developmental biology, 358 (2011) 44-55.
[43] F. Reifers, E.C. Walsh, S. Leger, D.Y. Stainier, M. Brand, Induction and differentiation of the zebrafish heart requires fibroblast growth factor 8 (fgf8/acerebellar), Development, 127 (2000) 225-235.
[44] N. Chen, J.L. Napoli, All-trans-retinoic acid stimulates translation and induces spine formation in hippocampal neurons through a membrane-associated RARalpha, Faseb J, 22 (2008) 236-245.
[45] J.L. de Jong, A.J. Davidson, Y. Wang, J. Palis, P. Opara, E. Pugach, G.Q. Daley, L.I. Zon, Interaction of retinoic acid and scl controls primitive blood development, Blood, 116 (2010) 201-209.
[46] A.J. Davidson, L.I. Zon, The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis., Oncogene, 23 (2004) 7233-7246.
[47] N. Hsia, L.I. Zon, Transcriptional regulation of hematopoietic stem cell development in zebrafish, Experimental hematology, 33 (2005) 1007-1014.
[48] D. Liang, W. Jia, J. Li, K. Li, Q. Zhao, Retinoic acid signaling plays a restrictive role in zebrafish primitive myelopoiesis, PloS one, 7 (2012) e30865.
[49] C. Papan, J.A. Campos-Ortega, On the formation of the neural keel and neural tube in the zebrafish Danio (Brachydanio) rerio, Roux's archives of developmental biology, 203 (1993) 178-186.
[50] A.K. Knecht, M. Bronner-Fraser, Induction of the neural crest: a multigene process, Nature reviews. Genetics, 3 (2002) 453-461.
[51] V.H. Nguyen, B. Schmid, J. Trout, S.A. Connors, M. Ekker, M.C. Mullins, Ventral and lateral regions of the zebrafish gastrula, including the neural crest progenitors, are established by a bmp2/swirl pathway of genes, Developmental Biology (1998) 103-110.
[52] M.W. Klymkowsky, C.C. Rossi, K.B. Artinger, Mechanisms driving neural crest induction and migration in the zebrafish and Xenopus laevis, Cell adhesion & migration, 4 (2010) 595-608.
[53] M.J. Aybar, R. Mayor, Early induction of neural crest cells: lessons learned from frog, fish and chick, Current opinion in genetics & development, 12 (2002) 452-458.

[54] S. Villanueva, A. Glavic, P. Ruiz, R. Mayor, Posteriorization by FGF, Wnt, and Retinoic Acid Is Required for Neural Crest Induction, Developmental biology, 241 (2002) 289-301.
[55] L. Bonstein, S. Elias, D. Frank, Paraxial-fated mesoderm is required for neural crest induction in Xenopus embryos, Developmental biology, 193 (1998) 156-168.
[56] J.W. Ragland, D.W. Raible, Signals derived from the underlying mesoderm are dispensable for zebrafish neural crest induction, Developmental biology, 276 (2004) 16-30.
[57] M. Minoux, F.M. Rijli, Molecular mechanisms of cranial neural crest cell migration and patterning in craniofacial development, Development, 137 (2010) 2605-2621.
[58] F. Santagati, F.M. Rijli, Cranial neural crest and the building of the vertebrate head, Nature reviews. Neuroscience, 4 (2003) 806-818.
[59] D. Kopinke, J. Sasine, J. Swift, W.Z. Stephens, T. Piotrowski, Retinoic Acid Is Required for Endodermal Pouch Morphogenesis and Not for Pharyngeal Endoderm Specification, Developmental Dynamics, 235 (2006) 2695-2709.
[60] P. Seritrakul, E. Samarut, T.T. Lama, Y. Gibert, V. Laudet, W.R. Jackman, Retinoic acid expands the evolutionarily reduced dentition of zebrafish, FASEB J, 26 (2012) 5014-5024.
[61] D.L. Ellies, R.M. Langille, C.C. Martin, M.A. Akimenko, M. Ekker, Specific craniofacial cartilage dysmorphogenesis coincides with a loss of dlx gene expression in retinoic acid-treated zebrafish embryos, Mechanisms of development, 61 (1997) 23-36.
[62] A. Linville, K. Radtke, J.S. Waxman, D. Yelon, T.F. Schilling, Combinatorial roles for zebrafish retinoic acid receptors in the hindbrain, limbs and pharyngeal arches, Developmental biology, 325 (2009) 60-70.
[63] A. Graham, A. Smith, Patterning the pharyngeal arches, BioEssays : news and reviews in molecular, cellular and developmental biology, 23 (2001) 54-61.
[64] I. Miletich, P.T. Sharpe, Neural crest contribution to mammalian tooth formation, Birth defects research. Part C, Embryo today : reviews, 72 (2004) 200-212.
[65] Y. Gibert, L. Bernard, M. Debiais-Thibaud, F. Bourrat, J.S. Joly, K. Pottin, A. Meyer, S. Retaux, D.W. Stock, W.R. Jackman, P. Seritrakul, G. Begemann, V. Laudet, Formation of oral and pharyngeal dentition in teleosts depends on differential recruitment of retinoic acid signaling, FASEB J, 24 (2010) 3298-3309.
[66] C.C. Cubbage, P.M. Mabee, Development of the cranium and paired fins in the zebrafish Danio rerio (Ostariophysi, Cyprinidae), Journal of Morphology, 229 (1996) 121-160.
[67] H.M. Kronenberg, Developmental regulation of the growth plate, Nature, 423 (2003) 332-336.
[68] N.C. Bird, P.M. Mabee, Developmental morphology of the axial skeleton of the zebrafish, Danio rerio (Ostariophysi: Cyprinidae), Developmental Dynamics, 228 (2003) 337-357.
[69] C.L. Hammond, S. Schulte-Merker, Two populations of endochondral osteoblasts with differential sensitivity to Hedgehog signalling, Development, 136 (2009) 3991-4000.
[70] P.E. Witten, A. Hansen, B.K. Hall, Features of mono-and multinucleated bone resorbing cells of the zebrafish Danio rerio and their contribution to skeletal development, remodeling, and growth, Journal of Morphology, 250 (2001) 197-207.
[71] N. Li, K. Felber, P. Elks, P. Croucher, H.H. Roehl, Tracking gene expression during zebrafish osteoblast differentiation, Developmental Dynamics, 238 (2009) 459-466.
[72] D. Kochhar, Limb development in mouse embryos. I. Analysis of teratogenic effects of retinoic acid, Teratology, 7 (1973) 289-298.
[73] D. Gazit, R. Ebner, A. Kahn, R. Derynck, Modulation of expression and cell surface binding of members of the transforming growth factor-beta superfamily during retinoic acid-induced osteoblastic differentiation of multipotential mesenchymal cells, Molecular Endocrinology, 7 (1993) 189-198.
[74] J. Skillington, L. Choy, R. Derynck, Bone morphogenetic protein and retinoic acid signaling cooperate to induce osteoblast differentiation of preadipocytes, The Journal of cell biology, 159 (2002) 135-146.
[75] H.J.M. Song, R.P. Nacamuli, W. Xia, A.S. Bari, Y.Y. Shi, T.D. Fang, M.T. Longaker, High-dose retinoic acid modulates rat calvarial osteoblast biology, Journal of cellular physiology, 202 (2005) 255-262.
[76] K. Iba, H. Chiba, T. Yamashita, S. Ishii, N. Sawada, Phase-independent inhibition by retinoic acid of mineralization correlated with loss of tetranectin expression in a human osteoblastic cell line, Cell structure and function, 26 (2001) 227-233.
[77] A. Cohen-Tanugi, N. Forest, Retinoic acid suppresses the osteogenic differentiation capacity of murine osteoblast-like 3/A/1D-1M cell cultures, Differentiation, 63 (1998) 115-123.
[78] N. Li, R.N. Kelsh, P. Croucher, H.H. Roehl, Regulation of neural crest cell fate by the retinoic acid and Pparg signalling pathways, Development, 137 (2010) 389-394.
[79] W. Kuri-Harcuch, Differentiation of 3T3-F442A cells into adipocytes is inhibited by retinoic acid, Differentiation, 23 (1982) 164-169.

[80] T. Murray, T.R. Russell, Inhibition of adipose conversion in 3T3-L2 cells by retinoic acid, Journal of supramolecular structure, 14 (1980) 255-266.
[81] I. Safonova, C. Darimont, E.Z. Amri, P. Grimaldi, G. Ailhaud, U. Reichert, B. Shroot, Retinoids are positive effectors of adipose cell differentiation, Molecular and cellular endocrinology, 104 (1994) 201-211.
[82] K. Morikawa, H. Hanada, K. Hirota, M. Nonaka, C. Ikeda, All-trans retinoic acid displays multiple effects on the growth, lipogenesis and adipokine gene expression of AML-I preadipocyte cell line, Cell biology international, 37 (2013) 36-46.
[83] B.W. Phillips, C. Vernochet, C. Dani, Differentiation of embryonic stem cells for pharmacological studies on adipose cells, Pharmacological research, 47 (2003) 263-268.
[84] E.J. Schwarz, M.J. Reginato, D. Shao, S.L. Krakow, M.A. Lazar, Retinoic acid blocks adipogenesis by inhibiting C/EBPbeta-mediated transcription, Molecular and cellular biology, 17 (1997) 1552-1561.
[85] J.C. Xue, E.J. Schwarz, A. Chawla, M.A. Lazar, Distinct stages in adipogenesis revealed by retinoid inhibition of differentiation after induction of PPARgamma, Molecular and cellular biology, 16 (1996) 1567-1575.
[86] F. Felipe, M.L. Bonet, J. Ribot, A. Palou, Modulation of resistin expression by retinoic acid and vitamin A status, Diabetes, 53 (2004) 882-889.
[87] D.C. Berry, D. DeSantis, H. Soltanian, C.M. Croniger, N. Noy, Retinoic Acid Upregulates Preadipocyte Genes to Block Adipogenesis and Suppress Diet-Induced Obesity, Diabetes, 61 (2012) 1112-1121.
[88] J. Mercader, J. Ribot, I. Murano, F. Felipe, S. Cinti, M.L. Bonet, A. Palou, Remodeling of white adipose tissue after retinoic acid administration in mice, Endocrinology, 147 (2006) 5325-5332.
[89] M.Á. ZULET, B. Puchau, H.H. Hermsdorff, C. Navarro, J.A. Martinez, Vitamin A intake is inversely related with adiposity in healthy young adults, Journal of nutritional science and vitaminology, 54 (2008) 347-352.
[90] E. Patatanian, D. Thompson, Retinoic acid syndrome: a review, Journal of clinical pharmacy and therapeutics, 33 (2008) 331-338.
[91] L. Sedova, O. Seda, D. Krenova, V. Kren, L. Kazdova, Isotretinoin and fenofibrate induce adiposity with distinct effect on metabolic profile in a rat model of the insulin resistance syndrome, International journal of obesity, 28 (2004) 719-725.
[92] A. Redonnet, C. Ferrand, C. Bairras, P. Higueret, C. Noël-Suberville, P. Cassand, C. Atgié, Synergic effect of vitamin A and high-fat diet in adipose tissue development and nuclear receptor expression in young rats, British Journal of Nutrition, 100 (2008) 722-730.
[93] E.J. Flynn, C.M. Trent, J.F. Rawls, Ontogeny and nutritional control of adipogenesis in zebrafish (Danio rerio), Journal of lipid research, 50 (2009) 1641-1652.
[94] J.D. Carten, M.K. Bradford, S.A. Farber, Visualizing digestive organ morphology and function using differential fatty acid metabolism in live zebrafish, Developmental biology, 360 (2011) 276-285.
[95] S.A. Farber, M. Pack, S.-Y. Ho, I.D. Johnson, D.S. Wagner, R. Dosch, M.C. Mullins, H.S. Hendrickson, E.K. Hendrickson, M.E. Halpern, Genetic analysis of digestive physiology using fluorescent phospholipid reporters, Science, 292 (2001) 1385-1388.
[96] A. Schlegel, D.Y. Stainier, Microsomal triglyceride transfer protein is required for yolk lipid utilization and absorption of dietary lipids in zebrafish larvae, Biochemistry, 45 (2006) 15179-15187.

Objectifs

Au vu de l'état actuel des connaissances sur les aspects moléculaires et développementaux de la voie de l'AR détaillés précédemment, les objectifs de ma thèse étaient les suivants :

RAR et PROCESSUS TRANSCRIPTIONNEL

Mieux comprendre la dynamique moléculaire autour de l'activité transcriptionnelle des RAR. Comment est régulé, au niveau moléculaire, le recrutement des RAR au niveau de l'ADN? Comment l'activité des RAR est-elle contrôlée au cours du processus transcriptionnel ? Etudier le comportement transcriptionnel des RAR endogènes *in vivo*.

PHOSPHORYLATION DES RAR

Quels sont les changements allostériques induits par la phosphorylation des RAR? Comment ont évolué les sites de phosphorylations des RAR et quelles sont les conséquences sur la régulation de leur activité ? Quel est le rôle des phosphorylations des RAR *in vivo* ?

EVO-DEVO de la VOIE de L'AR

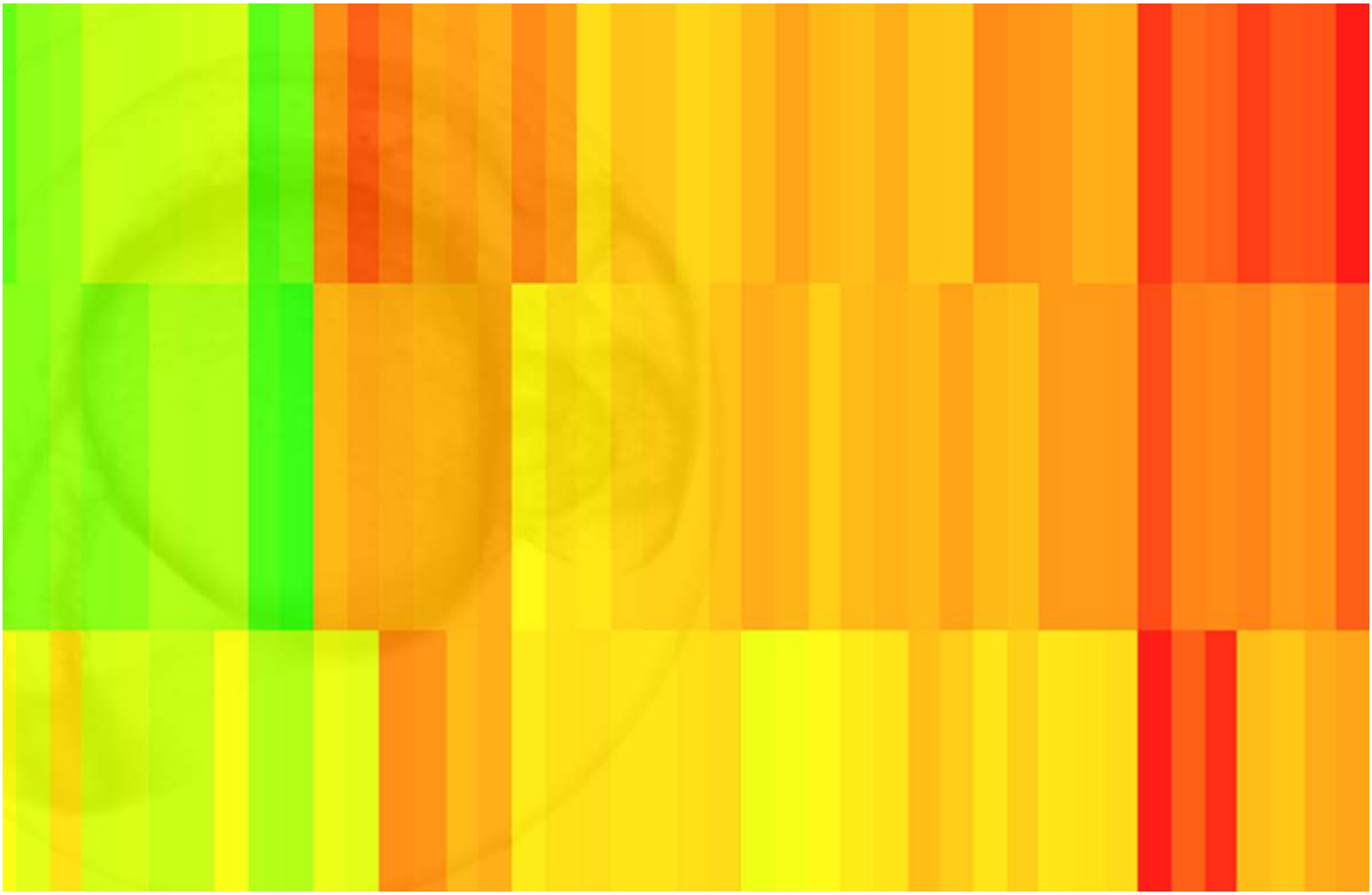
La voie de l'AR est-elle impliquée dans la genèse de nouveaux phénotypes ? Comment la voie de l'AR peut-elle être finement modulée sous la pression des contraintes développementales ?

COMMENT ?

Utiliser tout le potentiel de la simplicité de cultures de cellules et en complément, tirer profit du modèle poisson-zèbre pour étudier l'activité des RAR *in vivo* en travaillant sur les protéines endogènes. Utiliser les techniques de séquençage de dernière génération et de nouveaux outils bioinformatiques puissants. **Collaborer et échanger**. Etudier la voie de l'AR chez d'autres espèces non modèles. **Comparer et inférer**.

Chapitre#1

Activité transcriptionnelle spécifique des sous-types de RAR chez le poisson-zèbre



I. Contexte Scientifique

Alors qu'il existe trois sous-types de RAR chez les mammifères (α, β, γ), on retrouve quatre sous-types différents chez le poisson zèbre (α-A, α-B, γ-A, γ-B). En effet, le nombre de gènes paralogues codants pour les RAR a doublé chez les téléostéens à la suite d'une duplication du génome spécifique des poissons (FSGD) (Bertrand et al., 2007; Jaillon et al., 2004). Le poisson-zèbre semble avoir perdu les deux sous-types beta alors que d'autres poissons téléostéens comme le médaka ont conservé les six gènes paralogues. De nombreuses études chez la souris et chez le poisson-zèbre se sont intéressées au rôle des différents sous-types de RAR dans le développement. En effet, en invalidant un seul sous-type à la fois, par knock-out chez la souris ou par injection de morpholino (knock-down) chez le poisson-zèbre, des phénotypes développementaux spécifiques de chaque sous-type de RAR ont été mis en évidence.

Par exemple, il a été montré chez le poisson-zèbre que l'invalidation de RARα-B était suffisante pour inhiber la formation des nageoires pectorales ou encore que RARα-A est spécifiquement requis pour la détermination correcte du pharynx et du cerveau postérieur (Linville et al., 2009). De même, chez la souris, certaines anormalités comportementales ne sont observées que chez les mutants RARβ (Mark et al., 2009). Ces résultats suggèrent donc des rôles spécifiques pour chaque sous-type de RAR au cours du développement et il est concevable que cette spécificité d'action soit due à l'expression exclusive d'un sous-type donné dans un tissu en particulier. A l'inverse, certains phénotypes ne sont observés qu'après invalidation de plusieurs sous-types de RAR simultanément. C'est par exemple le cas des phénotypes typiques du syndrome lié à la déficience en vitamine A (VAD) qui ne sont retrouvés que chez des doubles mutants (KO) de RAR chez la souris (Mark et al., 2009). Ces résultats suggèrent qu'il peut exister une redondance fonctionnelle entre les différents RAR dans les tissus où plusieurs sous-types sont exprimés. Cependant, la spécificité d'action et la redondance fonctionnelle des différents sous-types de RAR n'ont jamais été étudiées au niveau transcriptionnel. Une étude récente sur le récepteur aux hormones thyroïdiennes (TR) a montré que le répertoire de gènes activés par les deux sous-types TRα et TRβ ainsi que leurs cistromes respectifs sont différents dans des cellules neurales (Chatonnet et al., 2013). Ainsi, cela suggère qu'il existe une spécificité d'action des différents TR au niveau transcriptionnel mais ces travaux reposent sur l'analyse de protéines surexprimées dans un

contexte cellulaire et ne reflètent pas nécessairement le comportement des protéines endogènes.

Dans notre étude, nous avons décidé d'étudier l'activité transcriptionnelle des différents sous-types de RAR endogènes *in vivo* chez l'embryon précoce de poisson-zèbre. Pour ce faire, nous avons combiné l'utilisation de morpholinos pour invalider différentiellement les sous-types de RAR dans l'embryon et l'analyse du transcriptome des différents morphants par séquençage haut-débit.

II. Principaux résultats

- **Description de différents isoformes pour chaque sous-type de RAR de poisson-zèbre**. Toutes les études sur les RAR de poisson-zèbre réalisées jusqu'à présent se concentrent sur un unique variant bien que deux isoformes soient prédites dans la base de données génomiques Ensembl. Nous avons validé l'existence de deux isoformes pour chacun des sous-types de RAR de poisson-zèbre et avons généré de nouveaux morpholinos pour leur invalidation.
- **Etablissement d'une liste de gènes répondants à l'AR dans l'embryon précoce de poisson-zèbre**. Nous avons établi, pour la première fois chez le poisson-zèbre, une liste de gènes répondant à l'AR chez l'embryon précoce (gastrula). Nos données constituent la première analyse transcriptomique par séquençage haut débit après traitement à l'AR chez le poisson-zèbre. Parmi ces gènes, certains gènes cibles canoniques comme *raldh2*, *cyp26a1* ou certains gènes *hox* sont retrouvés et valident notre approche.
- **Activité transcriptionnelle spécifique des sous-types de RAR**. Nous avons généré des morphants « alpha » dans lesquels les sous-types alpha de RAR ont été invalidés et des morphants « gamma » où les sous-types gamma sont invalidés. Nous avons également produit des morphants totaux où tous les RAR sont invalidés. Par comparaison de l'expression des gènes répondant à l'AR chez les différents morphants, nous avons mis en évidence qu'il existe des répertoires de gènes spécifiquement régulés par les sous-types alpha et d'autres par les sous-types gamma. De plus, certains gènes nécessitent les quatre sous-types pour être régulés par l'AR et d'autres semblent être indifféremment régulés par n'importe quels sous-types. En

conclusion, nos résultats suggèrent qu'il existe une activité transcriptionnelle spécifique des différents RAR de poisson-zèbre *in vivo* dans l'embryon précoce.

- **Différents niveaux de robustesse des gènes répondant à l'AR**. De manière inattendue, deux gènes de la voie de l'AR (*cyp26a1* et *raraa*) ne semblent pas affectés par le knock-down des RAR et leur régulation par l'AR résiste même chez les morphants « totaux ». Cela suggère qu'il existe un haut niveau de robustesse de la régulation de ces gènes par l'AR ce qui pourrait participer au maintien d'une voie de signalisation intègre et résistante.

III. Résultats - Manuscrit en préparation

Ce manuscrit présente l'ensemble de nos résultats à ce sujet et a été soumis pour publication dans le journal *Molecular Endocrinology*. Pour en faciliter la lecture, les figures principales et supplémentaires sont implémentées dans le texte et les tableaux supplémentaires sont présentés à la fin.

FULL TITLE: Retinoic Acid Receptor subtype-specific transcriptotypes in the early zebrafish embryo

ABBREVIATED TITLE: RAR-subtype specific transcriptotypes

Eric Samarut[1,2]*, Cyril Gaudin[1]*, Sandrine Hughes[1], Benjamin Gillet[1], Simon de Bernard[3], Pierre-Emmanuel Jouve[3], Laurent Buffat[3], Alexis Allot[2], Odile Lecompte[2], Liubov Berekelya[1], Cécile Rochette-Egly[2], Vincent Laudet[1]

[1] Institut de Génomique Fonctionnelle de Lyon, Université de Lyon, Université Lyon 1, Centre National de la Recherche Scientifique (CNRS), Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69364 Lyon CEDEX 07, France.
[2] IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire), Institut National de la Sante et de la Recherche Medicale (INSERM), U596, Centre National de la Recherche Scientifique (CNRS), UMR7104, Université de Strasbourg, 1 rue Laurent Fries, BP 10142, 67404 Illkirch Cedex, France.
[3] AltraBio SAS, Lyon, France.

* authors contributed equally to this work.

CONTACT: Vincent Laudet, IGFL-ENS de Lyon, 32-34 avenue Tony Garnier, 69007 LYON; +334 26 73 13 76; vincent.laudet@ens-lyon.fr.

KEY WORDS: Retinoic acid, Retinoic Acid Receptor, Transcriptome, Zebrafish
DISCLOSURE STATEMENT: The authors have nothing to disclose

ABSTRACT

Retinoic acid (RA) controls many aspect of embryonic development by binding to specific receptors (RARs) that regulate complex transcriptional network. Three different subtypes of RARs are present in vertebrates and play both common and specific roles in transducing RA signaling. Specific activity of each receptor subtype can be correlated with its exclusive expression pattern, and shared activities between different subtypes are generally assimilated to functional redundancy. However, the question remains if some subtype-specific activity still exists in regions or organs co-expressing multiple RAR subtypes. We tackled this issue at the transcriptional level using early zebrafish embryo as a model. Using morpholino knockdown, we specifically invalidated the zebrafish endogenous RAR subtypes in an *in vivo* context. After building up a list of RA-responsive genes in the zebrafish gastrula through a whole-transcriptome analysis, we compared this panel of genes with those that still respond to RA in embryos lacking one or another RAR subtype. Our work reveal that RAR subtypes do not have fully redundant functions at the transcriptional level but can transduce RA signal in a subtype-specific fashion. As a result, we define RAR subtype-specific transcriptotypes that correspond to different repertoires of genes activated by different RAR subtypes. Finally, we found genes of the RA pathway (*cyp26a1*, *raraa*) whose regulation by RA is highly robust and can even resist to the knockdown of all RARs. This suggests that RA-responsive genes are differentially sensitive to alterations in the RA pathway and that in particular *cyp26a1* and *raraa* are under a high pressure to maintain signaling integrity.

INTRODUCTION

Retinoic acid (RA) is the main active metabolite of vitamin A and has high pleiotropic effects being involved in major cellular processes such as proliferation and differentiation (1-3). It acts during development as a crucial morphogen for axes patterning and organogenesis and regulates adult homeostasis (4-6). Its wide action relies on the fact that it controls a complex gene regulatory network and connect to several other important signaling pathways (e.g. FGF, Hox etc.) controlling physiology and embryonic development. At the molecular level, RA acts through binding to nuclear retinoic acid receptors (RARs) that harbor a modular structure encompassing two main structured domains, a DNA Binding Domain (DBD) and a Ligand Binding Domain (LBD), and act as ligand-dependent transcription factors (7-10). Basically, RARs work as heterodimers with retinoid X receptors (RXRs) another nuclear receptor, following a classical ON/OFF binary model of action (11). In the absence of ligand (apo state), RAR/RXR heterodimer is thought to be bound to specific sequences located in the regulatory sequences of target genes, named RAREs for retinoic acid response elements, where it interacts with large

corepressor protein complexes that keep the chromatin in a silent state, thus preventing transcription (OFF) (12). Upon binding of RA (holo state), drastic changes in the conformation of the RAR LBD allow the dissociation of corepressors and favor the interaction with large coactivator complexes endowing enzymatic activities that are able to modify the surrounding chromatin to an active state (ON) (13,14). As a result, RA initiates the sequential and coordinated recruitment of the transcriptional machinery to activate transcription of the associated genes.

There are three RAR subtypes in mammals, α (NR1B1), β (NR1B2) and γ (NR1B3) that are encoded by distinct genes (15). The specific functions of each of them have been deeply investigated in mutant mice lacking RARα, RARβ and/or RARγ (16,17). Although RAR single KO mutants displayed congenital and post-natal abnormalities, they were viable and did not recapitulate the retinoic acid-deprived phenotypes defining the Vitamin A Deficiency (VAD) syndromes (18,19). Recapitulation of VAD phenotypes was only observed in double KO mutants in which at least two retinoid receptors were invalidated leading to the conclusion that RARs are, at least partially, functionally redundant.

As a result of a third-round of whole genome duplication (3R-WGD) in the teleost lineage, teleost fishes doubled their *rar* genes. However, in the zebrafish genome only four *rar* genes are present, *raraa*, *rarab*, *rarga*, *rargb*, because of the secondary loss of the beta subtypes, specifically in zebrafish (21,22,28). Knockdown of RARs using morpholino-oligonucleotides in zebrafish also revealed some shared and specific functions for each RAR subtype. Indeed, RARα-B is necessary for pectoral fin development whereas RARγ-A and RARγ-B are required for correct pharyngeal arches patterning (20). During zebrafish embryogenesis, RARs depict both exclusive and overlapping expression patterns (20-22). As an example, between 35 and 48 hours post-fertilization (hpf), the retina only express RARα-B subtype, the craniofacial mesoderm specifically express RARγ-A subtype while all four RARs are expressed in branchial arches. The specific function of one RAR subtype can therefore be correlated to its particular expression pattern. However, in regions in which different RAR subtypes are coexpressed, it has never been addressed if RAR function is fully redundant or if each of them have subtype-specific intrinsic properties that drive specific responses. For example, some embryonic regions such as the pectoral fin bud express several RAR subtypes (i.e RARα-B, RARγ-A, RARγ-B), but only one is required for the development of the organ suggesting that, even expressed in the same region, each RAR subtype could have specific functions.

Shedding light on such functional subtype-specificity would be of crucial interest for the understanding of the mechanisms controlling RAR activity *in vivo*. In this paper, we tackled this issue *in vivo* at the transcriptomic level using knockdown strategy in the early zebrafish embryo. Zebrafish gastrula presents many advantages as being an easily accessible *in vivo* model, but

with no highly differentiated structures or tissue. Moreover, it allows easy knockdown of a protein of interest by morpholino injection in an *in vivo* context, targeting endogenous proteins. Here we show that the different RAR subtypes regulate different repertoires of genes, providing evidence for subtype-specific transcriptional activity of RARs *in vivo*. Furthermore, we observed that the regulation of some RA-target genes is highly robust and could participate in the integrity maintenance of RA pathway *in vivo*.

MATERIAL AND METHODS

Fish stocks

AB-TU zebrafish strains were reared at 28.5°C and kept under a 10-hr dark/14-hr light cycle and staged as described (23).

Treatment of zebrafish embryos

All-trans retinoic acid (ATRA, Sigma) was diluted in 100% ethanol at a stock concentration of 10^{-2}M. Wild type embryos were treated from the stock solution to a final concentration of 10^{-7}M, diluted in embryo medium. Negative control embryos were treated with 100% ethanol as vehicle. The treatment occurred from the sphere stage (4hpf) until the 75%-epiboly stage (8hpf). Embryos were fixed in RNAlater (Ambion) and kept overnight at 4°C. They were processed the day after or frozen at -20°C.

Morpholino injection

Morpholino oligonucleotides (MO) have been purchased from GeneTools and resuspended in ultra-pure water at 2mM stock solution. 3.5ng of each translation-blocking MO were injected into one-cell stage embryos using a FemtoJet® Eppendorf micro-injector. Sequences of morpholinos targeting zebrafish RARs are provided in Figure 2C. Alpha morphants used for transcriptome analysis are injected with ATG-blocker morpholinos against raraa_001, raraa_201, rarab_001 and rarab_201, while gamma morphants are injected with ATG-blocker morpholinos against rarga_001, rarga_201, rargb_001, rargb_201. Full morphants are injected with a combination of ATG-blocker morpholinos against all RARs. Morpholinos against other nuclear receptors are listed here: PPARβa: AGCTGCCGCCGCTGCCCTCATCACT ; PPARβb: GGTTCTC CGGTCTGACTGCTAAATC ; NR2F1A : AGACGCTAACTACCATTGCCATATC ; NR2F1B : CATGGCTGGAATCTTTT ACGCGAAT ; NR2F2: AGCCTCTCCACACTACCA TTGCCAT; NR2F5: CACTGATTTACTACCAT TGCCATGC ; RARα-A_Ex2-Int2: TAAATCTAGTGTCTTACCTTGCAGC; RARα-B_Ex2-Int2: AGTTGGAAGCGTGGCCTT ACCTTAC. Splice-blocker morpholino efficiency was checked by RT-PCR from total RNA extracts (primer sequences available upon request).

Total RNA extraction and SOLiD library preparation

Two independent clutches were injected corresponding to experimental duplicates. Total

RNA from RNAlater-fixed embryos were extracted using RNAsolv reagent (Omega Biotek) following manufacturer standard protocol. About 10µg of total RNA (RIN>9) per samples (30-40 embryos) were purified onto Dynabeads® Oligo(dT)$_{25}$ (Life Technologies) yielding in between 100-500ng of polyA-purified RNA. Multiplexed-libraries were prepared following the SOLiD® total RNA-Seq kit protocol provided by Life Technologies, including a conversion step to adapt to our 5500Wildfire SOLiD sequencer. RNA quantitation and quality assessment was performed on a 2200 TapeStation (Agilent Technologies).

Sequencing and Data Analysis

Sequencing has been performed on a SOLiD 5500 sequencer upgraded with the Wildfire technology. Pre-analysis of the raw data and mapping have been achieved with Lifescope 2_5.1. Ensembl zebrafish genome Zv9 release 71 from April 2013 has been used as reference genome. Only genes with more than one count per million in at least two samples were kept for the statistical analysis. Raw library size normalization factors were computed using the TMM method (24) in the R package edgeR. Function voom from the R package limma (25) was then applied to the count data to yield log2 counts per million and associated observational-level weights. A linear model was fit to each gene by using function lmFit from limma to compute a clutch corrected average expression level for each treatment condition. Statistical contrasts where then computed for comparisons of interest. The empirical Bayes method was used to compute moderated p-values that were then corrected for multiple comparisons using the Benjamini and Hochberg's false discovery rate (FDR) controlling procedure (26). The Hierarchical Agglomerative Clusterings (HAC) of genes (associated with given heatmaps) were performed with the following parameters : (i) metric: euclidean distances between genes (genes are represented by their log(fold-changes) values), (ii) linkage criterion: complete linkage .

In situ hybridization and probe cloning

Specific probes for each RAR isoforms (two isoforms for each four zebrafish RAR subtypes) corresponding to their 5'UTR and first exon (sequences upon request) were cloned within the pCS2+ vector using TOPO TA cloning kit (Invitrogen). After sequencing, antisens probes were in vitro transcribed using SP6 or T7 RNA polymerases. Whole-mount in situ hybridization of zebrafish embryos was performed as described by Thisse et al., 2008 (27). Stained embryos were kept in 80% glycerol and photographed.

RT-qPCR

qPCR primers for each RAR isoforms were designed using the Universal Probe Library tool from Roche. Reverse Transcription was performed from 1µg total RNA using a mix of oligo(dT)$_{15}$ and random primers (Roche) and the Transcriptor Reverse Transcriptase (Roche). Quantitative PCR was performed on 3µl of 1:10-diluted cDNA using QuantiTect SYBR green (Qiagen) on a 96-

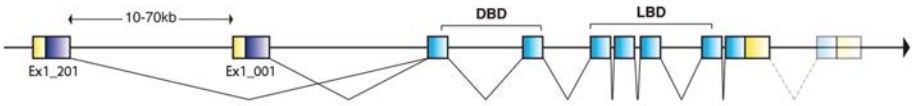
well plate on a MX3000P Stratagen system. *Polr2d* gene was used as a referenced gene for quantification.

RESULTS

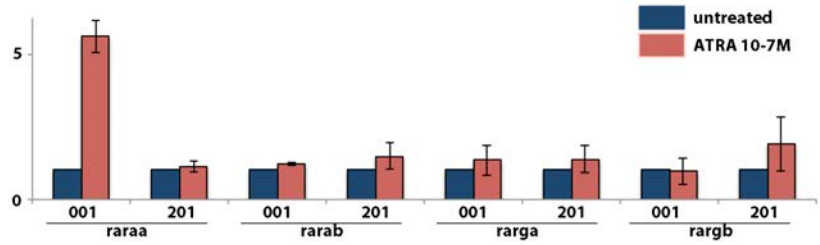
Zebrafish RAR genes encode two isoforms, including raraa.

Although different isoforms (generated by alternative promoter usage and/or alternative splicing) for each subtype are known in mouse (29,30), the existence of such isoforms for all zebrafish RARs has never been described. The Ensembl database predicts two variants for each subtype (except for *raraa*), all following the same splicing pattern (Figure 1A) which consists of the alternative use of the first exon (31). We first checked for their existence and validated their expression in early zebrafish embryo (8 hpf) by RT-qPCR and whole mount *in situ* hybridization (Figure 1B, Supplemental Figure 1). Moreover, using transcriptomic data from our laboratory, we described a second isoform for *raraa*, which follow the same genomic organization as for the other *rar* genes (Supplemental Figure 2). We cloned the full-length sequence of this newly identified isoform from total RNA extracts, and submitted the sequence to Genbank (accession# KF148618). As a result, we confirmed the existence of two isoforms for each zebrafish RAR subtypes (named *_001* and *_201* following the Ensembl nomenclature), which differ in their first exon encoding for the N-terminal domain of the receptor (Figure 1C).

A



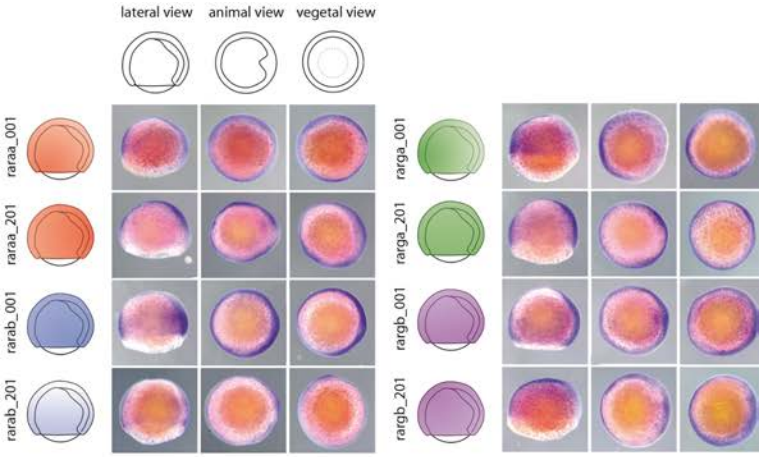
B



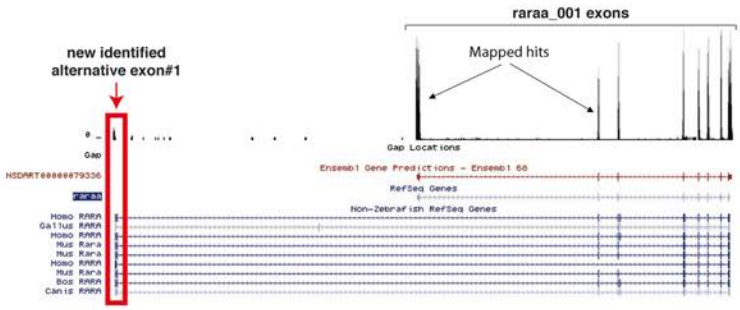
C

gene	chr	splice variant	alternative exon	ENSEMBL ID	NCBI ID	protein size (aa)	ATG MO sequence
raraa	chr 12	raraa_001	exon 1	ENSDART00000079336	NM_131406.2	444	GGTTCACATCCACACTCTCATACAT
		raraa_201	exon 1	N/A	KF148618	447	CACCTCTACAGTTCGCCCACCGTCA
rarab	chr 3	rarab_001	exon 1	ENSDART00000044677	NM_131399.1	458	CCACAACGTCCACGCTCTCGTACAT
		rarab_201	exon 1	ENSDART00000126022	N/A	460	GTTTCCAGACATGGCACGCTCAGAC
rarga	chr 23	rarga_001	exon 1	ENSDART00000049551	NM_131339.1	499	CCAGAGCCTCCATACAGTCGAACAT
		rarga_201	exon 1	ENSDART00000103158	N/A	507	GGCCAGTTGTGTGCTTCAACAGCCT
rargb	chr 11	rargb_001	exon 1 and 8	N/A	NM_001083310.1	489	CGTCTCTCATACCAAGTGGGCTGTT
		rargb_201	exon 1 and 8	ENSDART00000052546	N/A	482	GATGATGATGGCCCTTCCCCAGCGT

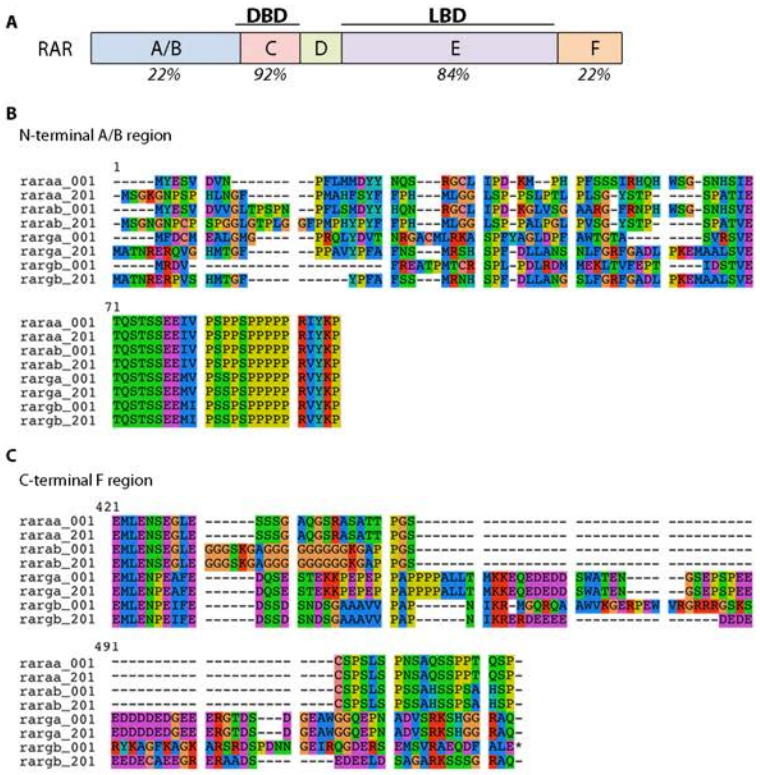
Figure 1: zebrafish RAR subtypes and isoforms. A. Splicing pattern of zebrafish RARs. Each rar subtype gene generate at least two different isoforms differing in their first exon encoding the N-term part of the receptor by alternative promoter usage. The more distal alternative exon 1 belongs to transcripts _201 and the more proximal to transcripts _001 (Ensembl nomenclature. The DBD of the receptor is encoded by exons 2 and 3, and the LBD is encoded by the 4th to 7th exon. Note that for *rargb* gene only, the deposited sequences show the last 8th exon also alternatively spliced between _001 and _201 isoforms leading to different C-terminal F domain at the protein level **B.** RT-qPCR detection of each RAR isoforms from whole mRNA extracts of untreated 8 hpf embryos (blue) or after treatment with ATRA at 10-7M from 4 to 8 hpf (red). The error bars correspond to standard deviation between at least two replicates from independent clutch. **C.** Table recapitulating for each RAR subtype, its bearing chromosome, the name and accession number of each splice variants, the corresponding size of the protein product, and the ATG-blocker morpholino sequence used for knockdown.



Supplemental Figure 1: Expression pattern of the eight zebrafish RARs and their isoforms in the zebrafish gastrula. All zebrafish RARs are expressed at 8hpf in the gastrula. For each probe, a lateral, an animal and a vegetal view is displayed. The overall expression pattern for each RAR is schematized with a color gradient on a lateral view. Briefly, all RARs are ubiquitously expressed in the gastrula with some weak gradients of expression observed: a minor dorso-ventral gradient of expression for *raraa_201*, a slight vegeto-animal gradient of *raraa_001* and *rarab_201* expressions and a modest ventro-dorsal gradient of *rarga_001* expression.



Supplemental Figure 2: Identification of a second isoform for *raraa*. Genome browser snapshot showing the hits from a RNA-seq assay, mapping to the exons of *raraa_001* isoform. Some hits mapped against a genomic region that align with exon 1 of *rara* gene in other species (red square). By RT-PCR on total zebrafish RNA extracts, we cloned the full-length newly identified raraa isoform *(raraa_201)* and submit the sequence GenBank (accession number: KF148618.1).



Supplemental Figure 3: Modular structures of RARs and sequence alignments of all zebrafish isoforms. A. RARs harbor a modular structure composed of 5 regions. The region C encompasses the DNA binding domain (DBD) and the region E corresponds to the Ligand Binding Domain (LBD). The region A/B corresponds to the unstructured N-terminal domain (NTD). Percentage of protein sequence identity between all eight RAR isoforms is provided. Sequence alignment is detailed for the A/B region (**B**) and the C-terminal F region (**C**).

We checked their expression pattern in the early zebrafish embryo and we found that all RAR isoforms were expressed almost ubiquitously at 8 hpf (Supplemental Figure 1). We also checked by RT-qPCR if some of the *rar* isoforms are responsive to RA in the early zebrafish gastrula (Figure 1B) and found that *raraa_001* is activated by ATRA exposure from 4 to 8 hpf, which is consistent

with previous observations (20). In our conditions at the early stage examined the other isoforms are virtually insensitive to RA. The alignment of the eight RAR protein sequences is provided in Supplemental Figure 3.

Designing new morpholinos and knockdown strategy

Previous studies describing the effects of RAR knockdown in zebrafish have used ATG blocker morpholinos targeting only one isoform (*_001*s) of each subtype (20). In order to perform a knockdown of all RAR subtypes, we designed new ATG blocker morpholinos (that avoid mRNA translation) specific of the *_201* isoforms (Figure 1C), and splice-blocking morpholinos targeting a common exon (that disrupt correct mRNA splicing). In order to test if some RA-responsive genes in the early zebrafish embryo are regulated in a RAR subtype-specific fashion, we co-injected morpholinos to knockdown (i) both α-A and α-B receptors, (ii) both γ-A and γ-B receptors and (iii) all four subtypes, in one-cell stage embryos (Figure 2A). We treated control (i.e uninjected) and injected embryos with all-trans retinoic acid (ATRA) at 10^{-7}M from 4 hpf (sphere stage) to 8 hpf (75% epiboly stage) and fixed them for RNA extraction. The knockdown strategy was tested by RT-qPCR on canonical target genes (e.g *hox* genes) to determine the amount of morpholino injection and to compare the effects of splice-blocker morpholinos *versus* ATG-blocker

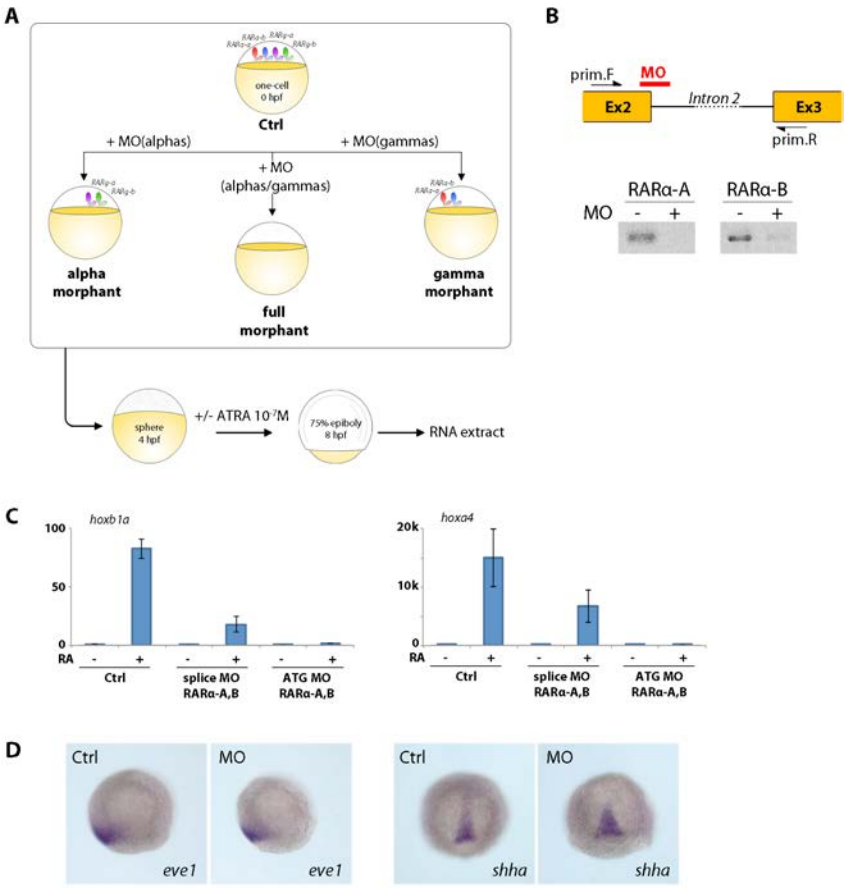


Figure 2: RAR knockdown strategy. A. Differential knockdown induced by morpholino injection. Combination of morpholinos are injected at the one-cell stage to invalidate both α-A and α-B subtypes (alpha morphants), both γ-A and γ-B subtypes (gamma morphants) or all four RARs (full morphants). **B.** For RARα-A and RARα-B, splicing blocker morpholinos spanning the junction between exon2 and intron2 were designed and their efficiency was checked by RT-PCR using primers in exon2 and exon3. The amplicon was barely detectable after morpholino injection validating the impaired splicing of the mRNA targeted. **C.** Comparison of the effects of splice-blocker *versus* ATG-blocker morpholinos targeting RARα-A and RARα-B on the expression of *hoxb1a* and *hoxa4* by RT-qPCR. The error bars correspond to standard deviation between at least two replicates from independent clutch. **D.** Whole mount *in situ* hybridization showing the expression of *eve1* and *shha* in uninjected embryos (Ctrl) or injected with ATG-blocker morpholinos against RARα-A and RARα-B (MO).

morpholinos (Figure 2B-C). We observed that splice-blocker morpholinos are less efficient than ATG morpholinos for the abolition of RA-induced target gene expression in the early embryo (Figure 2C). This is consistent with the fact that maternal RNA, that are not targeted by splice-blocker morpholinos, might still be present in the early blastula, still allowing a slight activation of target gene expression (32). We also checked in these morphants, the expression of ventral (*eve1*) and dorsal (*shha*) markers of the zebrafish gastrula (33). No change in the expression pattern of these markers were observed (Figure 2D) confirming that the gastrula patterning is not affected at this early stage by the knockdown of RAR. As a result, the changes in gene expression observed cannot be assimilated to changes of developmental territories induced by the lack of RARs but rather to transcriptional differences. Once all the experimental parameters were set-up, we widen our analysis to the whole transcriptome. We injected two independent spawning with the various morpholino combinations and prepared libraries from extracted RNA for transcriptome analysis using SOLiD next generation sequencer.

RA-responsive genes in the zebrafish gastrula

According to the results we obtained from the whole transcriptome analysis, we first established a list of genes responding to 4-hours ATRA exposure in the zebrafish embryo from 4 to 8 hpf (Figure 3A, Supplemental Table1). This list contains 311 genes differentially expressed, comparing pooled control (EtOH treated) and RA-treated embryos, selected on their p-value (adjusted p-value < 0,05) and their fold activation ($|logFC| > 1$). Among these genes, 70% were upregulated upon RA-treatment (i.e 219 genes) and 30% (i.e. 92) had their overall expression down-regulated. As expected, this list contained classical RA target genes that validate our approach. Among those genes, we found genes implicated in the RA metabolism pathway (e.g *cyp26a/b1*, *dhrs3a/b*, *aldh1a2),* genes from the WNT pathway, Hox genes and other homeobox transcription factors (e.g *sox*, *fox, meis*) (34)(Feng et al., 2010; Oliveira et al., 2013). Furthermore, we also found some new RA targets belonging to nuclear receptors (estrogen-related receptors: esrra, esrrgb), membrane receptors (ephrin receptor 7, retinal G protein, serotonin receptor) and genes of the extracellular matrix (collagen type XIV, type XXVII) (Figure 3A and Supplemental Table 1). This list of RA-responsive genes is of prime interest for a wide scientific community since no study has ever assayed RA-transcriptional response *in vivo,* in the early zebrafish embryo. In the present work, this list is used as a reference for further comparison with RAR-subtype specific morphants.

RAR subtype-specific basal regulation of RA-target genes.

We first observed the consequences of our RAR-subtype specific knockdown on the basal

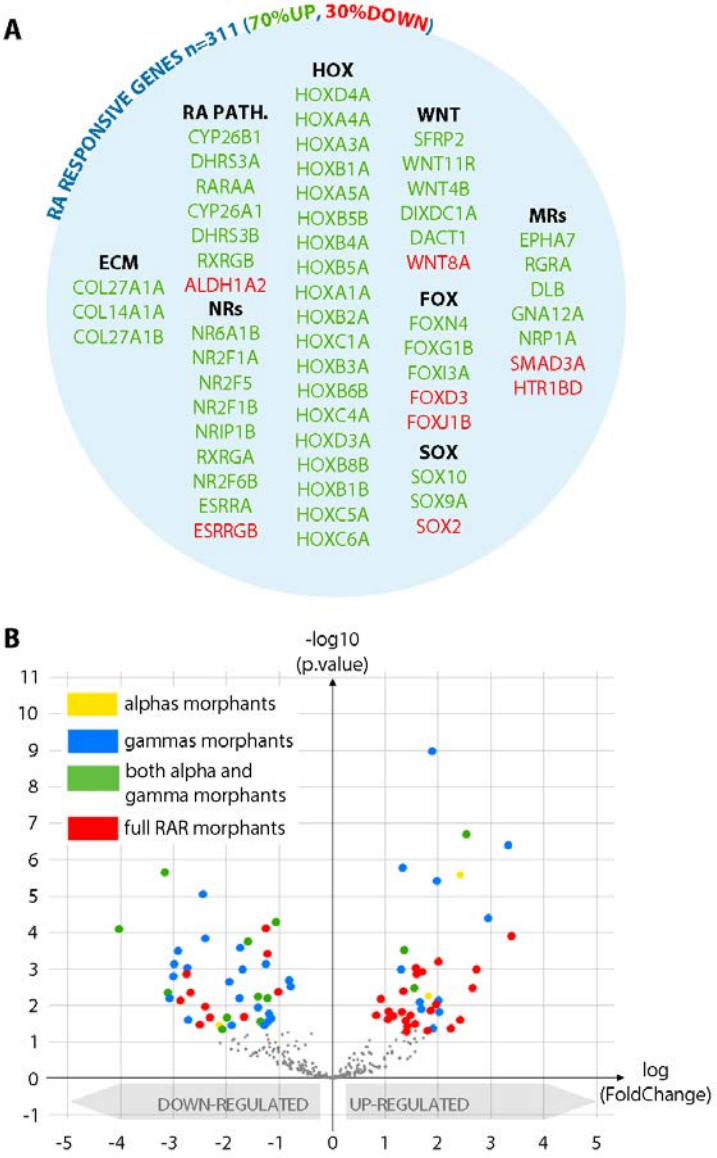


Figure 3: RA-responsive genes and their RAR-subtype dependency at the basal level. A. List of the main genes differentially expressed after a treatment of embryos from 4 to 8 hpf with 10^{-7}M ATRA. A total of 311 genes were identified after filtration on their p.value < 0,05 and |logFC| > 1. Among them, 70% are up-regulated (in green) and 30% are down-regulated (in red) after ATRA exposure. The full list of genes with their corresponding fold-change and p.value is provided in Supplemental Table 1. **B.** The expression of the 311 RA-responsive genes identified was checked in the different RAR morphants in regards to uninjected embryos at the basal level (i.e without ATRA exposure). Each gene colored in the volcano plot is differentially expressed (DE) in the full morphants and they are plotted according to its fold-change (x-axis) and its p.value (y-axis). Yellow genes are specifically DE in alpha morphants (also in full morphant but not in gamma morphants) while blue genes are specifically DE in gamma morphants (also in full morphant but not in alpha morphants). Green genes correspond to genes that are colored both in yellow and blue (DE both in alpha or gamma morphants) and red genes are DE only in the full morphant (i.e lacking all RARs). The full list of genes and their FC and p-value in each condition is provided in Supplemental Table 2.

expression (i.e without RA exposure) of the 311 RA-responsive genes (Figure 3B and Supplemental Table 2). We therefore tested if these genes, regulated after exogenous RA exposure, could indeed be under the control of endogenous RAR in the early embryo. Only 27% of these RA-responsive genes (n=85/311) exhibit a significant differential expression in the full-RAR morphants suggesting that many of them are not regulated by RARs at the basal level. Within these genes, some are only affected in alpha morphants (yellow genes in Figure 3B), others are affected only in gamma morphants (blue genes in Figure 3), some are affected in both situations (green genes in Figure 3) and lastly som e genes are only affected in the full RAR morphants (red genes in Figure 3). This result indicates that the basal expression of RA-target genes can be regulated specifically by one RAR subtype (either alphas or gammas) or can rely on both for their normal basal expression (e.g requires all four RARs). Of note, only 20 genes are differentially expressed in the alpha morphants (count yellow and green genes in Figure 3B) while 58 genes are affected in the gamma morphants (count blue and green genes in Figure 3B) suggesting that there are

more gene whose basal regulation is dependent on the gamma subtypes. Interestingly, 85% of the genes that are differentially expressed in the alpha or gamma morphants are also found in the full-RAR morphants, therefore emphasizing the validity of the effects observed in the various situations. Besides, nearly half of the affected genes (37 of 85) are only affected in the full-RAR morphant (red genes in Figure 3B) meaning that in some cases the activity of RAR subtypes is redundant to regulate basal expression of RA-target genes. Interestingly, at the basal level, we observed as much downregulated as upregulated genes in our RAR morphants. In the absence of ligand, RARs are thought to actively repress the transcription of their target (Samarut and Rochette-Egly, 2012; Waxman and Yelon, 2011). The fact that we observe many up-regulated genes after RAR knockdown suggest that RARs do have a basal repressive activity on some genes in the early zebrafish embryo contrarily to what has been recently suggested (35).

Differential regulation of RA-responsive genes by alpha or gamma RAR subtypes in vivo.

After comparing the list of RA-responsive genes in WT *versus* RAR-morphants embryos, we considered genes that are still significantly induced in the morphants as RA-target genes whose expression is independent of the knocked-down receptors (orange and red spots in Figure 4A-B). Conversely, genes whose expression in RAR-morphants is drastically decreased (less than 10% of the fold activation observed in controls) are considered as completely abolished and therefore as fully dependent of the targeted receptors (yellow and blue spots in Figure 4A-B). In the alpha morphants 42 genes are completely abolished although they are significantly responding to RA in controls, (yellow spots, Figure 4A, Supplemental Table 3) meaning that these genes are strictly requiring RAR alpha subtypes (either RARα-A and/or RARα-B) for their RA-dependent regulation. Similarly, we counted 21 genes that are strictly requiring RAR gamma subtypes (either RARγ-A and/or RARγ-B) for being regulated by RA (blue spots in Figure 4B, Supplemental Table 3). Among these, 3 genes are completely abolished both in alpha and gamma morphants (green spots in Figure 4A-B) meaning that they strictly require all four RAR subtypes to be regulated by RA. In these subtype-specific morphants, we also observed some genes that are not affected by the morpholino-induced RAR knockdown and that continue to be regulated in the same way as in controls (orange spots in Figure 4A-B). Among these non-affected genes, 6 continue to be normally regulated by RA both in alpha and gamma morphants (red spots in Figure 4A-B), suggesting that for these genes, the activity of the different RAR subtypes is redundant. In the full-RAR morphants situation (Figure 4C), 4 of these 6 genes (i.e *dhrs3b, meis2a, hob1b, her13*) are not anymore responding to RA. This observation indicates that these genes can be regulated either by alpha or gamma subtypes indifferently and are therefore regulated by RAR in a truly redundant manner. Altogether, these results indicate that in the early zebrafish, some RA-

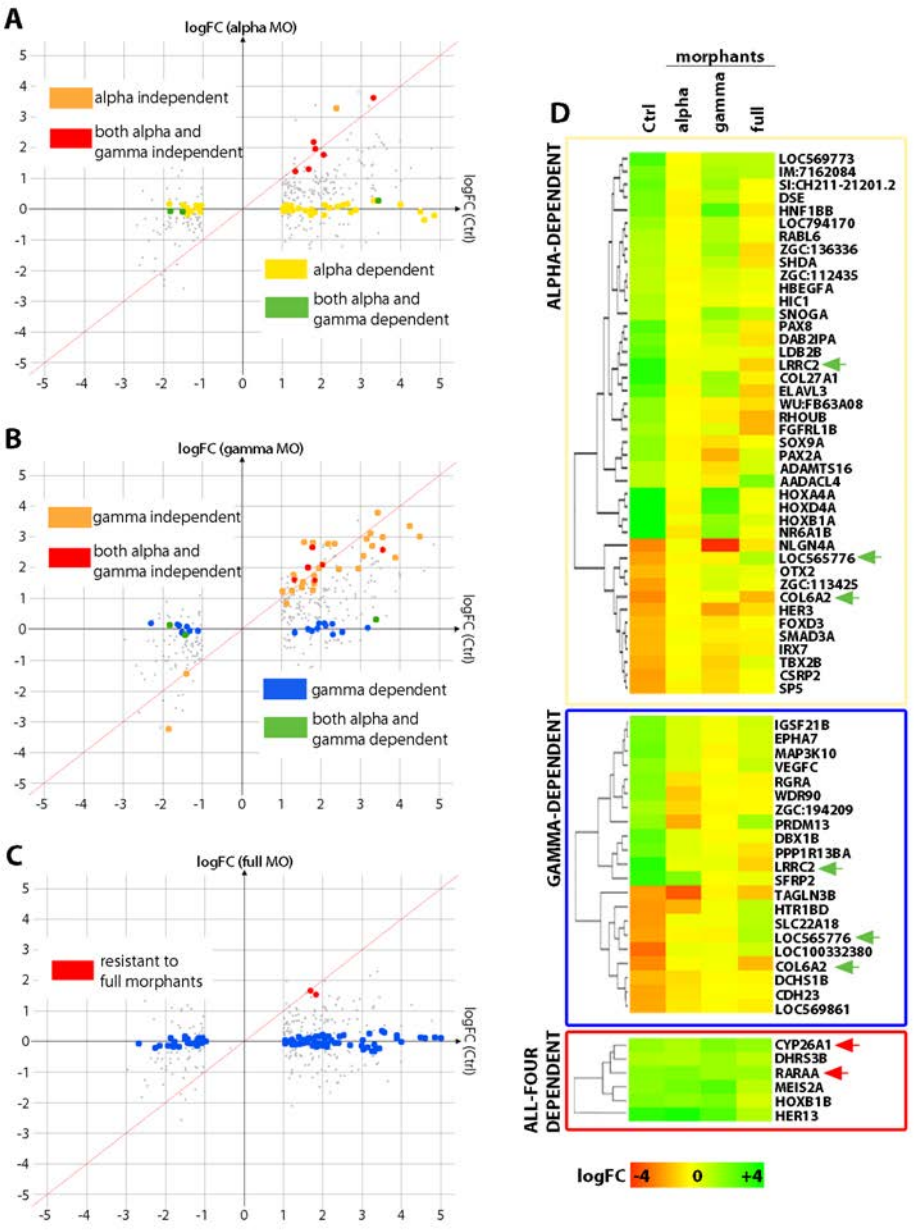
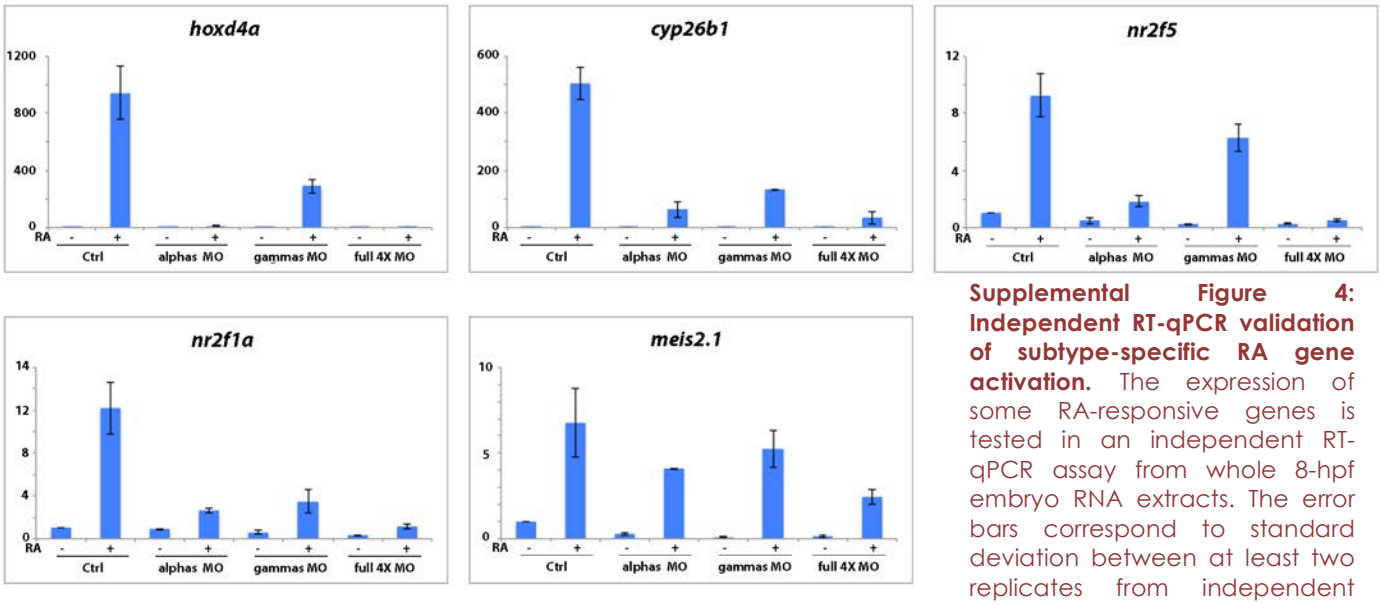


Figure 4: Comparison of the effect of RA on RA-responsive genes in control versus morphants embryos. Comparison of differentially expressed (DE) genes after ATRA exposure in control versus alpha morphants **(A)**, gamma morphants **(B)** or full morphants **(C)**. Each RA-responsive gene is plotted against its fold-change (FC) after ATRA exposure in control embryos (x-axis) and its FC in morphants (y-axis). If a gene has a same significant FC in control and morphants embryos, it is colored and follows the red diagonal. If the FC of a gene is abolished in morphants, it drops near the x-axis, and is colored if its FC in morphants is less than 10% of the FC in controls. In each situation the genes are differentially colored depending on their behavior (subtype-dependent or independent) in the different morphants. Green genes correspond to genes that are colored both in yellow and blue (merge). The full list of genes and their FC and p.value in each condition is provided in Supplemental Table 3. **D.** Alpha-dependent (yellow box), gamma-dependent (blue box) and all-four dependent (red box) genes are depicted on heatmaps. The color code is function of the logFC. Green arrows indicate the genes found both alpha- and gamma-dependent. Red arrows indicate the genes that are still DE in the full morphants.

responsive genes are specifically regulated by alpha or gamma subtypes independently (yellow and blue genes respectively) while others (green genes) require all four subtypes to be RA-responsive. Of note, almost twice as many genes appear sensitive to alpha morpholino knockdown (yellow and green genes in Figure 4A-B) compared with gamma morphants (blue and green genes in figure 4A-B). This suggests that the alpha RAR subtypes are the main subtypes involved in the regulation of RA-induced gene expression in the early zebrafish gastrula whereas as observed above (Figure 3B) gamma RAR subtypes appears to control the basal RA regulation of more genes.

The extreme cases of strict subtype specificity that we just described concern only a small fraction of all RA-responsive genes identified, and the majority of the genes (in grey in Figure 4A-

C) exhibit an intermediate specificity. Indeed, there is a continuum of transcriptional behaviors in response to RA in the different morphants, that we confirmed for some genes in an independent RT-qPCR assay (Supplemental Figure 4) suggesting that for many genes, proper RA regulation requires the availability of all RARs but do not rely on a full dependency on one specific RAR subtype but rather on a partial functional redundancy between subtypes.



Supplemental Figure 4: Independent RT-qPCR validation of subtype-specific RA gene activation. The expression of some RA-responsive genes is tested in an independent RT-qPCR assay from whole 8-hpf embryo RNA extracts. The error bars correspond to standard deviation between at least two replicates from independent clutch.

cyp26a1 and raraa are highly robust RA-responsive genes

During this analysis, we observed two genes that are still RA-regulated in the full-RAR morphants as they are in controls (red spots, Figure 4C). Interestingly, these two genes, *cyp26a1* and *raraa* are members of the RA signaling pathway and are known to be canonical RA target genes in zebrafish (36,37). As several other NRs such as COUP-TFII, TR4, RORβ and PPARβ have been recently proposed as being activated by RA and we therefore wonder if the apparent RAR-independant RA-regulation of *cyp26a1* and *raraa* could be mediated by one of these receptors (38-41). Members of the COUP-TFII group (NR2F1A, NR2F1B, NR2F2 and NR2F5) as well as the two PPARβ (PPARβ-A and PPARβ-B) are expressed in the zebrafish gastrula (21), we therefore knocked-down these receptors and checked if *cyp26a1* and *raraa* induction by RA was altered. We found that the level of RA-regulation was virtually unaltered after single or combined knock-down of these receptors (Figure 5A and data not shown) suggesting that these resistant genes are not under the control of these NRs. Therefore, to check if these genes are regulated by RA through a RAR-dependent mechanism, we monitored their expression by RT-qPCR after treatment with a pan-RAR inverse agonist BMS493 (Figure 5B). Interestingly, the antagonist treatment completely abolishes the RA activation of these genes thus confirming that

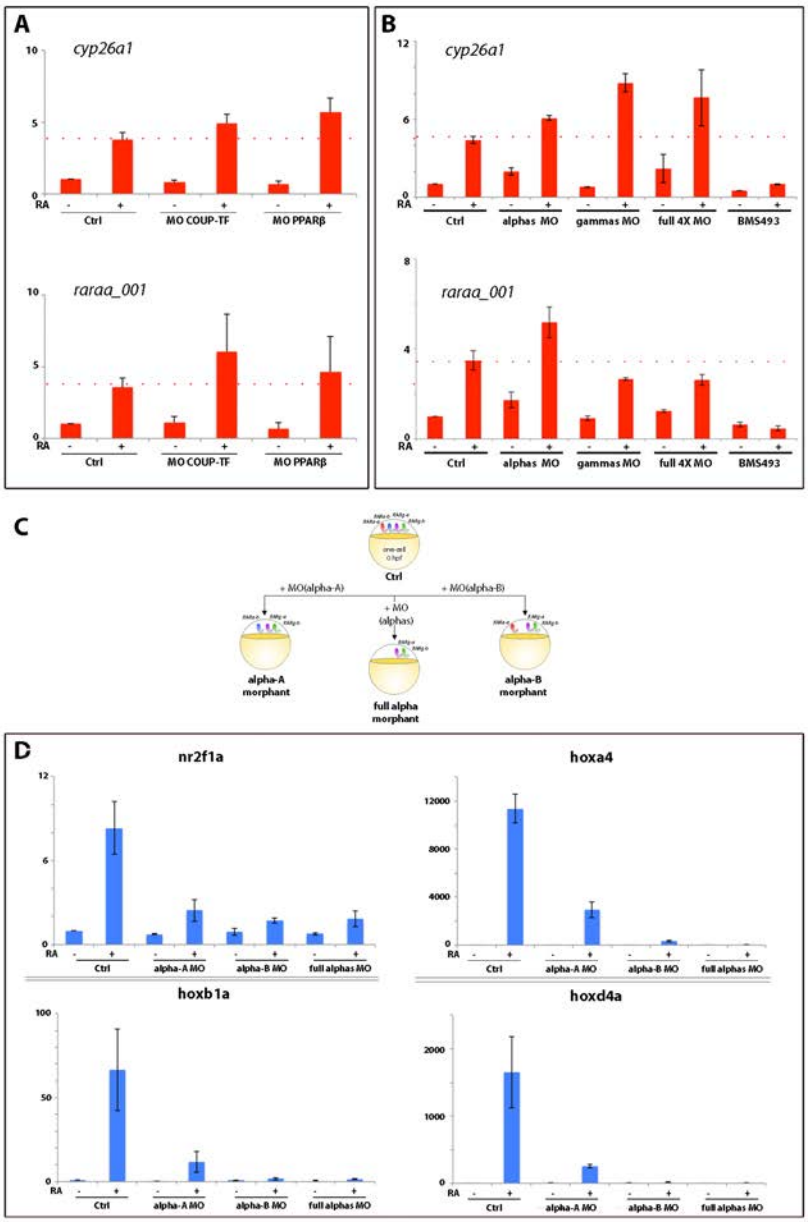


Figure 5: RAR antagonist treatment, knockdown of other NRs and of single RARα-A and RARα-B. A. The expression of *cyp26a1* and *raraa_001* checked by RT-qPCR is still activated by RA after knockdown of COUP-TFII receptors (NR2F1A, NR2F1B, NR2F2 and NR2F5) or PPARβ-A/β-B. **B.** Validation of the activation of *cyp26a1* and *raraa_001* by RA in the different RAR morphants. The activation is abolished by pan-RAR inverse agonist (BMS493) treatment one hour prior ATRA exposure. **C-D** The expression of alpha-dependent genes (i.e *nr2f1a*, *hoxa4*, *hoxb1a*, *hoxd4a*) after ATRA exposure (10-7M, 4-8 hpf) was checked by RT-qPCR in uninjected embryos (Ctrl) or after single knockdown of RARα-A or RARα-B compared with full alpha morphant (i.e both RARα-A and RARα-B) induced by ATG-blocker morpholino injection. The error bars correspond to standard deviation between at least two replicates from independent clutch.

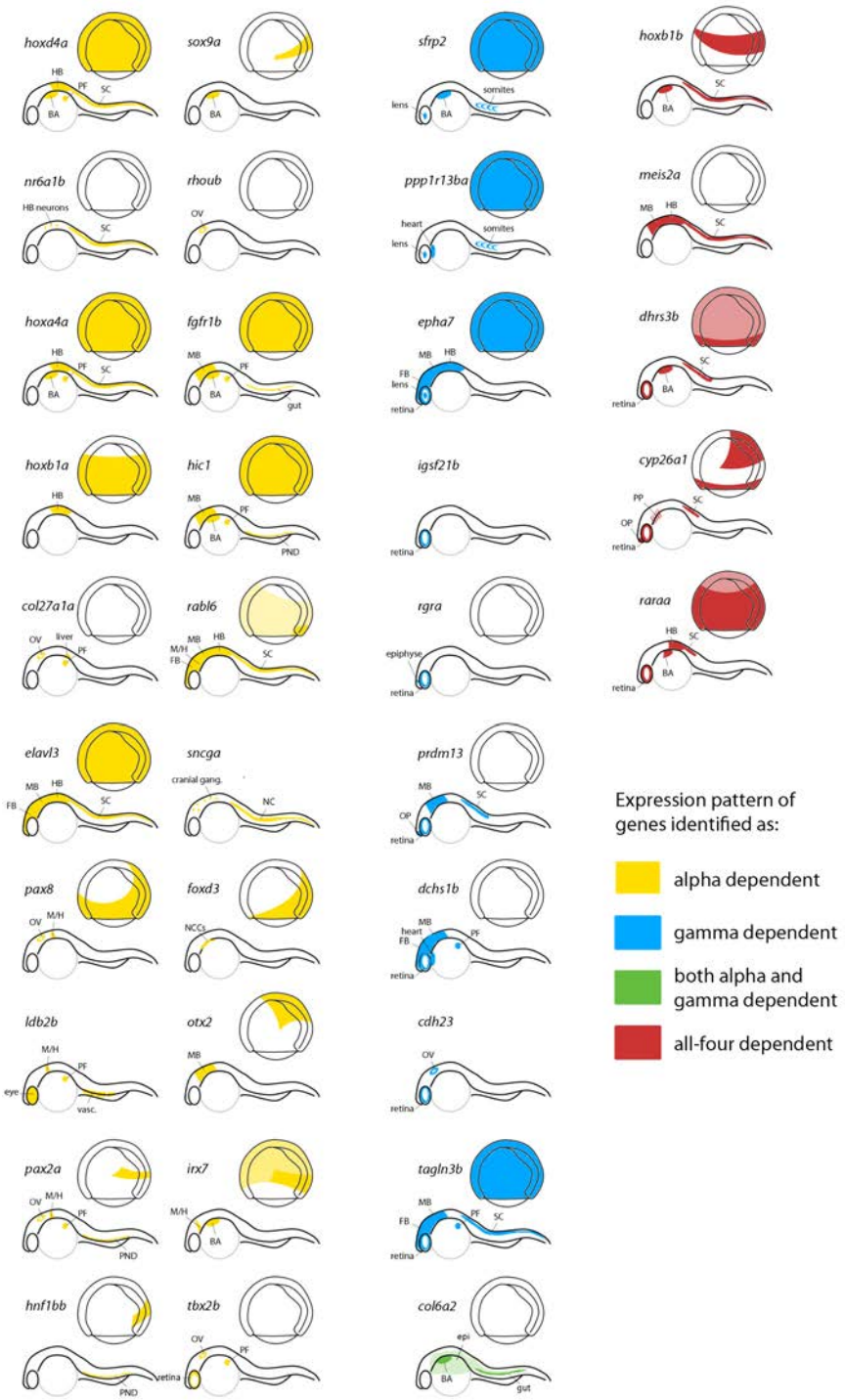
their RA regulation is mediated through RARs. Thus, their surprising resistance to full-RAR morphants can be explained by: (i) the fact that other undescribed isoforms of RARs could exist and are not targeted by our morpholinos or (ii) the fact that, it is known that morpholino injection only lead to incomplete knock-down (42). It is important to note in that respect that, these two genes represents only 0.6% of the 311 responsive genes validating that the knockdown induced by our morpholino is highly efficient. The fact that *cyp26a1* and *raraa* continue to be activated in these morphants as in controls suggest that their RA-induced activation is kept under a high robustness pressure in the early embryo. Besides, this would also suggest that different levels of robustness of RA-regulation exist among RA-target genes and that some genes such as *cyp26a1* and *raraa* could be considered as the most robust genes that can be regulated by RA even in the case of a drastic reduction of RARs availability.

The transcriptional activity of RARα-A and RARα-B is not redundant.

In order to decipher in more details the RAR-subtype-specific transcriptional activity, we decided to knockdown the RAR alpha subtypes (i.e RARα-A and RARα-B) separately. Therefore, we specifically knocked-down RARα-A or RARα-B by morpholino injection (Figure 5C) and we checked by RT-qPCR the expression of genes that have been categorized as alpha-dependent in our whole-transcriptome assay (Figure 4A). For the four genes tested, each single knockdown of RARα-A or RARα-B strongly reduces their RA-induced expression. This suggests that both RARα-A and RARα-B are required for proper RA regulation of these genes. Although the expression level in the single alpha morphants is not as reduced as in the full-alpha morphants these results suggest that, contrarily to what is commonly thought, there is not a full functional redundancy at the transcriptional level between each RAR alpha subtypes, at least at this stage. Of note is that the relative abundance of *raraa* and *rarab* transcripts in the early zebrafish embryo is the same (Oliveira et al., 2013), therefore the changes observed are not due to differences of expression level between both subtypes.

DISCUSSION

Many studies successfully deciphered the specific function of each RAR subtypes in mice and zebrafish using standard forward genetics approaches looking at developmental phenotypes that can be associated with the lack of one specific RAR subtype (19,20). However, it is still not understood why one developmental process could require only one specific RAR subtype although others are expressed in the same region of interest. This is expected to rely on differences in the recruitment and/or transcriptional activity of the different RAR subtypes. In this work, we provide evidences that the different RAR subtypes, even co-expressed *in vivo*, can drive RA transcriptional response in a subtype-specific fashion. Indeed, we were able to discriminate genes that (i) are specifically regulated by one or another RAR subtype (ii) strictly require all subtypes for their correct RA-regulation and (iii) can be regulated either by one or another subtype indifferently. Drawing the parallel with RAR-subtype specific phenotypes that have been characterized in zebrafish (20,43), our results revealed what we call RAR-subtype specific "transcriptotypes" that correspond to specific repertoire of genes activated in a subtype-specific fashion in response to RA. Moreover, we observed that the gamma subtypes of RARs appear to be more involved in the basal regulation of RA-responsive gene expression, whereas alpha subtypes are most implicated in the transcriptional RA-response in the early zebrafish embryo. These differences observed are unlikely to be due to different expression levels since in a 12hpf-embryo, the relative mRNA level of alpha (i.e α-A and α-B) and gamma (i.e γ-A ang γ-B) subtypes are equivalent (44).



More than simply providing exhaustive lists of genes regulated by distinct RARs *in vivo*, we also checked their expression pattern in the ZFIN database at 8hpf in the gastrula and later in development (Supplemental Figure 5). We did not observe any obvious correlation between the subtype-dependency of a gene and its specific expression pattern in the gastrula. In fact, many genes harbor a ubiquitous expression pattern at 8hpf whereas others present more specific, but variable expression patterns (Supplemental Figure 5). As a result, the subtype-specific transcriptotypes we described are more likely to rely on differences of RAR transcriptional activity at the chromatin level instead of differences of their activity in specific developmental territories. We also checked the expression of these genes later in development to see if their subtype-dependency at 8hpf

Supplemental Figure 5: Described expression pattern of RAR-subtype specific dependent genes. For each genes that we previously identified as alpha-dependent (in yellow), gamma-dependent (in blue), both alpha and gamma-dependent (in green) or only dependent of all-four RAR (i.e the full morphant is needed to alter the RA-regulation, in red), we looked for the expression pattern in the ZFIN database at 8 hpf and later during development (24 to 72 hpf). When expression data is available, it is illustrated on a schematized gastrula (dorsal right) and a young larvae (posterior right). For a gene, when no gastrula is presented, it corresponds to a lack of available data at this stage. When an uncolored-gastrula is shown, it means that the expression was assessed but not detected by whole mount *in situ* hybridization. The regions of expression are indicated as followed. OP: olfactory placode; FB: forebrain; MB: midbrain; HB: hindbrain; M/H: midbrain/hindbrain boundary; BA: branchial arches; PP: pharyngeal pouches; OV: otic vesicle; PF; pectoral fin bud; SC: spinal chord; vasc.: vascular system; PND: pro-nephritic duct; NCCs: neural crest cells; epi: epidermis.

correlates with their expression pattern later in development. Interestingly, the alpha-dependent genes at 8hpf are mainly expressed in the central nervous system (CNS) later in development, whereas our gamma-dependent genes are predominantly expressed in the eye (retina or lens), heart and somites of the zebrafish larvae (Supplemental Figure 5). Although all four RARs are expressed in the CNS, only RARα-A and RARα-B are expressed in the eye (21). Altogether, these observations suggest that the specific subtype-dependency of a gene is not necessarily the same at 8hpf and later in development. We therefore propose this dependency does not rely on fixed features but rather on dynamic mechanisms.

How such specificity can be driven at the transcriptional level? Since the DBD of the different RAR subtypes are highly conserved (92%), especially the P box (100% identity) that determines the specific interaction between RAR and DNA (Lee et al., 1993), we do not expect differences at the receptor level that could explain why one specific RAR subtype is specifically involved in the regulation of specific gene as we observed here. However, one can speculate that the specificity relies on specific features of the RAREs that would promote the recruitment of one specific RAR subtype. We exhaustively looked for canonical and non-classical RAREs within zebrafish genome using the PARSEC (PAtteRn Search and Contextualization) platform that have been recently developed to identify a specific sequence pattern within the whole genome (45). We look for direct-repeats (DR) of the consensus RARE nucleotide sequence RGKTSA (46) separated by 1 to 10 nucleotides (DR1-DR10) in the vicinity of the coding region of our RA-responsive genes (+/- 10kb). Unfortunately, neither specific consensus sequence nor special DR space enrichment was found to explain the specific recruitment of one specific RAR subtype. However, recent studies have revealed that the repertoire of RAR binding sites *in vivo* is highly diverse and that the classical RARE consensus does not recapitulate all true binding sites (47). As a result, an *in vivo* analysis of RAR-subtype specific binding sites by chomatin immunoprecipation (ChIP-seq) using antibodies recognizing specifically each subtypes would be necessary to have a better view in what could drive this RAR-subtype specificity at the DNA level, such as the presence of binding sites for other transcription factors in the vicinity of RAREs. Lastly, the unstructured N-terminal domain (NTD) of RARs is a good candidate for being involved in this specificity. In fact, thanks to its proximity with the DBD, proteins interacting with the NTD could regulate RAR recruitment onto DNA as recently hypothesized and/or the recruitment of specific coactivator complexes (48). Since NTDs are the most divergent parts of RARs between subtypes (Supplemental Figure 3B), one can speculate that differences in the NTDs can participate in driving RAR-subtype specific recruitment onto chromatin through specific protein interactions. This would require further work aiming at identifying RAR-subtype specific NTD interacting proteins *in vivo*.

Finally, our work also provides evidences that RA-responsive genes are not all sensitive to the same level of alterations in the RA pathway. In fact, we found genes of the RA pathway itself whose expression is not affected at all, even after a potent knockdown of all RARs. This suggests that the regulation of these genes by RA is extremely robust, suggesting that keeping their RA-regulation is subjected to a strong pressure that can participate to maintain the integrity of the pathway *in vivo*. This is consistent with many studies that highlighted the robustness of RA signaling pathway in the developing embryo, thanks to complex crosstalks and many feedback regulations controlling *cyp26a1* expression (49,50) (Figure 6A). Furthermore, our results show that *raraa* expression is also robustly induced by RA and this could be considered as a security path that can compensate any alteration of RAR availability to finally maintain the RA-regulation of *cyp26a1* (Figure 6). The high robustness of *cyp26a1* expression in response to RA might be associated with the fact that three canonical RAREs have been described within its promoter (36,37). Furthermore, using the PARSEC platform, we found four more canonical RAREs in the promoter region of *cyp26a1* and two DR5 in the vicinity of the *raraa* gene (data not show). We can wonder if the presence of multiple canonical RAREs in the cis-regulatory regions of these genes could be correlated to their robustness.

Our work therefore provides new insights in RAR activity at the transcriptional level in an *in vivo* context looking at endogenous proteins and open new prospects for further research to decipher the mechanisms regulating their activity *in vivo*.

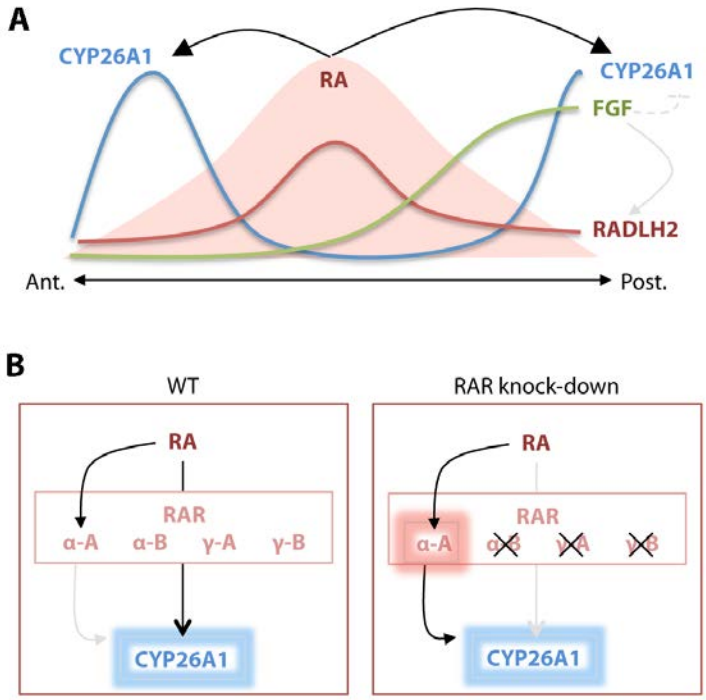


Figure 6: robust *cyp26a1* and *raraa* RA-regulation could participate in maintaining RA signaling *in vivo*. A. The proposed two-tailed gradient along the antero-posterior axis of the embryo is controlled by local degradation of RA by CYP26A1 whose expression is activated by RA and inhibited by FGF pathway (Shimozono et al., 2013). RALDH2 synthetizes RA and its expression is activated by FGF. **B.** In WT embryos, *cyp26a1* is activated by RA by any of the four RAR subtypes. However, RAR-αA is the only subtype to be strongly activated by RA. In the case of an impaired availability of RARs (e.g knockdown), the robust activation of *raraa* by RA ensure the subsequent activation of *cyp26a1*, necessary for the control of RA level *in vivo*.

ACKNOWLEDGMENTS

We warmly thank Laure Bernard for technical help and comprehensive support. We benefited of the zebrafish in-breeding facilities developed by the PRECI technical platform at the SFR Biosciences Gerland. This work was supported by La ligue contre le cancer and the French ministry of research, the CNRS and the Ecole Normale Supérieure de Lyon. Of note it was not funded by ANR who rejected the project.

REFERENCES

1. Samarut E, Rochette-Egly C. Nuclear retinoic acid receptors: conductors of the retinoic acid symphony during development. *Mol Cell Endocrinol.* 2012;348(2):348-360.
2. Clagett-Dame M, DeLuca HF. The role of vitamin A in mammalian reproduction and embryonic development. *Annu Rev Nutr.* 2002;22:347-381.
3. Zile MH. Function of vitamin A in vertebrate embryonic development. *J Nutr.* 2001;131(3):705-708.
4. Begemann G, Marx M, Mebus K, Meyer A, Bastmeyer M. Beyond the neckless phenotype: influence of reduced retinoic acid signaling on motor neuron development in the zebrafish hindbrain. *Dev Biol.* 2004;271(1):119-129.
5. Maves L, Kimmel CB. Dynamic and sequential patterning of the zebrafish posterior hindbrain by retinoic acid. *Dev Biol.* 2005;285(2):593-605.
6. Aulehla A, Pourquie O. Signaling gradients during paraxial mesoderm development. *Cold Spring Harb Perspect Biol.* 2010;2(2):a000869.
7. Chambon P. A decade of molecular biology ot retinoic acid receptors. *FASEBJ.* 1996;10:940-954.
8. Germain P, Altucci L, Bourguet W, Rochette-Egly C, Gronemeyer H. Nuclear receptor superfamily: Principles of signaling. In: Miyamoto J, Burger J, eds. *Pure Appl. Chem.,*. Vol 752003:1619-1664.
9. Germain P, Staels B, Dacquet C, Spedding M, Laudet V. Overview of nomenclature of nuclear receptors. *Pharmacol Rev.* 2006;58(4):685-704.
10. Gronemeyer H, Gustafsson JA, Laudet V. Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov.* 2004;3(11):950-964.
11. Rochette-Egly C, Germain P. Dynamic and combinatorial control of gene expression by nuclear retinoic acid receptors (RARs). *Nucl Recept Signal.* 2009;7:e005.
12. Chakravarti D. *Volume 87: Regulatory Mechanisms in Transcriptional Signaling, 1st Edition.* Elsevier ed2009.
13. Moras D, Gronemeyer H. The nuclear receptor ligand-binding domain: structure and function. *Curr Opin Cell Biol.* 1998;10(3):384-391.
14. Dilworth FJ, Chambon P. Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription. *Oncogene.* 2001;20(24):3047-3054.
15. Germain P, Chambon P, Eichele G, Evans RM, Lazar MA, Leid M, De Lera AR, Lotan R, Mangelsdorf DJ, Gronemeyer H. International Union of Pharmacology. LX. Retinoic acid receptors. *Pharmacol Rev.* 2006;58(4):712-725.
16. Dolle P. Developmental expression of retinoic acid receptors (RARs). *Nucl Recept Signal.* 2009;7:e006.
17. Niederreither K, Dolle P. Retinoic acid in development: towards an integrated view. *Nat Rev Genet.* 2008;9(7):541-553.
18. Mark M, Ghyselinck NB, Chambon P. Function of retinoid nuclear receptors: lessons from genetic and pharmacological dissections of the retinoic acid signaling pathway during mouse embryogenesis. *Annu Rev Pharmacol Toxicol.* 2006;46:451-480.
19. Mark M, Ghyselinck NB, Chambon P. Function of retinoic acid receptors during embryonic development. *Nucl Recept Signal.* 2009;7:e002.
20. Linville A, Radtke K, Waxman JS, Yelon D, Schilling TF. Combinatorial roles for zebrafish retinoic acid receptors in the hindbrain, limbs and pharyngeal arches. *Dev Biol.* 2009;325(1):60-70.
21. Bertrand S, Thisse B, Tavares R, Sachs L, Chaumot A, Bardet PL, Escriva H, Duffraisse M, Marchand O, Safi R, Thisse C, Laudet V. Unexpected novel relational links uncovered by extensive developmental profiling of nuclear receptor expression. *PLoS Genet.* 2007;3(11):e188.

22. Hale LA, Tallafuss A, Yan YL, Dudley L, Eisen JS, Postlethwait JH. Characterization of the retinoic acid receptor genes raraa, rarab and rarg during zebrafish development. *Gene Expr Patterns.* 2006;6(5):546-555.

23. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. Stages of embryonic development of the zebrafish. *Dev Dyn.* 1995;203(3):253-310.

24. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 2010;11(3):R25.

25. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol.* 2004;3:Article3.

26. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological).* 1995;57(1):289-300.

27. Thisse C, Thisse B. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nature Protocols.* 2008;3(1):59-69.

28. Waxman JS, Yelon D. Comparison of the expression patterns of newly identified zebrafish retinoic acid and retinoid X receptors. *Dev Dyn.* 2007;236(2):587-595.

29. Zelent A, Mendelsohn C, Kastner P, Krust A, Garnier JM, Ruffenach F, Leroy P, Chambon P. Differentially expressed isoforms of the mouse retinoic acid receptor beta generated by usage of two promoters and alternative splicing. *EMBO J.* 1991;10(1):71-81.

30. Mollard R, Viville S, Ward SJ, Decimo D, Chambon P, Dolle P. Tissue-specific expression of retinoic acid receptor isoform transcripts in the mouse embryo. *Mech Dev.* 2000;94(1-2):223-232.

31. Birney E, Andrews TD, Bevan P, Caccamo M, Chen Y, Clarke L, Coates G, Cuff J, Curwen V, Cutts T, Down T, Eyras E, Fernandez-Suarez XM, Gane P, Gibbins B, Gilbert J, Hammond M, Hotz HR, Iyer V, Jekosch K, Kahari A, Kasprzyk A, Keefe D, Keenan S, Lehvaslaiho H, McVicker G, Melsopp C, Meidl P, Mongin E, Pettett R, Potter S, Proctor G, Rae M, Searle S, Slater G, Smedley D, Smith J, Spooner W, Stabenau A, Stalker J, Storey R, Ureta-Vidal A, Woodwark KC, Cameron G, Durbin R, Cox A, Hubbard T, Clamp M. An overview of Ensembl. *Genome Res.* 2004;14(5):925-928.

32. Tadros W, Lipshitz HD. The maternal-to-zygotic transition: a play in two acts. *Development.* 2009;136(18):3033-3042.

33. Hammerschmidt M, Pelegri F, Mullins MC, Kane DA, van Eeden FJ, Granato M, Brand M, Furutani-Seiki M, Haffter P, Heisenberg CP, Jiang YJ, Kelsh RN, Odenthal J, Warga RM, Nusslein-Volhard C. dino and mercedes, two genes regulating dorsal development in the zebrafish embryo. *Development.* 1996;123:95-102.

34. Balmer JE, Blomhoff R. Gene expression regulation by retinoic acid. *J Lipid Res.* 2002;43(11):1773-1808.

35. Waxman JS, Yelon D. Zebrafish retinoic acid receptors function as context-dependent transcriptional activators. *Developmental Biology.* 2011;352(1):128-140.

36. Hu P, Tian M, Bao J, Xing G, Gu X, Gao X, Linney E, Zhao Q. Retinoid regulation of the zebrafish cyp26a1 promoter. *Developmental Dynamics.* 2008;237(12):3798-3808.

37. Li J, Hu P, Li K, Zhao Q. Identification and characterization of a novel retinoic acid response element in zebrafish cyp26a1 promoter. *Anat Rec (Hoboken).* 2012;295(2):268-277.

38. Zhou XE, Suino-Powell KM, Xu Y, Chan CW, Tanabe O, Kruse SW, Reynolds R, Engel JD, Xu HE. The Orphan Nuclear Receptor TR4 Is a Vitamin A-activated Nuclear Receptor. *Journal of Biological Chemistry.* 2011;286(4):2877-2885.

39. Stehlin-Gaon C, Willmann D, Zeyer D, Sanglier S, Van Dorsselaer A, Renaud JP, Moras D, Schule R. All-trans retinoic acid is a ligand for the orphan nuclear receptor ROR beta. *Nature Structural Biology.* 2003;10(10):820-825.

40. Kruse SW, Suino-Powell K, Zhou XE, Kretschman JE, Reynolds R, Vonrhein C, Xu Y, Wang L, Tsai SY, Tsai MJ, Xu HE. Identification of COUP-TFII orphan nuclear receptor as a retinoic acid-activated receptor. *PLoS Biol.* 2008;6(9):e227.

41. Shaw N, Elholm M, Noy N. Retinoic acid is a high affinity selective ligand for the peroxisome proliferator-activated receptor beta/delta. *Journal of Biological Chemistry.* 2003;278(43):41589-41592.

42. Bill BR, Petzold AM, Clark KJ, Schimmenti LA, Ekker SC. A Primer for Morpholino Use in Zebrafish. *Zebrafish.* 2009;6(1).

43. D'Aniello E, Rydeen AB, Anderson JL, Mandal A, Waxman JS. Depletion of Retinoic Acid Receptors Initiates a Novel Positive Feedback Mechanism that Promotes Teratogenic Increases in Retinoic Acid. *Plos Genetics.* 2013;9(8).

44. Oliveira E, Casado M, Raldua D, Soares A, Barata C, Pina B. Retinoic acid receptors' expression and function during zebrafish early development. *J Steroid Biochem Mol Biol.* 2013;138C:143-151.
45. Allot A, Anno YN, Poidevin L, Ripp R, Poch O, Lecompte O. PARSEC: PAtteRn SEarch and Contextualization. *Bioinformatics.* 2013.
46. Lalevee S, Anno YN, Chatagnon A, Samarut E, Poch O, Laudet V, Benoit G, Lecompte O, Rochette-Egly C. Genome-wide in silico identification of new conserved and functional retinoic acid receptor response elements (direct repeats separated by 5 bp). *J Biol Chem.* 2011;286(38):33322-33334.
47. Delacroix L, Moutier E, Altobelli G, Legras S, Poch O, Choukrallah MA, Bertin I, Jost B, Davidson I. Cell-specific interaction of retinoic acid receptors with target genes in mouse embryonic fibroblasts and embryonic stem cells. *Mol Cell Biol.* 2010;30(1):231-244.
48. Lalevee S, Bour G, Quinternet M, Samarut E, Kessler P, Vitorino M, Bruck N, Delsuc MA, Vonesch JL, Kieffer B, Rochette-Egly C. Vinexin, an atypical "sensor" of retinoic acid receptor signaling: union and sequestration, separation, and phosphorylation. *FASEB J.* 2010.
49. White RJ, Nie Q, Lander AD, Schilling TF. Complex regulation of cyp26a1 creates a robust retinoic acid gradient in the zebrafish embryo. *PLoS Biol.* 2007;5(11):e304.
50. Shimozono S, Iimura T, Kitaguchi T, Higashijima S, Miyawaki A. Visualization of an endogenous retinoic acid gradient across embryonic development. *Nature.* 2013;496(7445):363-366.

SUPPLEMENTAL TABLE LEGENDS

Supplemental Table 1: RA-responsive genes in the early zebrafish embryo. Full list of RA-responsive genes after treatment of zebrafish embryos from 4 to 8 hpf with 10^{-7}M ATRA. Only genes with a p.value<0,05 and a |logFC|>1 are selected as significant differentially expressed genes (n=311). Genes are classified within different gene families and are listed with decreasing logFC. Green up-arrows show up-regulated genes and red down-arrows show down-regulated genes.....page 114.

Supplemental Table 2: Basal expression of RA-responsive genes in RAR morphants. For each 311 RA-responsive genes, their relative fold-change at the basal level (i.e without exogenous RA exposure) between morphants (either alpha, gamma or full) and controls (uninjected) is presented. Genes are colored depending of their subtype dependency: alpha-dependent genes in yellow, gamma-dependent genes in blue, both alpha and gamma-dependent genes in green and all-four-dependent genes in red (i.e only differentially expressed in full morphant). .....page 118.

Supplemental Table 3: Effect of RAR knockdown on RA regulation of RA-responsive genes. For each RA-responsive gene, the fold-change between ATRA-treated and untreated embryos is shown in control (uninjected) or morphant background (either alpha, gamma or full). Genes are colored depending of their subtype dependency: alpha-dependent genes in yellow, gamma-dependent genes in blue, both alpha and gamma-dependent genes in green and all-four-dependent genes in red (i.e only differentially expressed in full morphant). .....page 124.

TABLE S1: RA RESPONSIVE GENES IN THE EARLY EMBRYO

RA PATHWAY			
Gene		logFC	adj_p.value
CYP26B1	↑	5,00	3,72E-20
DHRS3A	↑	4,78	6,26E-16
RARAA	↑	1,84	7,62E-12
CYP26A1	↑	1,68	2,64E-15
DHRS3B	↑	1,34	9,65E-10
RXRGB	↑	1,31	6,69E-04
ALDH1A2	↓	-1,38	4,89E-08
HOX GENES			
Gene		logFC	adj_p.value
HOXD4A	↑	4,86	6,68E-16
HOXA4A	↑	4,50	9,96E-21
HOXA3A	↑	4,25	8,37E-22
HOXB1A	↑	4,00	3,62E-19
HOXA5A	↑	3,98	1,66E-15
HOXB5B	↑	3,88	8,37E-22
HOXB4A	↑	3,58	2,27E-14
HOXB5A	↑	3,45	1,96E-14
HOXA1A	↑	3,32	1,63E-06
HOXB2A	↑	3,27	7,59E-06
HOXC1A	↑	3,26	1,59E-11
HOXB3A	↑	3,16	6,26E-16
HOXB6B	↑	3,12	1,77E-11
HOXC4A	↑	2,90	1,23E-04
HOXD3A	↑	2,70	1,96E-03
HOXB8B	↑	2,65	2,25E-06
HOXB1B	↑	2,05	7,00E-13
HOXC5A	↑	1,97	1,34E-05
HOXC6A	↑	1,01	1,75E-03
NUCLEAR RECEPTORS			
Gene		logFC	adj_p.value
NR6A1B	↑	4,61	8,87E-11
NR2F1A	↑	3,29	1,00E-10
NR2F5	↑	3,07	1,13E-14
NR2F1B	↑	2,52	9,63E-05
NRIP1B	↑	1,98	5,91E-15
RXRGA	↑	1,77	2,96E-05
NR2F6B	↑	1,74	4,70E-12
ESRRA	↑	1,19	7,14E-08
ESRRGB	↓	-1,41	3,80E-02
WNT PATHWAY			
Gene		logFC	adj_p.value
SFRP2	↑	3,19	3,03E-06
WNT11R	↑	2,56	7,56E-07
WNT4B	↑	2,40	1,69E-04
DIXDC1A	↑	1,55	1,47E-04
DACT1	↑	1,41	1,32E-12
WNT8A	↓	-1,02	1,13E-07
EXTRA-CELLULAR MATRIX			
Gene		logFC	adj_p.value
COL27A1A	↑	3,36	8,19E-16
COL14A1A	↑	2,14	2,23E-13
COL27A1B	↑	2,01	2,56E-02
SOX			

OTHERS			
Gene		logFC	adj_p.value
LOC568110	↑	4,00	1,65E-07
SEPT3	↑	3,88	1,28E-12
SKAP1	↑	3,57	1,35E-12
LOC100535433	↑	3,52	1,17E-14
LOC100003866	↑	3,44	1,63E-08
LRRC2	↑	3,40	3,37E-04
LOC566411	↑	3,39	5,37E-07
TMEM88B	↑	3,36	7,71E-07
ZGC:175176	↑	3,33	1,68E-10
RIPPLY3	↑	3,30	1,83E-08
ZGC:112466	↑	3,27	1,69E-04
LOC100537480	↑	3,20	1,59E-10
LOC100147851	↑	3,02	1,00E-04
LRIT1A	↑	3,02	2,52E-06
NHSA	↑	2,95	5,46E-13
PDZRN3	↑	2,94	1,67E-06
RBMS3	↑	2,87	1,08E-10
LOC569773	↑	2,87	3,76E-04
ELAVL3	↑	2,74	2,26E-08
LOC100332539	↑	2,74	5,81E-08
UGT5G1	↑	2,68	1,77E-13
LOC100537029	↑	2,58	1,03E-03
LDB2B	↑	2,56	2,72E-14
HAPLN1B	↑	2,45	5,35E-12
TMEM30C	↑	2,45	3,17E-04
SI:CH211-212O1.2	↑	2,41	3,93E-09
LOC563500	↑	2,40	9,21E-03
STARD13A	↑	2,37	1,58E-07
LOC100333973	↑	2,36	4,89E-06
ST3GAL4	↑	2,36	1,65E-03
LOC100538211	↑	2,36	9,16E-13
ENO4	↑	2,33	1,98E-09
LOC100331241	↑	2,33	6,50E-03
SI:DKEYP-80C12.7	↑	2,32	6,32E-09
MAP3K10	↑	2,31	8,36E-04
PPP1R13BA	↑	2,30	1,63E-06
LOC799595	↑	2,30	1,08E-11
FTR06	↑	2,23	5,96E-07
NAV3	↑	2,23	3,80E-08
PFKFB2A	↑	2,20	4,96E-05
DAB2IPA	↑	2,20	2,00E-06
NPTX2B	↑	2,19	2,94E-04
INSM1B	↑	2,16	7,00E-13
KIRRELB	↑	2,15	6,89E-03
LOC796490	↑	2,15	3,07E-07
TMEM88A	↑	2,14	7,97E-03
LOC100333111	↑	2,12	1,58E-08
LOC100148343	↑	2,12	1,37E-02
THSD7A	↑	2,11	2,30E-02
LOC100334928	↑	2,09	4,81E-03
ZGC:153720	↑	2,09	1,36E-09
IGSF21B	↑	2,08	1,19E-02
IM:7162084	↑	2,08	2,92E-09

Gene	logFC	adj_p.value
SOX10	1,76	8,67E-05
SOX9A	1,75	6,15E-10
SOX2	-1,14	1,02E-05
FOX		
Gene	logFC	adj_p.value
FOXN4	2,50	2,81E-02
FOXG1B	1,86	1,82E-04
FOXI3A	1,47	2,21E-03
FOXD3	-1,07	5,81E-08
FOXJ1B	-1,42	1,88E-02
TRANSCRIPTION		
Gene	logFC	adj_p.value
HAND2	3,65	2,04E-07
HER13	3,30	2,79E-13
NFIA	3,28	5,47E-10
VAX2	3,00	2,24E-03
IRX3B	2,79	9,45E-07
NEUROG1	2,67	7,84E-10
PAX8	2,65	8,84E-09
DBX1B	2,53	1,69E-04
HNF1BB	2,07	5,80E-05
MEIS1	1,96	2,80E-10
MEIS3	1,83	4,38E-10
MEIS2A	1,79	9,75E-09
PAX2A	1,78	1,20E-03
HNF1BA	1,58	1,77E-13
NFE2L1A	1,40	1,19E-04
MNX1	1,34	5,60E-03
PAX3A	1,32	1,25E-05
ZGC:158824	1,16	5,98E-07
IRX5A	1,15	7,58E-03
ZNF503	1,14	1,59E-10
TP63	1,09	7,33E-07
ZNF703	1,04	5,08E-10
TSHZ1	1,03	8,37E-09
SIX3A	-1,06	1,70E-02
OTX2	-1,07	4,06E-05
IRX7	-1,10	3,96E-11
LHX5	-1,17	5,74E-05
HER3	-1,40	5,03E-06
DBX1A	-1,48	2,30E-09
DMRT2B	-1,79	1,75E-02
SIX7	-2,16	6,47E-10
HESX1	-2,17	7,42E-07
MYOD1	-2,53	3,77E-04
MEMBRANE RECEPTORS		
Gene	logFC	adj_p.value
EPHA7	2,12	1,26E-08
RGRA	1,97	8,63E-04
DLB	1,95	1,21E-07
GNA12A	1,22	4,24E-05
NRP1A	1,02	4,81E-06
SMAD3A	-1,13	2,89E-02
HTR1BD	-1,60	1,82E-03

DSE	2,06	9,90E-11
FAM101A	2,06	6,18E-04
SGK2B	2,05	2,17E-04
LOC562665	2,05	5,76E-09
SEMA3BL	2,05	7,58E-06
KLF12B	1,99	2,96E-05
POU2F2A	1,96	2,48E-12
LRRN1	1,92	3,25E-13
CAPN12	1,92	6,62E-06
JAK2A	1,90	2,90E-09
WDR90	1,90	2,99E-06
LOC100538119	1,87	1,73E-02
PDGFAA	1,85	5,88E-06
DYRK4	1,84	8,21E-04
BCAR1	1,82	1,08E-10
IM:7137453	1,80	2,91E-13
ZIC1	1,80	3,60E-10
LOC100330160	1,77	3,35E-07
VEGFC	1,76	7,04E-03
LOC100538245	1,75	2,61E-03
WU:FB63A08	1,75	2,89E-03
RHOUB	1,73	1,17E-06
HOOK1	1,73	1,91E-04
DYSF	1,71	6,87E-04
PRDM13	1,71	3,94E-02
MXRA8A	1,70	5,81E-08
LOC100330657	1,69	8,23E-08
KANK1	1,67	5,75E-11
B3GNT7L	1,65	1,05E-02
LOC794170	1,61	1,66E-06
DCN	1,61	5,27E-03
ZIC4	1,58	1,97E-05
TNS1	1,56	1,82E-04
PCDH18A	1,55	2,55E-07
MPPED2	1,52	1,37E-11
SP7	1,52	1,13E-05
SCOSPONDIN	1,49	3,23E-03
LOC100333651	1,47	8,51E-11
CUX2A	1,46	5,03E-06
LOC100535716	1,45	6,72E-03
MMP14A	1,45	4,24E-04
SI:CH211-103A14.4	1,44	4,17E-05
ABCB11A	1,42	1,48E-05
FGFRL1B	1,42	3,44E-05
MYO10	1,42	1,86E-06
BCR	1,38	1,68E-04
LOC556476	1,38	1,11E-04
OLIG2	1,35	3,14E-02
ZGC:194209	1,34	4,02E-06
PGP	1,32	2,82E-06
LOC566479	1,32	7,05E-04
LOC569124	1,32	1,02E-03
CERCAM	1,32	1,91E-08
PRDM12B	1,32	7,04E-04
LOC100148864	1,30	1,88E-05
INHBB	1,29	8,80E-05
SKIB	1,28	1,98E-09

HIC1	↑	1,25	1,40E-05
CHPFA	↑	1,24	8,33E-09
MET	↑	1,24	1,59E-03
ILDR2	↑	1,23	1,02E-03
OLIG4	↑	1,23	9,71E-07
RABL6	↑	1,23	1,00E-07
WU:FB16H09	↑	1,22	2,25E-06
B3GNT9	↑	1,21	3,28E-02
SI:CH211-137A8.2	↑	1,21	1,65E-02
BMI1A	↑	1,20	9,25E-06
HAO2	↑	1,20	1,38E-02
SHDA	↑	1,18	2,87E-07
PLOD2	↑	1,18	3,62E-06
GRHL2B	↑	1,17	5,10E-04
SNCGA	↑	1,16	2,22E-03
RGMD	↑	1,16	6,08E-09
ZGC:136336	↑	1,15	2,41E-02
SI:DKEY-261E22.1	↑	1,15	1,00E-04
ADARB1A	↑	1,14	3,55E-06
GADD45AB	↑	1,12	3,06E-02
GSTM3	↑	1,12	4,91E-03
TIPARP	↑	1,11	9,39E-10
KITLGA	↑	1,10	5,21E-04
MTMR8	↑	1,09	6,39E-04
AADACL4	↑	1,08	3,41E-02
SI:CH211-217A12.1	↑	1,08	2,10E-02
REM1	↑	1,08	6,50E-05
ADAMTS16	↑	1,08	2,33E-02
AHCYL1	↑	1,07	2,64E-04
RHBDL3	↑	1,07	7,04E-03
PTPRDB	↑	1,07	1,22E-04
CYP2P6	↑	1,06	1,68E-03
ZGC:112435	↑	1,03	7,38E-04
HBEGFA	↑	1,02	2,89E-02
LOC100537717	↑	1,01	4,43E-02
LOC100334713	↓	-0,35	9,09E-01
NPR3	↓	-1,00	3,07E-03
ZGC:123298	↓	-1,04	5,93E-03
LOC100536197	↓	-1,04	4,04E-03
YWHAG1	↓	-1,04	3,87E-02
FAM167AB	↓	-1,08	2,11E-02
SALL3B	↓	-1,09	8,41E-04
ZGC:123339	↓	-1,09	4,52E-02
DCHS1B	↓	-1,11	7,08E-03
ZGC:153154	↓	-1,12	1,01E-04
LOC560775	↓	-1,13	1,25E-05
CHRNA5	↓	-1,14	1,57E-02
TNFSF10L3	↓	-1,14	3,27E-02
SYN1	↓	-1,19	1,65E-02
CA9	↓	-1,19	1,68E-07
MAP1B	↓	-1,23	2,74E-03
EIF2AK2	↓	-1,24	3,44E-02
MASP1	↓	-1,24	1,13E-03
LOC100004296	↓	-1,25	6,49E-05
SEBOX	↓	-1,25	3,90E-02
ST8SIA2	↓	-1,28	4,95E-02
TBX2B	↓	-1,29	2,23E-02
LOC569861	↓	-1,32	4,67E-04

ZGC:112375	-1,34	2,79E-02
PTGDSB	-1,36	2,35E-07
CDH23	-1,38	3,16E-02
LOC793240	-1,38	4,13E-02
RASL11B	-1,39	2,01E-06
SP5	-1,41	3,00E-12
B2M	-1,43	2,94E-04
LOC565776	-1,44	2,83E-02
TAGLN3B	-1,49	4,64E-03
SOCS3A	-1,49	2,49E-02
ZGC:113425	-1,50	1,78E-02
CSRP2	-1,50	9,71E-07
LOC556872	-1,50	1,24E-02
SEPT9B	-1,54	3,86E-02
SLC22A18	-1,59	1,49E-03
IM:7154842	-1,60	3,32E-02
SCN8AA	-1,63	4,78E-02
FLRT2	-1,66	1,50E-03
RAB42A	-1,67	2,24E-03
LOC100001239	-1,67	2,09E-03
DSEL	-1,68	2,89E-02
SP8B	-1,68	9,18E-10
GABRP	-1,71	6,90E-03
ASB5B	-1,81	1,71E-04
COL6A2	-1,83	3,41E-02
SLC35F3A	-1,83	2,24E-03
GUCY1A3	-1,84	1,10E-02
TTLL3	-1,84	1,56E-02
NLGN4A	-1,85	4,28E-03
SI:CH211-55A7.1	-1,85	2,21E-02
ZGC:114118	-1,87	1,15E-07
KLHL5	-1,88	2,04E-02
ZGC:162623	-1,91	1,72E-02
SI:DKEY-27I16.2	-1,93	7,62E-07
SI:DKEY-26G8.4	-1,96	1,75E-02
SLIT3	-1,99	1,22E-02
F8	-1,99	1,13E-02
LOC561565	-2,00	7,91E-03
LOC567790	-2,05	7,41E-03
NAB1A	-2,08	9,23E-03
LRRTM1	-2,09	3,68E-02
LOC561355	-2,20	2,17E-04
SI:DKEY-269I1.3	-2,24	3,48E-03
CPNE8	-2,25	5,23E-04
LOC100332380	-2,30	7,22E-03
APOA4	-2,37	3,97E-03
GDF10B	-2,46	3,32E-04
LINGO1B	-2,52	2,10E-03
LOC100334188	-2,60	1,47E-03
FEZF2	-2,67	7,03E-08
LOC562320	-2,96	1,22E-05

TABLE S2: BASAL EXPRESSION OF RA-RESPONSIVE GENES IN RAR MORPHANTS

	alpha morphants		gamma morphants		full-RAR morphants	
Gene	logFC	adj_P.value	logFC	adj_P.value	logFC	adj_P.value
AADACL4	1,46	1,20E-01	-0,25	8,47E-01	-0,41	7,00E-01
ABCB11A	-0,27	8,16E-01	0,35	6,48E-01	0,19	7,97E-01
ADAMTS16	0,74	5,66E-01	1,63	4,95E-02	1,36	9,61E-02
ADARB1A	0,06	9,60E-01	-0,26	7,05E-01	-0,24	6,85E-01
AHCYL1	-0,81	3,10E-01	-1,83	1,80E-02	-1,07	1,05E-01
ALDH1A2	-0,21	6,91E-01	-0,40	2,70E-01	-1,22	3,86E-04
APOA4	-1,27	5,02E-01	-2,70	7,78E-02	-1,91	1,60E-01
ASB5B	-0,50	5,98E-01	-1,60	3,95E-02	-0,50	4,64E-01
B2M	-0,63	5,22E-01	-1,50	6,90E-02	-0,93	2,18E-01
B3GNT7L	-1,18	5,09E-01	-1,16	3,93E-01	-1,03	4,12E-01
B3GNT9	-0,52	7,28E-01	-1,04	3,37E-01	-0,52	6,02E-01
BCAR1	1,00	4,37E-02	0,20	7,54E-01	0,56	2,05E-01
BCR	-0,95	4,18E-01	0,40	6,63E-01	-0,25	7,93E-01
BMI1A	-0,01	9,95E-01	0,16	8,30E-01	0,59	2,21E-01
CA9	-0,12	8,37E-01	-0,55	1,02E-01	-0,11	7,54E-01
CAPN12	-0,33	8,64E-01	0,32	8,15E-01	1,57	5,02E-02
CDH23	1,49	1,23E-01	0,92	3,38E-01	1,05	2,09E-01
CERCAM	-0,97	5,67E-02	-1,62	1,59E-03	-2,40	1,46E-04
CHPFA	0,29	5,94E-01	-0,06	9,25E-01	0,32	4,09E-01
CHRNA5	0,23	8,19E-01	-1,10	1,25E-01	-0,90	1,84E-01
COL14A1A	0,83	9,40E-02	-0,68	2,46E-01	-0,50	3,67E-01
COL27A1A	0,71	3,45E-01	0,75	2,26E-01	0,47	4,55E-01
COL27A1B	-0,15	9,68E-01	0,50	8,16E-01	-0,33	8,61E-01
COL6A2	-0,53	7,94E-01	-0,37	8,23E-01	-0,28	8,36E-01
CPNE8	-0,16	9,40E-01	-1,14	3,17E-01	-0,66	5,49E-01
CSRP2	-0,42	5,16E-01	-1,95	1,29E-03	-1,21	1,74E-02
CUX2A	1,10	7,71E-02	0,55	3,87E-01	1,61	9,69E-04
CYP26A1	0,13	7,63E-01	-0,26	3,66E-01	-0,02	9,58E-01
CYP26B1	0,50	7,07E-01	1,14	1,42E-01	1,33	5,38E-02
CYP2P6	-0,05	9,80E-01	-0,55	6,08E-01	0,07	9,43E-01
DAB2IPA	0,34	8,06E-01	1,12	1,56E-01	1,44	3,88E-02
DACT1	-0,10	8,42E-01	-0,30	3,43E-01	0,16	5,79E-01
DBX1A	-0,27	5,26E-01	0,32	3,19E-01	0,15	6,55E-01
DBX1B	1,18	4,09E-01	2,03	3,94E-02	1,90	4,41E-02
DCHS1B	-0,62	4,11E-01	-0,93	1,31E-01	-0,36	5,38E-01
DCN	-0,23	9,08E-01	0,03	9,89E-01	0,77	4,42E-01
DHRS3A	-0,13	9,55E-01	0,84	3,93E-01	0,61	5,49E-01
DHRS3B	0,06	9,27E-01	-0,50	1,68E-01	-0,51	1,23E-01
DIXDC1A	-0,37	7,80E-01	-1,06	2,49E-01	-2,40	1,06E-02
DLB	1,57	7,37E-03	1,05	5,39E-02	2,42	2,60E-06
DMRT2B	-1,88	2,14E-01	-1,30	3,06E-01	-2,50	3,39E-02
DSE	-0,85	3,55E-01	-0,71	3,62E-01	-1,02	1,66E-01
DSEL	-1,17	4,63E-01	0,17	9,11E-01	-1,45	1,98E-01
DYRK4	-0,85	6,41E-01	-0,58	7,00E-01	-0,24	8,62E-01
DYSF	0,49	8,01E-01	1,12	3,09E-01	1,83	4,96E-02
EIF2AK2	-3,59	2,36E-03	-2,70	6,46E-03	-1,98	2,27E-02
ELAVL3	0,22	8,96E-01	-0,24	8,56E-01	1,41	4,84E-02
ENO4	0,65	4,38E-01	-0,48	5,90E-01	1,69	1,29E-03
EPHA7	0,23	8,75E-01	0,05	9,67E-01	-1,10	2,67E-01
ESRRA	-0,56	2,94E-01	-0,81	6,99E-02	-0,84	5,56E-02
ESRRGB	0,00	9,99E-01	-0,20	8,96E-01	0,27	8,14E-01
F8	0,13	9,61E-01	-0,54	7,40E-01	0,43	7,27E-01
FAM101A	-1,38	4,17E-01	-0,36	8,31E-01	0,04	9,75E-01
FAM167AB	-0,32	7,63E-01	-1,99	1,37E-02	-0,10	9,00E-01

FEZF2	-1,47	3,21E-02	-2,76	6,04E-04	-4,02	8,24E-05
FGFRL1B	0,18	8,89E-01	-0,36	7,12E-01	-0,17	8,43E-01
FLRT2	-0,72	5,56E-01	-0,36	7,44E-01	0,12	9,04E-01
FOXD3	0,60	5,52E-02	0,97	3,01E-04	1,88	1,06E-09
FOXG1B	-0,67	6,98E-01	0,29	8,38E-01	-1,43	1,99E-01
FOXI3A	-0,71	5,92E-01	-2,08	5,97E-02	-1,24	1,97E-01
FOXJ1B	0,18	9,16E-01	0,01	9,97E-01	0,34	7,33E-01
FOXN4	-0,27	9,51E-01	-0,01	9,99E-01	0,42	8,49E-01
FTR06	0,24	8,81E-01	-0,71	5,35E-01	-0,79	4,53E-01
GABRP	-1,21	2,65E-01	-3,51	3,65E-03	-3,08	5,52E-03
GADD45AB	0,74	5,00E-01	0,85	3,30E-01	0,61	4,66E-01
GDF10B	0,94	5,51E-01	-1,51	3,01E-01	-0,43	7,61E-01
GNA12A	-0,87	2,60E-01	-0,93	1,51E-01	-1,66	2,03E-02
GRHL2B	-1,15	2,57E-01	-1,70	5,34E-02	-0,81	3,00E-01
GSTM3	-0,06	9,63E-01	-0,36	6,48E-01	0,70	1,95E-01
GUCY1A3	-1,14	5,17E-01	-3,30	2,70E-02	-2,12	1,08E-01
HAND2	-2,18	2,16E-01	0,32	8,48E-01	0,81	4,85E-01
HAO2	1,00	4,43E-01	-0,44	7,68E-01	0,61	5,48E-01
HAPLN1B	-0,46	6,28E-01	-0,26	7,55E-01	-0,34	6,36E-01
HBEGFA	1,38	1,37E-01	1,13	1,84E-01	1,57	3,26E-02
HER13	-1,05	4,11E-01	-0,68	5,35E-01	-0,46	6,45E-01
HER3	-0,02	9,86E-01	0,66	1,24E-01	0,33	4,47E-01
HESX1	-0,31	6,88E-01	-1,08	4,33E-02	-1,18	2,57E-02
HIC1	1,16	3,95E-02	1,99	5,71E-05	2,54	2,04E-07
HNF1BA	0,08	8,80E-01	0,74	7,05E-03	1,33	1,69E-06
HNF1BB	0,96	6,00E-01	-0,97	5,47E-01	1,27	2,85E-01
HOOK1	0,47	7,95E-01	1,89	3,64E-02	2,02	1,52E-02
HOXA1A	0,29	9,02E-01	1,70	1,46E-01	0,05	9,75E-01
HOXA3A	-0,61	5,83E-01	-0,07	9,50E-01	-0,74	3,73E-01
HOXA4A	0,22	8,73E-01	0,01	9,91E-01	0,10	9,22E-01
HOXA5A	-0,96	5,59E-01	1,28	2,02E-01	1,43	1,03E-01
HOXB1A	-0,01	9,95E-01	-0,09	9,37E-01	-0,91	2,59E-01
HOXB1B	-1,06	3,73E-02	-1,01	2,41E-02	-1,22	6,41E-03
HOXB2A	-0,47	8,76E-01	-0,20	9,36E-01	1,00	4,99E-01
HOXB3A	-0,78	3,48E-01	-0,67	3,34E-01	-1,03	1,48E-01
HOXB4A	-0,94	4,86E-01	0,04	9,80E-01	0,05	9,66E-01
HOXB5A	-1,75	1,75E-01	-1,39	1,95E-01	-0,89	3,61E-01
HOXB5B	-0,27	8,05E-01	0,67	2,40E-01	0,28	6,66E-01
HOXB6B	0,27	8,75E-01	0,29	8,24E-01	-0,85	4,23E-01
HOXB8B	-1,39	4,54E-01	-0,37	8,38E-01	-0,77	5,64E-01
HOXC1A	-0,98	5,41E-01	-0,71	6,00E-01	1,13	2,03E-01
HOXC4A	-1,92	3,75E-01	-0,81	6,77E-01	-0,46	7,87E-01
HOXC5A	-2,90	2,77E-02	-1,76	1,05E-01	-2,11	3,86E-02
HOXC6A	-1,89	1,54E-02	-2,17	3,53E-03	-1,36	2,92E-02
HOXD3A	-0,60	8,16E-01	-0,33	8,74E-01	0,01	9,95E-01
HOXD4A	-0,28	8,91E-01	0,00	9,99E-01	-0,05	9,73E-01
HTR1BD	0,18	9,06E-01	-0,40	7,07E-01	-1,29	1,42E-01
IGSF21B	-1,32	4,69E-01	-1,12	4,60E-01	-0,79	5,44E-01
ILDR2	0,22	8,77E-01	-1,02	2,99E-01	0,04	9,66E-01
IM:7137453	-0,06	9,37E-01	-0,36	4,08E-01	-0,68	9,67E-02
IM:7154842	-2,06	1,45E-01	-1,80	1,32E-01	0,99	3,04E-01
IM:7162084	-0,48	5,71E-01	-2,35	4,48E-03	-2,98	7,32E-04
INHBB	-0,46	7,25E-01	-0,12	9,18E-01	-1,12	1,85E-01
INSM1B	0,70	1,93E-01	-0,12	8,74E-01	0,48	3,10E-01
IRX3B	0,75	6,32E-01	0,64	6,30E-01	1,48	1,12E-01
IRX5A	-0,06	9,78E-01	0,02	9,89E-01	0,72	4,02E-01
IRX7	-0,54	3,72E-02	-0,57	1,16E-02	-0,39	6,05E-02
JAK2A	0,66	3,87E-01	1,05	6,23E-02	1,60	1,39E-03

KANK1	0,44	4,76E-01	0,38	4,81E-01	0,64	1,39E-01
KIRRELB	0,04	9,90E-01	-0,83	6,31E-01	0,25	8,71E-01
KITLGA	0,38	7,07E-01	-0,08	9,44E-01	-0,78	3,32E-01
KLF12B	0,53	6,97E-01	0,15	9,13E-01	0,49	5,98E-01
KLHL5	0,55	7,63E-01	0,58	6,81E-01	0,06	9,65E-01
LDB2B	0,32	7,18E-01	-0,10	9,14E-01	-0,63	3,64E-01
LHX5	0,46	4,67E-01	0,75	1,00E-01	0,71	9,09E-02
LINGO1B	-1,55	3,12E-01	-3,90	5,15E-03	-2,73	2,59E-02
LOC100001239	-0,24	8,66E-01	-1,30	1,54E-01	-2,67	4,52E-03
LOC100003866	0,85	5,94E-01	-0,48	7,65E-01	-0,79	5,38E-01
LOC100004296	0,67	2,41E-01	-0,04	9,57E-01	-0,53	3,18E-01
LOC100147851	-0,89	6,82E-01	-0,62	7,27E-01	0,58	6,73E-01
LOC100148343	-0,59	8,01E-01	-0,32	8,67E-01	0,04	9,82E-01
LOC100148864	-1,79	4,11E-02	-1,48	4,45E-02	-1,23	6,59E-02
LOC100330160	0,36	7,78E-01	-0,21	8,56E-01	0,30	7,27E-01
LOC100330657	0,27	8,46E-01	-0,23	8,40E-01	0,47	5,56E-01
LOC100331241	-0,73	7,62E-01	-0,47	8,12E-01	0,70	6,24E-01
LOC100332380	-1,07	5,97E-01	-2,01	1,82E-01	-0,16	9,22E-01
LOC100332539	-0,95	6,20E-01	-1,45	3,18E-01	-0,25	8,65E-01
LOC100333111	0,15	9,22E-01	-0,61	5,68E-01	0,22	8,19E-01
LOC100333651	0,92	4,27E-02	0,81	4,87E-02	1,36	3,04E-04
LOC100333973	-0,28	8,80E-01	-1,06	3,96E-01	0,19	8,69E-01
LOC100334188	0,48	8,14E-01	0,44	7,81E-01	-0,47	7,33E-01
LOC100334713	0,03	9,94E-01	-2,24	1,29E-01	-0,30	8,40E-01
LOC100334928	-1,20	5,24E-01	-0,94	5,52E-01	-0,55	7,01E-01
LOC100535433	-0,57	6,58E-01	0,14	8,96E-01	-0,54	5,68E-01
LOC100535716	1,41	1,81E-01	0,90	3,85E-01	0,99	2,72E-01
LOC100536197	0,51	5,41E-01	-0,60	4,39E-01	-2,76	1,29E-03
LOC100537029	0,64	7,80E-01	0,92	5,68E-01	0,89	5,12E-01
LOC100537480	-1,83	2,17E-01	0,52	6,49E-01	-0,83	4,55E-01
LOC100537717	-0,40	8,42E-01	-0,34	8,23E-01	0,41	7,21E-01
LOC100538119	-1,95	3,00E-01	-0,28	8,84E-01	-0,39	7,98E-01
LOC100538211	0,81	3,10E-01	0,52	4,93E-01	1,33	1,61E-02
LOC100538245	0,10	9,65E-01	-0,22	8,80E-01	1,97	9,97E-03
LOC556476	1,83	2,02E-02	0,62	5,13E-01	1,82	5,66E-03
LOC556872	-0,16	9,37E-01	-0,10	9,45E-01	-0,17	8,89E-01
LOC560775	-0,48	3,82E-01	0,31	4,81E-01	0,92	6,69E-03
LOC561355	0,44	6,37E-01	0,51	4,80E-01	0,74	2,15E-01
LOC561565	0,15	9,55E-01	1,38	2,05E-01	-0,19	8,93E-01
LOC562320	-3,85	4,75E-03	-3,42	4,07E-03	-3,09	4,49E-03
LOC562665	0,87	2,86E-01	-0,10	9,36E-01	0,18	8,36E-01
LOC563500	-0,75	7,43E-01	-0,27	8,91E-01	1,28	2,78E-01
LOC565776	-0,98	3,88E-01	-1,43	1,29E-01	-1,30	1,53E-01
LOC566411	0,10	9,75E-01	1,30	3,19E-01	0,53	7,25E-01
LOC566479	0,46	7,09E-01	0,53	5,77E-01	0,48	5,71E-01
LOC567790	-0,84	5,87E-01	-1,90	1,14E-01	-1,77	1,13E-01
LOC568110	0,17	9,60E-01	1,66	2,25E-01	1,01	4,74E-01
LOC569124	-1,26	2,79E-01	-1,24	1,84E-01	-1,65	7,54E-02
LOC569773	1,23	5,51E-01	0,39	8,59E-01	0,73	6,48E-01
LOC569861	0,26	7,18E-01	-1,20	3,05E-02	-1,41	1,18E-02
LOC793240	-0,46	7,79E-01	0,51	6,42E-01	-0,43	6,88E-01
LOC794170	0,41	6,67E-01	-0,40	6,53E-01	0,22	7,73E-01
LOC796490	0,16	9,20E-01	0,98	2,04E-01	1,20	8,93E-02
LOC799595	0,13	9,31E-01	-0,06	9,58E-01	0,21	8,05E-01
LRIT1A	-0,57	8,10E-01	0,46	7,91E-01	0,89	4,77E-01
LRRC2	0,72	7,71E-01	0,58	7,69E-01	1,42	2,96E-01
LRRN1	0,32	5,52E-01	-0,55	2,15E-01	-0,12	7,96E-01
LRRTM1	-0,94	6,36E-01	-2,90	4,94E-02	-1,78	1,75E-01

MAP1B	-0,19	8,71E-01	0,24	7,65E-01	-0,23	7,68E-01
MAP3K10	-2,27	1,33E-01	0,55	6,86E-01	-1,63	1,60E-01
MASP1	-0,27	7,57E-01	-2,06	3,16E-03	-2,91	3,19E-04
MEIS1	-0,61	3,68E-01	-1,42	1,64E-02	-1,75	6,30E-03
MEIS2A	-1,18	7,67E-02	-1,60	9,29E-03	-1,23	2,97E-02
MEIS3	-1,36	1,89E-02	-1,35	8,08E-03	-1,39	6,03E-03
MET	-0,23	8,82E-01	-0,20	8,70E-01	0,33	7,11E-01
MMP14A	-0,47	7,56E-01	1,70	1,54E-02	1,68	1,28E-02
MNX1	-2,74	3,95E-02	-2,47	2,92E-02	-2,09	4,46E-02
MPPED2	0,48	3,34E-01	0,65	8,54E-02	0,83	1,88E-02
MTMR8	-0,84	3,71E-01	-0,55	4,89E-01	-0,58	4,31E-01
MXRA8A	0,12	9,28E-01	0,81	2,13E-01	1,33	1,60E-02
MYO10	0,14	8,88E-01	0,53	3,68E-01	0,16	8,14E-01
MYOD1	-0,17	9,38E-01	-1,04	4,19E-01	-1,04	3,53E-01
NAB1A	-1,58	4,00E-01	-1,99	1,83E-01	-0,02	9,89E-01
NAV3	1,23	2,64E-01	2,81	1,67E-04	2,95	4,01E-05
NEUROG1	0,09	9,65E-01	-0,51	6,91E-01	-0,89	4,05E-01
NFE2L1A	0,55	5,64E-01	0,34	6,98E-01	-0,57	4,96E-01
NFIA	0,35	8,69E-01	-0,57	7,37E-01	0,61	6,18E-01
NHSA	-0,04	9,83E-01	0,56	4,78E-01	0,26	7,59E-01
NLGN4A	-0,88	5,71E-01	1,43	1,06E-01	0,06	9,63E-01
NPR3	0,04	9,69E-01	-0,95	6,04E-02	-1,42	6,37E-03
NPTX2B	-0,51	8,14E-01	1,75	9,23E-02	1,59	1,02E-01
NR2F1A	-1,11	5,05E-01	-0,85	5,25E-01	-0,11	9,35E-01
NR2F1B	1,37	3,73E-01	1,35	2,88E-01	0,85	4,98E-01
NR2F5	-0,90	4,35E-01	-0,55	5,73E-01	-0,27	7,66E-01
NR2F6B	0,32	5,41E-01	-0,19	7,06E-01	0,26	5,11E-01
NR6A1B	2,10	7,93E-02	1,64	1,51E-01	3,38	1,27E-04
NRIP1B	0,14	8,18E-01	0,48	1,67E-01	0,41	2,11E-01
NRP1A	0,25	6,80E-01	-0,76	1,04E-01	-0,64	1,56E-01
OLIG2	-1,42	2,04E-01	-2,95	5,61E-03	-1,89	3,64E-02
OLIG4	-0,42	4,98E-01	-1,01	4,19E-02	-0,70	1,30E-01
OTX2	-0,55	2,81E-01	-1,79	1,89E-04	-2,44	8,91E-06
PAX2A	-0,18	9,50E-01	-0,71	6,71E-01	0,11	9,43E-01
PAX3A	-0,61	5,47E-01	-0,92	2,52E-01	-1,01	1,90E-01
PAX8	-1,33	3,55E-01	-0,43	7,46E-01	-0,93	3,88E-01
PCDH18A	-0,74	2,89E-01	-0,31	6,29E-01	-0,52	3,50E-01
PDGFAA	-2,24	7,96E-02	0,41	6,77E-01	-0,91	3,33E-01
PDZRN3	-1,18	6,08E-01	-0,18	9,34E-01	0,71	6,28E-01
PFKFB2A	0,76	6,58E-01	1,39	2,25E-01	2,66	3,38E-03
PGP	-0,21	8,46E-01	-0,62	4,13E-01	1,35	4,04E-03
PLOD2	-0,97	1,17E-01	-2,17	1,99E-03	-1,95	2,30E-03
POU2F2A	-0,25	7,97E-01	0,54	3,17E-01	0,36	5,14E-01
PPP1R13BA	-1,31	3,77E-01	0,75	4,44E-01	1,86	1,41E-02
PRDM12B	0,05	9,83E-01	2,65	4,46E-04	2,00	7,66E-03
PRDM13	1,97	2,26E-01	1,78	2,04E-01	1,47	2,56E-01
PTGDSB	0,18	7,37E-01	-1,55	1,61E-04	-1,26	7,63E-04
PTPRDB	-0,11	9,35E-01	-0,22	8,12E-01	0,50	4,25E-01
RAB42A	-0,64	5,46E-01	-1,04	2,07E-01	-1,10	1,62E-01
RABL6	-0,22	8,32E-01	0,20	7,92E-01	1,08	1,62E-02
RARAA	0,44	3,76E-01	0,08	8,91E-01	0,31	4,22E-01
RASL11B	0,61	2,43E-01	0,01	9,94E-01	0,41	3,19E-01
RBMS3	0,94	3,34E-01	0,89	2,75E-01	1,09	1,35E-01
REM1	0,47	5,28E-01	0,65	2,33E-01	1,05	2,33E-02
RGMD	0,14	8,03E-01	0,20	6,00E-01	0,33	2,80E-01
RGRA	-1,39	3,40E-01	-1,83	1,17E-01	-1,85	8,57E-02
RHBDL3	0,57	5,01E-01	0,50	4,73E-01	-0,01	9,94E-01
RHOUB	-0,15	9,37E-01	0,60	5,48E-01	0,13	9,14E-01

RIPPLY3	0,62	7,50E-01	1,17	3,42E-01	1,72	7,83E-02
RXRGA	-0,68	5,34E-01	-0,44	6,43E-01	-0,61	4,49E-01
RXRGB	-0,70	5,90E-01	-1,33	2,24E-01	-1,89	9,05E-02
SALL3B	0,70	2,09E-01	0,95	3,11E-02	1,98	3,94E-06
SCN8AA	-1,66	2,98E-01	-1,52	2,39E-01	0,11	9,35E-01
SCOSPONDIN	0,38	7,80E-01	0,37	7,08E-01	1,31	5,76E-02
SEBOX	1,28	1,07E-01	2,29	4,54E-04	3,33	4,13E-07
SEMA3BL	-0,85	6,78E-01	0,02	9,93E-01	0,07	9,64E-01
SEPT3	0,17	9,25E-01	0,48	6,65E-01	0,28	7,92E-01
SEPT9B	-0,59	7,03E-01	-1,54	1,56E-01	0,09	9,41E-01
SFRP2	-1,76	2,86E-01	0,98	4,28E-01	-0,24	8,69E-01
SGK2B	-0,14	9,51E-01	0,45	7,33E-01	-0,26	8,32E-01
SHDA	0,16	8,34E-01	-0,36	5,13E-01	-0,94	6,34E-02
SI:CH211-103A14,4	0,47	5,68E-01	-1,39	6,52E-02	-1,01	1,45E-01
SI:CH211-137A8,2	1,42	1,63E-01	1,05	2,76E-01	1,13	1,80E-01
SI:CH211-212O1,2	0,50	6,53E-01	1,01	1,56E-01	2,01	6,34E-04
SI:CH211-217A12,1	0,38	8,00E-01	-1,63	1,34E-01	-0,60	5,52E-01
SI:CH211-55A7,1	0,06	9,83E-01	1,49	1,23E-01	1,56	8,24E-02
SI:DKEY-261E22,1	1,01	2,64E-01	1,41	3,59E-02	1,65	8,14E-03
SI:DKEY-269I1,3	0,47	8,03E-01	-2,69	4,73E-02	-1,50	2,16E-01
SI:DKEY-26G8,4	0,56	7,64E-01	-2,61	5,66E-02	-1,41	2,51E-01
SI:DKEY-27I16,2	0,33	6,52E-01	-0,03	9,72E-01	0,67	1,41E-01
SI:DKEYP-80C12,7	-0,30	8,45E-01	-1,20	2,55E-01	-0,50	6,14E-01
SIX3A	0,10	9,42E-01	-1,62	7,84E-02	-1,61	6,63E-02
SIX7	-1,00	4,43E-02	-2,19	4,66E-05	-3,16	2,18E-06
SKAP1	-0,97	4,13E-01	-1,38	1,65E-01	-0,79	3,88E-01
SKIB	0,32	5,25E-01	0,00	9,96E-01	-0,03	9,51E-01
SLC22A18	-0,73	4,98E-01	-2,37	1,28E-02	-2,99	1,65E-03
SLC35F3A	0,23	8,86E-01	-0,62	5,75E-01	-0,03	9,83E-01
SLIT3	-1,31	3,14E-01	-2,06	5,66E-02	-2,86	7,47E-03
SMAD3A	0,28	8,35E-01	0,66	4,45E-01	1,48	1,92E-02
SNCGA	-0,64	4,88E-01	-0,43	5,98E-01	0,17	8,14E-01
SOCS3A	0,41	8,19E-01	1,01	3,56E-01	-0,03	9,83E-01
SOX10	-1,35	3,33E-01	-0,37	7,72E-01	-1,38	1,74E-01
SOX2	0,04	9,61E-01	0,10	8,67E-01	-0,37	3,91E-01
SOX9A	0,75	2,93E-01	0,59	3,62E-01	0,56	3,36E-01
SP5	-0,61	3,21E-02	-0,87	8,07E-04	-1,06	5,35E-05
SP7	-0,97	4,18E-01	-0,38	7,13E-01	0,08	9,32E-01
SP8B	-0,85	4,36E-02	-1,00	8,27E-03	-1,59	1,78E-04
ST3GAL4	-0,82	6,79E-01	-0,55	7,38E-01	0,20	8,89E-01
ST8SIA2	-0,40	8,04E-01	-0,43	7,25E-01	0,04	9,73E-01
STARD13A	-1,65	1,45E-01	0,33	7,51E-01	0,20	8,29E-01
SYN1	-0,23	8,29E-01	-0,87	2,06E-01	-0,84	1,92E-01
TAGLN3B	-0,88	2,97E-01	-3,88	1,61E-04	-2,74	1,02E-03
TBX2B	0,10	9,58E-01	-1,49	1,51E-01	-1,18	2,32E-01
THSD7A	0,20	9,43E-01	0,15	9,42E-01	2,43	2,55E-02
TIPARP	0,30	4,09E-01	-0,05	9,11E-01	0,05	8,68E-01
TMEM30C	-1,19	5,14E-01	0,07	9,72E-01	0,49	6,93E-01
TMEM88A	-0,49	8,26E-01	0,72	6,11E-01	0,61	6,21E-01
TMEM88B	0,98	6,27E-01	1,05	5,06E-01	1,56	1,86E-01
TNFSF10L3	0,32	7,41E-01	-1,92	2,31E-02	-1,21	8,77E-02
TNS1	0,07	9,78E-01	1,10	2,45E-01	0,54	5,93E-01
TP63	-0,50	3,47E-01	-1,35	4,23E-03	-1,69	1,04E-03
TSHZ1	-0,15	7,83E-01	-0,67	5,42E-02	-1,02	4,50E-03
TTLL3	-0,14	9,54E-01	0,36	8,00E-01	-0,59	6,10E-01
UGT5G1	-0,14	9,01E-01	-0,23	7,93E-01	-0,12	8,80E-01
VAX2	0,10	9,81E-01	0,37	8,85E-01	0,74	6,87E-01
VEGFC	-0,34	8,86E-01	0,97	4,61E-01	-0,14	9,30E-01

WDR90	0,66	5,31E-01	0,90	2,38E-01	1,41	2,72E-02
WNT11R	0,07	9,80E-01	1,19	2,62E-01	2,72	1,02E-03
WNT4B	0,30	8,87E-01	-0,17	9,23E-01	-0,64	6,13E-01
WNT8A	-0,18	6,96E-01	-0,54	7,55E-02	-1,25	7,46E-05
WU:FB16H09	0,95	5,25E-02	0,92	3,30E-02	1,30	1,03E-03
WU:FB63A08	2,20	1,07E-01	1,81	1,50E-01	2,25	4,32E-02
YWHAG1	-0,64	5,83E-01	-0,22	8,43E-01	-1,18	1,42E-01
ZGC:112375	-0,03	9,88E-01	-0,45	7,02E-01	-2,31	2,12E-02
ZGC:112435	0,59	4,39E-01	0,47	4,77E-01	1,15	1,94E-02
ZGC:112466	-0,29	9,21E-01	0,35	8,65E-01	0,33	8,40E-01
ZGC:113425	-0,71	5,04E-01	-1,43	1,02E-01	-0,09	9,18E-01
ZGC:114118	-0,33	5,83E-01	-0,52	2,69E-01	-0,31	4,97E-01
ZGC:123298	-0,44	5,84E-01	-1,35	3,29E-02	-1,30	3,35E-02
ZGC:123339	-0,20	8,76E-01	-1,28	1,07E-01	-1,18	1,09E-01
ZGC:136336	0,26	8,78E-01	-1,17	3,16E-01	0,44	6,61E-01
ZGC:153154	-0,42	4,16E-01	-1,73	2,69E-04	-1,73	2,61E-04
ZGC:153720	-0,24	8,26E-01	0,22	7,70E-01	-0,55	4,35E-01
ZGC:158824	-0,15	8,78E-01	0,01	9,87E-01	-0,13	8,36E-01
ZGC:162623	-2,16	2,33E-01	-1,90	2,06E-01	-1,54	2,47E-01
ZGC:175176	-0,05	9,82E-01	-0,27	8,55E-01	-0,01	9,92E-01
ZGC:194209	1,46	2,36E-02	1,33	1,88E-02	1,56	3,34E-03
ZIC1	-0,06	9,53E-01	-1,25	3,21E-02	-0,83	1,26E-01
ZIC4	-1,25	2,26E-01	-1,31	1,23E-01	-1,88	2,86E-02
ZNF503	-0,56	8,06E-02	-0,70	1,10E-02	-0,78	3,00E-03
ZNF703	-0,43	1,84E-01	-0,99	5,40E-04	-0,80	2,18E-03

TABLE S3: EFFECT OF RAR KNOCKDOWN ON RA REGULATION OF RA-RESPONSIVE GENES

	Ctrl		alpha morphants		gamma morphants		full RAR morphants	
Gene	logFC	adj_p,value	logFC	adj_p,value	logFC	adj_p,value	logFC	adj_p,value
AADACL4	1,08	3,41E-02	0,10	1,00E+00	-0,29	1,00E+00	1,98	6,11E-01
ABCB11A	1,42	1,48E-05	-0,53	1,00E+00	-1,14	7,58E-01	-0,45	9,98E-01
ADAMTS16	1,08	2,33E-02	0,08	1,00E+00	-0,53	1,00E+00	0,63	9,98E-01
ADARB1A	1,14	3,55E-06	-0,17	1,00E+00	-0,34	1,00E+00	0,56	9,98E-01
AHCYL1	1,07	2,64E-04	0,81	1,00E+00	0,74	1,00E+00	0,35	9,98E-01
ALDH1A2	-1,38	4,89E-08	-0,56	1,00E+00	-0,91	1,76E-01	-0,50	9,98E-01
APOA4	-2,37	3,97E-03	-1,04	1,00E+00	-0,36	1,00E+00	-0,55	9,98E-01
ASB5B	-1,81	1,71E-04	-0,59	1,00E+00	-0,20	1,00E+00	-0,02	9,99E-01
B2M	-1,43	2,94E-04	0,20	1,00E+00	-0,48	1,00E+00	-0,58	9,98E-01
B3GNT7L	1,65	1,05E-02	1,29	1,00E+00	1,47	9,64E-01	-0,76	9,98E-01
B3GNT9	1,21	3,28E-02	-0,21	1,00E+00	-0,36	1,00E+00	-2,11	7,57E-01
BCAR1	1,82	1,08E-10	0,59	1,00E+00	1,49	2,79E-02	0,75	9,14E-01
BCR	1,38	1,68E-04	0,69	1,00E+00	0,69	1,00E+00	-0,08	9,98E-01
BMI1A	1,20	9,25E-06	-0,19	1,00E+00	0,41	1,00E+00	0,03	9,99E-01
CA9	-1,19	1,68E-07	-0,16	1,00E+00	-0,86	2,17E-01	-0,11	9,98E-01
CAPN12	1,92	6,62E-06	0,84	1,00E+00	0,37	1,00E+00	-1,49	9,71E-01
CDH23	-1,38	3,16E-02	-0,42	1,00E+00	0,05	1,00E+00	-0,06	9,98E-01
CERCAM	1,32	1,91E-08	0,28	1,00E+00	1,78	5,14E-03	1,10	8,10E-01
CHPFA	1,24	8,33E-09	0,52	1,00E+00	1,10	7,91E-02	0,84	5,86E-01
CHRNA5	-1,14	1,57E-02	0,41	1,00E+00	-0,37	1,00E+00	0,10	9,98E-01
COL14A1A	2,14	2,23E-13	1,16	3,12E-01	0,75	8,73E-01	0,71	9,98E-01
COL27A1A	3,36	8,19E-16	0,29	1,00E+00	1,56	9,58E-02	-0,24	9,98E-01
COL27A1B	2,01	2,56E-02	1,04	1,00E+00	1,02	1,00E+00	1,14	9,98E-01
COL6A2	-1,83	3,41E-02	-0,05	1,00E+00	0,14	1,00E+00	-1,07	9,98E-01
CPNE8	-2,25	5,23E-04	-0,69	1,00E+00	0,29	1,00E+00	-0,68	9,98E-01
CSRP2	-1,50	9,71E-07	0,13	1,00E+00	-0,58	1,00E+00	0,08	9,98E-01
CUX2A	1,46	5,03E-06	0,18	1,00E+00	0,71	9,45E-01	0,48	9,98E-01
CYP26B1	5,00	3,72E-20	1,03	1,00E+00	2,41	1,32E-02	0,11	9,98E-01
CYP2P6	1,06	1,68E-03	0,45	1,00E+00	1,15	8,87E-01	-0,46	9,98E-01
DAB2IPA	2,20	2,00E-06	0,05	1,00E+00	0,60	1,00E+00	-0,45	9,98E-01
DACT1	1,41	1,32E-12	0,66	7,47E-01	0,82	8,56E-02	0,12	9,98E-01
DBX1A	-1,48	2,30E-09	-0,28	1,00E+00	-0,74	2,69E-01	-0,22	9,98E-01
DBX1B	2,53	1,69E-04	0,48	1,00E+00	-0,07	1,00E+00	0,20	9,98E-01

DCHS1B	-1,11	7,08E-03	-0,52	1,00E+00	-0,05	1,00E+00	-0,12	9,98E-01
DCN	1,61	5,27E-03	-0,70	1,00E+00	0,38	1,00E+00	0,29	9,98E-01
DHRS3A	4,78	6,26E-16	1,98	8,94E-01	2,46	4,71E-02	0,68	9,98E-01
DIXDC1A	1,55	1,47E-04	-0,65	1,00E+00	0,73	1,00E+00	0,76	9,98E-01
DLB	1,95	1,21E-07	0,46	1,00E+00	0,79	8,57E-01	0,71	9,98E-01
DMRT2B	-1,79	1,75E-02	1,35	1,00E+00	-1,28	1,00E+00	-0,13	9,98E-01
DSE	2,06	9,90E-11	-0,17	1,00E+00	1,07	7,61E-01	0,02	9,99E-01
DSEL	-1,68	2,89E-02	0,33	1,00E+00	-1,21	1,00E+00	1,58	9,98E-01
DYRK4	1,84	8,21E-04	1,09	1,00E+00	1,50	9,18E-01	-0,14	9,98E-01
DYSF	1,71	6,87E-04	-0,40	1,00E+00	0,88	1,00E+00	0,28	9,98E-01
EIF2AK2	-1,24	3,44E-02	1,67	1,00E+00	0,19	1,00E+00	-1,56	9,72E-01
ELAVL3	2,74	2,26E-08	-0,11	1,00E+00	1,43	7,61E-01	-0,79	9,98E-01
ENO4	2,33	1,98E-09	1,29	8,94E-01	2,78	1,48E-03	1,21	5,92E-01
EPHA7	2,12	1,26E-08	0,60	1,00E+00	0,15	1,00E+00	0,66	9,98E-01
ESRRA	1,19	7,14E-08	0,79	1,00E+00	1,00	2,64E-01	0,34	9,98E-01
ESRRGB	-1,41	3,80E-02	-0,43	1,00E+00	-1,26	9,94E-01	-0,32	9,98E-01
F8	-1,99	1,13E-02	-2,49	1,00E+00	-0,85	1,00E+00	-0,81	9,98E-01
FAM101A	2,06	6,18E-04	0,68	1,00E+00	1,31	1,00E+00	-0,98	9,98E-01
FAM167AB	-1,08	2,11E-02	-0,58	1,00E+00	0,46	1,00E+00	-1,02	9,98E-01
FEZF2	-2,67	7,03E-08	-0,52	1,00E+00	-1,17	8,89E-01	-0,05	9,99E-01
FGFRL1B	1,42	3,44E-05	0,05	1,00E+00	0,48	1,00E+00	-1,13	9,98E-01
FLRT2	-1,66	1,50E-03	-0,28	1,00E+00	-0,22	1,00E+00	-0,15	9,98E-01
FOXD3	-1,07	5,81E-08	-0,09	1,00E+00	-0,43	8,11E-01	0,06	9,98E-01
FOXG1B	1,86	1,82E-04	0,88	1,00E+00	0,86	1,00E+00	-0,14	9,98E-01
FOXI3A	1,47	2,21E-03	-0,21	1,00E+00	0,89	1,00E+00	-0,48	9,98E-01
FOXJ1B	-1,42	1,88E-02	-0,17	1,00E+00	-0,40	1,00E+00	-0,58	9,98E-01
FOXN4	2,50	2,81E-02	0,69	1,00E+00	0,85	1,00E+00	1,07	9,98E-01
FTR06	2,23	5,96E-07	0,86	1,00E+00	2,63	6,20E-02	0,72	9,98E-01
GABRP	-1,71	6,90E-03	-1,14	1,00E+00	-0,34	1,00E+00	-0,36	9,98E-01
GADD45AB	1,12	3,06E-02	-0,53	1,00E+00	-1,39	7,99E-01	0,00	1,00E+00
GDF10B	-2,46	3,32E-04	-0,31	1,00E+00	-1,62	9,86E-01	-1,44	9,98E-01
GNA12A	1,22	4,24E-05	0,12	1,00E+00	0,36	1,00E+00	-0,92	9,98E-01
GRHL2B	1,17	5,10E-04	0,89	1,00E+00	0,81	1,00E+00	-0,02	9,99E-01
GSTM3	1,12	4,91E-03	0,84	1,00E+00	1,00	7,78E-01	0,43	9,98E-01
GUCY1A3	-1,84	1,10E-02	-0,48	1,00E+00	0,38	1,00E+00	-0,13	9,98E-01
HAND2	3,65	2,04E-07	1,96	1,00E+00	0,43	1,00E+00	-1,30	9,98E-01
HAO2	1,20	1,38E-02	-0,94	1,00E+00	-0,32	1,00E+00	-1,05	9,98E-01

STARD13A	2,37	1,58E-07	3,30	1,88E-02	1,33	6,70E-01	1,38	8,54E-01
HAPLN1B	2,45	5,35E-12	1,45	6,73E-01	1,94	1,96E-02	0,16	9,98E-01
HBEGFA	1,02	2,89E-02	0,01	1,00E+00	0,48	1,00E+00	0,25	9,98E-01
HER3	-1,40	5,03E-06	0,03	1,00E+00	-1,44	3,52E-02	-0,52	9,98E-01
HESX1	-2,17	7,42E-07	-0,30	1,00E+00	-1,02	6,97E-01	-0,37	9,98E-01
HIC1	1,25	1,40E-05	-0,04	1,00E+00	0,29	1,00E+00	-0,02	9,99E-01
HNF1BA	1,58	1,77E-13	0,55	1,00E+00	0,51	6,35E-01	0,17	9,98E-01
HNF1BB	2,07	5,80E-05	-0,20	1,00E+00	2,75	3,82E-01	-0,35	9,98E-01
HOOK1	1,73	1,91E-04	-0,63	1,00E+00	-1,24	9,50E-01	-0,59	9,98E-01
HOXA1A	3,32	1,63E-06	0,68	1,00E+00	0,72	1,00E+00	-0,13	9,98E-01
HOXA3A	4,25	8,37E-22	1,34	1,00E+00	3,35	6,24E-05	1,20	9,46E-01
HOXA4A	4,50	9,96E-21	-0,09	1,00E+00	3,01	5,59E-04	0,13	9,98E-01
HOXA5A	3,98	1,66E-15	2,23	9,15E-01	1,82	4,34E-01	-0,66	9,98E-01
HOXB1A	4,00	3,62E-19	0,16	1,00E+00	1,70	1,39E-01	0,51	9,98E-01
HOXB2A	3,27	7,59E-06	0,68	1,00E+00	1,74	1,00E+00	-0,29	9,98E-01
HER13	3,30	2,79E-13	3,63	4,06E-03	2,71	2,70E-02	1,16	9,98E-01
HOXB1B	2,05	7,00E-13	1,79	6,34E-03	2,10	1,42E-04	0,63	9,98E-01
RARAA	1,84	7,62E-12	1,97	1,74E-04	1,56	1,88E-03	1,54	2,68E-02
MEIS2A	1,79	9,75E-09	2,18	9,35E-03	2,69	2,74E-04	1,03	8,69E-01
CYP26A1	1,68	2,64E-15	1,30	2,25E-04	2,02	8,89E-09	1,64	3,65E-06
DHRS3B	1,34	9,65E-10	1,25	1,93E-02	1,62	1,98E-04	1,06	4,15E-01
HOXB3A	3,16	6,26E-16	1,01	1,00E+00	2,94	8,59E-05	-0,18	9,98E-01
HOXB4A	3,58	2,27E-14	2,08	7,47E-01	2,57	3,52E-02	0,31	9,98E-01
HOXB5A	3,45	1,96E-14	2,35	6,22E-01	2,20	2,57E-01	-0,24	9,98E-01
HOXB5B	3,88	8,37E-22	1,26	8,94E-01	2,33	1,98E-04	0,09	9,98E-01
HOXB6B	3,12	1,77E-11	-1,07	1,00E+00	1,23	8,92E-01	0,53	9,98E-01
HOXB8B	2,65	2,25E-06	2,26	1,00E+00	1,62	9,12E-01	0,57	9,98E-01
HOXC1A	3,26	1,59E-11	1,42	1,00E+00	3,12	4,45E-02	-0,48	9,98E-01
HOXC4A	2,90	1,23E-04	1,93	1,00E+00	0,36	1,00E+00	-0,93	9,98E-01
HOXC5A	1,97	1,34E-05	1,98	1,00E+00	0,39	1,00E+00	-0,92	9,98E-01
HOXC6A	1,01	1,75E-03	0,80	1,00E+00	0,89	9,36E-01	0,43	9,98E-01
HOXD3A	2,70	1,96E-03	0,68	1,00E+00	0,45	1,00E+00	-0,13	9,98E-01
HOXD4A	4,86	6,68E-16	-0,19	1,00E+00	2,76	6,20E-02	0,13	9,98E-01
HTR1BD	-1,60	1,82E-03	-1,20	1,00E+00	0,15	1,00E+00	1,00	9,98E-01
IGSF21B	2,08	1,19E-02	0,64	1,00E+00	0,17	1,00E+00	0,72	9,98E-01
ILDR2	1,23	1,02E-03	0,62	1,00E+00	-0,58	1,00E+00	-1,01	9,98E-01
IM:7137453	1,80	2,91E-13	0,49	1,00E+00	1,26	1,96E-02	0,38	9,98E-01

IM:7154842	-1,60	3,32E-02	0,68	1,00E+00	1,55	8,95E-01	-2,56	5,33E-01
IM:7162084	2,08	2,92E-09	0,13	1,00E+00	1,32	7,01E-01	1,06	9,98E-01
INHBB	1,29	8,80E-05	1,03	1,00E+00	1,13	8,11E-01	-0,44	9,98E-01
INSM1B	2,16	7,00E-13	1,10	5,52E-01	2,77	7,49E-06	1,49	3,79E-01
IRX3B	2,79	9,45E-07	0,31	1,00E+00	2,67	1,39E-01	1,33	9,98E-01
IRX5A	1,15	7,58E-03	0,74	1,00E+00	-0,59	1,00E+00	-1,21	9,98E-01
IRX7	-1,10	3,96E-11	0,06	1,00E+00	-0,49	4,42E-01	-0,26	9,98E-01
JAK2A	1,90	2,90E-09	0,85	1,00E+00	0,32	1,00E+00	-0,49	9,98E-01
KANK1	1,67	5,75E-11	0,39	1,00E+00	0,28	1,00E+00	0,09	9,98E-01
KIRRELB	2,15	6,89E-03	1,14	1,00E+00	1,27	1,00E+00	-0,81	9,98E-01
KITLGA	1,10	5,21E-04	1,12	1,00E+00	0,94	8,66E-01	0,64	9,98E-01
KLF12B	1,99	2,96E-05	1,76	8,94E-01	1,98	3,02E-01	0,20	9,98E-01
KLHL5	-1,88	2,04E-02	-0,41	1,00E+00	0,46	1,00E+00	-0,56	9,98E-01
LDB2B	2,56	2,72E-14	0,14	1,00E+00	0,29	1,00E+00	0,31	9,98E-01
LHX5	-1,17	5,74E-05	-0,27	1,00E+00	-0,40	1,00E+00	0,01	9,99E-01
LINGO1B	-2,52	2,10E-03	0,59	1,00E+00	2,05	8,32E-01	1,33	9,98E-01
LOC100001239	-1,67	2,09E-03	-0,31	1,00E+00	-2,55	1,76E-01	-0,56	9,98E-01
LOC100003866	3,44	1,63E-08	-0,78	1,00E+00	3,79	1,96E-02	1,45	9,98E-01
LOC100004296	-1,25	6,49E-05	-0,14	1,00E+00	-0,45	1,00E+00	-0,09	9,98E-01
LOC100147851	3,02	1,00E-04	1,88	1,00E+00	0,48	1,00E+00	-0,25	9,98E-01
LOC100148343	2,12	1,37E-02	0,68	1,00E+00	-0,38	1,00E+00	-0,13	9,98E-01
LOC100148864	1,30	1,88E-05	1,48	9,15E-01	0,39	1,00E+00	-0,27	9,98E-01
LOC100330160	1,77	3,35E-07	0,63	1,00E+00	1,19	8,11E-01	0,57	9,98E-01
LOC100330657	1,69	8,23E-08	-1,15	1,00E+00	1,34	6,73E-01	0,24	9,98E-01
LOC100331241	2,33	6,50E-03	0,68	1,00E+00	-0,33	1,00E+00	-0,92	9,98E-01
LOC100332380	-2,30	7,22E-03	0,29	1,00E+00	0,20	1,00E+00	0,74	9,98E-01
LOC100332539	2,74	5,81E-08	2,47	9,15E-01	-0,39	1,00E+00	-0,98	9,98E-01
LOC100333111	2,12	1,58E-08	0,96	1,00E+00	1,30	8,11E-01	-0,05	9,99E-01
LOC100333651	1,47	8,51E-11	0,61	1,00E+00	1,06	1,39E-01	0,63	9,20E-01
LOC100333973	2,36	4,89E-06	-0,86	1,00E+00	2,38	2,92E-01	-0,50	9,98E-01
LOC100334188	-2,60	1,47E-03	-1,83	1,00E+00	-2,70	4,21E-01	-1,82	9,98E-01
LOC100334713	-0,35	9,09E-01	-1,85	1,00E+00	0,26	1,00E+00	-0,23	9,98E-01
LOC100334928	2,09	4,81E-03	0,68	1,00E+00	1,24	1,00E+00	1,01	9,98E-01
LOC100535433	3,52	1,17E-14	2,17	3,12E-01	1,99	7,91E-02	0,34	9,98E-01
LOC100535716	1,45	6,72E-03	-1,43	1,00E+00	0,77	1,00E+00	0,05	9,99E-01
LOC100536197	-1,04	4,04E-03	-0,69	1,00E+00	0,13	1,00E+00	-0,83	9,98E-01
LOC100537029	2,58	1,03E-03	1,12	1,00E+00	1,67	9,34E-01	0,26	9,98E-01

LOC100537480	3,20	1,59E-10	3,36	1,99E-01	1,52	6,75E-01	2,16	6,40E-01
LOC100537717	1,01	4,43E-02	-0,22	1,00E+00	1,64	7,78E-01	0,31	9,98E-01
LOC100538119	1,87	1,73E-02	1,04	1,00E+00	0,39	1,00E+00	-1,04	9,98E-01
LOC100538211	2,36	9,16E-13	0,55	1,00E+00	1,22	4,90E-01	-0,54	9,98E-01
LOC100538245	1,75	2,61E-03	-1,06	1,00E+00	1,13	9,92E-01	-0,12	9,98E-01
LOC556476	1,38	1,11E-04	0,36	1,00E+00	1,66	3,60E-01	0,89	9,98E-01
LOC556872	-1,50	1,24E-02	-0,46	1,00E+00	-2,62	2,47E-01	-0,86	9,98E-01
LOC560775	-1,13	1,25E-05	-0,34	1,00E+00	-0,79	5,21E-01	-0,03	9,98E-01
LOC561355	-2,20	2,17E-04	-1,60	9,46E-01	-1,47	4,35E-01	-0,56	9,98E-01
LOC561565	-2,00	7,91E-03	-0,52	1,00E+00	-2,30	5,05E-01	1,88	9,10E-01
LOC562320	-2,96	1,22E-05	1,56	1,00E+00	-0,32	1,00E+00	-0,92	9,98E-01
LOC562665	2,05	5,76E-09	0,91	1,00E+00	2,09	6,20E-02	1,13	9,30E-01
LOC563500	2,40	9,21E-03	1,26	1,00E+00	2,06	8,03E-01	-1,34	9,98E-01
LOC565776	-1,44	2,83E-02	-0,10	1,00E+00	-0,16	1,00E+00	1,19	9,98E-01
LOC566411	3,39	5,37E-07	2,45	1,00E+00	2,20	5,08E-01	2,30	8,16E-01
LOC566479	1,32	7,05E-04	-1,42	1,00E+00	-0,57	1,00E+00	-0,38	9,98E-01
LOC567790	-2,05	7,41E-03	-2,14	1,00E+00	-3,18	3,05E-01	-1,41	9,98E-01
LOC568110	4,00	1,65E-07	0,92	1,00E+00	1,79	8,70E-01	-0,01	1,00E+00
LOC569124	1,32	1,02E-03	0,78	1,00E+00	-0,22	1,00E+00	-0,11	9,98E-01
LOC569773	2,87	3,76E-04	-0,01	1,00E+00	1,04	1,00E+00	1,02	9,98E-01
LOC569861	-1,32	4,67E-04	-0,32	1,00E+00	-0,06	1,00E+00	-0,25	9,98E-01
LOC793240	-1,38	4,13E-02	-0,51	1,00E+00	-1,14	9,54E-01	0,44	9,98E-01
LOC794170	1,61	1,66E-06	-0,03	1,00E+00	1,18	6,91E-01	-0,02	9,99E-01
LOC796490	2,15	3,07E-07	0,54	1,00E+00	-0,42	1,00E+00	-0,41	9,98E-01
LOC799595	2,30	1,08E-11	0,89	1,00E+00	1,69	2,07E-01	0,28	9,98E-01
LRIT1A	3,02	2,52E-06	1,48	1,00E+00	2,24	6,37E-01	0,03	9,99E-01
LRRC2	3,40	3,37E-04	0,29	1,00E+00	0,33	1,00E+00	-0,71	9,98E-01
LRRN1	1,92	3,25E-13	0,52	1,00E+00	0,85	4,34E-01	0,43	9,98E-01
LRRTM1	-2,09	3,68E-02	-0,82	1,00E+00	0,49	1,00E+00	-0,89	9,98E-01
MAP1B	-1,23	2,74E-03	-0,76	1,00E+00	-1,67	2,63E-01	-0,83	9,98E-01
MAP3K10	2,31	8,36E-04	0,68	1,00E+00	-0,15	1,00E+00	0,63	9,98E-01
MASP1	-1,24	1,13E-03	-0,17	1,00E+00	-0,55	1,00E+00	0,26	9,98E-01
MEIS1	1,96	2,80E-10	0,90	1,00E+00	2,00	1,02E-02	0,17	9,98E-01
MEIS3	1,83	4,38E-10	1,24	5,84E-01	1,72	1,28E-02	0,10	9,98E-01
MET	1,24	1,59E-03	-0,84	1,00E+00	0,40	1,00E+00	-0,10	9,98E-01
MMP14A	1,45	4,24E-04	0,88	1,00E+00	0,47	1,00E+00	0,35	9,98E-01
MNX1	1,34	5,60E-03	2,90	4,50E-01	1,64	8,52E-01	0,68	9,98E-01

MPPED2	1,52	1,37E-11	-0,29	1,00E+00	1,36	6,11E-03	0,04	9,98E-01
MTMR8	1,09	6,39E-04	0,94	1,00E+00	0,57	1,00E+00	-0,15	9,98E-01
MXRA8A	1,70	5,81E-08	0,63	1,00E+00	0,47	1,00E+00	-0,40	9,98E-01
MYO10	1,42	1,86E-06	0,62	1,00E+00	0,50	1,00E+00	0,20	9,98E-01
MYOD1	-2,53	3,77E-04	-1,58	1,00E+00	-1,42	9,83E-01	-1,26	9,98E-01
NAB1A	-2,08	9,23E-03	1,56	1,00E+00	0,36	1,00E+00	-0,57	9,98E-01
NAV3	2,23	3,80E-08	0,34	1,00E+00	-0,36	1,00E+00	0,04	9,99E-01
NEUROG1	2,67	7,84E-10	-0,48	1,00E+00	1,79	5,10E-01	0,73	9,98E-01
NFE2L1A	1,40	1,19E-04	0,38	1,00E+00	1,25	5,76E-01	0,69	9,98E-01
NFIA	3,28	5,47E-10	1,11	1,00E+00	2,10	6,86E-01	-0,31	9,98E-01
NHSA	2,95	5,46E-13	1,58	7,50E-01	1,96	4,71E-02	1,39	7,37E-01
NLGN4A	-1,85	4,28E-03	0,16	1,00E+00	-3,23	3,98E-02	-0,36	9,98E-01
NPR3	-1,00	3,07E-03	-0,13	1,00E+00	-0,53	1,00E+00	0,00	1,00E+00
NPTX2B	2,19	2,94E-04	0,53	1,00E+00	0,71	1,00E+00	-1,67	9,98E-01
NR2F1A	3,29	1,00E-10	0,78	1,00E+00	1,05	1,00E+00	-0,46	9,98E-01
NR2F1B	2,52	9,63E-05	-1,20	1,00E+00	-0,44	1,00E+00	0,69	9,98E-01
NR2F5	3,07	1,13E-14	1,42	1,00E+00	2,38	2,70E-02	-0,78	9,98E-01
NR2F6B	1,74	4,70E-12	0,25	1,00E+00	0,96	2,03E-01	0,07	9,98E-01
NR6A1B	4,61	8,87E-11	-0,35	1,00E+00	1,90	6,27E-01	0,12	9,98E-01
NRIP1B	1,98	5,91E-15	0,38	1,00E+00	0,73	3,58E-01	-0,21	9,98E-01
NRP1A	1,02	4,81E-06	0,27	1,00E+00	1,11	1,84E-01	0,76	9,10E-01
OLIG2	1,35	3,14E-02	1,35	1,00E+00	1,49	8,66E-01	-0,11	9,98E-01
OLIG4	1,23	9,71E-07	0,74	1,00E+00	1,33	1,03E-01	-0,26	9,98E-01
OTX2	-1,07	4,06E-05	0,10	1,00E+00	0,57	9,45E-01	0,34	9,98E-01
PAX2A	1,78	1,20E-03	-0,11	1,00E+00	-1,19	1,00E+00	0,56	9,98E-01
PAX3A	1,32	1,25E-05	1,65	5,62E-01	1,11	8,11E-01	0,21	9,98E-01
PAX8	2,65	8,84E-09	0,17	1,00E+00	0,98	1,00E+00	-0,30	9,98E-01
PCDH18A	1,55	2,55E-07	1,23	7,47E-01	1,52	4,71E-02	0,78	9,98E-01
PDGFAA	1,85	5,88E-06	3,39	6,42E-02	1,56	4,34E-01	1,54	8,92E-01
PDZRN3	2,94	1,67E-06	3,81	4,50E-01	3,54	1,42E-01	0,31	9,98E-01
PFKFB2A	2,20	4,96E-05	-0,29	1,00E+00	1,30	9,55E-01	-0,24	9,98E-01
PGP	1,32	2,82E-06	0,54	1,00E+00	1,24	4,90E-01	-0,47	9,98E-01
PLOD2	1,18	3,62E-06	0,36	1,00E+00	0,98	8,21E-01	-0,32	9,98E-01
POU2F2A	1,96	2,48E-12	0,93	1,00E+00	0,95	4,90E-01	0,85	9,10E-01
PPP1R13BA	2,30	1,63E-06	0,49	1,00E+00	0,16	1,00E+00	-0,43	9,98E-01
PRDM12B	1,32	7,04E-04	0,69	1,00E+00	0,36	1,00E+00	0,86	9,98E-01
PRDM13	1,71	3,94E-02	-1,28	1,00E+00	0,00	1,00E+00	1,46	9,98E-01

PTGDSB	-1,36	2,35E-07	-0,25	1,00E+00	-0,66	8,14E-01	-0,08	9,98E-01
PTPRDB	1,07	1,22E-04	-0,14	1,00E+00	0,58	1,00E+00	-0,13	9,98E-01
RAB42A	-1,67	2,24E-03	-1,80	9,75E-01	-0,77	1,00E+00	0,53	9,98E-01
RABL6	1,23	1,00E-07	-0,03	1,00E+00	1,20	2,70E-01	0,17	9,98E-01
RASL11B	-1,39	2,01E-06	-0,28	1,00E+00	-0,36	1,00E+00	-0,04	9,98E-01
RBMS3	2,87	1,08E-10	0,83	1,00E+00	1,14	8,27E-01	0,74	9,98E-01
REM1	1,08	6,50E-05	0,55	1,00E+00	0,38	1,00E+00	0,58	9,98E-01
RGMD	1,16	6,08E-09	0,68	8,94E-01	0,22	1,00E+00	0,26	9,98E-01
RGRA	1,97	8,63E-04	-0,47	1,00E+00	0,14	1,00E+00	-0,14	9,98E-01
RHBDL3	1,07	7,04E-03	-1,05	1,00E+00	0,95	8,01E-01	0,02	9,99E-01
RHOUB	1,73	1,17E-06	-0,02	1,00E+00	-0,30	1,00E+00	-1,17	9,98E-01
RIPPLY3	3,30	1,83E-08	1,46	1,00E+00	2,30	3,53E-01	0,36	9,98E-01
RXRGA	1,77	2,96E-05	0,51	1,00E+00	1,55	3,84E-01	0,86	9,98E-01
RXRGB	1,31	6,69E-04	0,87	1,00E+00	1,24	9,20E-01	-0,24	9,98E-01
SALL3B	-1,09	8,41E-04	-0,49	1,00E+00	-0,96	4,64E-01	-0,02	9,99E-01
SCN8AA	-1,63	4,78E-02	1,08	1,00E+00	1,70	8,73E-01	0,44	9,98E-01
SCOSPONDIN	1,49	3,23E-03	-0,37	1,00E+00	0,61	1,00E+00	-0,15	9,98E-01
SEBOX	-1,25	3,90E-02	-1,02	1,00E+00	-0,85	9,36E-01	0,13	9,98E-01
SEMA3BL	2,05	7,58E-06	0,49	1,00E+00	2,67	2,41E-01	1,98	8,69E-01
SEPT3	3,88	1,28E-12	0,72	1,00E+00	1,24	8,52E-01	0,19	9,98E-01
SEPT9B	-1,54	3,86E-02	1,26	1,00E+00	0,42	1,00E+00	-1,19	9,98E-01
SFRP2	3,19	3,03E-06	1,97	1,00E+00	0,06	1,00E+00	0,26	9,98E-01
SGK2B	2,05	2,17E-04	-0,46	1,00E+00	0,26	1,00E+00	-0,27	9,98E-01
SHDA	1,18	2,87E-07	0,09	1,00E+00	1,04	2,64E-01	-0,45	9,98E-01
SI:CH211-103A14.4	1,44	4,17E-05	0,67	1,00E+00	0,69	1,00E+00	-0,48	9,98E-01
SI:CH211-137A8.2	1,21	1,65E-02	-0,63	1,00E+00	0,15	1,00E+00	0,31	9,98E-01
SI:CH211-212O1.2	2,41	3,93E-09	0,08	1,00E+00	1,58	2,47E-01	-0,02	9,99E-01
SI:CH211-217A12.1	1,08	2,10E-02	-0,11	1,00E+00	1,31	9,31E-01	1,24	9,98E-01
SI:CH211-55A7.1	-1,85	2,21E-02	-0,25	1,00E+00	-0,88	1,00E+00	0,09	9,98E-01
SI:DKEY-261E22.1	1,15	1,00E-04	-0,28	1,00E+00	1,36	4,20E-01	1,14	8,69E-01
SI:DKEY-269I1.3	-2,24	3,48E-03	-2,02	1,00E+00	-0,36	1,00E+00	-0,21	9,98E-01
SI:DKEY-26G8.4	-1,96	1,75E-02	-2,02	1,00E+00	-0,36	1,00E+00	-0,20	9,98E-01
SI:DKEY-27I16.2	-1,93	7,62E-07	-0,64	1,00E+00	-1,02	6,13E-01	-0,36	9,98E-01
SI:DKEYP-80C12.7	2,32	6,32E-09	0,87	1,00E+00	2,36	1,39E-01	-0,04	9,99E-01
SIX3A	-1,06	1,70E-02	-0,45	1,00E+00	1,10	9,18E-01	0,44	9,98E-01
SIX7	-2,16	6,47E-10	-0,75	1,00E+00	-0,53	1,00E+00	-0,15	9,98E-01
SKAP1	3,57	1,35E-12	2,12	4,99E-01	3,00	1,72E-02	0,54	9,98E-01

SKIB	1,28	1,98E-09	0,32	1,00E+00	1,24	1,96E-02	0,66	8,98E-01
SLC22A18	-1,59	1,49E-03	0,18	1,00E+00	0,13	1,00E+00	1,07	9,98E-01
SLC35F3A	-1,83	2,24E-03	-0,29	1,00E+00	-0,48	1,00E+00	0,66	9,98E-01
SLIT3	-1,99	1,22E-02	-2,29	1,00E+00	-2,31	5,41E-01	-0,92	9,98E-01
SMAD3A	-1,13	2,89E-02	0,10	1,00E+00	-0,24	1,00E+00	-0,18	9,98E-01
SNCGA	1,16	2,22E-03	-0,02	1,00E+00	1,71	1,25E-01	0,95	9,70E-01
SOCS3A	-1,49	2,49E-02	-0,16	1,00E+00	-2,11	4,99E-01	-1,90	9,10E-01
SOX10	1,76	8,67E-05	1,90	1,00E+00	1,33	8,66E-01	-0,02	9,99E-01
SOX2	-1,14	1,02E-05	-0,36	1,00E+00	-0,85	5,08E-01	0,26	9,98E-01
SOX9A	1,75	6,15E-10	-0,13	1,00E+00	-0,49	1,00E+00	0,07	9,98E-01
SP5	-1,41	3,00E-12	-0,13	1,00E+00	-0,53	4,99E-01	0,03	9,98E-01
SP7	1,52	1,13E-05	0,55	1,00E+00	0,35	1,00E+00	-0,05	9,99E-01
SP8B	-1,68	9,18E-10	-0,41	1,00E+00	-1,13	1,28E-01	-0,20	9,98E-01
ST3GAL4	2,36	1,65E-03	1,07	1,00E+00	0,35	1,00E+00	-1,35	9,98E-01
ST8SIA2	-1,28	4,95E-02	-0,48	1,00E+00	-0,50	1,00E+00	-0,22	9,98E-01
SYN1	-1,19	1,65E-02	-0,64	1,00E+00	-1,07	8,57E-01	-0,86	9,98E-01
TAGLN3B	-1,49	4,64E-03	-2,59	3,10E-01	-0,14	1,00E+00	-0,93	9,98E-01
TBX2B	-1,29	2,23E-02	-0,04	1,00E+00	-0,72	1,00E+00	0,49	9,98E-01
THSD7A	2,11	2,30E-02	0,92	1,00E+00	2,28	6,91E-01	-1,30	9,98E-01
TIPARP	1,11	9,39E-10	0,44	1,00E+00	0,73	1,57E-01	0,34	9,98E-01
TMEM30C	2,45	3,17E-04	-0,50	1,00E+00	-0,34	1,00E+00	-1,12	9,98E-01
TMEM88A	2,14	7,97E-03	1,75	1,00E+00	-0,59	1,00E+00	0,18	9,98E-01
TMEM88B	3,36	7,71E-07	0,62	1,00E+00	1,03	1,00E+00	-0,74	9,98E-01
TNFSF10L3	-1,14	3,27E-02	-0,93	1,00E+00	-0,28	1,00E+00	-0,31	9,98E-01
TNS1	1,56	1,82E-04	0,43	1,00E+00	-0,78	1,00E+00	-0,77	9,98E-01
TP63	1,09	7,33E-07	0,52	1,00E+00	0,57	9,50E-01	-0,08	9,98E-01
TSHZ1	1,03	8,37E-09	0,21	1,00E+00	0,31	1,00E+00	0,21	9,98E-01
TTLL3	-1,84	1,56E-02	0,33	1,00E+00	0,39	1,00E+00	1,11	9,98E-01
UGT5G1	2,68	1,77E-13	1,65	3,12E-01	0,64	1,00E+00	-0,35	9,98E-01
VAX2	3,00	2,24E-03	1,51	1,00E+00	-0,35	1,00E+00	-0,13	9,98E-01
VEGFC	1,76	7,04E-03	0,56	1,00E+00	-0,02	1,00E+00	0,87	9,98E-01
WDR90	1,90	2,99E-06	-0,89	1,00E+00	0,04	1,00E+00	-0,07	9,98E-01
WNT11R	2,56	7,56E-07	-0,50	1,00E+00	1,24	9,31E-01	-1,12	9,98E-01
WNT4B	2,40	1,69E-04	-0,88	1,00E+00	-0,34	1,00E+00	1,44	9,98E-01
WNT8A	-1,02	1,13E-07	-0,14	1,00E+00	-0,33	1,00E+00	-0,18	9,98E-01
WU:FB16H09	1,22	2,25E-06	0,55	1,00E+00	0,74	7,07E-01	0,66	9,46E-01
WU:FB63A08	1,75	2,89E-03	-0,04	1,00E+00	-0,21	1,00E+00	-0,52	9,98E-01

YWHAG1	-1,04	3,87E-02	0,49	1,00E+00	-0,95	9,86E-01	0,82	9,98E-01
ZGC:112375	-1,34	2,79E-02	-0,68	1,00E+00	0,40	1,00E+00	1,07	9,98E-01
ZGC:112435	1,03	7,38E-04	-0,10	1,00E+00	0,60	1,00E+00	0,25	9,98E-01
ZGC:112466	3,27	1,69E-04	0,42	1,00E+00	1,58	9,94E-01	1,10	9,98E-01
ZGC:113425	-1,50	1,78E-02	-0,06	1,00E+00	0,63	1,00E+00	0,14	9,98E-01
ZGC:114118	-1,87	1,15E-07	-0,57	1,00E+00	-1,01	4,34E-01	-0,70	9,98E-01
ZGC:123298	-1,04	5,93E-03	0,69	1,00E+00	-0,66	1,00E+00	0,24	9,98E-01
ZGC:123339	-1,09	4,52E-02	-0,25	1,00E+00	0,52	1,00E+00	0,59	9,98E-01
ZGC:136336	1,15	2,41E-02	0,09	1,00E+00	1,50	8,53E-01	-0,55	9,98E-01
ZGC:153154	-1,12	1,01E-04	-0,28	1,00E+00	-0,76	8,47E-01	-0,30	9,98E-01
ZGC:153720	2,09	1,36E-09	0,61	1,00E+00	0,87	8,11E-01	0,11	9,98E-01
ZGC:158824	1,16	5,98E-07	-0,17	1,00E+00	-0,16	1,00E+00	-0,51	9,98E-01
ZGC:162623	-1,91	1,72E-02	0,68	1,00E+00	0,90	1,00E+00	-0,13	9,98E-01
ZGC:175176	3,33	1,68E-10	2,05	7,47E-01	2,13	3,05E-01	-0,30	9,98E-01
ZGC:194209	1,34	4,02E-06	-0,60	1,00E+00	-0,10	1,00E+00	-0,03	9,99E-01
ZIC1	1,80	3,60E-10	0,94	1,00E+00	2,80	6,24E-05	0,98	8,52E-01
ZIC4	1,58	1,97E-05	1,37	1,00E+00	2,83	6,77E-03	1,59	8,16E-01
ZNF503	1,14	1,59E-10	0,83	2,52E-01	0,84	4,77E-02	0,34	9,98E-01
ZNF703	1,04	5,08E-10	0,87	1,14E-01	1,23	4,82E-04	0,56	7,07E-01

IV. Travaux complémentaires - Discussion – Perspectives

Des expériences en cours dans le laboratoire de Cécile Rochette-Egly sur des cellules ES de souris traitées à l'AR ont également montré que les différents sous-types de RAR activent des gènes différents. En effet, les cellules ES expriment les sous-types RARα et RARγ et parmi l'ensemble des gènes répondant à l'AR, la moitié nécessite spécifiquement RARγ pour leur régulation par l'AR (données de RNAseq couplées à du ChIPseq (Al Tanoury et al., en préparation). De manière intéressante, il semble que les gènes régulés par RARγ soient impliqués dans la différenciation neuronale induite par l'AR, alors que RARα régule les autres gènes cibles de l'AR. Ainsi, ces données récentes appuient nos travaux en montrant que dans un système cellulaire exprimant plusieurs sous-types de RAR, il existe une activité transcriptionnelle propre à chacun.

1. Comment est régulée la spécificité d'action des sous-types de RAR ?

Les transcriptotypes spécifiques des différents sous-types de RAR que nous avons décrit dans nos résultats sont la conséquence de différences transcriptionnelles (page 101 et figure 2 page 100). Quels sont les mécanismes permettant le recrutement spécifique d'un sous-type de RAR au niveau du promoteur d'un gène cible en particulier ? Comment les différents sous-types de RAR, bien qu'exprimés dans les mêmes cellules, régulent des répertoires de gènes différents ?

1.1. Régulation par la séquence d'ADN ?

C'en en premier lieu la présence d'un RARE dans le promoteur d'un gène qui est à l'origine du recrutement des RAR pour la régulation de son expression. Bien qu'un consensus de RARE canonique ait été établi, à savoir la répétition directe du motif RGKTSA séparé par 5 nucléotides, la description des RARE fonctionnels *in vivo* montre une bien plus grande diversité de motifs et d'espacements (section A-III-2, page 12). Parmi ces différences, existe-t-il un motif ou un espacement en particulier recrutant spécifiquement un sous-type de récepteur ? L'analyse *in silico* de l'ensemble des RARE consensus trouvés dans le génome du poisson-zèbre, aux abords des gènes répondant à l'AR, ne montre aucune corrélation entre la spécificité d'action d'un sous-type et l'enrichissement d'un espacement en particulier ou d'une séquence consensus spécifique. En d'autres termes, ces résultats

suggèrent que le recrutement spécifique des différents sous-types de RAR ne repose pas sur des différences dans la séquence brute des éléments de réponse au niveau de l'ADN. Toutefois, les RARE *in silico* pris en compte dans notre étude ont été obtenus par la recherche de motifs RGKTSA du type DR1-10 ou IR1-10 dans l'ensemble du génome du poisson-zèbre par la plateforme PARSEC. Ainsi, ces éléments de réponse ne correspondent pas nécessairement aux RARE effectivement fonctionnels régulant l'expression des gènes cibles de l'AR. Pour avoir accès à ces derniers, il serait nécessaire d'analyser par ChIP-seq le recrutement différentiel des différents sous-types de RAR *in vivo* et ainsi décrire des cistromes spécifiques pour chacun des sous-types. Ces travaux nécessitent des anticorps spécifiques de chaque sous-type de RAR de poisson-zèbre et capables d'immunoprécipiter la protéine *in vivo*. Ces travaux permettraient d'avoir un aperçu de l'ensemble des sites de recrutement différentiels des différents sous-types de RAR (cistromes spécifiques de chaque sous-type) et ainsi de corréler d'éventuelles spécificités de séquence et/ou d'espacement au niveau des RARE et le recrutement spécifique d'un sous-type de RAR.

1.2. Régulation au niveau du récepteur ?

i. RAR ?

Les différences majeures dans la séquence protéique des sous-types de RAR sont retrouvées dans le NTD et la partie C-terminale. Mais comment de telles régions non structurées peuvent influer sur le recrutement du récepteur à l'ADN ? Des études récentes suggèrent que de part sa proximité avec le DBD, le NTD pourrait influencer l'interaction du récepteur avec l'ADN. En particulier, l'équipe de Cécile Rochette-Egly a montré que certaines protéines du cytosquelette interagissent spécifiquement avec un sous-type de RAR et peuvent réguler leur recrutement au niveau de l'ADN. C'est le cas de la vinexine beta qui interagit spécifiquement avec le NTD de RARγ et empêche son interaction *in vitro* avec un RARE, et cette interaction est abolie par la phosphorylation du NTD de RARγ, (publication en annexe, page 195). D'autre part, des études récentes du laboratoire ont montré que dans des cellules ES, RARγ phosphorylé était spécifiquement recruté au niveau de RARE de type DR5 et DR7. Ainsi, il est possible que la phosphorylation du récepteur et l'interaction de protéines au niveau du NTD agissent de manière synergique et coordonnée pour réguler le recrutement du récepteur au niveau d'un RARE spécifique. Une telle hypothèse suggère une régulation complexe de la spécificité d'action des différents sous-types de RAR.

ii. RXR ?

Le partenaire d'hétérodimérisation RXR est aussi susceptible de jouer un rôle dans l'activité transcriptionnelle spécifique des différents sous-types de RAR. Il existe six gènes paralogues codants pour les RXR chez le poisson-zèbre et *rxrβa* est le sous-type le plus fortement exprimé dans les premières étapes du développement du poisson-zèbre (Bertrand et al., 2007; Oliveira et al., 2013; Tallafuss et al., 2006; Waxman and Yelon, 2007). Ainsi, il peut se former pas moins de 24 différents hétérodimères RAR-RXR sans prendre en compte les différentes isoformes au sein de chaque sous-type. De ce fait, il est fortement envisageable que les différents hétérodimères puissent présenter une spécificité d'interaction avec certains cofacteurs. De plus, il existe une relation entre la topologie d'hétéromidérisation des DBD des deux partenaires RAR et RXR et l'espacement entre les demi-sites des RARE. Ainsi, les différents hétérodimères peuvent présenter des différences topologiques subtiles favorisant leur recrutement sur un type de RARE donné.

1.3. Régulation par l'environnement chromatinien ?

Les facteurs de transcription peuvent agir en synergie pour contrôler l'expression de gènes cibles au sein de réseaux cis-régulateurs dans les promoteurs des gènes régulés. En particulier, il a été montré que l'association de certains récepteurs nucléaires au niveau de leurs éléments de réponse nécessite l'interaction concomitante d'autres facteurs de transcription dans l'environnement chromatinien proche. C'est le cas des facteurs FoxA1 ou GATA3 dont l'interaction à la chromatine est requise pour l'interaction des RAR sur les RARE de certains gènes dans des cellules de cancer du sein (Hua et al., 2009). En effet, dans ces cellules plus de 40% des régions occupées par les RAR sont également occupées par FoxA1 et 20% sont aussi occupées par le récepteur des œstrogènes ERα. Ainsi, il est tout à fait envisageable que le recrutement spécifique d'un sous-type de RAR au niveau des séquences régulatrices d'un gène cible puisse être corrélé à la liaison d'autres facteurs au niveau de la même région de chromatine. Cette hypothèse pourrait être testée par l'analyse des sites de liaisons de facteurs de transcription comme FoxA1 ou d'autres facteurs aux alentours des RARE recensés par ChiP-seq.

2. Génération de nouveaux outils

2.1. Lignées transgéniques KO

Dans le laboratoire de Vincent Laudet, deux lignées KO pour RARα-A et RARα-B sont en cours de croissance. Un couple de ZFN (Zinc Finger Nuclease) a été généré pour cibler le début du LBD de RARα-A. Nous avons sélectionné une mutation induite par ces ZFN correspondant à une insertion de 5 nucléotides décalant le cadre de lecture (Figure 32). Les embryons hétérozygotes pour la mutation se développent normalement et la lignée homozygote RARα-A$^{-/-}$ est en cours de croissance et sera prête pour accouplement en décembre 2013. D'autre part, un projet de mutagénèse dirigée par TALEN (Transcription Activator-Like Effector Nuclease) ciblant la région NTD de RARα-B est en cours. Nous avons identifié deux mutations correspondant à de courtes délétions décalant le cadre de lecture. Les embryons hétérozygotes RARα-B$^{+/-}$ se développent normalement et sont en cours de croissance pour la génération de la lignée homozygote RARα-B$^{-/-}$ (décembre 2013).

Une fois établie, ces lignées homozygotes seront d'un grand intérêt car elles limiteront les problèmes liés à l'injection de morpholinos (variabilité entre injections et knockdown incomplet). Ainsi, ils permettront une confirmation robuste de nos résultats et serviront d'outils solides pour comprendre comment est contrôlée la spécificité d'action des différents sous-types de RAR au niveau transcriptionnel.

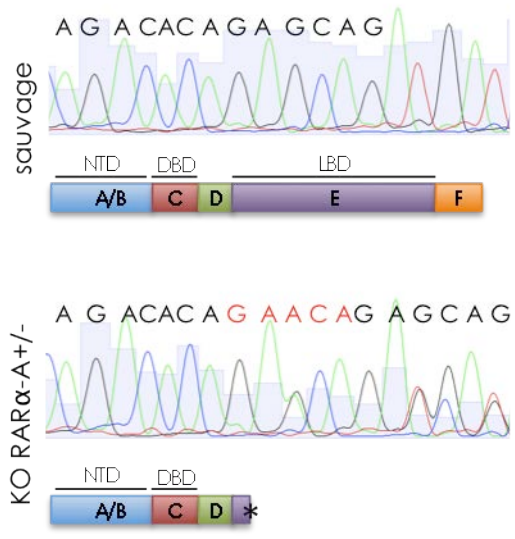


Figure 32 : Génération de lignées transgéniques mutantes pour RARα-A et RARα-B. Chromatogrammes des poissons sauvages ou mutants obtenus par mutagénèse ZFN (Zinc-Finger Nuclease). La mutation correspond à une insertion de 5 nucléotides décalant le cadre de lecture et stoppant la traduction du récepteur au début du LBD. Les poissons hétérozygotes pour la mutation ne présentent pas de phénotype apparent et la lignée homozygote est en cours de croissance.

2.2. Anticorps

Grâce aux services de la plateforme de production d'anticorps monoclonaux de l'IGBMC, nous avons déjà généré des anticorps monoclonaux reconnaissant spécifiquement les sous-types RARα-A ou RARα-B de poisson-zèbre dans leur région F (C-

terminale) (Figure 33). Nous avons également généré des anticorps monoclonaux reconnaissant indifféremment les deux sous-types RARα-A et RARα-B. La génération de tels anticorps pour les sous-types RARγ-A et RARγ-B fait partie d'une demande de financement en cours. Grâce à ces anticorps, il sera envisageable de réaliser des ChIP différentiels *in vivo* pour chacun des sous-types de RAR. Ainsi, il sera possible d'établir un cistrome spécifique de chaque sous-type et de le corréler aux « transcriptotypes» définis dans notre étude. De plus l'analyse des séquences d'ADN liant spécifiquement un sous-type en particulier permettra de déterminer si la régulation du recrutement spécifique d'un sous-type de RAR nécessite des caractéristiques particulières au niveau de la séquence d'ADN.

Experimental Name	Usual Name	Region	H RARα	ZF RARα-A	ZF RARα-B	ZF RARγ-A	ZF RARγ-B
31α 2G5	Ab(31α) α-RARαs	A	N.D	+	+	-	-
34α 5F2	Ab(34α) α-RARαA	F	-	++	-	-	-
35α 4B4	Ab(35α) α-RARαB	F	-	-	++	-	-

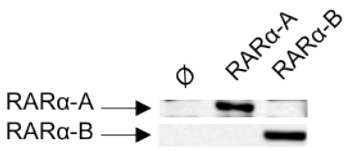
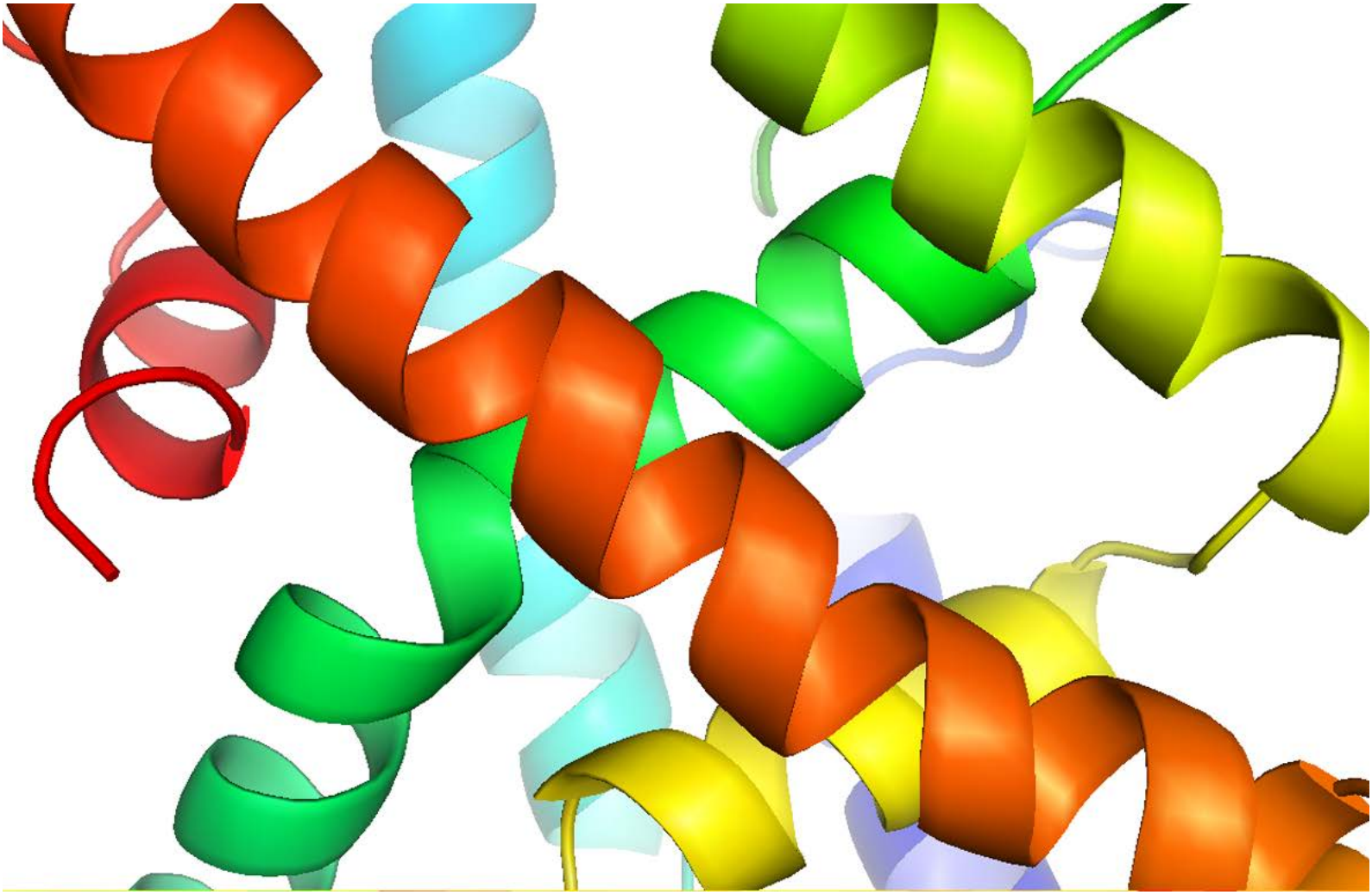


Figure 33: Génération d'anticorps reconnaissant spécifiquement les RARα de poisson-zèbre. Les anticorps généré ont été testés par western blot sur des extraits de cellules surexprimant la protéine RARα humaine ou l'un des quatre RAR de poisson-zèbre.

Chapitre#2

La phosphorylation des RAR : structure – fonction – évolution



I. Contexte Scientifique

Les travaux du laboratoire de Cécile Rochette-Egly ont mis en évidence que les RAR étaient la cible de kinases activées en réponse à l'AR et ont décrit plus particulièrement une cascade de phosphorylation coordonnée ciblant RARα. En effet, la kinase MSK1 activée par l'AR phosphoryle RARα au niveau de S(LBD) et cette phosphorylation est requise pour l'interaction du récepteur avec la cycline H. Cette cycline est associée à la kinase cdk7 au sein du facteur général de la transcription TFIIH et phosphoryle RARα au niveau de S(NTD) (section A-IV-2.1, page 18). L'étude fonctionnelle de ces sites de phosphorylation dans des modèles cellulaires de mammifères a mis en évidence que l'intégrité de ces deux sérines est requise pour le recrutement du récepteur à la chromatine au niveau du RARE de gènes cibles comme *cyp26a1* et est donc nécessaire pour la régulation de leur expression par l'AR. En conclusion, la phosphorylation des RAR joue un rôle crucial dans leur activité transcriptionnelle.

De manière surprenante, le site de phosphorylation du LBD n'est pas présent chez les RAR de poisson-zèbre alors que la sérine au sein du NTD est conservée. Il apparaît donc que la cascade de phosphorylation décrite pour les RAR de mammifères ne semble pas s'appliquer telle quelle chez le poisson-zèbre. Cette observation nous a conduit à étudier l'évolution des sites de phosphorylation des RAR chez les chordés. De plus, pour mieux comprendre comment l'absence de S(LBD) est compensée chez les RAR de poisson-zèbre pour maintenir l'intégrité de leur activité transcriptionnelle, nous avons élucidé les conséquences structurales de la phosphorylation de S(LBD) chez RARα de mammifères. Ainsi, notre étude propose un scénario évolutif de la régulation de l'activité des RAR en lien avec l'acquisition de sites de phosphorylation chez RARα au sein des chordés.

II. Principaux résultats

- **La phosphorylation de S(LBD) chez les RARα de mammifères induit des changements allostériques au sein du LBD.** La modélisation de la phosphorylation de S(LBD) a montré des changements locaux de la dynamique du LBD, en particulier une augmentation de la flexibilité de la boucle entre les hélices 8 et 9 qui interagit avec la cycline H. Ainsi, ces résultats permettent de comprendre comment la phosphorylation de S(LBD) favorise l'interaction avec le complexe cycline H- cdk7.

- **Les RARα de poisson-zèbre sont phosphorylés au niveau du NTD mais pas dans le LBD.** Les alignements de séquence entre les RARα humain et de poisson-zèbre indiquent que S(NTD) est conservée alors que S(LBD) est substituée par un résidu non phosphorylable chez le poisson zèbre. Nous avons confirmé par des expériences de phosphorylations *in vitro* et *in cellulo* que S(NTD) est bien phosphorylée par cdk7 chez le poisson-zèbre mais qu'aucune phosphorylation n'est présente dans le LBD.
- **L'acquisition de S(LBD) permet une modulation plus fine de l'activité de RARα.** Nos travaux ont permis d'établir un scénario évolutif de la phosphorylation des RAR où S(LBD) est acquise quatre fois indépendamment chez les vertébrés et permet une régulation plus fine de l'activité du récepteur. Chez les espèces où S(LBD) est absente, nous avons identifié certains résidus du LBD qui compensent cette absence pour maintenir l'intégrité de l'activité de RARα.

III. Publication

Nos travaux ont été publiés dans le journal *Molecular Biology and Evolution* en 2011.

Molecular Biology and Evolution

Research article

Evolution of Nuclear Retinoic Acid Receptor Alpha (RARα) Phosphorylation Sites. Serine Gain Provides Fine-Tuned Regulation

Eric Samarut,[1,2] Ismail Amal,[1] Gabriel V. Markov,[2,3] Roland Stote,[1] Annick Dejaegere,[1] Vincent Laudet,[2] and Cécile Rochette-Egly*,[1]

[1]IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire), Institut National de la Sante et de la Recherche Medicale (INSERM), U596, Centre National de la Recherche Scientifique (CNRS), UMR7104, Université de Strasbourg, 1 rue Laurent Fries, BP 10142, 67404 Illkirch Cedex, France

[2]Institut de Génomique Fonctionelle de Lyon, UMR 5242, Institut National de la Recherche Agronomique (INRA), Université de Lyon, Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364 Lyon, France

[3]UMR 7221, Museum d'Histoire Naturelle, Paris, France

***Corresponding author:** E-mail: cegly@igbmc.fr.

Associate editor: Claudia Schmidt-Dannert

Abstract

The human nuclear retinoic acid (RA) receptor alpha (hRARα) is a ligand-dependent transcriptional regulator, which is controlled by a phosphorylation cascade. The cascade starts with the RA-induced phosphorylation of a serine residue located in the ligand-binding domain, S(LBD), allowing the recruitment of the cdk7/cyclin H/MAT1 subcomplex of TFIIH through the docking of cyclin H. It ends by the subsequent phosphorylation by cdk7 of an other serine located in the N-terminal domain, S(NTD). Here, we show that this cascade relies on an increase in the flexibility of the domain involved in cyclin H binding, subsequently to the phosphorylation of S(LBD). Owing to the functional importance of RARα in several vertebrate species, we investigated whether the phosphorylation cascade was conserved in zebrafish (*Danio rerio*), which expresses two RARα genes: *RARα*-A and *RARα*-B. We found that in zebrafish RARαs, S(LBD) is absent, whereas S(NTD) is conserved and phosphorylated. Therefore, we analyzed the pattern of conservation of the phosphorylation sites and traced back their evolution. We found that S(LBD) is most often absent outside mammalian RARα and appears late during vertebrate evolution. In contrast, S(NTD) is conserved, indicating that the phosphorylation of this functional site has been under ancient high selection constraint. This suggests that, during evolution, different regulatory circuits control RARα activity.

Key words: nuclear retinoic acid receptor, phosphorylation, evolution.

Introduction

Retinoic acid (RA), the main active metabolite of vitamin A, plays a critical role in many biological processes, such as cell proliferation and differentiation, embryonic development, and adult homeostasis (Bour et al. 2006; Mark et al. 2009; Theodosiou et al. 2010). RA acts through nuclear receptors, RARs, which have been identified in a wide variety of animals.

There is one unique RAR ancestral gene for which an ortholog is known in some prostosomes, such as mollusks (*Lottia gigantea*) and annelids (*Capitella capitata*) (Campo-Paysaa et al. 2008; Albalat and Canestro 2009), and in some invertebrate deusterostomes, such as echinoderms (*Strongylocentrotus purpuratus*; Canestro et al. 2006; Marletaz et al. 2006), cephalochordates (*Branchiostoma floridae*; Escriva et al. 2002a; Fujiwara 2006), and urochordates (*Ciona intestinalis and Polyandrocarpa misakiensis*; Hisata et al. 1998; Fujiwara 2006). Early during vertebrates evolution, the total number of genes markedly increased by two rounds (2R) of whole-genome duplication (Dehal and Boore 2005). This is why vertebrates have three RAR paralogous genes that encode the three known subtypes of receptors: α (NR1B1), β (NR1B2), and γ (NR1B3) (Escriva et al. 2006; Germain et al. 2006). Note that in teleost fishes, a third round (3R) of whole-genome duplication combined to gene losses occurred (Amores et al. 1998; Postlethwait et al. 1998), giving rise to 4 RAR genes in zebrafish (Bertrand et al. 2007). The history of RARs in regard to genome duplications has been addressed in several phylogenetic studies that clearly validated this evolutionary scenario (Escriva et al. 2002b; Jaillon et al. 2004; Robinson-Rechavi et al. 2004; Bertrand et al. 2007; Kuraku et al. 2009). Note that the timing of the genome duplications inferred from a recent analysis of the RAR synteny group by Kuraku et al. (2009).

RARs are ligand-dependent transcriptional regulators (for review, see Rochette-Egly and Germain (2009) and references therein), which bind to specific sequence elements located in the promoters of target genes. They have a well-defined domain organization, consisting mainly of a central DNA-binding domain (DBD) linked to a C-terminal ligand-binding domain (LBD) and a N-terminal domain (NTD) (fig. 1A). Although the NTDs are naturally not structured and not conserved (Dyson and Wright 2005; Lavery and McEwan 2005), DBDs and LBDs are highly structured and depict a significant degree of conservation

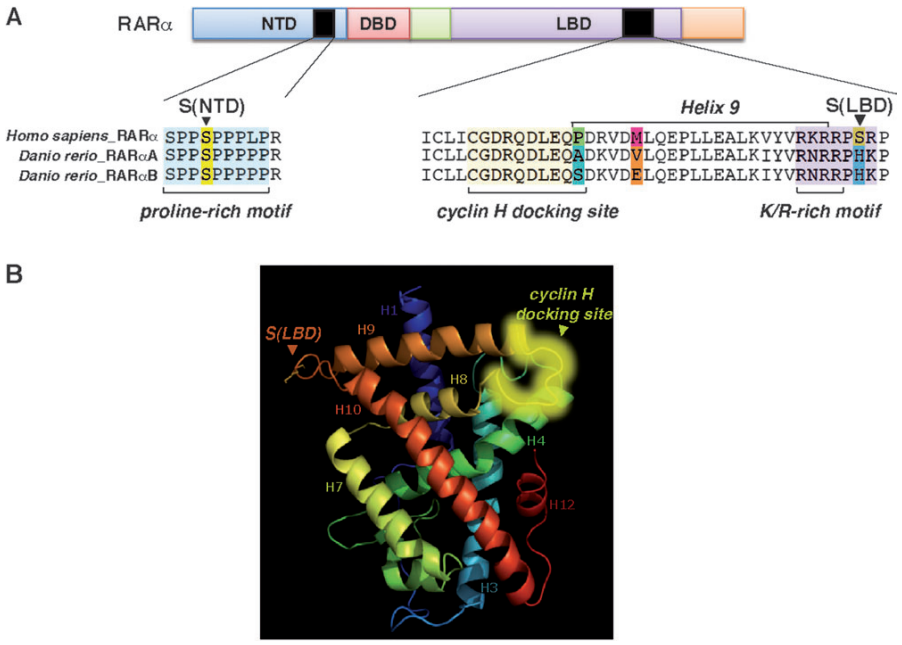


Fig. 1. Schematic representation of RARα. (*A*) Human, mouse, and zebrafish RARα modular structure with the phosphorylation sites (arrowheads) located in the N-terminal proline-rich domain (S[NTD]) and in L9–10 (S[LBD]). (*B*) Structure of the LBD in the presence of ligand (PDB2LBD). The serine located in loop L9–10 of RARα is shown as well as the cyclin H–binding site, which encompasses loop L8–9.

between vertebrate species (Escriva et al. 2006; Campo-Paysaa et al. 2008; Theodosiou et al. 2010). Briefly, the DBD contains two typical cysteine-rich zinc–binding motifs and two alpha helices, which cross at right angles, folding into a globular conformation to form the core of the DBD. Concerning the LBD, it shows a common fold comprising 12 conserved alpha helices and a short beta turn, arranged in three layers to form an antiparallel «alpha-helical sandwich» (Renaud et al. 1995) (fig. 1*B*). Ligand binding triggers conformational changes in the LBD that direct the dissociation/association of several coregulator protein complexes and thereby the transcription of target genes (Rochette-Egly and Germain 2009).

In addition to this scenario, a new concept emerged according to which RARs are also subjected to rapid phosphorylation cascades. Recent studies from our laboratory (Bruck et al. 2009) demonstrated that, via nongenomic effects, RA activates rapidly the p38MAPK/MSK1 pathway, which in turn leads to the phosphorylation of the RARα subtype (mouse and human) at two serine residues located in solvent-accessible regions of the receptor. One serine is located in the LBD (S[LBD]), in a loop between helices 9 and 10 (L9–10) (fig. 1*A* and *B*) and belongs to an arginine-lysine-rich motif that corresponds to a consensus phosphorylation motif for MSK1 (fig. 1*A*). The other serine residue is located in the NTD (S[NTD]), in a proline-rich motif (fig. 1*A*) and is phosphorylated by cdk7 (Rochette-Egly et al. 1997; Bastien et al. 2000), which forms with cyclin H and MAT1 the CAK subcomplex of the general transcription factor TFIIH. Most interestingly, the correct positioning of cdk7 and thereby the efficiency of the NTD phosphorylation rely on the docking of cyclin H at a specific site of the LBD located in loop L8–9 and the N-terminal part of helix 9 (H9) (fig. 1*A* and *B*) (Bour et al. 2005).

In the case of human and mouse RARα, we previously demonstrated that the phosphorylation of the two serines results from a coordinated phosphorylation cascade starting with the phosphorylation by MSK1 of S(LBD) (fig. 1*B*) (Bruck et al. 2009). Phosphorylation of this residue increases the binding efficiency of cyclin H to the nearby loop L8–9 (fig. 1*B*), allowing the right positioning of cdk7 and the phosphorylation of the serine located in the NTD (fig. 1*A*) by this kinase (Gaillard et al. 2006). Finally, phosphorylation of S(NTD) leads to the recruitment of RARα to promoters (Bruck et al. 2009). Whether it also controls the association/dissociation of specific coregulators as described for the other RAR subtypes (Vucetic et al. 2008; Lalevee et al. 2010) is still unknown.

In RARs, ligand binding is conserved at least in chordates (Escriva et al. 2006), indicating that the ligand-triggered conformational changes are a common feature of all chordate species. Interestingly, the high regulatory potential of the phosphorylation cascade also makes phosphorylations prime candidates for evolutionary studies. It must be noted that S(LBD) and S(NTD) are conserved in the different human and mouse RAR subtypes α, β, and γ (Rochette-Egly 2003; Rochette-Egly and Germain 2009), but the above cascade has been described only in the context of RARα (Bruck et al. 2009), which has ubiquitous or quite widespread expression patterns. There are still no indications whether this cascade also occurs in the context of the other RAR paralogs (RARβ and RARγ), which show rather complex tissue-specific expression (Dolle 2009). Therefore, we analyzed the pattern of conservation of the S(NTD) and S(LBD) phosphorylation sites focusing on the RARα subtype.

First, we demonstrated that in nonmammalian vertebrates exemplified by zebrafish, the S(NTD) of RARα is conserved, whereas S(LBD), the phosphorylation of which increases the flexibility of L8–9 that is required for cyclin H-binding, is absent. However, this process was compensated by changes in the sequence of L8–9 mimicking the conformation/flexibility changes induced by phosphorylation. Then, we traced back the evolution of chordate RARα phosphorylation sites. This work led to the conclusion that in RARα, S(NTD) is evolutionary conserved, indicating that the phosphorylation of this functional site has been under ancient strong selection constraint. However, S(LBD) is most often absent outside mammalian RARα. This indicates that the fine-tuned phosphorylation cascade of RARα, starting at S(LBD), appears late during vertebrate evolution. Thus, the evolution of phosphorylation sites appears to provide a reservoir of changes in order to provide additional levels of regulation of critical functional proteins.

Materials and Methods

Sequences Alignment and Ancestral Sequences Reconstructions

RAR protein sequences were found in the nuclear receptor database (NureXbase) (http://nurexbase.prabi.fr) and by Blast and gene homology (NCBI). Multiple sequence alignments were performed by the MUSCLE software (Edgar 2004) and analyzed with ClustalX. Sequence assignment was verified by phylogenetical reconstruction as in Escriva et al. (2006). Ancestral sequences were estimated with PAML (Yang 1997) under the JTT + γ substitution model from a data set of 71 sequences containing a 232 amino acid long portion of the LBD. Some sequences containing obvious predictions errors or indels at unambiguous positions were manually corrected by parsimony. Other sequences with too many uncertainties were excluded from the reconstruction data set. PhyML (Guindon and Gascuel 2003) generated the starting tree.

Molecular Dynamics Simulations

Structure Preparation

Given the lack of an experimental structure for the LDB of human (h) RARα in an agonist form (holo) at the start of this work, a model structure was assembled from closely related structures available in the Protein Data Bank (Berman et al. 2000). The majority of the structure that includes helix 1 (H1) to helix 10 (H10) was taken from the structure of human nuclear RA receptor alpha (hRARα) (PDBID 1DKF) bound to the selective antagonist BMS614 (Bourguet et al. 2000). Structural information for an agonist conformation of H11 and H12 was taken from the structures of hRARγ (PDBID 1FCZ) and RARβ (PDBID 1XAP) in the agonist forms (Klaholz et al. 2000; Germain et al. 2004). Side chains specific to hRARα were positioned using the Scwrl3.0 software (Canutescu et al. 2003). The structures of 9-cis RA and of a fragment of the TRAP220 coactivator were obtained from the structure of RARβ (PDBID 1XDK) in an agonist conformation (Pogenberg et al. 2005). The protonation states of all titratable groups at physiological pH (7.4) were determined as described in Schaefer et al. (1998), and all were found to favor their standard protonation states. Our model shows a very high degree of correspondence with an experimental structure of hRARα in an agonist form, which has been recently deposited in the Protein Data Bank (3A9E) (Sato et al. 2010), after the termination of this work.

Structural models were also constructed for apo hRARα, that is, in the absence of ligand and coactivator peptide. Under these conditions, the C-terminal end of the LBD, in particular H12, extends toward the solvent where it displays significant conformational flexibility (Renaud and Moras 2000). In the absence of any experimental apo hRARα structure, a model was constructed that kept H1–H10 in the same conformation as the holo structure but repositioned the C-terminal end based on the apo RXR structure (PDB 3A9E) (Sato et al. 2010). This model was constructed using the Modeler 9v8 program (Sali and Blundell 1993). Given that in the apo structures, H12 is conformationally mobile, variants of this apo model were constructed with different initial positions of H12. As all the initial apo models gave similar simulation results, we presented the data corresponding to the initial structure (Sato et al. 2010).

The LBDs of zebrafish (zf) RARα(-A and -B) were constructed by modifying all residues that differ from hRARα, maintaining the backbone conformation and modifying the side chains using the SCRWL4 program (Canutescu et al. 2003). For zfRARα-A, this involved the following side chain modifications: E183D, V184T, G185E, E186Q, L187M, E189D, K190R, A201S, N211S, Q216R, S219A, I222V, I335L, P345A, R347K, M350V, V361I, K365N, S369H, and R370K. For zfRARα-B, the modifications were E183D, V184T, G185E, E186K, L187M, K190Q, A201S, S214A, E215D, Q216H, S219A, I222V, E280D, I335L, P345S, R347K, M350E, V361I, K365N, S369H, and R370K. A similar protocol was used to construct hRARα mutants (hRARαP345G/D346A and hRARαP345A). The models for the apo forms of zfRARαs and the hRARα mutants were constructed as above.

Molecular Simulations

All molecular dynamic simulations were done using the CHARMM program (Brooks et al. 1983) and the all atom parameter set of CHARMM27 (MacKerell et al. 1998), with CMAP corrections (Mackerell et al. 2004). Hydrogen atoms were added using the HBUILD module (Brunger and Karplus 1988). Bonds between heavy atoms and hydrogen atoms were constrained using SHAKE (Ryckaert et al. 1977). We employed a shift-type cutoff at 14 Å for electrostatic interactions and a switch-type cutoff at 12.0 Å for the van der Waals energy terms.

The system was energy minimized using the steepest descent algorithm after placing harmonic constraints on the backbone and side chain heavy atoms with force constants of 50 and 100 kcal $mol^{-1}Å^{-2}$, respectively. The force constants were systematically scaled by a factor of 0.65, and

minimization was repeated until there were no constraints on the protein. The protein was then solvated with a shell of explicit TIP3P water molecules (Gaillard et al. 2009) extending 12 Å from the protein surface. The system was equilibrated in two phases. In the first phase, a 20 ps molecular dynamics simulation of the water around the fixed protein was performed with a time step of 2 fs. In the second phase, the entire solvated protein was heated to 300 K and equilibrated. During heating, velocities were assigned every 50 steps from a Gaussian distribution function. During equilibration, velocities were scaled by a single factor only when the average temperature was lying outside the 300 ± 10 K window. This was followed by a 10 ns production phase without any further intervention.

Simulations were stable as measured by the backbone root-mean-square coordinate differences (RMSD), which were all less than 1.24 Å with respect to the initial starting structure. The phosphate group was assigned a charge of −2 based on the pKa of 6.5 (Kast et al. 2010).

Using the above protocol, simulations were run for the LBD of hRARα unphosphorylated or phosphorylated at S(LBD) either in the apo or holo- forms. Simulations were also run for hRARαP345G/D346A, hRARαP345A, and zfRARα-A and -B in the apo forms. Upon completion of the simulations, the root-mean-square fluctuations (RMSfl) as well as the RMSD were calculated from the trajectories.

Plasmids

The pSG5- and pGEX-2T-based expression vectors for hRARα1 have been previously described (Bour et al. 2005). The full-length or truncated cDNAs of zfRARα-A and zfRARα-B were amplified by polymerase chain reaction (PCR) and inserted into pSG5-hER-B10-tag or pGEX-2T vectors. The cDNA of hcyclin H (a gift from D. Busso, Institut de Génétique et de Biologie Moléculaire et Cellulaire [IGBMC]) was inserted into pCX-HA-FLAG. The cDNA of zfcyclin H (Liu et al. 2007) was inserted into pCX-HA or pET-15b. All constructs were generated using standard cloning procedures and were verified by PCR, restriction enzyme analysis, and DNA sequencing. The sequence of primers used for PCR amplifications are available upon request.

Antibodies

Mouse monoclonal antibodies recognizing hRARα phosphorylated at S77, cyclin H, and the epitope B of the estrogen receptor (B10) were previously described (Ali et al. 1993; Bruck et al. 2009). Rabbit polyclonal recognizing the N-terminal part of cyclin H and anti-FLAG monoclonal antibodies were from Sigma.

Mouse monoclonal antibodies recognizing zfRARα-B phosphorylated at the conserved serine residue located in the N-terminal proline-rich domain (S72) were generated by immunization of Balb/c mice with a synthetic phosphopeptide (EEMVPSSPS(p)PPPPPRVYKPC). Six-week-old female BALB/c mice were injected intraperitoneally (thrice at 2-week intervals) with 100 μg of peptide coupled to ovalbumin and 100 μg of poly I/C as adjuvant. Mice with positive sera were reinjected 4 days prior to hybridoma fusion and spleens were fused with Sp2/0.Ag14 myeloma cells. After hybridoma cell selection and cloning (de StGroth and Scheidegger 1980), the culture supernatants were tested by differential enzyme-linked immunosorbent assay with the phosphopeptide, the corresponding nonphosphopeptide, and an irrelevant phosphopeptide. Positive clones were confirmed by immunoblotting and cloned twice on soft agar. Ascites fluids were prepared by injection of 2×10^6 hybridoma cells into pristane-primed BALB/c mice.

Cell Lines, Transfections, and Immunoprecipitation Experiments

COS-1 cells were grown and transiently transfected as described (Bour et al. 2005). ZF13 cells were grown at 27 °C in Leibovitz L-15 medium (Invitrogen) supplemented with 5% fetal calf serum and 15 mM acid 4-(2-hydroxy ethyl)-1 piperazine ethane sulfonic acid and transiently transfected by using FuGene 6 reagent (Roche). Immunoprecipitations were performed with cell extracts prepared from paraformaldehyde-fixed cells (Bruck et al. 2009).

In vitro Binding and Phosphorylation Experiments

Glutathione S-transferase (GST) and GST-fusion proteins expressed in *Escherichia coli* were immobilized onto glutathione–Sepharose beads and incubated with recombinant human cyclin H over expressed in insect Sf9 cells (Bour et al. 2005) or with purified bacterially expressed zfcyclin H. Bound proteins were immunoprobed and quantified by using the Chemigenius XE imaging system as described (Bour et al. 2005). Data were analyzed according to standard statistical procedures using Graph Pad Prism 5.0 and compared using the Tukey's test in conjunction with analysis of variance.

In vitro phosphorylation experiments were performed with equimolar amounts of immobilized GST-RARα proteins (5 μg). Phosphorylation by the purified cdk7/cyclin H complex (Bour et al. 2005) was performed as in (Bruck et al. 2009) and detected by immunoblotting with antibodies recognizing specifically the phosphorylated forms. Phosphorylation by recombinant active MSK1 (Millipore Upstate Chemicon) (30 ng) was performed in the presence of $\gamma[^{32}P]$ as described (Rochette-Egly et al. 1995) and visualized by autoradiography.

Results

In Mammalian RARα, Phosphorylation of S(LBD) Increases the Dynamics/Flexibility of the Cyclin H-Binding Domain

In hRARα, the upstream serine residue of the phosphorylation cascade, S369 [S(LBD)], is located in the LBD, in loop L9–10 within an arginine–lysine-rich motif (fig. 1*A* and *B*). This serine is in the vicinity of a specific domain of the LBD, encompassing loops L8–9 and the N-terminal tip of H9 and involved in the binding of cyclin H (fig. 1*A* and *B*) (Bour et al. 2005). Phosphorylation of S(LBD) has been shown to increase the ability of hRARα to interact with cyclin H, with a characteristic downstream consequence on

the phosphorylation by cdk7 of the serine located in the NTD (S[NTD]) (Gaillard et al. 2006; Bruck et al. 2009, fig. 1A). This is a typical model of substrate recognition by a protein kinase through association via another substrate-binding subunit.

To further investigate the consequences of phosphorylation on the LBD of hRARα, molecular dynamic simulations (MD) were performed (see Materials and Methods) to analyze whether phosphorylation of S(LBD) generates conformational changes affecting the cyclin H-binding domain located at a 30 Å distance, in L8–9 (fig. 1B).

Simulations were first performed with the holo form of hRARα (i.e., in the presence of RA and of a coactivator peptide), which is closest to the in vivo experimental phosphorylation studies (Bruck et al. 2009).

Average structures of the native and phosphorylated forms of hRARα were calculated and the RSMD from the initial structures were measured (fig. 2A). RMSD of the backbone atoms forming secondary structure was less than 1.0 Å, indicating that the overall structure of the LBD is conserved upon phosphorylation. However, local conformational changes of loops L8–9 were observed with RMSD values in the order of 4Å. These conformational changes are shown in figure 2A where the average structures of unphosphorylated and phosphorylated hRARα are superposed. An upward displacement of L8–9 is clearly visible, linked to an upward bending of helix H9. These changes likely result from the locally enhanced electrostatic environment due to the −2 charge of the phosphate moiety.

More significant, however, is the increase in the local conformational dynamics or flexibility of L8–9 as measured by the atomic RMSfl averaged by residue (Fidelak et al. 2010). RMSfl are directly related to the temperature factors determined during an X-ray crystallography structural study and are a direct calculation of local short-time scale dynamics of L8–9. As shown in figure 2B, in the absence of S(LBD) phosphorylation, L8–9 was generally more flexible than the neighboring helices, with an RMSfl in the order of 0.9 Å. However, with S(LBD) phosphorylated, the average RMSfl of L8–9 increased by a factor of 2. Thus, the simulations clearly indicate that S(LBD) phosphorylation affects the cyclin H-binding domain. In the absence of an experimental structure of the hRARα–cyclin H complex, the investigation of the detailed molecular mechanism of this allosteric signaling was, however, beyond the scope of this study.

Then, molecular dynamics simulations were repeated without or with S(LBD) phosphorylated but with the apo form of hRARα in order to assess whether phosphorylation can still affect L8–9 structural dynamics in the absence of RA and of a coactivator peptide and with H12 in an extended conformation (see Materials and Methods). RMSfl analysis shows that L8–9 exhibits an increased flexibility when the apo form was phosphorylated at S(LBD) (fig. 2C). This suggests that S(LBD) phosphorylation by itself can affect the conformational dynamics of the cyclin H–-binding domain of hRARα in the apo form.

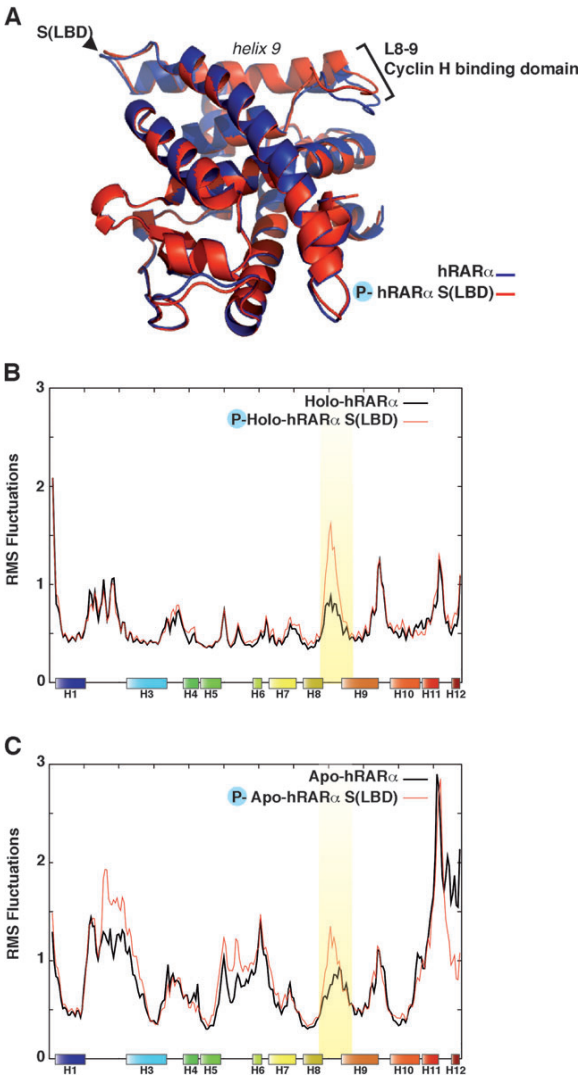


FIG. 2. In hRARα, phosphorylation of S(LBD) increases the flexibility of loops L8–9: molecular dynamic simulations. (*A*) Superposition of the average structures of hRARα LBD unphosphorylated (blue) and phosphorylated (red) in an holo conformation, that is, complexed with RA and a coactivator peptide. The positions of phosphorylated S(LBD) and of L8–9 are indicated. The average structures were computed from the final 5 ns of MD simulations. (*B*) Fluctuations of the backbone atoms of hRARα LBDs unphosphorylated (black) and phosphorylated at S(LBD) (red) in an holo conformation. Global motion was removed by reorienting trajectory frames onto H3, H5, and H10 of the initial LBD structure. Fluctuations were calculated from the final 5 ns of molecular dynamics simulations and given in Angstrom. Positions of the α-helices are schematized along the *x* axis and the cyclin H-binding domain (L8–9) is highlighted in yellow. (*C*). Same as in B but with the LBDs of hRARα under the apo- form, that is, in the absence of RA and of a coactivator peptide and with H12 extended in solution. Note that in this apo form, the flexibility of the C-terminal end of the hRARα LBD is higher compared with that of the holo form shown in panel B, due to the extended conformation of H12. In contrast, in the holo form, H12 is stabilized against the core of the LBD.

In conclusion, from these results and our previous experimental results (Gaillard et al. 2006; Bruck et al. 2009), one can suggest that the increase in the flexibility of loops L8–9

observed upon phosphorylation of S(LBD) might facilitate the binding of cyclin H to this domain.

In zebrafish RARα, S(LBD) Is Absent but S(NTD) Is Conserved and Phosphorylated

Given the functional importance of RARα not only in mouse and human but also in other vertebrate species such as zebrafish (Dolle 2009; Linville et al. 2009), we investigated whether the phosphorylation cascade was conserved in zebrafish (*Danio rerio*), which expresses two RARα genes: *RARα*-A and *RARα*-B.

Sequence alignment revealed that in zf RARα-A and RARα-B, the S(LBD) residue was not present (fig. 1A). Instead, a histidine residue was found. However, the arginine–lysine-rich motif flanking this residue was well conserved (fig. 1A). Accordingly, in vitro phosphorylation experiments indicated that the LBDs of zfRARαs were not phosphorylated by MSK1, the upstream kinase involved in the phosphorylation cascade (fig. 3A). As phosphorylation sites that are not strictly conserved at a specific position can be compensated by others, driving the same functions (Nguyen Ba and Moses 2010), we investigated whether other phosphorylation sites are present in the nearby region. However, an in silico prediction of potential phosphorylation sites in the LBDs of zfRARα-A and zfRARα-B did not reveal any other compensatory phosphorylation sites.

In contrast, in zfRARα-A and zfRARα-B, the serine residue located in the NTD [S(NTD)] was conserved as well as its flanking region, that is, the proline-rich motif (fig. 1A). Then, the question was whether, in zfRARαs, the conserved S(NTD) could be phosphorylated even in the absence of a phosphorylatable S(LBD). Most interestingly, in vitro, the S(NTD) of zfRARα, was phosphorylated by the purified cdk7/cyclin H complex, as assessed by immunoblotting with antibodies recognizing specifically this phosphorylated residue (fig. 3B, lanes 6 and 7). No signal was obtained with zfRARα deleted for the NTD, confirming the specificity of the antibodies (fig. 3B, lane 8). These results were confirmed in vivo, with B10-tagged zfRARα over expressed in zebrafish (ZF13) or mammalian (COS-1) cells and immunoprecipitated with an antibody recognizing specifically the phosphorylated receptor (fig. 3C). Collectively, these results indicate that, in zebrafish, RARα can be phosphorylated at S(NTD) even in the absence of phosphorylation of the LBD. This raises the question of how the phosphorylation cascade starting at the LBD can be bypassed in zebrafish.

In zebrafish RARα, the Cyclin H-Binding Domain Is More Flexible than in hRARα

Given that in hRARα, phosphorylation of S(NTD) by cdk7 relies on the binding of the associated cyclin H, we investigated whether zfRARα could interact with cyclin H despite the absence of phosphorylation of the LBD.

The ability of zfRARα to interact with cyclin H was compared with that of hRARα in in vitro protein–protein interaction assays. In GST-pulldown assays that use non-

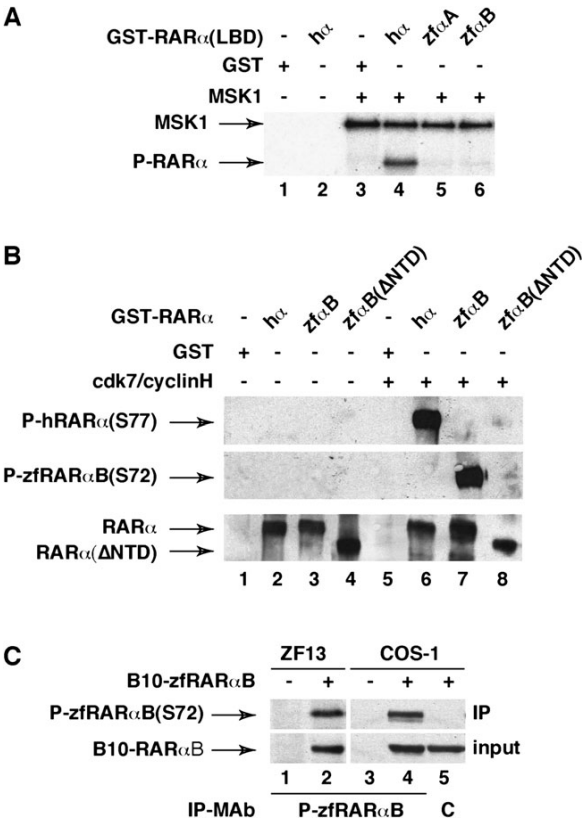


Fig. 3. Comparison of h and zfRARα phosphorylation (*A*) In vitro phosphorylation of the hRARα and zfRARα LBDs fused to GST by MSK1 and analysis by autoradiography (*B*) In vitro phosphorylation of GST-hRARα and zfRARα (WT and △NTD) with cyclin H/cdk7, analyzed by immunoblotting with antibodies recognizing specifically the receptors phosphorylated at S(NTD). (*C*) In transfected ZF13 and COS-1 cells, B10-tagged zfRARα is phosphorylated at the N-terminal serine (S72). B10-zfRARα was immunoprecipitated with antibodies recognizing specifically the form phosphorylated at S(NTD) and immunoblotted with B10 antibodies. The lower panel corresponds to the inputs.

phosphorylated bacterially expressed fusion proteins, both zfRARα-A and zfRARα-B interacted with cyclin H but more efficiently than did hRARα (fig. 4A, lanes 4 and 5 and fig. 4B). zebrafish RARαs also interacted with cyclin H in coimmunoprecipitation experiments performed with extracts from transfected COS-1 cells (supplementary fig. S1A, Supplementary Material online). In this case, the interaction was as efficient as with hRARα in line with the fact that hRARα is phosphorylated at S369 in transfected cells (supplementary fig. S1B, Supplementary Material online). Similar results were obtained whatever cyclin H was from human (supplementary fig. S1A, Supplementary Material online) or zefrafish (supplementary fig. S1C, Supplementary Material online). Altogether, these observations suggest that the conformation of zfRARα favors cyclin H binding.

Given that the efficiency of cyclin H binding to hRARα relies on the conformational features of L8–9 and the N-terminal tip of H9 (Bour et al. 2005), we compared this domain with that of zfRARα-A and zfRARα-B (fig. 1A).

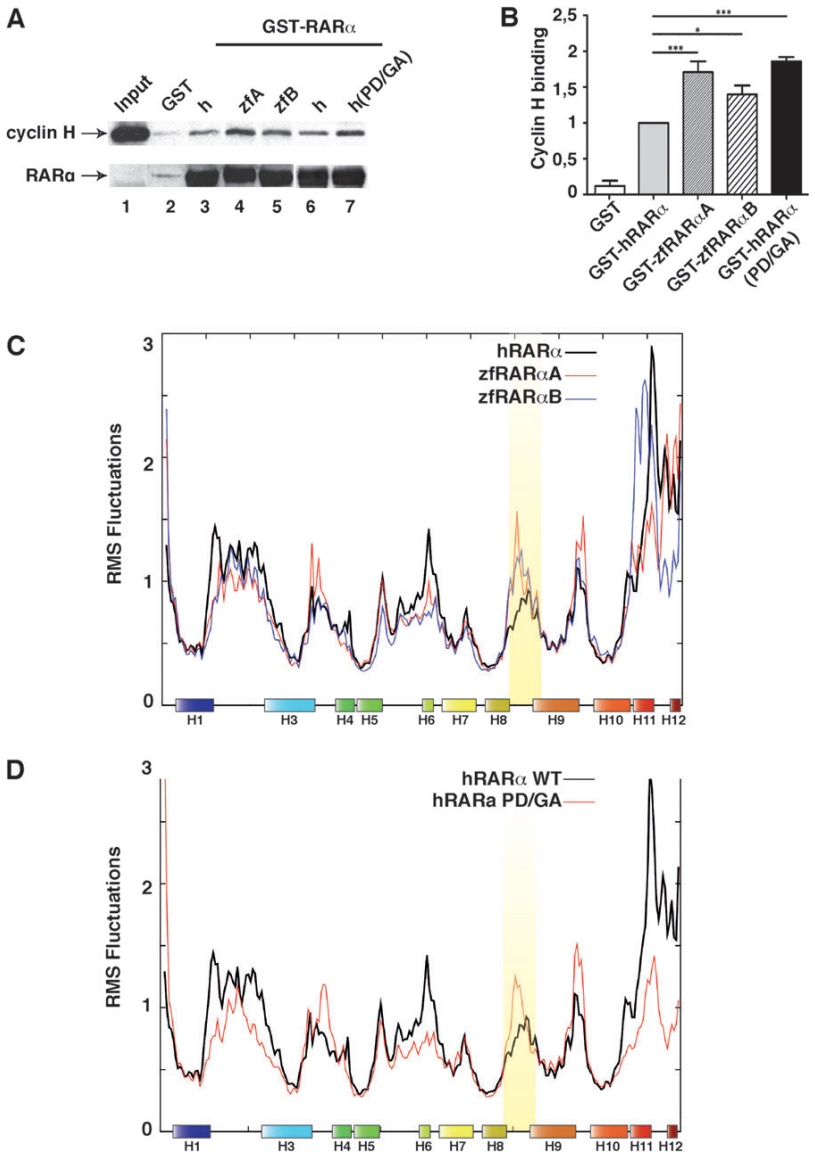


FIG. 4. Comparison of h and zfRARα interaction with cyclin H and flexibility of the cyclin H-binding domain (*A*) In vitro, zfRARα interacts more efficiently than hRARα with human cyclin H. The hRARαP345G/D346A mutant (hPD/GA) also interacts more efficiently. (*B*) Mean ± standard deviation of five individual experiments after quantification and normalization to cyclin H binding to hRARαWT. Significantly different data are shown by asterisks: * *P* value < 0,05, *** *P* value < 0,01 (*C*) Fluctuations of the backbone atoms of the LBDs of hRARα (black), zfRARα-A (red), and zfRARα-B (blue) under the apo form, that is, in the absence of RA and of a coactivator peptide and with H12 extended in solution. Fluctuations were calculated from the final 5 ns of the molecular dynamics simulations and are given in Angstorm. Global motion was removed by reorienting trajectory frames onto H3, H5, and H10 of the initial LBD structure. (*D*) Fluctuations of the backbone atoms of the LBDs of hRARα WT (black) and P345G-D346A (red) under the apo form and calculated as in C.

The overall sequence of the cyclin H–-binding domain is well conserved, but the proline residue located at the N-terminal tip of H9 in hRARα (P345) is replaced by an alanine and a serine in zfRARα-A and zfRARα-B, respectively. Similarly, the methionine residue found in helix 9 at position 350 is replaced by a valine or a glutamic acid in zfRARαs. Altogether, these observations pinpoint some specific mutations that may be functionally relevant to affect the conformation of zfRARα L8–9 in order to favor cyclin H binding.

This led us to explore the conformational dynamics of zfRARαs in molecular dynamics simulations carried out with the apo forms that best match the experimental GST-pulldown conditions (i.e., in the absence of RA and of a coactivator peptide and with H12 extended into solution). Comparison of the RMSfl calculated from these simulations showed that both zfRARα-A and zfRARα-B exhibited an increased flexibility of L8–9 compared with hRARα also in the apo form (fig. 4C). Such results indicate

that in zebrafish RARα, L8–9 is natively more flexible than in the human counterpart.

Then, given that in hRARα, substitution of P345 and D346 with a glycine and an alanine respectively, significantly increases cyclin H binding (fig. 4A, lane 7; Bour et al. 2005), we analyzed the consequences of these changes in MD simulations. As shown in figure 4D, the hRARαP345G/D346A mutant in the apo form showed an increased flexibility of L8–9 compared with WT RARα. Finally, the single substitution of P345 with an alanine in hRARα (as in zfRARα-A) was sufficient to increase the flexibility of L8–9 (supplementary fig. S2, Supplementary Material online). Collectively, these results highlight the importance of the amino acid sequence in the flexibility of the cyclin H-binding domain.

During Vertebrate Evolution, S(NTD) Is Conserved but not S(LBD)

From the comparison of human and zebrafish RARαs, one can suggest that the serine located in the NTD would be conserved across vertebrates, whereas it would not be the case for the serine in the LBD. Therefore, we investigated whether there is a constraint on these two RARα phosphorylation sites during evolution. Sequence comparison and prediction of ancestral sequences at all nodes of the chordate RAR tree were performed.

Figure 5 provides a global overview of the evolution of S(NTD) and S(LBD) in all known full length and functional chordate RARαs with available complete sequences, from invertebrate chordates such as amphioxus (*B. floridae*) and ascidian tunicates (*C. savignyi*) through basal vertebrates such as lampreys (*Lethenteron japonicum*) and teleost fishes (e.g., *D. rerio*) to mammals.

It appeared that S(NTD) is strictly conserved in all available complete chordate sequences (fig. 5), even in amphioxus (*B. floridae*). Note, however, that for some species such as *Eptatretus burgeri*, *Mordacia mordax*, *Callorhinchus callorynchus*, and *Lepisosteus platyrincus*, the RARα sequences are incomplete (hyphens in fig. 5), making difficult the introduction of these species in our evolutionary study. The situation was very similar for the RARβ and RARγ paralogs (fig. 5). Interestingly, the flanking region, that is, the proline-rich motif, was also conserved (fig. 6). This indicates that S(NTD) has been under a high selective pressure through chordate evolution.

In contrast, S(LBD) was not present in RAR from cephalochordates (*B. floridae*) and urochordates (*C. savignyi* and *P. misakensis*) (fig. 5) despite the conservation of the flanking arginine–lysine-rich motif (fig. 6 and supplementary fig. S3, Supplementary Material online). It was not present either in RARα from most vertebrates (teleost fish, amphibians, and birds) (fig. 5). Instead, aspartic acid, glutamic acid, asparagine, or histidine residues were found. Most interestingly, the presence of a serine in L9–10 seems to be specific to the main clades of mammals with an exception in the case of *Anolis carolinensis* (fig. 5 and supplementary fig. S3, Supplementary Material online). The situation was very similar for RARγ (fig. 5). However, in the case of RARβ, a serine was present in the LBD not only in mammals but also in teleost fishes and in *Lepisosteus platyrhincus* (fig. 5).

Given this complex pattern, it was difficult to infer directly (using simple parsimony reasoning) which amino acid was at the position of S(LBD) before the two duplications (A1 and A2 in fig. 7) that led to RARα in gnathostomes. Thus, we reconstructed the ancestors using maximum likelihood. This allowed us to propose an evolutionary scenario (fig. 7) in which the inferred ancestor harbors an asparagine in L9–10. Then in vertebrates, this asparagine is replaced by a serine in teleost fishes RARβ as well as in mammalian RARα, RARβ, and RARγ. This suggests that, late during vertebrate evolution, a change in selective pressure allowed the acquisition of an easily phosphorylatable residue in L9–10.

Note that S(NTD) was also found in TR (fig. 7), a nuclear receptor belonging to the same NR subfamily. Though part of a distinct motif, this serine is known as a functional phosphorylation site (Glineur et al. 1990), corroborating the ancient high selection constraint acting on this residue.

Discussion

Protein phosphorylation is crucial for the regulation of many cellular events, and it has been hypothesized that it would serve as a transcriptional clock, orchestrating rapid, and dynamic exchanges of coregulators so that at the end, the right proteins are present with the right activity, at the right place, and at the right time (Rochette-Egly and Germain 2009; Lalevee et al. 2010). Most interestingly, similar to changes in genes cis-regulatory modules that modify specific aspects of the expression pattern of a gene without affecting the function of the encoded protein (Hoekstra and Coyne 2007), specific changes of phosphorylation processes can modify a regulatory cascade without affecting the overall function of the protein (Basu et al. 2008). Therefore, phosphorylation sites may be important targets of evolutionary processes. Now, with the availability of high throughput data sets, it becomes possible to examine and test experimentally the evolution of large sets of proteins and phosphorylation sites.

Human and mouse RARα are typical transcriptional regulators that are modulated by RA binding and rapid concomitant RA-induced phosphorylation cascades starting at a serine located in the LBD (S[LBD]) and ending at an other serine in the NTD (S[NTD]).

The present work indicates that there is a strong conservation of S(NTD) in all chordate RARα sequences known to date, comprising cephalochordates (amphioxus), urochordates (*Ciona*), and vertebrates. Such an evolutionary constraint correlates with the importance of this phosphorylation site for RARα binding to DNA and RARα-mediated transcription (Bruck et al. 2009; Rochette-Egly and Germain 2009). It is worth noting that S(NTD) is also highly conserved in the other RAR paralogs, RARγ and RARβ, all along the chordate phylum, confirming the functional importance of this phosphorylation site. However, out of chordates, only a few RAR sequences have been identified, one in another deuterostome such as

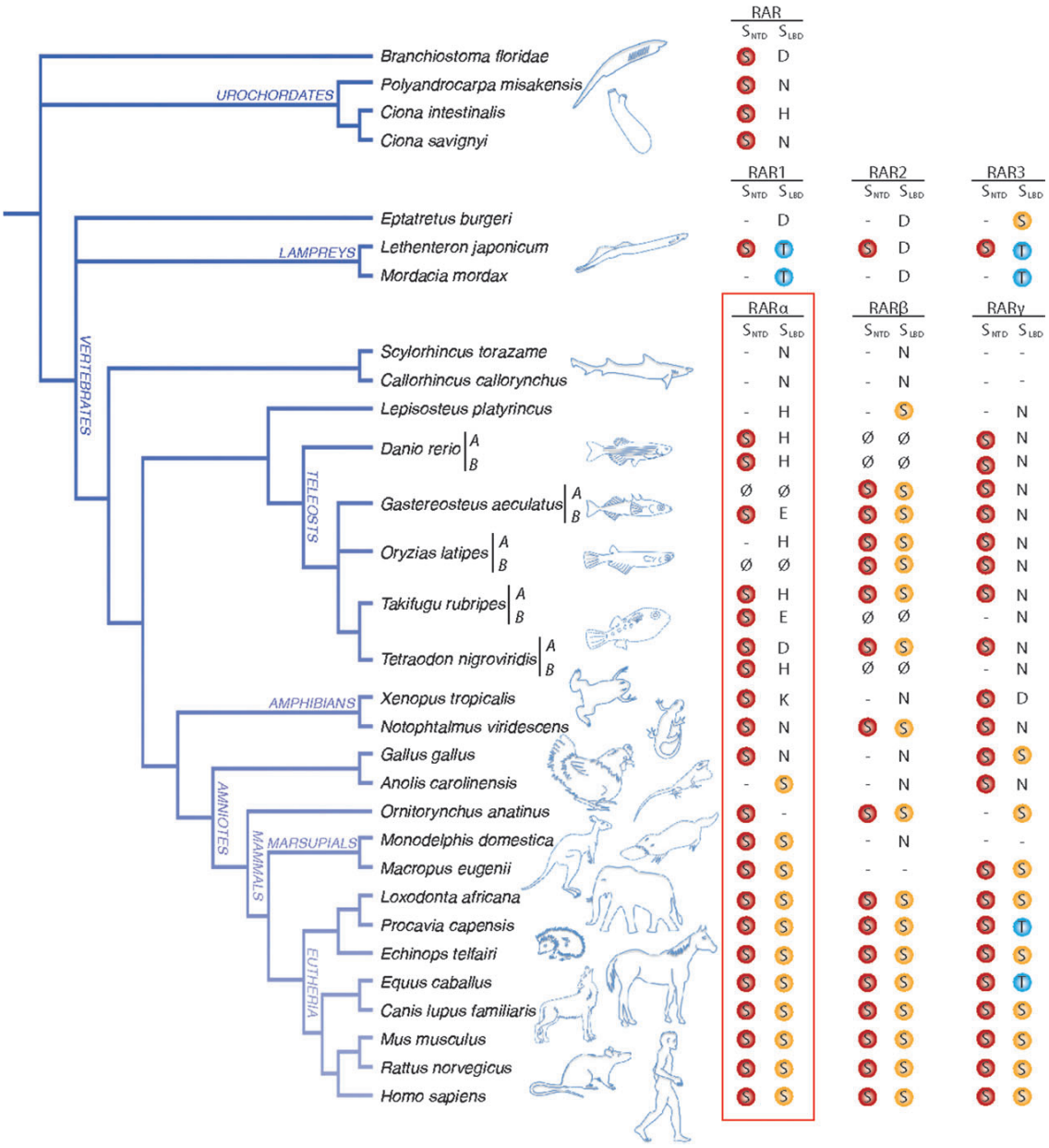


Fig. 5. Schematic phylogenetic tree of chordates showing the evolution of RAR phosphorylation sites. The conservation of S(NTD) and S(LBD) is represented after multiple alignment of the RAR sequences from different species. Ø Corresponds to gene loss. Hyphens correspond to the lack of validated sequences. Sequences are named with the nomenclature code used in the nuclear receptor database NUREXBASE (http://nurexbase.prabi.fr).

S. purpuratus (Canestro et al. 2006; Marletaz et al. 2006) and only two in prostosomes (one in a mollusk and one in an annelid) (Campo-Paysaa et al. 2008; Albalat and Canestro 2009), and functional data are still lacking. Therefore, the phylogenic coverage is not sufficient yet to generalize our data out of chordates and to protostomes on a safe basis. Finally, given the overall conservation of the flanking proline-rich motif, S(NTD) appears to be phosphorylated by similar kinases in all species. In line with this, cdk7 and cyclin H are highly conserved and functional homologs have been described even in yeast (Damagnez et al. 1995).

The original aspect of human RARα phosphorylation resides in the fact that the cdk7 kinase involved in the phosphorylation of S(NTD), recognizes the receptor through the binding of cyclin H at a specific domain located in a disordered loop of the LBD (L8–9). This process is controlled by the phosphorylation of S(LBD), a nearby serine residue located in an other disordered loop, L9–10 (Bour et al. 2005; Gaillard et al. 2006). Due to the importance of this phosphorylation cascade, we have combined evolutionary studies with molecular dynamics computer simulations and experimental analysis, to predict phosphorylation of

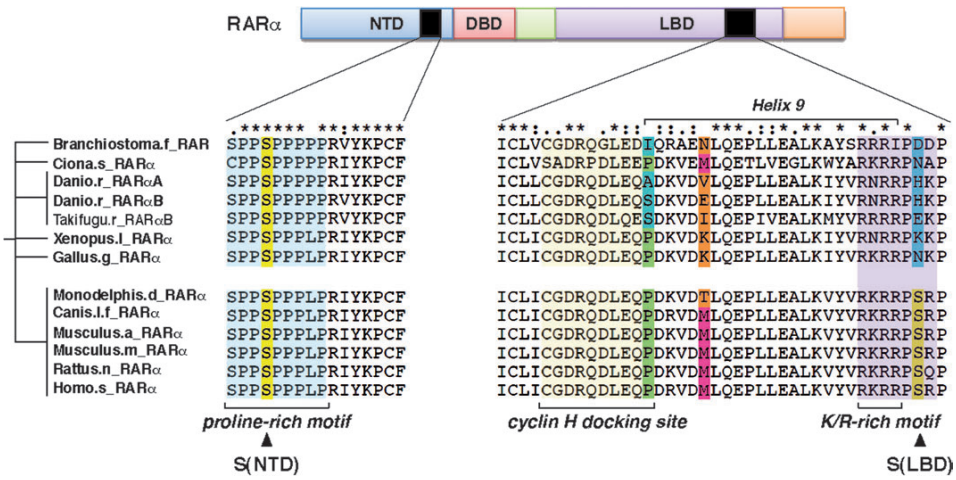


Fig. 6. Alignment of the proline-rich domain, the cyclin H-binding domain (Loop L8–9 and the N-terminal tip of H9) and loops L9–10 in RARαs from different species and for which complete full-length sequences are available. The strongly conserved positions are marked over the alignments, according to the Clustal software program. "Asterisk" indicates positions that have a single fully conserved residue. "colon" indicates that one of the following "strong" groups is fully conserved: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, and FYW. "Full point" indicates that one of the following "weaker" groups is fully conserved: CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEQHRK, FVLIM, and HFY.

S(LBD), flexibility of the cyclin H-binding domain, cyclin H-binding, and the evolution of these processes.

The present study demonstrates that in human RARα, phosphorylation of S(LBD) increases the flexibility of the nearby L8–9 involved in cyclin H binding and thereby in the phosphorylation of S(NTD), whatever RARα is under an apo or holo form. Thus, one can suggest that this process cooperates with the conformational changes induced by RA-binding for hRARα transcriptional activity (fig. 8*B*).

However, despite the good conservation of the flanking basic K/R-rich motif, located at the end of the highly structured helix 9, our evolutionary analysis points out that S(LBD), located in L9–10, a disordered loop, is not universally conserved throughout vertebrates. Indeed, S(LBD) is present mostly in mammalian RARα, whereas an asparagine is present at this position at the basis of the chordate RAR tree, as exemplified by amphioxus. The same conclusion was made for the paralog RARγ but not for RARβ because S(LBD) was also found in teleost RARβ. From these observations, two major different mechanisms of evolution can be proposed. In the first one, S(LBD) might have appeared in the three RAR paralogs at the basis of vertebrates during the second round (2R) of duplication. Then, this residue might have been lost independently in the three RARs from several vertebrates and during the third round (3R) of duplication in teleost RARα and RARγ. In the second mechanism, the asparagine might change to a serine in each of the three mammalian RARs (α, β, and γ) after the two rounds of duplications and in teleost RARβ during the third round of duplication. This latter mechanism might be the most probable one, in line with the functional role of S(LBD). Nevertheles, whatever the mechanism is, our observations suggest a convergent evolution of a similar regulatory mechanisms that occured four times independently. However, as we still have no evidence whether the phosphorylation cascade described for RARα also occurs in the context of RARγ and RARβ, it is still premature to speculate on this convergence. Nevertheless, it is worth noting that due to its unstructured nature, loops

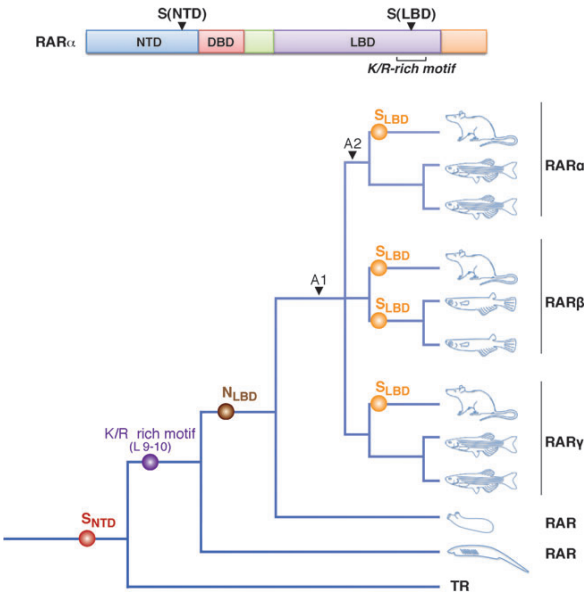


Fig. 7. Model for the evolution of RAR phosphorylation sites. S(NTD): presence of a phosphorylatable serine residue in the N-terminal proline-rich motif of the common ancestor of RAR and TR. K/R: acquisitions of an arginine-/lysine-rich motif in H9 of the common ancestor of chordate RARs. N(LBD), acquisition of an asparagine in the LBD (L9–10). S(LBD): acquisition of a phosphorylatable serine in the the LBD (L9–10). A1 designate the duplication leading to the three RARs and A2 the diversification of RARα in gnathostomes.

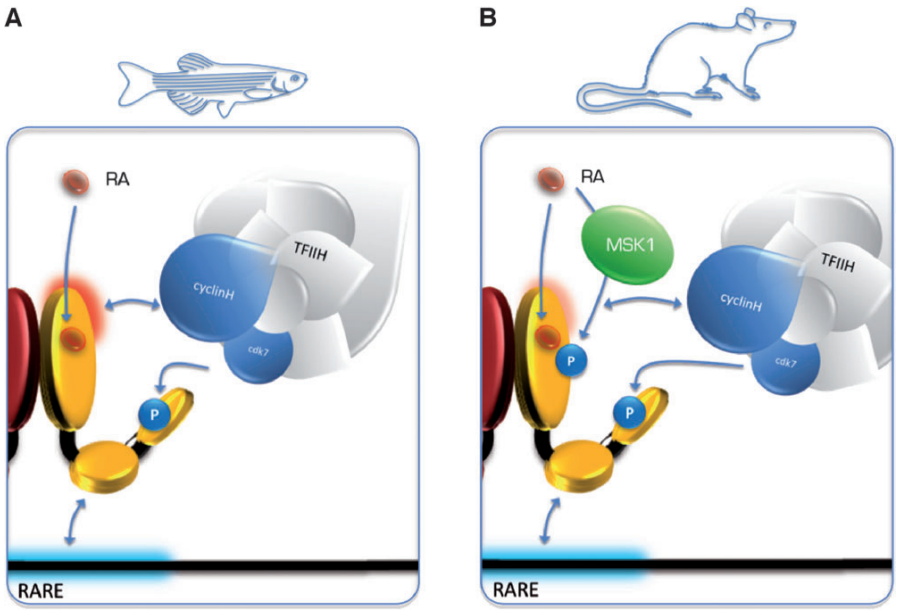


Fig. 8. Model for the evolution of the fine-tuned phosphorylation of RARα. (*A*) In RAR***α*** from nonmammalian species, exemplified by zebrafish, L8–9 is naturally highly flexible as shown by the intense red halo, allowing cyclin H binding and S(NTD) phosphorylation by cdk7. RAR***α*** activity is switched on by ligand binding. (*B*) In mammalian RAR***α***, L8–9 is naturally rigid making necessary a fine-tuned regulation of cyclin H recruitment by the phosphorylation of S(LBD). Subsequently, RAR***α*** activity requires not only ligand binding but also this evolved fine-tuning phosphorylation cascade.

L9–10 should respond rapidly and accurately to changing environmental conditions, that is, requirement of a phosphorylation or not (see below).

Besides, the cyclin H docking site of RARα evolved subtly in parallel to S(LBD). Indeed, in zebrafish RARα, which did not acquire S(LBD), L8–9 harbors a flexible conformation that favors cyclin H-binding without any requirement for a phosphorylation process in L9–10 (fig. 8*A*). Interestingly, amphoxius RAR also depicts a mutation in the cyclin H-binding domain (fig. 6), suggesting that the flexibility of this domain might be also increased. In contrast, in mammalian RARα, the acquisition of a serine in L9–10 is associated to a drastic reduction in the dynamics of L8–9 and in its ability to interact with cyclin H. The appearance of such a rigidity makes necessary a fine-tuned regulation by the phosphorylation of S(LBD) (fig. 8*B*). It is worth noting that both loops L8–9 and L9–10 correspond to disordered domains that evolve faster than ordered ones (Schaefer et al. 2010). In line with this, phosphosites frequently appear in such disordered regions, thus facilitating the evolution of kinase-signaling circuits (Beltrao et al. 2009; Holt et al. 2009; Landry et al. 2009).

Thus, we believe that during evolution, a selective pressure might push for rapid changes in the disordered loops L8–9 and L9–10 of the LBD, in order to maintain in a changing environment, the phosphorylation of the NTD that is essential for RARα transcriptional activity (fig. 8). Indeed, when L8–9 lost its flexibility, there was a strong pressure for compensation, that is, the appearance of a phosphorylatable serine in L9–10. Of note, MSK1, the kinase involved in the phosphorylation of S(LBD) is a vertebrate kinase, but an ortholog has been identified in drosophila (Jin et al. 1999), suggesting that the phosphorylation machinery predates chordates RAR diversification.

In conclusion, the present work highlights the evolutionary potential of the RARα phosphorylation network, especially at the level of the kinase–substrate interaction. As the complex combinatorial control of hRARα phosphorylation by multiple kinases is a readily evolved network, one can predict that its deregulation might be at the basis of disease. In support of such an hypothesis, we have shown that in Xeroderma Pigmentosum patients, RARα is not efficiently phosphorylated by cdk7 with characteristic downstream consequences on the expression of RAR target genes (Keriel et al. 2002). This has been correlated at least in part to the clinical abnormalities of the patients but also to their high risk of skin cancer in response to UV.

Supplementary Material

Supplementary figures S1–S3 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We thank Dr Yiping Liu (Shanghai Institute for Biological Science, China) for the gift of the zebrafish cyclin H cDNA, M. Oulad Abdelghani (Institut de Génétique et de Biologie Moléculaire et Cellulaire [IGBMC]) for the mouse monoclonal antibodies, and members of the cell culture facilities for help. Special thanks to all teams members for fruitful discussions and suggestions and to L. Azzab (IGBMC),

A. Rodriguez and A. Perret (Universite de Strasbourg) for help in the development of simulation protocols. This work was supported by funds from Centre National de la Recherche Scientifique (CNRS), INSERM, the Association pour la Recherche sur le Cancer (ARC 3169), the Agence Nationale pour la Recherche (ANR-05-BLAN-0390-02 and ANR-09-BLAN-0127-01), the Fondation pour la Recherche Médicale (DEQ20090515423), and the Institut National du Cancer (INCa-PL09-194). The Institut du Developpement et des Ressources en Informatique Scientifique (IDRIS), the Centre Informatique National de l'Enseignement Supérieur (CINES), and the Centre d'Etude du Calcul Parallèle de Strasbourg (Université de Strasbourg) are aknowledged for generous allocations of computer time. I.A. was supported by the Ligue Nationale contre le cancer.

References

Albalat R, Canestro C. 2009. Identification of Aldh1a, Cyp26 and RAR orthologs in protostomes pushes back the retinoic acid genetic machinery in evolutionary time to the bilaterian ancestor. *Chem Biol Interact*. 178:188–196.

Ali S, Lutz Y, Bellocq JP, Chenard-Neu MP, Rouyer N, Metzger D. 1993. Production and characterization of monoclonal antibodies recognising defined regions of the human oestrogen receptor. *Hybridoma* 12:391–405.

Amores A, Force A, Yan YL, et al. (13 co-authors). 1998. Zebrafish hox clusters and vertebrate genome evolution. *Science* 282:1711–1714.

Bastien J, Adam-Stitah S, Riedl T, Egly JM, Chambon P, Rochette-Egly C. 2000. TFIIH interacts with the retinoic acid receptor gamma and phosphorylates its AF-1-activating domain through cdk7. *J Biol Chem*. 275:21896–21904.

Basu U, Wang Y, Alt FW. 2008. Evolution of phosphorylation-dependent regulation of activation-induced cytidine deaminase. *Mol Cell*. 32:285–291.

Beltrao P, Trinidad JC, Fiedler D, Roguev A, Lim WA, Shokat KM, Burlingame AL, Krogan NJ. 2009. Evolution of phosphoregulation: comparison of phosphorylation patterns across yeast species. *PLoS Biol*. 7:e1000134.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. *Nucleic Acids Res*. 28:235–242.

Bertrand S, Thisse B, Tavares R, et al. (12 co-authors). 2007. Unexpected novel relational links uncovered by extensive developmental profiling of nuclear receptor expression. *PLoS Genet*. 3:e188.

Bour G, Gaillard E, Bruck N, Lalevee S, Plassat JL, Busso D, Samama JP, Rochette-Egly C. 2005. Cyclin H binding to the RAR {alpha} activation function (AF)-2 domain directs phosphorylation of the AF-1 domain by cyclin-dependent kinase 7. *Proc Natl Acad Sci U S A*. 102:16608–16613.

Bour G, Taneja R, Rochette-Egly C. 2006. Mouse embryocarcinoma F9 cells and retinoic acid. A model to study the molecular mechanisms of endodermal differentiation. In: Taneja R, editor. Nuclear receptors in development. New York: Elsevier Press Inc. p. 211–253.

Bourguet W, Vivat V, Wurtz JM, Chambon P, Gronemeyer H, Moras D. 2000. Crystal structure of a heterodimeric complex of RAR and RXR ligand-binding domains. *Mol Cell*. 5:289–298.

Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S. 1983. CHARMM: a program for macromolecular energy minimization and dynamics calculations. *J Comp Chem*. 4:187–217.

Bruck N, Vitoux D, Ferry C, Duong V, Bauer A, de The H, Rochette-Egly C. 2009. A coordinated phosphorylation cascade initiated by p38MAPK/MSK1 directs RARalpha to target promoters. *EMBO J*. 28:34–47.

Brunger AT, Karplus M. 1988. Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. *Proteins* 4:148–156.

Campo-Paysaa F, Marletaz F, Laudet V, Schubert M. 2008. Retinoic acid signaling in development: tissue-specific functions and evolutionary origins. *Genesis* 46:640–656.

Canestro C, Postlethwait JH, Gonzalez-Duarte R, Albalat R. 2006. Is retinoic acid genetic machinery a chordate innovation? *Evol Dev*. 8:394–406.

Canutescu AA, Shelenkov AA, Dunbrack RL Jr. 2003. A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci*. 12:2001–2014.

Damagnez V, Makela TP, Cottarel G. 1995. Schizosaccharomyces pombe Mop1-Mcs2 is related to mammalian CAK. *EMBO J*. 14:6164–6172.

de StGroth SF, Scheidegger D. 1980. Production of monoclonal antibodies: strategy and tactics. *J Immunol Methods*. 35:1–21.

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

Dolle P. 2009. Developmental expression of retinoic acid receptors (RARs). *Nucl Recept Signal*. 7:e006.

Dyson HJ, Wright PE. 2005. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol*. 6:197–208.

Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*. 5:113.

Escriva H, Bertrand S, Germain P, Robinson-Rechavi M, Umbhauer M, Cartry J, Duffraisse M, Holland L, Gronemeyer H, Laudet V. 2006. Neofunctionalization in vertebrates: the example of retinoic acid receptors. *PLoS Genet*. 2:e102.

Escriva H, Holland ND, Gronemeyer H, Laudet V, Holland LZ. 2002a. The retinoic acid signaling pathway regulates anterior/posterior patterning in the nerve cord and pharynx of amphioxus, a chordate lacking neural crest. *Development* 129:2905–2916.

Escriva H, Manzon L, Youson J, Laudet V. 2002b. Analysis of lamprey and hagfish genes reveals a complex history of gene duplications during early vertebrate evolution. *Mol Biol Evol*. 19:1440–1450.

Fidelak J, Ferrer S, Oberlin M, Moras D, Dejaegere A, Stote RH. 2010. Dynamic correlation networks in human peroxisome proliferator-activated receptor-gamma nuclear receptor protein. *Eur Biophys J*. 39:1503–1512.

Fujiwara S. 2006. Retinoids and nonvertebrate chordate development. *J Neurobiol*. 66:645–652.

Gaillard E, Bruck N, Brelivet Y, Bour G, Lalevee S, Bauer A, Poch O, Moras D, Rochette-Egly C. 2006. Phosphorylation by protein kinase A potentiates retinoic acid repeptor a activity by means of increasing interaction with and phosphorylation by cyclin H/cdk7. *Proc Natl Acad Sci U S A*. 103:9548–9553.

Gaillard T, Dejaegere A, Stote RH. 2009. Dynamics of beta3 integrin I-like and hybrid domains: insight from simulations on the mechanism of transition between open and closed forms. *Proteins* 76:977–994.

Germain P, Kammerer S, Perez E, et al. (13 co-authors). 2004. Rational design of RAR-selective ligands revealed by RARbeta crystal stucture. *EMBO Rep*. 5:877–882.

Germain P, Staels B, Dacquet C, Spedding M, Laudet V. 2006. Overview of nomenclature of nuclear receptors. *Pharmacol Rev*. 58:685–704.

Glineur C, Zenke M, Beug H, Ghysdael J. 1990. Phosphorylation of the v-erbA protein is required for its function as an oncogene. *Genes Dev*. 4:1663–1676.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*. 52:696–704.

Hisata K, Fujiwara S, Tsuchida Y, Ohashi M, Kawamura K. 1998. Expression and function of a retinoic acid receptor in budding ascidians. *Dev Genes Evol*. 208:537–546.

Hoekstra HE, Coyne JA. 2007. The locus of evolution: evo devo and the genetics of adaptation. *Evolution* 61:995–1016.

Holt LJ, Tuch BB, Villen J, Johnson AD, Gygi SP, Morgan DO. 2009. Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science* 325:1682–1686.

Jaillon O, Aury JM, Brunet F, et al. (61 co-authors). 2004. Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature* 431:946–957.

Jin Y, Wang Y, Walker DL, Dong H, Conley C, Johansen J, Johansen KM. 1999. JIL-1: a novel chromosomal tandem kinase implicated in transcriptional regulation in Drosophila. *Mol Cell*. 4:129–135.

Kast D, Espinoza-Fonseca LM, Yi C, Thomas DD. 2010. Phosphorylation-induced structural changes in smooth muscle myosin regulatory light chain. *Proc Natl Acad Sci U S A*. 107: 8207–8212.

Keriel A, Stary A, Sarasin A, Rochette-Egly C, Egly JM. 2002. XPD mutations prevent TFIIH-dependent transactivation by nuclear receptors and phosphorylation of RARalpha. *Cell* 109:125–135.

Klaholz BP, Mitschler A, Moras D. 2000. Structural basis for isotype selectivity of the human retinoic acid nuclear receptor. *J Mol Biol*. 302:155–170.

Kuraku S, Meyer A, Kuratani S. 2009. Timing of genome duplications relative to the origin of the vertebrates: did cyclostomes diverge before or after? *Mol Biol Evol*. 26:47–59.

Lalevee S, Bour G, Quinternet M, et al. (11 co-authors). 2010. Vinexin{beta}, an atypical "sensor" of retinoic acid receptor {gamma} signaling: union and sequestration, separation, and phosphorylation. *FASEB J*. 24:4523–4534.

Lalevee S, Ferry C, Rochette-Egly C. 2010. Phosphorylation control of nuclear receptors. *Methods Mol Biol*. 647:251–266.

Landry CR, Levy ED, Michnick SW. 2009. Weak functional constraints on phosphoproteomes. *Trends Genet*. 25:193–197.

Lavery DN, McEwan IJ. 2005. Structure and function of steroid receptor AF1 transactivation domains: induction of active conformations. *Biochem J*. 391:449–464.

Linville A, Radtke K, Waxman JS, Yelon D, Schilling TF. 2009. Combinatorial roles for zebrafish retinoic acid receptors in the hindbrain, limbs and pharyngeal arches. *Dev Biol*. 325:60–70.

Liu QY, Wu ZL, Lv WJ, Yan YC, Li YP. 2007. Developmental expression of cyclin H and Cdk7 in zebrafish: the essential role of cyclin H during early embryo development. *Cell Res*. 17:163–173.

MacKerell AD, Bashford D, Bellott M, et al. (29 co-authors). 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B*. 102:3586–3616.

Mackerell AD Jr, Feig M, Brooks CL. 2004. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem*. 25:1400–1415.

Mark M, Ghyselinck NB, Chambon P. 2009. Function of retinoic acid receptors during embryonic development. *Nucl Recept Signal*. 7:e002.

Marletaz F, Holland LZ, Laudet V, Schubert M. 2006. Retinoic acid signaling and the evolution of chordates. *Int J Biol Sci*. 2:38–47.

Nguyen Ba AN, Moses AM. Forthcoming 2010. Evolution of characterized phosphorylation sites in budding yeast. *Mol Biol Evol*. 27:2027–2037.

Pogenberg V, Guichou JF, Vivat-Hannah V, Kammerer S, Perez E, Germain P, de Lera AR, Gronemeyer H, Royer CA, Bourguet W. 2005. Characterization of the interaction between retinoic acid receptor/retinoid X receptor (RAR/RXR) heterodimers and transcriptional coactivators through structural and fluorescence anisotropy studies. *J Biol Chem*. 280:1625–1633.

Postlethwait JH, Yan YL, Gates MA, et al. (29 co-authors). 1998. Vertebrate genome evolution and the zebrafish gene map. *Nat Genet*. 18:345–349.

Renaud JP, Moras D. 2000. Structural studies on nuclear receptors. *Cell Mol Life Sci*. 57:1748–1769.

Renaud JP, Rochel N, Ruff M, Vivat V, Chambon P, Gronemeyer H, Moras D. 1995. Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid. *Nature* 378:681–689.

Robinson-Rechavi M, Boussau B, Laudet V. 2004. Phylogenetic dating and characterization of gene duplications in vertebrates: the cartilaginous fish reference. *Mol Biol Evol*. 21:580–586.

Rochette-Egly C. 2003. Nuclear receptors: integration of multiple signalling pathways through phosphorylation. *Cell Signal*. 15:355–366.

Rochette-Egly C, Adam S, Rossignol M, Egly JM, Chambon P. 1997. Stimulation of RAR alpha activation function AF-1 through binding to the general transcription factor TFIIH and phosphorylation by CDK7. *Cell* 90:97–107.

Rochette-Egly C, Germain P. 2009. Dynamic and combinatorial control of gene expression by nuclear retinoic acid receptors. *Nucl Recep Signal*. 7:e005.

Rochette-Egly C, Oulad-Abdelghani M, Staub A, Pfister V, Scheuer I, Chambon P, Gaub MP. 1995. Phosphorylation of the retinoic acid receptor-alpha by protein kinase A. *Mol Endocrinol*. 9:860–871.

Ryckaert JP, Ciccotti G, Berendsen HJC. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comp Phys*. 23:327–341.

Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*. 234:779–815.

Sato Y, Ramalanjaona N, Huet T, et al. (13 co-authors). 2010. The "phantom effect" of the rexinoid LG100754: structural and functional insights. *PloS One*. 5:e15119.

Schaefer C, Schlessinger A, Rost B. 2010. Protein secondary structure appears to be robust under in silico evolution while protein disorder appears not to be. *Bioinformatics* 26:625–631.

Schaefer M, van Vlijmen HW, Karplus M. 1998. Electrostatic contributions to molecular free energies in solution. *Adv Protein Chem*. 51:1–57.

Theodosiou M, Laudet V, Schubert M. 2010. From carrot to clinic: an overview of the retinoic acid signaling pathway. *Cell Mol Life Sci*. 67:1423–1445.

Vucetic Z, Zhang Z, Zhao J, Wang F, Soprano KJ, Soprano DR. 2008. Acinus-S' represses retinoic acid receptor (RAR)-regulated gene expression through interaction with the B domains of RARs. *Mol Cell Biol*. 28:2549–2558.

Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci*. 13:555–556.

IV. Discussion - Perspectives

1. L'acquisition de modifications post-traductionnelles (PTM) par mutations dans les séquences codantes des gènes.

Nos résultats apportent un exemple supplémentaire de la sélection positive d'une mutation dans la séquence codante des gènes au cours de l'évolution. En effet, comme nous le discutions dans l'introduction (section B-III-1.2, page 140), il est généralement admis que des mutations dans les régions *cis*-régulatrices des gènes sont plus fréquemment impliquées dans l'évolution de l'expression et de la régulation des gènes du développement. En effet, à l'inverse des mutations dans les régions codantes, elles sont supposées être moins délétères car ne touchent pas directement la fonction et l'activité de la protéine codée (Figure 34), Toutefois, nos résultats suggèrent que l'acquisition de résidus pouvant être modifiés en post-traductionnel (PTM) peut permettre une modification subtile de la régulation de la protéine en impliquant de nouvelles voies de signalisation pour le contrôle de son activité. S(LBD) est présente au sein d'une boucle du LBD qui est une région désordonnée et il a été montré que de telles régions semblent évoluer plus rapidement du fait qu'elles sont moins directement impliquées dans le maintien strict d'une structure. Ainsi, il serait intéressant d'étudier l'évolution de la séquence de ces régions désordonnées car elles peuvent être le siège de mutations ayant participé à moduler l'activité de la protéine tout en contournant les contraintes sélectives maintenant sa structure. Dans le cas de RARα, il est intéressant de noter que les régions les moins conservées au sein des vertébrés correspondent à des zones fortement exposées et souvent à des régions désordonnées comme les boucles entre hélices du LBD (Figure 35). Cette « cartographie » des divergences sur la structure du LBD permet d'identifier des régions « points chauds » de divergence qui pourraient être impliquées dans l'évolution de la

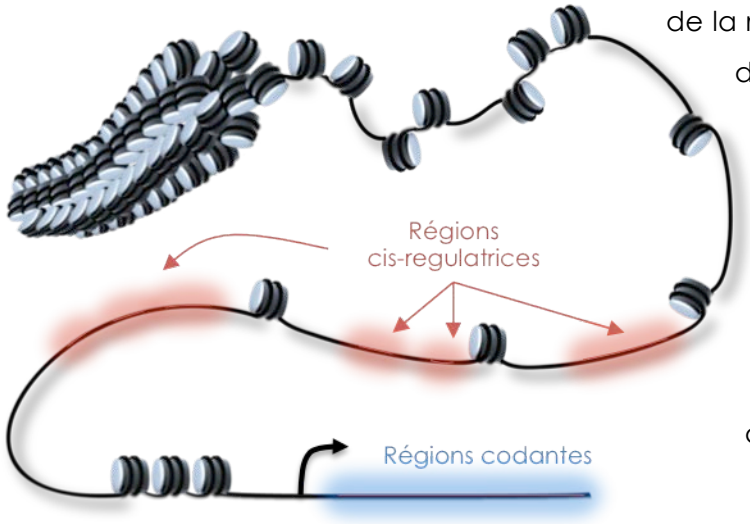


Figure 34: Mutations en cis- ou dans les régions codantes?

régulation de l'activité du récepteur, entre autre en hébergeant des sites modifiés par PTM. La boucle oméga entre les hélices 2 et 3 du LBD est particulièrement intéressante car elle présente un site de phosphorylation chez RXR (Bruck et al., 2005). A noter que la S(LBD) et le domaine de liaison de la cycline H dans la boucle entre les hélices 8 et 9 constituent deux « points chauds » de divergence (Figure 35).

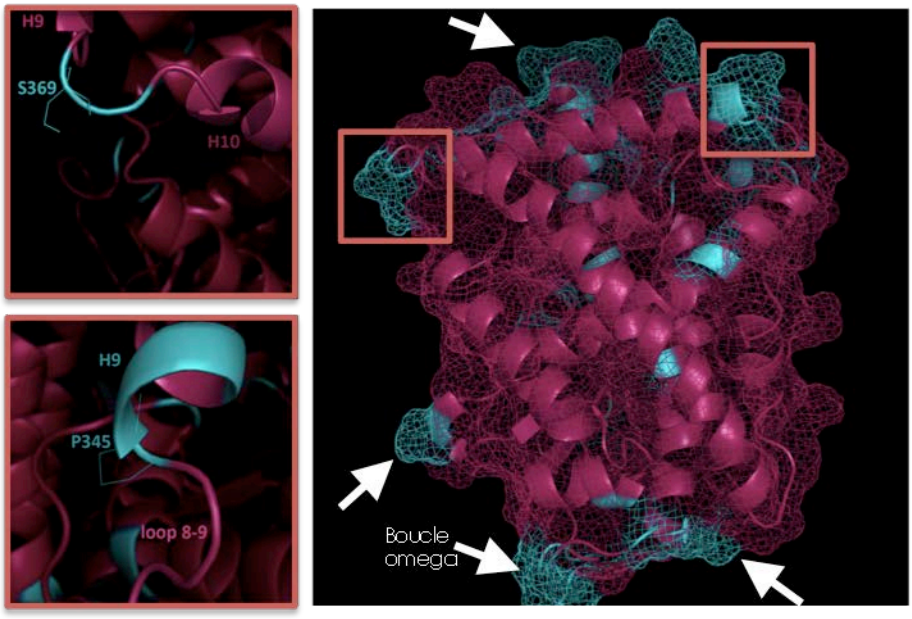


Figure 35: Points chauds de divergence au sein du LBD des RAR. La conservation de la séquence protéique du LBD des RAR entre différentes espèces de vertébrés est superposée à la structure du LBD de RAR humain (3LBD) par le serveur Consurf. Les zones de forte conservation sont en rose alors que les zones de forte divergence dans la séquence protéique sont en bleu. Ces points chauds de divergence (pointés par des flèches blanches) sont fréquemment retrouvés dans des régions désordonnées et exposées au solvant comme des boucles, et pourraient correspondre à des régions impliquées dans l'évolution de la régulation de l'activité du récepteur (comme la boucle oméga). A noter que la boucle entre les hélices 9 et 10 contenant S(LBD) et la boucle entre les hélices 8 et 9 correspondant au site de liaison de la cycline H sont des points chauds de divergence (cadres rouges).

2. Régulation allostérique et changements dynamiques induits par la phosphorylation.

Pour compléter nos modélisations des conséquences de la phosphorylation de S(LBD) sur la dynamique de flexibilité du LBD, en particulier de la boucle 8-9, le laboratoire de notre collaboratrice Annick Dejaegere a récemment disséqué les changements conformationnels induits par la phosphorylation sur le LBD de RARα (Chebaro et al., 2013). Cette étude montre que la structure globale du LBD n'est pas modifiée après

phosphorylation de S(LBD). Cependant, elle a une conséquence subtile sur la courbure de l'hélice H9 qui joue un rôle majeur dans la « liberté conformationnelle » de la boucle 8-9. En effet, la phosphorylation de S(LBD) module un réseau local de ponts salins entre les hélices H4, H8 et H9, dont la conséquence structurale est la réduction de la courbure et donc l'extension de l'hélice H9. Ces changements subtiles sont à l'origine de changements dynamiques au sein du LBD comme illustré dans la figure 36 menant principalement à une augmentation de la flexibilité et donc de la « liberté de conformation » de la boucle 8-9 propice à une interaction protéique.

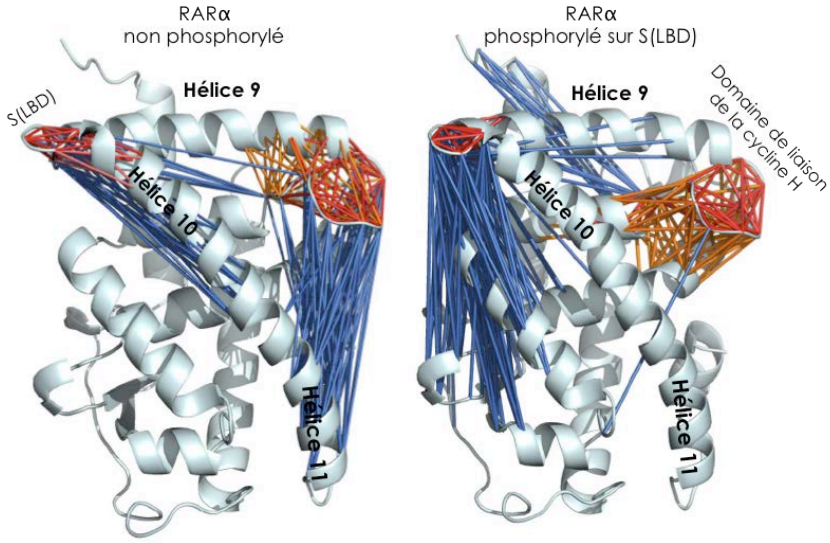


Figure 36: Changements de la dynamique au sein du LBD après phosphorylation de S(LBD). Images de Chebaro et al., 2013. La simulation de la phosphorylation de S(LBD) indique un changement de la dynamique autour de l'hélice 9. Les liens colorés entre les régions du LBD indiquent les zones dont les mouvements sont corrélés positivement (en bleu) ou négativement (orange et rouge). On remarque une perte des corrélations positives entre les hélices 9 et 11 et leur augmentation autour de la boucle entre les hélices 9 et 10 contenant S(LBD).

3. Vers l'étude du rôle de la phosphorylation des RAR *in vivo*

Nos travaux mettent en avant une conservation forte de S(NTD) chez tous les RAR de vertébrés suggérant que la phosphorylation de ce site a été soumis à une forte pression de sélection au cours de l'évolution. Cela suggère également un rôle crucial de la phosphorylation des RAR au niveau de S(NTD) et les travaux précédents du laboratoire ont montré leur importance pour le recrutement de RARα au niveau des promoteurs de gènes cibles dans des cellules de mammifères (Bruck et al., 2009). Il apparaît maintenant important de montrer le rôle de ces phosphorylations *in vivo* en particulier dans des processus développementaux régulés par l'AR.

3.1. La différenciation de cellules souches en neurones induite par l'AR requiert la phosphorylation de RARγ.

Le traitement de cellules souches de souris (cellule ES) à l'AR conduit à leur différentiation en progéniteurs de neurones en quelques jours (Bibel et al., 2007). Des travaux en cours dans le laboratoire de Cécile Rochette-Egly ont montré que c'est le sous-type gamma de RAR qui est requis lors de cette différentiation, et plus particulièrement l'isoforme 2 de RARγ (RARγ2). En effet, alors que les cellules ES invalidées pour les RAR (KO) ne peuvent plus se différencier en neurones, la différenciation est restaurée lorsqu'on exprime à nouveau RARγ2, mais ne l'est pas avec RARα ou RARγ1 (Al Tanoury et al. En préparation). Il a aussi été montré que dans les cellules ES, la voie des ERK (p42/p44MAPK) était rapidement activée par l'AR et phosphoryle RARγ au niveau de S(NTD) (en position 68 chez RARγ murin). De manière très intéressante, la phosphorylation de RARγ est nécessaire pour permettre la différentiation des cellules ES en neurones. En effet, alors que l'expression de RARγ2 dans des cellules ES KO permet de restaurer la différentiation en neurones, l'expression d'un phospho-mutant de RARγ où la S(NTD) est substituée par une alanine (RARγ S-A) ne permet pas cette restauration. En particulier, des expériences de ChIP ont montré que le recrutement de RARγ au niveau du promoteur de gènes impliqués dans la différentiation de cellules ES en neurones était abolit avec un phospho-mutant RARγ S-A. De plus, il semble que la phosphorylation de RARγ soit particulièrement nécessaire pour le recrutement du récepteur au niveau de RARE précis dont les DR5 et DR7. A noter qu'à l'inverse, certains gènes cibles de l'AR non impliqués dans la différentiation neuronale ne requièrent pas la phosphorylation de RARγ pour être régulés par l'AR.

L'ensemble de ces résultats confirme que la phosphorylation des RAR joue un rôle essentiel dans leur activité transcriptionnelle mais montre également pour la première fois leur rôle dans un processus biologique impliqué dans le développement. De plus, ces données révèlent qu'il existe des gènes cibles de l'AR phospho-dépendants car ne pouvant pas être régulés par un RAR phospho-mutant, et des gènes phospho-indépendants répondant à l'AR même en l'absence de phosphorylation des RAR. Bien que les mécanismes précis de cette dépendance reste flous, il semble que RARγ phosphorylé soit spécifiquement recruté au niveau de RARE de type DR5 et DR7.

3.2. Etudier le rôle de la phosphorylation des RAR au cours du développement du poisson zèbre.

i. Sauvetage phénotypique

Un des objectifs initiaux de mes travaux de thèse était d'étudier le rôle de la phosphorylation des RAR au cours du développement du poisson-zèbre. Nous avons reproduit le phénotype décrit de perte des nageoires pectorales chez les embryons injectés avec un morpholino contre RARα-B (Figure 37) (Linville et al., 2009). La stratégie consistait à restaurer le développement normal des nageoires chez ces morphants par sauvetage phénotypique après injection d'ARNm codant pour RARα-B sous la forme sauvage (WT) ou phospho-mutante (S-A). La séquence de l'ARNm injecté correspondait à celle du gène de RARα-B de poisson-zèbre mais avec des mutations silencieuses qui assuraient que l'ARNm injecté ne soit pas sensible au morpholino. Ainsi, on pouvait étudier si la forme phospho-mutante de RARα-B était ou non capable de restaurer le développement des nageoires chez des embryons morphants. La réponse a cette question permettrait de comprendre le rôle de la phosphorylation des RAR *in vivo* dans un processus développemental. Cependant, malgré de très nombreuses tentatives, nous n'avons observé aucun sauvetage phénotypique après l'injection de la forme sauvage du récepteur. En cherchant à comprendre cette absence de sauvetage, j'ai observé une dégradation très rapide de la protéine RARα-B traduite à partir de l'ARNm injecté (Figure 38) : la présence de cette protéine n'était plus détectable 16 heures après l'injection. Or l'induction du développement de la nageoire, entre autre via l'expression de *tbx5* ne commence qu'à partir de 23 hpf (Begemann and Ingham, 2000; Grandel and Schulte-Merker, 1998). Ainsi, il est probable que la protéine RARα-B traduite à partir de l'ARNm injecté ne soit plus présente en assez grande quantité lorsqu'elle est requise pour le développement de la nageoire.

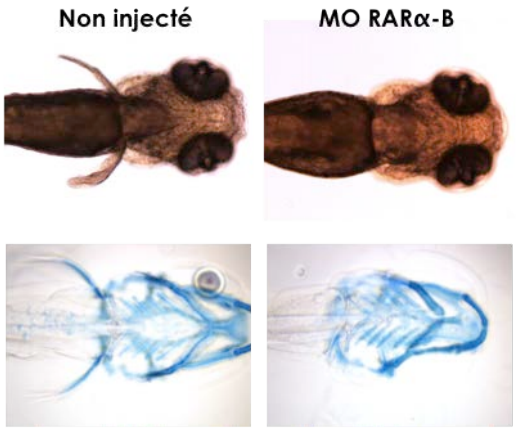


Figure 37: Phénotype de perte des nageoires pectorales par invalidation de RARα-B. Nous avons reproduit avec succès le phénotype décrit par Linville et al., 2009.

Dans le but de mieux comprendre les mécanismes de cette dégradation rapide, nous avons par la suite construit des récepteurs mutants, tronqués au niveau de différentes

régions et avons suivi leur cinétique de dégradation après injection de l'ARNm correspondant (Figure 38). Les constructions dont l'expression protéique perdure le plus longtemps après injection correspondent au récepteur tronqué dans la partie A/B ou au LBD seul, avec une expression de la protéine détectable jusqu'à 20 hpf (Figure 38). Comme cela a déjà été décrit pour les RAR de mammifères (Kopf et al., 2000), cela suggère que la région A/B intègre des signaux de dégradation, possiblement via la phosphorylation de S(NTD). En conclusion, les constructions permettant une expression plus longue de la protéine ne contiennent pas le site de phosphorylation S(NTD) dont nous souhaitons étudier le rôle *in vivo*.

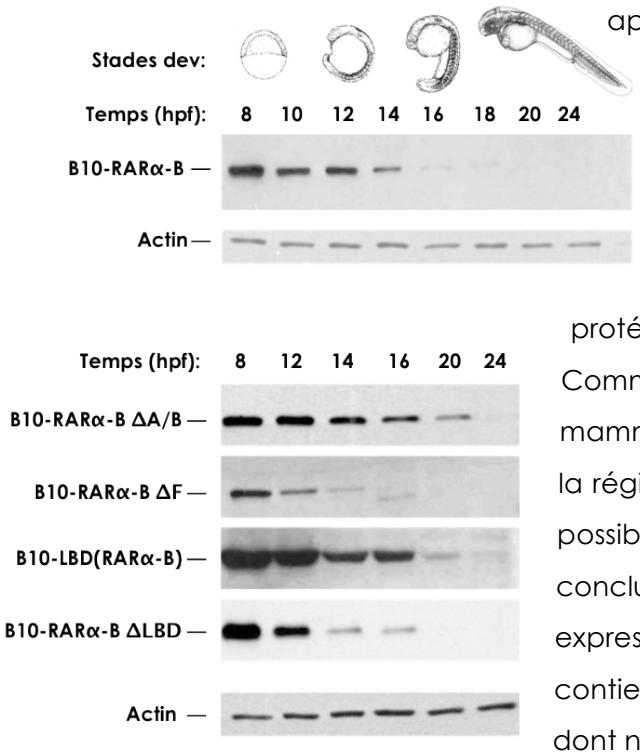


Figure 38 –Cinétique d'expression de la protéine RARα-B après injection d'ARNm. L'ARNm codant pour RARα-B auquel un tag B10 a été ajouté en N-terminal, est injecté au stade 1-cellule. Après 8hpf, l'expression de la protéine est contrôlée toutes les deux heures par western blot en utilisant un anticorps contre le tag B10. Différents mutants correspondant au récepteur RARα-B tronqué de (i) la région A/B, (ii) la région F, (iii) au LBD seul ou (iv) au récepteur tronqué du LBD sont également injectés.

ii. Génération de lignées transgéniques

Pour circonvenir ce problème, il est donc nécessaire de générer une lignée transgénique exprimant notre transgène de manière stable. L'une des perspectives envisagée et qui fait actuellement l'objet d'une demande de financement, est d'utiliser la technique de mutagénèse par TALEN (Transcription Activator-Like Effector Nuclease) qui est établit dans le laboratoire de Vincent Laudet. Les TALEN correspondent à des enzymes de restriction artificielles issues de la fusion entre un domaine de liaison à l'ADN de type TALE pouvant cibler une séquence d'ADN d'intérêt et un domaine de clivage de l'ADN (Boch, 2011). Elles clivent la double hélice d'ADN au niveau de la séquence reconnue et permettent l'insertion ou la délétion de quelques nucléotides par *non-homologous end-joining*. Ainsi, l'injection d'un couple de TALEN ciblant la S(NTD) de RARα-B permettrait de générer simultanément des embryons KO pour RARα-B (dans le cas d'une insertion ou

délétion décalant le cadre de lecture) et des embryons exprimant la protéine RARα-B phospho-mutante (si l'insertion-délétion ne modifie pas le cadre de lecture) (Figure 39). La comparaison des phénotypes et des transcriptomes entre les embryons KO et phospho-mutants permettrait d'appréhender le rôle de la phosphorylation de S(NTD) des RAR au cours du développement, en étudiant la protéine endogène.

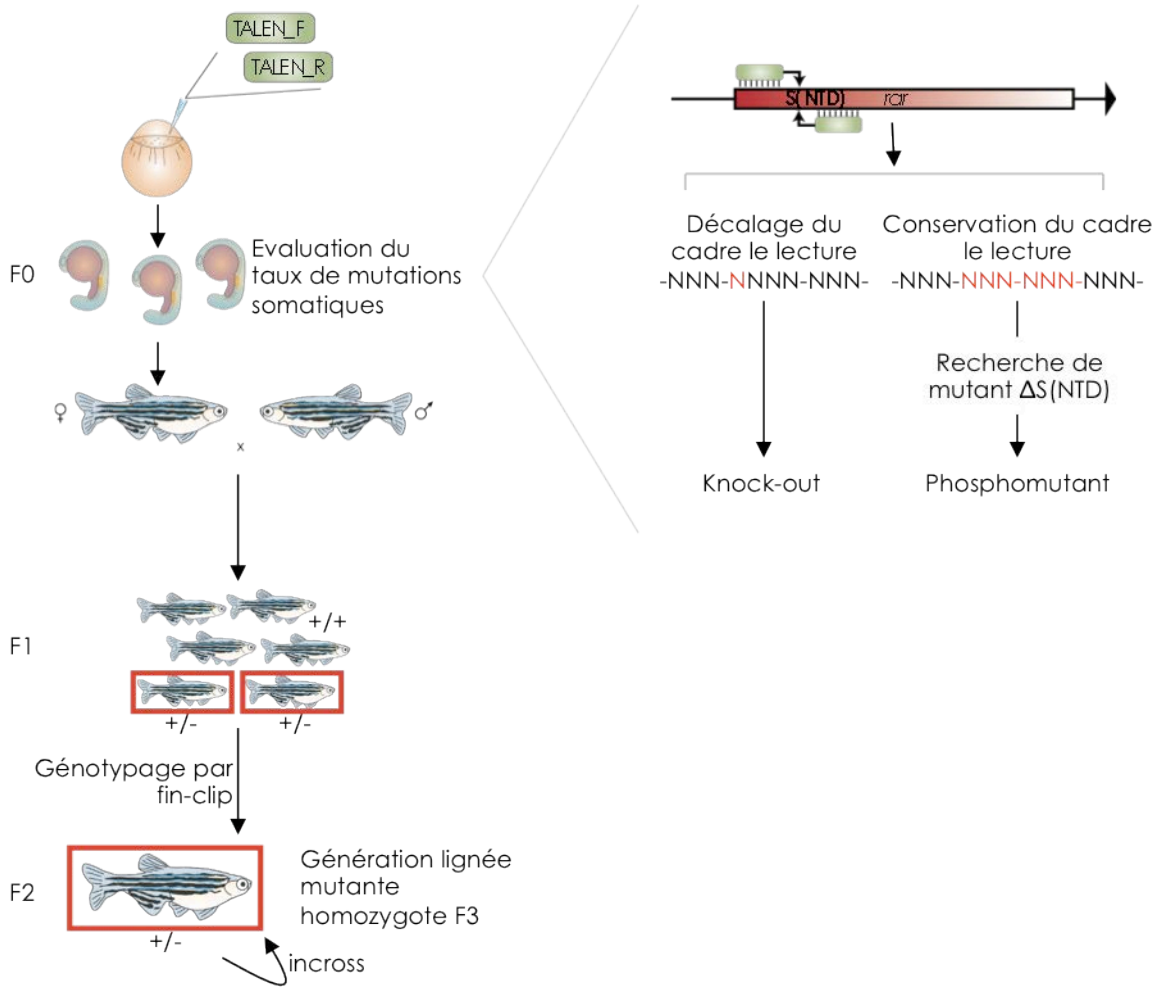


Figure 39: Mutagénèse par TALEN pour la génération de lignées KO et phospho-mutante ΔS(NTD). L'utilisation de TALEN ciblant le site S(NTD) d'un RAR donné permettrait de générer dans un même temps une lignée KO et une lignée phospho-mutante où S(NTD) est substituée ou deletée. La lignée homozygote est générée en F3 (temps estimé: 9 mois).

Chapitre#3

Le rôle de l'AR dans la diversification de la denture chez les poissons

I. Contexte Scientifique

Il existe une très grande diversité de denture chez les vertébrés et particulièrement chez les poissons. Ainsi, c'est un modèle de choix pour l'étude des mécanismes développementaux à l'origine d'une telle diversité (section B-III-4.1, page 47). Des travaux du laboratoire de Vincent Laudet ont mis en évidence un rôle de l'AR dans l'induction des dents chez le poisson-zèbre. En effet, l'inhibition de la synthèse endogène d'AR inhibe l'expression de marqueur dentaire comme *dlx2a* et empêche le développement des dents (Gibert et al., 2010). Toutefois, l'effet d'un excès d'AR sur le développement des dents n'a jamais été testé. De ce fait, nous avons entrepris d'étudier les effets d'un excès exogène ou endogène d'AR chez le poisson-zèbre et d'autres espèces de cypriniformes.

II. Principaux résultats

- **Un traitement précoce à l'AR induit une expansion antérieure de la denture pharyngienne chez le poisson-zèbre.** L'expression de marqueurs dentaires comme *pitx2* ou *dlx2b* est également étendue antérieurement chez des embryons traités. L'expansion de la denture induite par l'AR ressemble à la denture d'autres poissons téléostéens comme le médaka.
- **Les enzymes *raldh2* et *cyp26b1* contrôlent le métabolisme de l'AR dans la région de formation des dents chez le poisson-zèbre**. La dynamique d'expression temporelle de l'enzyme de synthèse (*raldh2*) et de dégradation (*cyp26b1*) de l'AR définit une fenêtre d'action temporelle et locale au niveau de la région de formation des dents dans le pharynx.
- **La mutation de *cyp26b1* induit la formation d'une dent supplémentaire chez le poisson-zèbre.** Alors que 5 dents sont observées sur la rangée ventrale du 5ème arc branchial chez le poisson-zèbre adulte, les mutants hétérozygotes *cyp26b1*$^{+/-}$ développent une 6ème dent qui est attachée normalement.
- **Un traitement tardif à l'AR induit un changement de la forme des dents chez l'embryon de poisson-zèbre.** Dans de rares cas, le traitement à l'AR induit la formation de dents bicuspides ce qui n'est jamais observé de manière naturelle chez le poisson-zèbre. De plus, un phénotype plus pénétrant observé est un allongement et un affinement de la forme des dents après traitement à l'AR.

- **Les phénotypes au niveau de la denture induits par l'AR sont retrouvés chez d'autres espèces de cypriniformes.** Plusieurs espèces développent 6 dents sur la rangée ventrale du 5ème arc branchial comme observé chez les mutants *cyp26b1*$^{+/-}$ de poisson zèbre. D'autre part, *Tanichthys albonubes* est un cypriniforme dont les dents sont plus allongées et plus fines que le poisson-zèbre et sa denture ressemble à celle induite par un traitement tardif à l'AR. Nous avons corrélé l'acquisition de ce phénotype à une différence de l'expression temporelle de *cyp26b1* dans le pharynx postérieur entre les deux espèces.

III. Publications

L'ensemble de nos résultats est présenté dans les deux articles à suivre.

1. Retinoic acid expands the evolutionary reduced dentition of zebrafish

Ce travail est issu du laboratoire de William Jackman du *Bowdoin College* aux Etats-Unis avec qui nous avons collaboré. Ma participation a été la description comparative de l'expression des enzymes de synthèse de l'AR chez le poisson-zèbre et le médaka. Ces travaux ont été publiés en 2012 dans *The FASEB Journal*.

The FASEB Journal article fj.12-209304. Published online August 31, 2012.

The FASEB Journal • Research Communication

Retinoic acid expands the evolutionarily reduced dentition of zebrafish

Pawat Seritrakul,* Eric Samarut,† Tenzing T. S. Lama,* Yann Gibert,‡ Vincent Laudet,† and William R. Jackman*,[1]

*Department of Biology, Bowdoin College, Brunswick, Maine, USA; †Molecular Zoology Group, Institut de Génomique Fonctionnelle de Lyon, Université de Lyon, Centre National de la Recherche Scientifique, Institut National de la Recherche Agronomique, Université Claude Bernard Lyon 1, Ecole Normale Supérieure de Lyon, Lyon, France; and ‡Deakin University School of Medicine, Waurn Ponds, Victoria, Australia

ABSTRACT **Zebrafish lost anterior teeth during evolution but retain a posterior pharyngeal dentition that requires retinoic acid (RA) cell-cell signaling for its development. The purposes of this study were to test the sufficiency of RA to induce tooth development and to assess its role in evolution. We found that exposure of embryos to exogenous RA induces a dramatic anterior expansion of the number of pharyngeal teeth that later form and shifts anteriorly the expression patterns of genes normally expressed in the posterior tooth-forming region, such as *pitx2* and *dlx2b*. After RA exposure, we also observed a correlation between cartilage malformations and ectopic tooth induction, as well as abnormal cranial neural crest marker gene expression. Additionally, we observed that the RA-induced zebrafish anterior teeth resemble in pattern and number the dentition of fish species that retain anterior pharyngeal teeth such as medaka but that medaka do not express the *aldh1a2* RA-synthesizing enzyme in tooth-forming regions. We conclude that RA is sufficient to induce anterior ectopic tooth development in zebrafish where teeth were lost in evolution, potentially by altering neural crest cell development, and that changes in the location of RA synthesis correlate with evolutionary changes in vertebrate dentitions.—Seritrakul, P., Samarut, E., Lama, T. T. S., Gibert, Y., Laudet, V., Jackman, W. R. Retinoic acid expands the evolutionarily reduced dentition of zebrafish. *FASEB J.* 26, 000–000 (2012). www.fasebj.org**

Key Words: teeth · neural crest cells · retinaldehyde dehydrogenase

Primitive ray-finned fish possessed numerous teeth in their pharyngeal regions, and the generally reduced dentitions of modern species are thought to be largely the result of evolutionary tooth loss (1, 2). The zebrafish (*Danio rerio*) lineage, as a member of the order Cypriniformes, lost anterodorsal pharyngeal teeth 65 or more million years ago (3, 4), and only retain part of the original posteroventral pharyngeal dentition. A more extensive dentition has not reevolved in any of the nearly 3000 described cypriniform species, despite diverse feeding modes, including piscivory (5). If there is indeed selective pressure to increase the dentition but it is no longer possible due to developmental constraints, this may represent a modern example of Dollo's law of the irreversibility of evolution (6–8). Such irreversibility may be the result of pleiotropic interactions during development (9, 10).

Retinoic acid (RA), an oxidized form of vitamin A, is a cell-cell signaling molecule with numerous functions in developing vertebrate embryos. Evidence of the crucial roles RA plays in development was established over half a century ago, when the offspring of female rodents fed a vitamin A-deficient diet during pregnancy showed congenital defects in many organ systems (11). Synthesized from carotenoids, the active all-*trans* form of RA acts as a ligand for nuclear RA receptors, which in turn function as transcription factors to control the expression of numerous targets, including Hox genes (12). The amphiphilic nature of RA, with both polar and nonpolar moieties, allows this molecule to pass relatively easily through cell membranes and act as a positional cue for cells during embryogenesis (13).

In vertebrates, RA is particularly important in anteroposterior (AP) patterning and in the development of cranial neural crest (CNC) cells. During the developmental specification of the AP axis, RA can directly regulate Hox gene expression, and alterations in RA levels result in homeotic changes in AP identity (14–16). Such RA-induced homeotic transformations include the duplication of skeletal elements, such as vertebrae (17, 18). RA is also required for the proper specification of CNC cells, a class of vertebrate pluripotent migratory cells that give rise to cephalic structures, including cartilage, bone, and teeth (19, 20). Disrup-

Abbreviations: AP, anteroposterior; CB5, 5th ceratobranchial element; CNC, cranial neural crest; dpf, days postfertilization; hpf, hours postfertilization; RA, retinoic acid

[1] Correspondence: Bowdoin College, 6500 College Station, Brunswick, ME 04011, USA. E-mail: wjackman@bowdoin.edu

doi: 10.1096/fj.12-209304

tions in RA signaling cause malformations of CNC-derived tissues in many vertebrates (21, 22).

Teeth exhibit conservation in their early development across vertebrate species and begin their formation in both mammals and teleost fishes as an epithelial placode (23, 24). Subsequent stages of tooth morphogenesis and differentiation are coordinated by reciprocal cell signaling between this dental epithelium and an underlying mesenchyme, which employs members of several evolutionarily conserved vertebrate gene families, including the fibroblast growth factor and hedgehog signaling pathways (25–27). The mammalian dental mesenchyme is known to be largely derived from CNC cells (28), and the dental mesenchyme of zebrafish pharyngeal teeth also express CNC cell markers (25).

RA signaling has been studied previously in relation to tooth development. In mice, mutants lacking RA receptors develop deformed skulls without teeth (29). Conversely, exposing embryonic mouse mandibles to tissue culture medium containing excess RA alters dental epithelial morphology, increases tooth bud size in the diastema region, and possibly switches tooth identity between molars and incisors (30). These studies provide some clues to the potential roles of RA in mammalian tooth development, although there has been no gene expression data or direct functional tests at a molecular level to explain the mechanisms underlying these phenotypes. In zebrafish, it has been shown that tooth induction depends on an endogenous source of RA from retinaldehyde dehydrogenase-expressing cells and that chemically blocking RA synthesis results in the complete absence of teeth (31). However, effects of RA overexpression on zebrafish tooth development have not yet been described.

Here we report that exposure of zebrafish embryos to exogenous RA induces the formation of an extensive symmetrical pattern of supernumerary teeth. Corresponding with this induction, we find up-regulation of genes normally expressed in developing teeth and in the nearby posterior pharyngeal region such as *pitx2*, *dlx2b*, and *hoxb5a*. After RA treatment, CNC cells are present but with disrupted arrangements and abnormal gene expression, and cartilage development is correspondingly disrupted. We compare the expanded RA-induced zebrafish pharyngeal dentition with the wild-type dentitions of other teleost species and find similarities especially with the teeth of medaka. However, comparisons of retinaldehyde dehydrogenase expression between these species suggest that evolutionary changes in RA signaling were not likely responsible for the reduction of the zebrafish pharyngeal dentition.

MATERIALS AND METHODS

Animal husbandry

Embryos of wild-type zebrafish (*Danio rerio*, Hamilton 1822) were obtained from in-house stocks derived originally from a commercial supplier (LiveAquaria.com, Rhinelander, WI, USA). Embryos were raised at 28.5°C in an embryo medium consisting of 30% Danieu's medium and 0.003% phenylthiourea to prevent pigmentation (25). Phenylthiourea use is important to note, as it has recently been reported to enhance certain RA-related craniofacial phenotypes in zebrafish (32). Zebrafish staging was done as described previously (33), with developmental time reported in hours post-fertilization (hpf) or days postfertilization (dpf). The fli1:GFP reporter fish were of the *Tg(fli1:EGFP)*^{y1} line (34), and the dlx2b:GFP reporters were from the *Tg(dlx2b:EGFP)*^{cs1} line (35). *Astyanax mexicanus* (De Filippi 1853) stages are reported in dpf, and medaka (*Oryzias latipes*, Temminck and Schlegel 1846) stages are reported as dpf or stage numbers (36). Animal use was in accordance with protocols approved by the Institutional Animal Care and Use Committee at Bowdoin College and the Université de Lyon.

Exogenous RA treatment and bead implantation

Embryos were manually dechorionated and added to embryo medium containing either 0.05% DMSO alone as a control or DMSO plus 6×10^{-7} M all-*trans* RA (Biomol International, Plymouth, PA, USA). RA treatments were from 24 hpf onward unless indicated otherwise. Embryos were kept in the dark or with minimal lighting to avoid RA degradation (31). To inactivate RA, embryos were washed into new medium and exposed to bright light for 5 min. AG 1-X8 beads (Bio-Rad, Hercules, CA, USA) were incubated in fresh solution of 6×10^{-4} M RA and implanted as described previously (37).

Histology and *in situ* hybridization

Whole-mount mRNA *in situ* hybridization was performed as described previously (25). Riboprobe sources: zebrafish *dlx2b*, *dlx2a*, and *pitx2* (25); *crestin* (38); *hoxb5a* (39); *sox9a* (40); *foxd3* (41); *mitfa* (42), *lef1* (43); *prdm1a* (44); *aldh1a2* (31); and medaka *aldh1a2* (sequences 1367–1830 of Genbank accession number NM_001104821). Optical confocal microscopy sections of *in situ* hybridization specimens were obtained using 633-nm fluorescence of the BM Purple (Roche, Indianapolis, IN, USA) staining product (45). Double fluorescence *in situ* hybridization was performed as described previously (46). Alizarin red and alcian blue skeletal single and double staining were performed as described previously (47).

Microscopy and statistical analysis

Embryos were visualized and photographed on a Leica DMI3000B inverted microscope with DFC420C camera (Leica Microsystems, Solms, Germany), a Leica MZ16F stereomicroscope with DFC300FX camera (Leica Microsystems), or a Zeiss 510 Meta confocal laser-scanning microscope (Carl Zeiss, Jena, Germany). Three-dimensional images of teeth were generated from confocal data using Volocity (Perkin Elmer, San Jose, CA, USA). Photograph files were processed with Adobe Photoshop (Adobe Systems, San Jose, CA, USA) and minimally adjusted as a whole for brightness, contrast, and/or color balance. Statistical significance of collected data was calculated using a 2-tailed Fisher's exact test of independence (48).

RESULTS

Exogenous RA expands zebrafish tooth development anteriorly and dorsally

Wild-type zebrafish larvae normally develop a single pair of fully formed teeth on each side of the ventral,

posterior pharynx by 4 dpf (**Fig. 1*A***), with 2–3 younger pairs of tooth germs also present nearby but not as easily visualized (49). We found that embryos treated with 6×10^{-7} M exogenous RA from 24 hpf onward or in a window from 24 to 36 hpf exhibited a dramatic supernumerary tooth phenotype by 4 dpf, with tooth formation extending into ectopic anterior and dorsal regions (Fig. 1*B*; n=119/119 relative to 0/66 controls; P<0.01, 2-tailed Fisher's exact test). These RA-treated larvae possessed ~10–16 supernumerary teeth arranged in a bilaterally symmetrical pattern. Most of the ectopic teeth were oriented with the tip toward the midline; however, we often also observed 2–4 teeth at the anterior end of the pattern with tips pointing caudally. The apparent fusion of adjacent teeth was also occasionally observed (n=23). We also employed a dlx2b:GFP reporter transgenic line that marks developing dental epithelium and mesenchyme (25, 35). DMSO-treated control fish showed 2–3 closely associated spots of bright GFP expression on each side of the posterior pharynx (Fig. 1*C*; n=20), while RA-treated embryos exhibited 8–12 supernumerary spots of GFP expression, correlating to the number and location of developing supernumerary teeth (Fig. 1*D*; n=20; P<0.01, 2-tailed Fisher's exact test). Confocal microscopy additionally revealed the presence of teeth forming in positions dorsal to the pharynx (Fig. 1*E–H*). Embryos treated with lower RA concentrations or beginning at later stages did not exhibit this supernumerary tooth phenotype as has previously been reported (50), and those treated with higher concentrations or at earlier time points did not survive long enough to develop teeth.

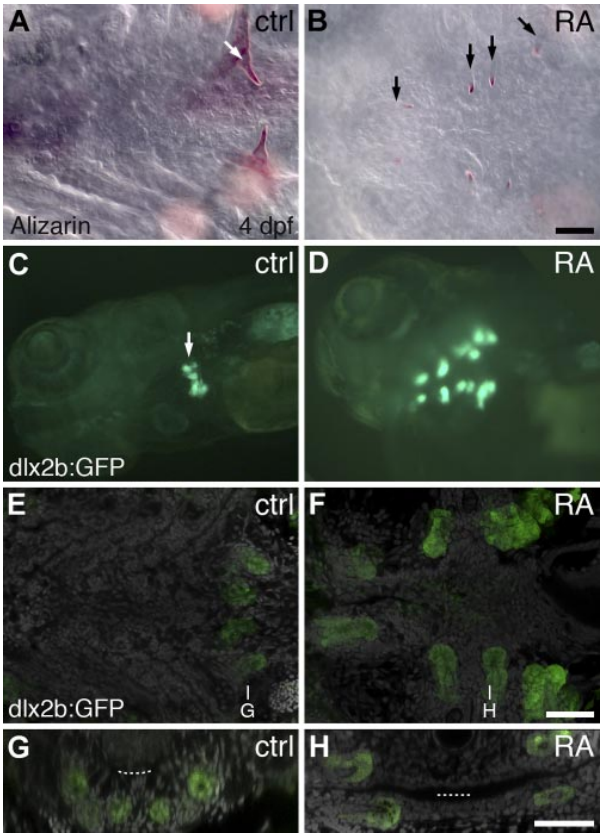
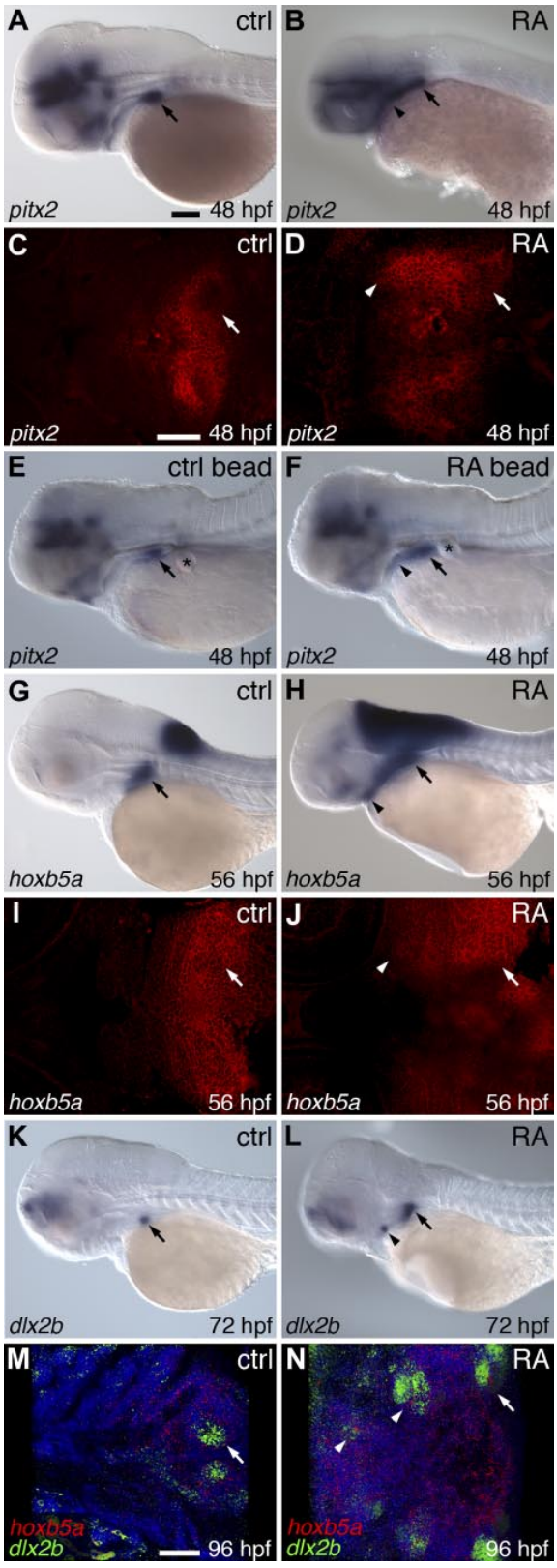


Figure 1. Exogenous RA treatment expands tooth development. *A*, *B*) Ventral view of the pharyngeal region of 4-dpf alizarin red-stained zebrafish larvae, anterior to the left. *A*) DMSO control exhibiting a pair of well-formed teeth on each side of the ventral, posterior pharynx. *B*) Specimen treated exogenously with 6×10^{-7} M RA between 24 and 36 hpf, with ectopic teeth positioned more anteriorly than in wild type. Additional teeth located more dorsally are not visible in this focal plane. *C*, *D*) Oblique ventrolateral view of the head of 4-dpf dlx2b:GFP reporter larvae. *C*) Control embryo with GFP expression in developing tooth germs. *D*) RA-treated larva with anteriorly expanded supernumerary foci of GFP expression. *E*, *F*) Confocal micrographs oriented as in *A* and *B*, with a single plane of labeled cell nuclei (gray) in the ventral pharynx superimposed on an extended-focus composite of dlx2b:GFP reporter expression (green). *G*, *H*) Transverse sections in planes indicated in *E* and *F*. *G*) Control larva with dlx2b:GFP expression in tooth germs developing normally, ventral to the lumen of the pharynx (dotted line). *H*) RA-treated specimen with tooth germs located both dorsal and ventral to the pharyngeal lumen. Arrows indicate selected teeth or tooth germs. Scale bars = 50 μm.

RA expands expression of tooth-associated genes

We characterized gene expression changes associated with RA exposure beginning at 24 hpf at several tissue and spatial levels. Expression of the homeodomain transcription factor *pitx2* has been reported in both the dental epithelium and in nearby nondental pharyngeal epithelium and represents the earliest marker of tooth germs in both mammals and fish (25, 51). We found that relative to control embryos that express *pitx2* mRNA strongly in the posterior pharyngeal region where teeth are developing by 48 hpf (**Fig. 2*A*, *C***; n=10), in RA-treated embryos *pitx2* expression was expanded anteriorly (Fig. 2*B*, *D*; n=17; P<0.01, 2-tailed Fisher's exact test). To apply RA in a more localized manner, we implanted RA-coated beads at 24-25 hpf adjacent to the posterior pharynx. In RA bead-implanted larvae, we also observed an expansion of pharyngeal *pitx2* expression near the site of implantation by 48 hpf (Fig. 2*E*; n=5/10), a result significantly different from the normal *pitx2* expression observed in control bead-implanted embryos (Fig. 2*F*; n=0/15; P<0.01, 2-tailed Fisher's exact test).

The homeobox transcription factor *hoxb5a* exhibits somewhat broader posterior pharyngeal expression than does *pitx2*, being strongly expressed in the region where teeth are forming by 56 hpf (Fig. 2*G*, *I*; n=9). We found that in RA-treated embryos, pharyngeal *hoxb5a* expression was expanded both anteriorly and dorsally, correlating to the regions where we find supernumerary teeth (Fig. 2*H*, *J*; n=16; P<0.01, 2-tailed Fisher's exact test). We also examined early *dlx2b* expression as a more restricted marker of the dental epithelium and mesenchyme (25). We found in RA-treated embryos that *dlx2b* mRNA expression patterns appeared normal at 48 and 60 hpf but at 72 hpf were expanded anteriorly (Fig. 2*L*; n=11) relative to controls (Fig. 2*K*; n=8;



$P<0.01$, 2-tailed Fisher's exact test). Simultaneously labeling *dlx2b* and *hoxb5a* mRNA at 96 hpf revealed that the ectopic anterior expansion of *dlx2b* correlated with ectopic anterior *hoxb5a* expression (Fig. 2*N*; $n=6$).

RA tooth expansion correlates with cartilage disruption

We observed that the occurrence of supernumerary teeth in RA-treated embryos was often associated with cartilage malformation or loss. Zebrafish teeth do not require skeletal support for their initial development (52), but because dental mesenchyme cells and cartilage cells both derive from CNC cell precursors, we further examined this correlation. By 4 dpf, the first pair of pharyngeal teeth has attached to the fifth ceratobranchial cartilage and younger teeth are developing nearby (**Fig. 3*A***). In embryos treated with exogenous RA starting at 24 hpf, cartilage morphology was severely disrupted in the region of supernumerary tooth formation (Fig. 3*B*). In localized treatments using RA-coated beads implanted at 24 hpf, at 4 dpf we observed cartilage malformation accompanied by supernumerary tooth formation nearby to where the beads were implanted (Fig. 3*D–F*; $n=15/23$), relative to controls (Fig. 3*C*; $n=0/13$; $P<0.01$, 2-tailed Fisher's exact test).

CNC cells are present but disorganized after RA exposure

Because CNC cells contribute to both tooth and cartilage development, and because the specification of CNC cells has been shown to be regulated by RA (19),

Figure 2. RA expands *pitx2*, *dlx2b*, and *hoxb5a* expression. Lateral (*A*, *B*, *E–H*, *K*, *L*) or ventral (*C*, *D*, *I*, *J*, *M*, *N*) view of mRNA *in situ* hybridizations, anterior to the left, with normal site of tooth formation indicated (arrow). *A*) Control embryo expressing *pitx2* in the posterior pharyngeal region at 48 hpf. *B*) Embryo treated with exogenous RA exhibits an anterior expansion of *pitx2* (arrowhead). *C*) Horizontal confocal fluorescence section of pharyngeal epithelial and tooth germ *pitx2* expression in control. *D*) Identical section after RA treatment with *pitx2* expression expanded anteriorly in the pharyngeal epithelium (arrowhead). *E*) Embryo implanted with a control bead (asterisk) and a wild-type pattern of *pitx2* expression. *F*) RA bead-implanted embryo with *pitx2* expression extended more anteriorly (arrowhead). *G*) Expression of *hoxb5a* mRNA in the posterior brain and pharyngeal region (arrow) at 56 hpf. *H*) RA-treated embryo with anterior expansion of pharyngeal *hoxb5a* expression (arrowhead). *I*) Horizontal section of *hoxb5a* expression surrounding control tooth germs (arrow). *J*) RA-exposed embryo with *hoxb5a* expression expanded anteriorly (arrowhead). *K*) *dlx2b* expression in a control embryo at 72 hpf. *L*) RA-treated embryo showing *dlx2b* expression expanded anteriorly (arrowhead). *M*) 96 hpf larva with *dlx2b* (green) and *hoxb5a* (red) expression in the posterior pharynx in and around developing teeth. *N*) Larva after RA treatment with both *dlx2b* and *hoxb5a* expression expanded anteriorly (arrowheads). Scale bars = 100 μm.

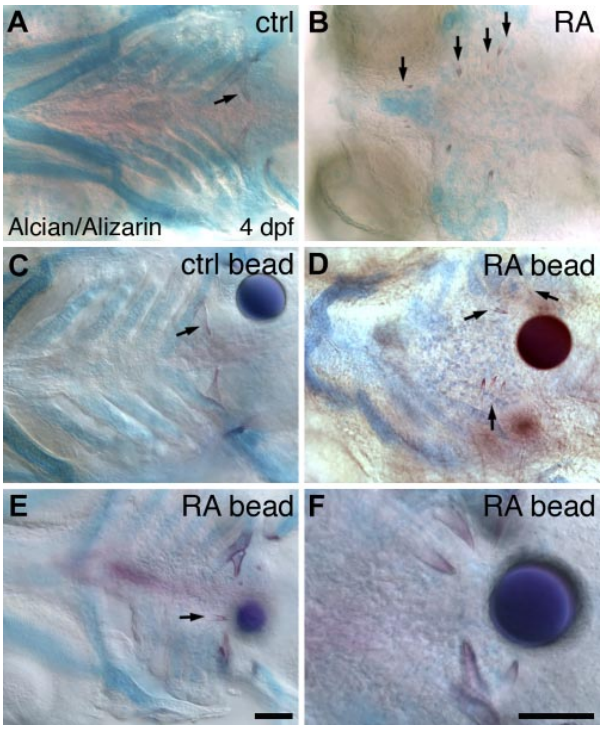


Figure 3. RA tooth expansion correlates with cartilage disruption. Ventral views of the 4-dpf pharyngeal region in alcian blue, alizarin red double-stained larvae, anterior to the left (arrows indicate selected teeth). *A*) Control zebrafish larva with a pair of well-formed teeth attached to the ossifying fifth ceratobranchial cartilages. *B*) Larva treated with RA from 24 hpf exhibiting supernumerary teeth and severe cartilage loss. *C*) Larva showing normal tooth development with a control bead positioned nearby. *D–F*) Larvae with RA-coated bead implantation at 24 hpf exhibit a range of cartilage deformation and supernumerary tooth phenotypes ranging from relatively severe (*D*) to relatively mild (*E*, arrow indicates ectopic midline tooth). *F*) Closeup view showing supernumerary teeth in proximity to an RA bead. Scale bars = 50 μm.

we investigated CNC-cell-associated gene expression after RA exposure. After RA exposure starting at 24 hpf, we surveyed early postmigratory CNC marker expression from 27 to 36 hpf and focused on 30 hpf as having maximal effects. We first employed a fli1:GFP reporter transgenic line that labels CNC cells (34, 53). Relative to controls (**Fig. 4*A*, *C***; n=12), we found that after 24–30 hpf RA exposure, fli1:GFP reporter expression in CNC cells was present at similar levels but that CNC cell arrangements, especially in the most posterior pharyngeal arch stream near where teeth later arise, appeared less compactly organized (Fig. 4*B*, *D*; n=12; P<0.01, 2-tailed Fisher's exact test). *dlx2a* marks CNC cells during their migration into the pharyngeal arches (54). Similarly to what has previously been reported (55), we found that RA-treated embryos exhibited down-regulation of *dlx2a* (Fig. 4*E*; n=25/25) relative to controls (Fig. 4*F*; n=4/26; P<0.01, 2-tailed Fisher's exact test). *crestin* marks premigratory and migrating CNC cells (38). Relative to controls (Fig. 4*G*; n=9), after RA treatment, embryos at 30 hpf showed general up-regulation of *crestin* (Fig. 4*L*; n=10; P<0.01, 2-tailed Fisher's exact test). Together, these results suggest that pharyngeal CNC cells are present after RA treatment, but that they are disorganized in arrangement and exhibit abnormal gene expression.

To further assess RA effects on CNC cell fate near the tooth-forming region, we examined later mRNA expression of markers of specific differentiating CNC cell types. *sox9a* is expressed in differentiating CNC-derived skeleton and has been shown to directly regulate type-II collagen during cartilage development (40, 56). We found that relative to controls (**Fig. 5*A***; n=4), after 24–56 hpf RA exposure *sox9a* expression was upregulated in the posterior pharyngeal region (Fig. 5*B*; n=5). Similarly, we examined *foxd3*, as at later stages its

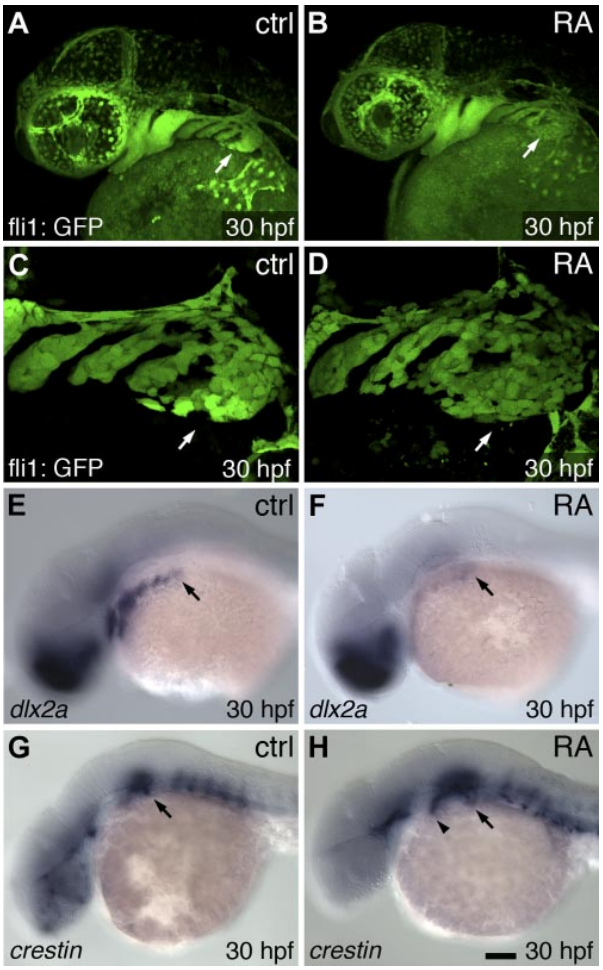


Figure 4. RA alters late migratory/early postmigratory cranial neural crest organization and gene expression. Lateral views of 30-hpf embryos, anterior to the left (arrows indicate normal tooth-forming region). *A*) Control fli1:GFP reporter expression. *B*) RA-treated embryos at with disorganized cell arrangements at 24–30 hpf. *C*, *D*) Magnified views from embryos in *A* and *B*. *E*) Control *dlx2a* mRNA expression. *F*) RA-treated embryo with down-regulated *dlx2a*. *G*) Control *crestin* mRNA expression. *H*) RA-treated embryo with *crestin* expression upregulated in the anterior pharyngeal region (arrowhead). Scale bar = 100 μm.

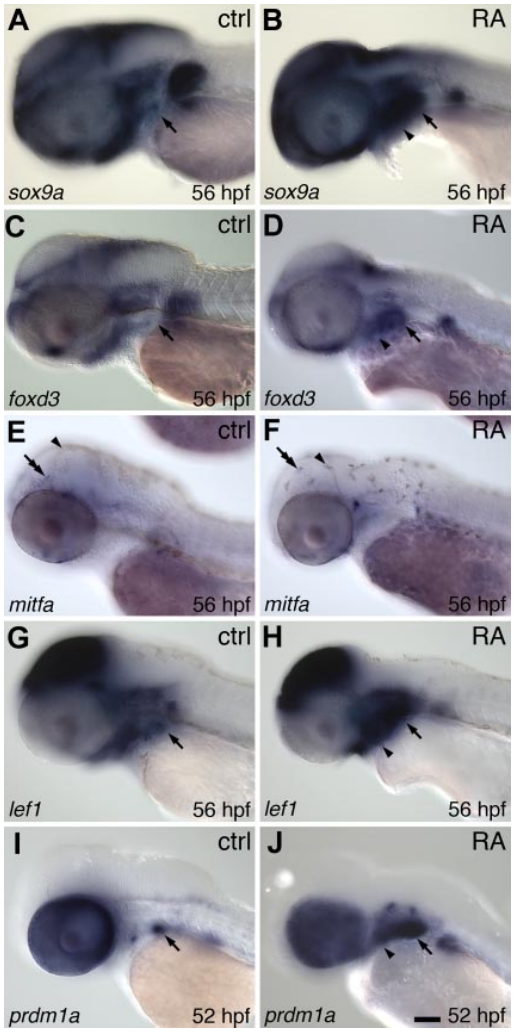


Figure 5. Gene expression analysis of CNC cell subtype markers and potential RA target genes. Lateral views of 52- to 56-hpf embryos, anterior to the left (arrows indicate normal tooth-forming region). *A*) Control *sox9a* mRNA expression at 56 hpf. *B*) RA-treated embryo with up-regulation of *sox9a* expression in the pharyngeal region (arrowhead) at 24–56 hpf. *C*) Control *foxd3* mRNA expression. *D*) RA-treated embryo with more widespread pharyngeal *foxd3* expression (arrowhead). *E*) Control *mitfa* mRNA expression marking developing melanocytes (double arrow). Differentiated melanocytes are also visible (arrowhead). *F*) RA-treated embryo with similar levels of *mitfa* expression but a disorganized-appearing melanocyte pattern (arrowhead). *G*) Control *lef1* mRNA expression. *H*) RA-treated embryo with up-regulation of pharyngeal *lef1* expression (arrowhead). *I*) Control *prdm1a* mRNA expression, restricted primarily to the developing tooth germ. *J*) RA-treated embryo showing strong *prdm1a* up-regulation (arrowhead). Scale bar = 100 μm.

expression pattern includes developing CNC-derived glial cells (57). Similarly we found that compared with controls (Fig. 5*C*; $n=6$), *foxd3* mRNA was present in a more extensive area of the pharyngeal region after RA exposure (Fig. 5*D*; $n=6$). In contrast, the expression of *mitfa*, a marker of differentiating melanocytes (42), appeared similar between control (Fig. 5*E*; $n=6$) and RA-exposed embryos (Fig. 5*F*; $n=6$), although the pattern of differentiated melanocytes appeared abnormal. Expression of these markers suggests that 24 hpf RA exposure is affecting the differentiation and/or distribution of both ectomesenchymal and nonectomesenchymal CNC cell types.

We also examined RA effects on the mRNA expression of the canonical wingless pathway transcription factor *lef1* and the zinc-finger protein *prdm1a*, as both of these genes are required for zebrafish tooth development (58, 59) and both have been characterized to lie downstream of an RA signal during fin development (60). We found *lef1* mRNA to be expressed throughout the pharyngeal region at 56 hpf (Fig. 5*G*; $n=8$) and that after 24–56 hpf RA exposure, expression levels appeared to increase (Fig. 5*H*; $n=6/7$). Pharyngeal expression of *prdm1a* mRNA was more limited in control embryos, mostly to the developing tooth germs (Fig. 5*I*; $n=11$) but appeared greatly expanded after RA treatment (Fig. 5*J*; $n=22$). Taken together, these expression data suggest that RA may regulate tooth development with a similar mechanism to its role in fin formation.

Phylogenetic comparisons

To visualize the overall arrangement of teeth in RA-treated zebrafish embryos, and make comparisons with other species, we generated three-dimensional, confocal microscope images of alizarin-red-stained specimens (**Fig. 6**). Simultaneously viewing teeth from all focal planes revealed a bilaterally symmetrical pattern of zebrafish RA-induced teeth (Fig. 6*C*). Teeth at particular anteroposterior levels also typically shared the same relative dorsal/ventral positioning with their contralateral counterparts (Fig. 6*D*). We next examined similarly staged Mexican tetra (*Astyanax mexicanus*) and medaka (*Oryzias latipes*) larvae (Fig. 6*E–H*). Mexican tetras (order Characiformes) are relatively closely related to zebrafish (order Cypriniformes), with medaka (order Beloniformes) as a more distant relative of both (61). To match the developmental stage in all three species as closely as possible, we chose stages when pharyngeal tooth development had commenced but nearby bones were not extensively calcified. At this stage, Mexican tetra possess 2 bilateral pairs of dorsal pharyngeal teeth located rostrally to a pair of ventral teeth, the latter of which correspond to the teeth in wild-type zebrafish (Fig. 6*E*, *F*). Medaka at this stage share a similar posteroventral dentition with 2 pairs of teeth but have a much more extensive dorsal pharyngeal dentition with five bilateral pairs of teeth extending further anteriorly (Fig. 6*G*, *H*). The supernumerary RA induced teeth in zebrafish appeared similar to the pharyngeal dentition in medaka in tooth number, anteroposterior location, and dorsoventral arrangement, but the patterns are not identical.

Because of the similarities between the medaka wild-type and zebrafish RA-induced pharyngeal dentitions, we wondered whether the medaka condition may be the result of more extensive RA signaling relative to

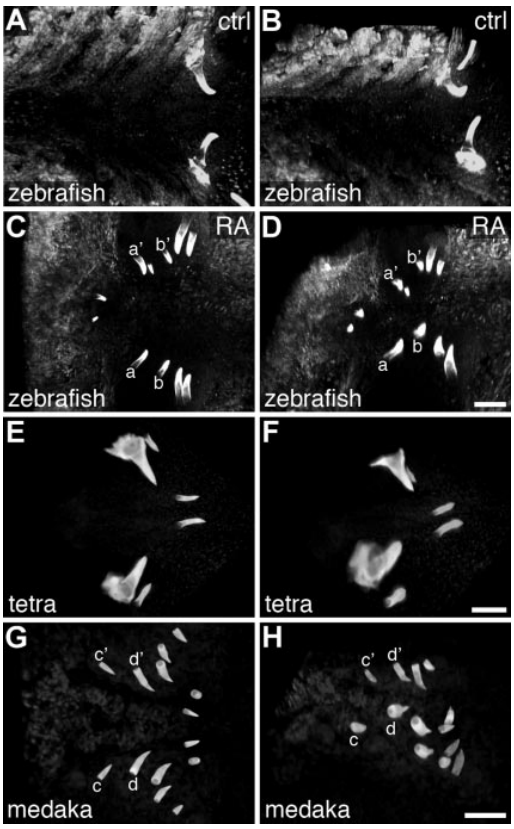


Figure 6. Multispecies comparison of the early larval pharyngeal dentition. Ventral (*A*, *C*, *E*, *G*) and 45°-rotated, ventrolateral (*B*, *D*, *F*, *H*) views of alizarin red-stained larva, visualized as 3D projections of confocal *z* stacks, anterior to the left. *A*, *B*) Zebrafish at 4 dpf with a bilateral pair of well-formed teeth located in the ventral, posterior pharynx. *C*) Zebrafish larva at 4 dpf, after exposure to RA beginning at 24 hpf, exhibiting supernumerary teeth that are anterior to normal positions, yet retaining a pattern of growth that is close to bilaterally symmetrical. For example, one bilateral pair of teeth (a/a′) is located closer to the midline than a nearby pair (b/b′). *D*) A ventrolateral view of the individual from *C* reveals dorsoventral bilateral symmetry as well: *e.g.*, the more anterior labeled pair (a/a′) is located more ventrally than the more posterior pair (b/b′). Teeth near the b/b′ position were often found dorsal to the pharyngeal lumen (Fig. 1*H*). *E*, *F*) Mexican tetra larva at 3 dpf with 2 pairs of dorsal pharyngeal teeth located anteriorly to a single pair of ventral teeth. *G*, *H*) Medaka larva at 5 dpf with 5 pairs of anterodorsal teeth and 2 pairs of posteroventral teeth. Scale bars = 50 μm.

zebrafish. To address this question, we examined mRNA expression of the retinaldehyde dehydrogenase gene *aldh1a2* (formerly *raldh2*) between zebrafish and medaka, as this gene has been implicated in the RA synthesis necessary for zebrafish tooth development (31). However, we found that unlike zebrafish which express *aldh1a2* in the posterior pharyngeal region adjacent to the position where teeth will later form (**Fig. 7*A***; n=66; Fig. 7*C*; n=28), we were unable to detect *aldh1a2* expression at all in the pharyngeal region of comparably staged medaka embryos (Fig. 7*B*, n=21; Fig. 7*D*, n=20). This result suggests that RA is not directly involved with medaka pharyngeal tooth specification and is consistent with similar, previously-reported *aldh1a2* expression comparisons between zebrafish and Mexican tetras (31).

DISCUSSION

Mechanisms of RA-mediated supernumerary tooth induction

We have considered several potential mechanisms to address how exogenous RA exposure is able to induce supernumerary tooth development in zebrafish. One possible explanation is that RA may be initiating a homeotic transformation of anterior and dorsal pharyngeal tissues to a more posterior and ventral identity and thus indirectly inducing tooth development. RA is known to regulate the expression of Hox genes (62), and differential Hox expression is a fundamental and evolutionarily conserved mechanism for patterning the vertebrate AP axis (12). Previous studies of zebrafish craniofacial development have shown that altering the expression of Hox genes can result in the homeotic transformation of CNC derived skeletal elements (63, 64). We examined *hoxb5a* expression in this study because of its strong and specific expression in the normal tooth-forming region (Fig. 2*G*) and because it has been shown specifically to be positively regulated by RA signaling (65). The anterior and dorsal expansion of *hoxb5a* expression we observed after RA treatment (Fig. 2*H*, *J*, *N*) is consistent with the idea of RA-induced homeotic transformation of the pharyngeal region. We attempted to phenocopy RA-mediated tooth induction

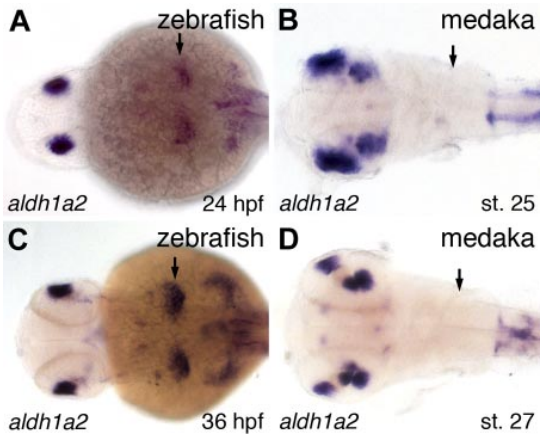


Figure 7. Comparison of retinaldehyde dehydrogenase expression between zebrafish and medaka. Dorsal views of approximately stage-matched zebrafish and medaka embryos, anterior to the left, labeled for *aldh1a2* mRNA expression by *in situ* hybridization. *A*, *C*) Zebrafish *aldh1a2* mRNA expression visualized at 24 hpf (*A*) and 36 hpf (*C*) in the posterior pharyngeal region near the location where pharyngeal tooth germs will later form (arrows). *B*, *D*) Stage 25 (*B*) and 27 (*D*) medaka embryos with no detectable *aldh1a2* mRNA expression in any part of the pharyngeal region (arrows).

with *hoxb5a* overexpression by mRNA injection but did not observe a tooth phenotype (unpublished results). However, there are many potential explanations for this lack of result, including that homeotic transformation is not responsible for RA-induced ectopic tooth formation, that RA is acting through more than one Hox gene and overexpressing a single gene has no effect, or that RA is acting through other factors to alter AP identity such as Cdx genes (17).

An alternative, although not necessarily mutually exclusive, way RA could be influencing tooth development is through CNC cell specification. One of the defining features of CNC cells is high developmental plasticity: CNC cell fate is largely determined by molecular signals in the microenvironment and generally not preprogrammed (66). CNC cells have long been known to contribute to the development of both cranial cartilage and teeth (20), raising the possibility that a switch in CNC cell fate could lead to the overproduction of one organ type at the expense of the other. CNC cells in the zebrafish pharyngeal region express the homeodomain transcription factor *dlx2a* (54), and inhibition of *dlx2a* function results in reduced and abnormal development of pharyngeal cartilages (67). In addition, in chickens, misexpression of *Dlx2* leads to ectopic cranial cartilage formation (68), and exogenous RA exposure in zebrafish suppresses Dlx gene expression, resulting in pharyngeal cartilage malformation and loss (55). Together, these data suggest that RA may be involved in directing CNC cells toward a cartilage fate by influencing the expression of Dlx genes. The presence of fli1:GFP-expressing cells in the posterior pharyngeal region suggests that CNC cells are present at late migration stages after RA exposure (Fig. 4*D*), but down-regulation of *dlx2a* and up-regulation of *crestin* suggest that the remaining CNC cells are already not developing normally (Fig. 4*F, H*). However, inhibition of *dlx2a* function alone does not result in supernumerary tooth formation (35), suggesting that if RA is altering CNC cell fate specification, it is doing so by changing the expression of multiple downstream targets. In addition, our analysis of later CNC marker expression suggests that multiple CNC cell types may be overrepresented after 24 hpf RA exposure, including chondrocytes (*sox9a*; Fig. 5*B*) and glia (*foxd3*; Fig. 5*D*), but examining melanocytes suggest that not all CNC cell types are overrepresented after RA treatment (*mitfa*; Fig. 5*F*). Together, these data suggest that RA may be promoting the development of certain CNC cell types, such as both cartilage and tooth odontoblasts, rather than directing a switch in fate between CNC cell subtypes.

It is also possible that RA induction of tooth development may be relatively direct. A study using chromatin immunoprecipitation and gel mobility shift assays in mice has shown that RA nuclear receptor complexes directly bind to *cis*-regulatory regions of the *Pitx2* gene (69), the earliest marker of tooth formation (51). A study using similar methods has, in turn, demonstrated direct regulation of *Dlx2* by *Pitx2* (70). Our observation of expansion of *pitx2* expression after RA exposure (Fig. 2*B, D, F*) is consistent with the direct up-regulation of *pitx2* by RA. We also do not observe ectopic *dlx2b* expression until many hours after the time of initial RA exposure (Fig. 2*L*), consistent with RA having an indirect effect on *dlx2b* expression. However, our data from *lef1* and *prdm1a* (Fig. 5*G–J*), genes known to be downstream of RA signaling in limb development and required for zebrafish tooth development (58–60), are consistent with RA exhibiting a more indirect control of tooth development than by direct up-regulation of *pitx2*. Direct tests of *pitx2* function during tooth development could help establish whether this gene may indeed be acting as an intermediary between RA signaling and tooth morphogenesis.

RA in the evolution of teleost dentitions

The order Cypriniformes is thought to have greatly reduced its dentition early in its evolution such that teeth remained only associated with the posterior and ventral 5th ceratobranchial element (CB5), possibly as an adaptation for suction feeding (1, 71). It is interesting to consider whether the evolutionary reduction or loss of RA signaling in the anterior and dorsal pharynx may have been part of a developmental mechanism by which the Cypriniform ancestor reduced its dentition. If this were the case, one might expect to find similarities between the zebrafish RA-induced dental patterns with the normal tooth arrangements of other teleost fishes that have retained a less modified dentition, as well as more extensive RA signaling in the anterior pharyngeal regions of such species.

The bowfin *Amia clava*, a species immediately basal to teleosts phylogenetically (61), possesses an extensive larval dentition associated with both anterior and posterior pharyngeal skeletal elements (72). As the teleost sister group, the dentition of *Amia* may represent the ancestral condition for teleosts. Pharyngeal tooth reduction is common in other orders of the teleost radiation, including the belaniform medaka (*Oryzias latipes*) and the characiform Mexican tetra (*Astyanax mexicanus*), but is typically less severe than what is found in the cypriniformes (2, 73). The medaka and zebrafish lineages diverged early in the teleost radiation (61); thus, the teeth of *Amia clava* may represent the closest proxy for the ancestral pharyngeal dentition of these model species. By this logic, the more extensive pharyngeal dentition of medaka is more similar to the ancestral condition than is the more reduced dentition of zebrafish.

Of the species we examined, we interpret the RA-induced zebrafish tooth pattern as most resembling the wild-type pharyngeal dentition of medaka (Fig. 6). In medaka, pharyngeal teeth attach to 2 dorsal parabranchial bones and ventrally to CB5 (73). Similarities between RA-induced zebrafish teeth and wild-type medaka dentitions include tooth number, orientation, the anteroposterior extent of tooth formation, and symmetry in the left-right and dorsoventral axes. It is

also notable that RA-exposed zebrafish often develop teeth dorsal to the pharyngeal lumen, a position where teeth never form in zebrafish or any other cypriniform species (5). These similarities prompted us to wonder whether RA signaling may have been ancestrally more prevalent in the pharyngeal region, and whether the evolutionary reduction of the cypriniform dentition may have been the consequence of reduction in anterior pharyngeal RA signaling. This scenario would predict that a species with a more extensive pharyngeal dentition, such as medaka, would have more RA signaling associated with the development of the anterior pharynx. To test this idea, we compared the zebrafish and medaka mRNA expression of *aldh1a2*, as this gene has been implicated as a key regulator of RA levels required for zebrafish tooth and craniofacial development (21, 31). However, in contrast to the above scenario, our data suggest instead that medaka have no retinaldehyde dehydrogenase expression associated with pharyngeal tooth development (Fig. 7). This result is consistent with previous work where RA inhibition had a much lesser effect on tooth development in medaka than it had in zebrafish (31). We therefore speculate that the RA-induced ectopic dentition in zebrafish may not be recapitulating an ancestral condition; however, phylogenetic comparisons of RA expression and function during the tooth development of other species will shed further light on this question.

However, even if raising RA signaling levels in zebrafish does not actually restore a more ancestral-like dentition, the number of teeth and area in which they reside is clearly increased by increasing RA, so why has this not occurred during cypriniform evolution to bring back teeth in species that would benefit from a more extensive dentition? Examples of cypriniforms where adaptive pressure may be present to produce more teeth include piscivorous pikeminnows (74) and *Danionella dracula*, which has evolved fang-like jawbone projections used in male displays (75). The answer to this question may lie with the corresponding cartilage disruptions we have observed to accompany RA-induced tooth expansion. It is possible to induce a more extensive dentition by gross RA overexpression, but to increase tooth number without adversely affecting nearby skeletal elements may require more sophisticated control. Such fine control over RA levels, distribution pattern, or the simultaneous coexpression of modulatory cofactors may be difficult to evolve; thus constraining how cypriniforms could possibly reacquire a more extensive dentition. Further comparative study of the mechanisms of RA action on tooth and skeletal development in zebrafish and medaka and other non-cypriniform species, as well as similar comparative analyses of other developmental signaling pathways, will help to explain this apparent case of evolutionary developmental constraint. FJ

The authors are very grateful to Kristin Artinger (University of Colorado, Aurora, CO, USA) for sharing ideas, prepublication data, and reagents; Vicki Prince (University of Chicago, Chicago, IL, USA) and Paul Henion (Ohio State University, Columbus, OH, USA) for reagents; and Mike Palopoli, Andrea Jowdry, and David Rivers for helpful suggestions. This work was supported by U.S. National Institutes of Health grants 5P20-RR-016463-12 and 8P20-GM-103423-12 (to W.R.J.) and Agence Nationale de la Recherche grant ANR-09-BLAN-0127-01 (to V.L.).

REFERENCES

1. Gosline, W. (1973) Considerations regarding the phylogeny of cypriniform fishes, with special reference to structures associated with feeding. *Copeia* **1973,** 761–776
2. Stock, D. W. (2001) The genetic basis of modularity in the development and evolution of the vertebrate dentition. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **356,** 1633–1653
3. Patterson, C. (1993) Osteichthyes: teleostei. In *The Fossil Record 2* (Benton, M., ed.) pp. 621–656, Chapman & Hall, London
4. Sibbing, F. (1991) Food capture and oral processing. In *Cyprinid Fishes: Systematics, Biology, and Exploitation* (Winfield, I., and Nelson, J., eds.) pp. 377–412, Chapman & Hall, New York
5. Stock, D. W. (2007) Zebrafish dentition in comparative context. *J. Exp. Zoolog. B Mol. Dev. Evol.* **308B,** 523–549
6. Dollo, L. (1893) Les lois de l'évolution. *Bull. Soc. Bel. Géol. Pal. Hydr.* **7,** 164–166
7. Goldberg, E. E., and Igić, B. (2008) On phylogenetic tests of irreversible evolution. *Evolution* **62,** 2727–2741
8. Marshall, C. R., Raff, E. C., and Raff, R. A. (1994) Dollo's law and the death and resurrection of genes. *Proc. Natl. Acad. Sci. U. S. A.* **91,** 12283–12287
9. Galis, F., Arntzen, J. W., and Lande, R. (2010) Dollo's law and the irreversibility of digit loss in *Bachia*. *Evolution* **64,** 2466–2485
10. Lande, R. (1978) Evolutionary mechanisms of limb loss in tetrapods. *Evolution* **32,** 73–92
11. Wilson, J., Roth, C., and Warkany, J. (1953) An analysis of the syndrome of malformations induced by maternal vitamin A deficiency. Effects of restoration of vitamin A at various times during gestation. *Am. J. Anat.* **92,** 189–217
12. Mallo, M., Wellik, D. M., and Deschamps, J. (2010) Hox genes and regional patterning of the vertebrate body plan. *Dev. Biol.* **344,** 7–15
13. Niederreither, K., and Dollé, P. (2008) Retinoic acid in development: towards an integrated view. *Nat. Rev. Genet.* **9,** 541–553
14. Allan, D., Houle, M., Bouchard, N., Meyer, B. I., Gruss, P., and Lohnes, D. (2001) RARγ and Cdx1 interactions in vertebral patterning. *Dev. Biol.* **240,** 46–60
15. Manzanares, M., Wada, H., Itasaki, N., Trainor, P. A., Krumlauf, R., and Holland, P. W. (2000) Conservation and elaboration of Hox gene regulation during evolution of the vertebrate head. *Nature* **408,** 854–857
16. Kessel, M., and Gruss, P. (1991) Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell* **67,** 89–104
17. Houle, M., Sylvestre, J.-R., and Lohnes, D. (2003) Retinoic acid regulates a subset of Cdx1 function in vivo. *Development* **130,** 6555–6567
18. Yamaguchi, M., Nakamoto, M., Honda, H., Nakagawa, T., Fujita, H., Nakamura, T., Hirai, H., Narumiya, S., and Kakizuka, A. (1998) Retardation of skeletal development and cervical abnormalities in transgenic mice expressing a dominant-negative retinoic acid receptor in chondrogenic cells. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 7491–7496
19. Minoux, M., and Rijli, F. M. (2010) Molecular mechanisms of cranial neural crest cell migration and patterning in craniofacial development. *Development* **137,** 2605–2621
20. De Beer, G. (1947) The differentiation of neural crest cells into visceral cartilages and odontoblasts in *Amblystoma*, and a re-examination of the germ-layer theory. *Proc. R Soc. Lond. B Biol. Sci.* **134,** 377–398
21. Begemann, G., Schilling, T. F., Rauch, G. J., Geisler, R., and Ingham, P. W. (2001) The zebrafish neckless mutation reveals a requirement for *raldh2* in mesodermal signals that pattern the hindbrain. *Development* **128,** 3081–3094

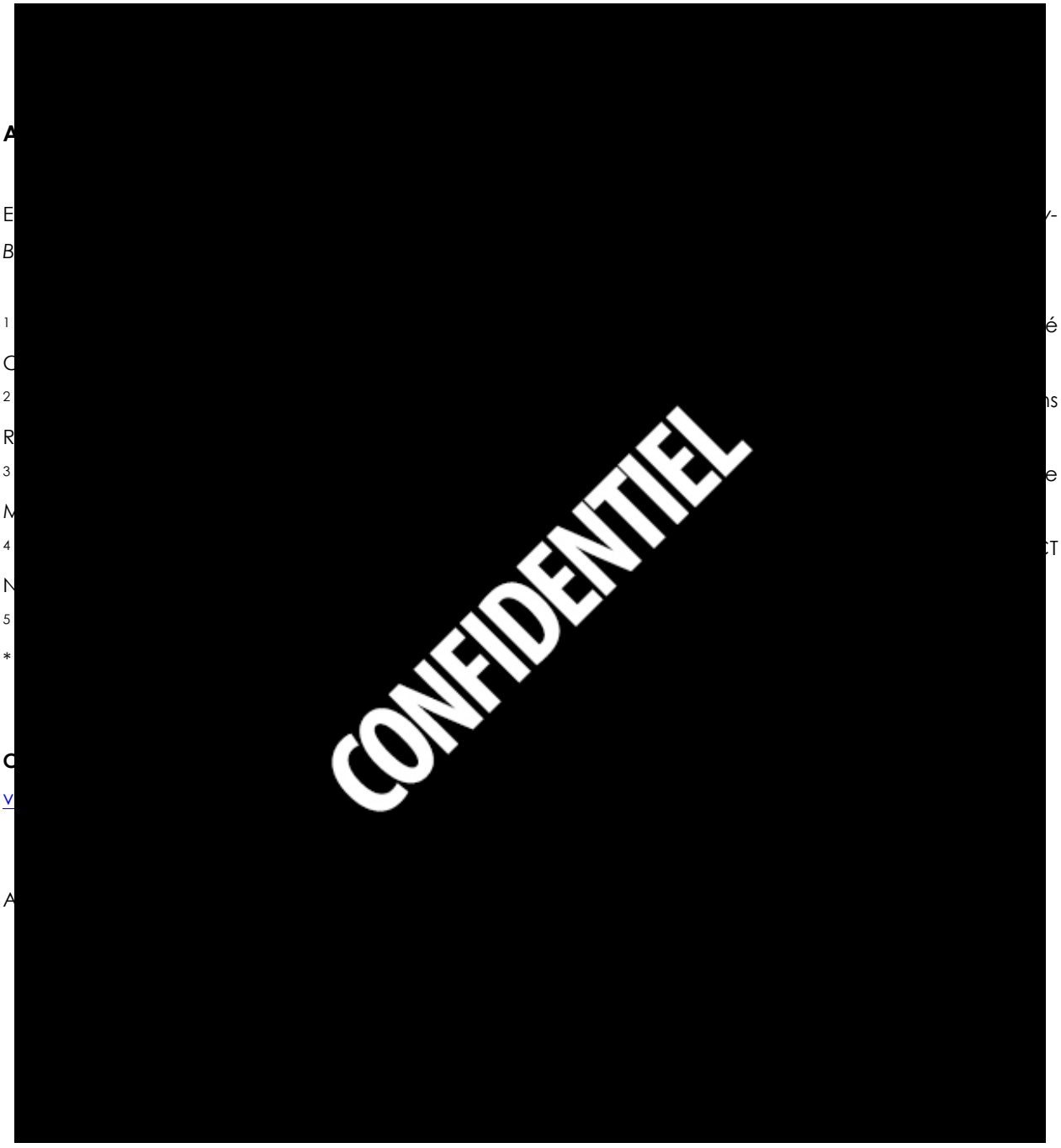
22. Lohnes, D., Mark, M., Mendelsohn, C., Dollé, P., Dierich, A., Gorry, P., Gansmuller, A., and Chambon, P. (1994) Function of the retinoic acid receptors (RARs) during development (I). Craniofacial and skeletal abnormalities in RAR double mutants. *Development* **120,** 2723–2748
23. Huysseune, A., Van der heyden, C., and Sire, J.-Y. (1998) Early development of the zebrafish (*Danio rerio*) pharyngeal dentition (Teleostei, *Cyprinidae*). *Anat. Embryol.* **198,** 289–305
24. Peters, H., and Balling, R. (1999) Teeth. Where and how to make them. *Trends Genet.* **15,** 59–65
25. Jackman, W. R., Draper, B. W., and Stock, D. W. (2004) Fgf signaling is required for zebrafish tooth development. *Dev. Biol.* **274,** 139–157
26. Jackman, W. R., Yoo, J. J., and Stock, D. W. (2010) Hedgehog signaling is required at multiple stages of zebrafish tooth development. *BMC Dev. Biol.* **10,** 119
27. Klein, O. D., Lyons, D. B., Balooch, G., Marshall, G. W., Basson, M. A., Peterka, M., Boran, T., Peterkova, R., and Martin, G. R. (2008) An FGF signaling loop sustains the generation of differentiated progeny from stem cells in mouse incisors. *Development* **135,** 377–385
28. Chai, Y., Jiang, X., Ito, Y., Bringas, P., Han, J., Rowitch, D. H., Soriano, P., McMahon, A. P., and Sucov, H. M. (2000) Fate of the mammalian cranial neural crest during tooth and mandibular morphogenesis. *Development* **127,** 1671–1679
29. Mark, M., Lohnes, D., Mendelsohn, C., Dupé, V., Vonesch, J. L., Kastner, P., Rijli, F., Bloch-Zupan, A., and Chambon, P. (1995) Roles of retinoic acid receptors and of Hox genes in the patterning of the teeth and of the jaw skeleton. *Int. J. Dev. Biol.* **39,** 111–121
30. Kronmiller, J. E., Nguyen, T., and Berndt, W. (1995) Instruction by retinoic acid of incisor morphology in the mouse embryonic mandible. *Arch. Oral Biol.* **40,** 589–595
31. Gibert, Y., Bernard, L., Debiais-Thibaud, M., Bourrat, F., Joly, J.-S., Pottin, K., Meyer, A., Rétaux, S., Stock, D. W., Jackman, W. R., Seritrakul, P., Begemann, G., and Laudet, V. (2010) Formation of oral and pharyngeal dentition in teleosts depends on differential recruitment of retinoic acid signaling. *FASEB J.* **24,** 3298–3309
32. Bohnsack, B. L., Gallina, D., and Kahana, A. (2011) Phenothiourea sensitizes zebrafish cranial neural crest and extraocular muscle development to changes in retinoic acid and IGF signaling. *PLoS ONE* **6,** e22991
33. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., and Schilling, T. F. (1995) Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203,** 253–310
34. Lawson, N. D., and Weinstein, B. M. (2002) In vivo imaging of embryonic vascular development using transgenic zebrafish. *Dev. Biol.* **248,** 307–318
35. Jackman, W. R., and Stock, D. W. (2006) Transgenic analysis of Dlx regulation in fish tooth development reveals evolutionary retention of enhancer function despite organ loss. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 19390–19395
36. Iwamatsu, T. (2004) Stages of normal development in the medaka *Oryzias latipes*. *Mech. Dev.* **121,** 605–618
37. White, R. J., Nie, Q., Lander, A. D., and Schilling, T. F. (2007) Complex regulation of *cyp26a1* creates a robust retinoic acid gradient in the zebrafish embryo. *PLoS Biol.* **5,** e304
38. Luo, R., An, M., Arduini, B. L., and Henion, P. D. (2001) Specific pan-neural crest expression of zebrafish *Crestin* throughout embryonic development. *Dev. Dyn.* **220,** 169–174
39. Prince, V. E., Joly, L., Ekker, M., and Ho, R. K. (1998) Zebrafish hox genes: genomic organization and modified colinear expression patterns in the trunk. *Development* **125,** 407–420
40. Chiang, E. F., Pai, C. I., Wyatt, M., Yan, Y. L., Postlethwait, J., and Chung, B. (2001) Two *sox9* genes on duplicated zebrafish chromosomes: expression of similar transcription activators in distinct sites. *Dev. Biol.* **231,** 149–163
41. Odenthal, J., and Nüsslein-Volhard, C. (1998) Fork head domain genes in zebrafish. *Dev. Genes Evol.* **208,** 245–258
42. Lister, J. A., Robertson, C. P., Lepage, T., Johnson, S. L., and Raible, D. W. (1999) *nacre* encodes a zebrafish microphthalmia-related protein that regulates neural-crest-derived pigment cell fate. *Development* **126,** 3757–3767
43. Dorsky, R. I., Snyder, A., Cretekos, C. J., Grunwald, D. J., Geisler, R., Haffter, P., Moon, R. T., and Raible, D. W. (1999) Maternal and embryonic expression of zebrafish *lef1*. *Mech. Dev.* **86,** 147–150
44. Hernandez-Lagunas, L., Choi, I. F., Kaji, T., Simpson, P., Hershey, C., Zhou, Y., Zon, L., Mercola, M., and Artinger, K. B. (2005) Zebrafish *narrowminded* disrupts the transcription factor *prdm1* and is required for neural crest and sensory neuron specification. *Dev. Biol.* **278,** 347–357
45. Trinh, L. A., McCutchen, M. D., Bonner-Fraser, M., Fraser, S. E., Bumm, L. A., and McCauley, D. W. (2007) Fluorescent in situ hybridization employing the conventional NBT/BCIP chromogenic stain. *Biotechniques* **42,** 756–759
46. Talbot, J. C., Johnson, S. L., and Kimmel, C. B. (2010) Hand and Dlx genes specify dorsal, intermediate and ventral domains within zebrafish pharyngeal arches. *Development* **137,** 2507–2517
47. Walker, M. B., and Kimmel, C. B. (2007) A two-color acid-free cartilage and bone stain for zebrafish larvae. *Biotech. Histochem.* **82,** 23–28
48. McDonald, J. (2009) Fisher's exact test of independence. In *Handbook of Biological Statistics*, pp. 70–75, Sparky House, Baltimore
49. Van der Heyden, C., and Huysseune, A. (2000) Dynamics of tooth formation and replacement in the zebrafish (*Danio rerio*) (Teleostei, *Cyprinidae*). *Dev. Dyn.* **219,** 486–496
50. Laue, K., Jänicke, M., Plaster, N., Sonntag, C., and Hammerschmidt, M. (2008) Restriction of retinoic acid activity by Cyp26b1 is required for proper timing and patterning of osteogenesis during zebrafish development. *Development* **135,** 3775–3787
51. Mucchielli, M. L., Mitsiadis, T. A., Raffo, S., Brunet, J. F., Proust, J. P., and Goridis, C. (1997) Mouse *Otlx2/RIEG* expression in the odontogenic epithelium precedes tooth initiation and requires mesenchyme-derived signals for its maintenance. *Dev. Biol.* **189,** 275–284
52. Schilling, T. F., Piotrowski, T., Grandel, H., Brand, M., Heisenberg, C. P., Jiang, Y. J., Beuchle, D., Hammerschmidt, M., Kane, D. A., Mullins, M. C., van Eeden, F. J., Kelsh, R. N., Furutani-Seiki, M., Granato, M., Haffter, P., Odenthal, J. Warga, R. M., Trowe, T., and Nusslein-Volhard, C. (1996) Jaw and branchial arch mutants in zebrafish I: branchial arches. *Development* **123,** 329–344
53. Crump, J. G., Maves, L., Lawson, N. D., Weinstein, B. M., and Kimmel, C. B. (2004) An essential role for Fgfs in endodermal pouch formation influences later craniofacial skeletal patterning. *Development* **131,** 5703–5716
54. Akimenko, M. A., Ekker, M., Wegner, J., Lin, W., and Westerfield, M. (1994) Combinatorial expression of three zebrafish genes related to distal-less: part of a homeobox gene code for the head. *J. Neurosci.* **14,** 3475–3486
55. Ellies, D. L., Langille, R. M., Martin, C. C., Akimenko, M. A., and Ekker, M. (1997) Specific craniofacial cartilage dysmorphogenesis coincides with a loss of *dlx* gene expression in retinoic acid-treated zebrafish embryos. *Mech. Dev.* **61,** 23–36
56. Bell, D. M., Leung, K. K., Wheatley, S. C., Ng, L. J., Zhou, S., Ling, K. W., Sham, M. H., Koopman, P., Tam, P. P., and Cheah, K. S. (1997) SOX9 directly regulates the type-II collagen gene. *Nat. Genet.* **16,** 174–178
57. Ignatius, M.S., Moose, H. E., El-Hodiri, H. M., and Henion, P. D. (2008) *Colgate/hdac1* repression of *foxd3* expression is required to permit mitfa-dependent melanogenesis. Dev. Biol. **313,** 568–583
58. Birkholz, D. A., Olesnicky Killian, E. C., George, K. M., and Artinger, K. B. (2009) Prdm1a is necessary for posterior pharyngeal arch development in zebrafish. *Dev. Dyn.* **238,** 2575–2587
59. McGraw, H. F., Drerup, C. M., Culbertson, M. D., Linbo, T., Raible, D. W., and Nechiporuk, A. V. (2011) Lef1 is required for progenitor cell identity in the zebrafish lateral line primordium. *Development* **138,** 3921–3930
60. Mercader, N., Fischer, S., and Neumann, C. J. (2006) Prdm1 acts downstream of a sequential RA, Wnt and Fgf signaling cascade during zebrafish forelimb induction. *Development* **133,** 2805–2815
61. Nelson, J. (2006). *Fishes of the World, 4th Ed.*, John Wiley & Sons, Hoboken, NJ, USA
62. Marshall, H., Morrison, A., Studer, M., Pöpperl, H., and Krumlauf, R. (1996) Retinoids and Hox genes. *FASEB J.* **10,** 969–978

63. Hunter, M. P., and Prince, V. E. (2002) Zebrafish hox paralogue group 2 genes function redundantly as selector genes to pattern the second pharyngeal arch. *Dev. Biol.* **247,** 367–389
64. Alexandre, D., Clarke, J. D., Oxtoby, E., Yan, Y. L., Jowett, T., and Holder, N. (1996) Ectopic expression of *Hoxa-1* in the zebrafish alters the fate of the mandibular arch neural crest and phenocopies a retinoic acid-induced phenotype. *Development* **122,** 735–746
65. Feng, L., Hernandez, R. E., Waxman, J. S., Yelon, D., and Moens, C. B. (2010) Dhrs3a regulates retinoic acid biosynthesis through a feedback inhibition mechanism. *Dev. Biol.* **338,** 1–14
66. Baker, C. V., and Bronner-Fraser, M. (1997) The origins of the neural crest. Part II: an evolutionary perspective. *Mech. Dev.* **69,** 13–29
67. Sperber, S. M., Saxena, V., Hatch, G., and Ekker, M. (2008) Zebrafish *dlx2a* contributes to hindbrain neural crest survival, is necessary for differentiation of sensory ganglia and functions with *dlx1a* in maturation of the arch cartilage elements. *Dev. Biol.* **314,** 59–70
68. Gordon, C. T., Brinas, I. M. L., Rodda, F. A., Bendall, A. J., and Farlie, P. G. (2010) Role of *Dlx* genes in craniofacial morphogenesis: *Dlx2* influences skeletal patterning by inducing ectomesenchymal aggregation in ovo. *Evol. Dev.* **12,** 459–473
69. Kumar, S., and Duester, G. (2010) Retinoic acid signaling in perioptic mesenchyme represses Wnt signaling via induction of *Pitx2 and Dkk2. Dev. Biol.* **340,** 67–74
70. Venugopalan, S. R., Li, X., Amen, M. A., Florez, S., Gutierrez, D., Cao, H., Wang, J., and Amendt, B. A. (2011) Hierarchical interactions of homeodomain and forkhead transcription factors in regulating odontogenic gene expression. *J. Biol. Chem.* **286,** 21372–21383
71. Nichols, J. T. (1930) Speculation on the history of the ostariophysi. *Copeia* **1930,** 148–151
72. Grande, L., and Bemis, W. E. (1998) A comprehensive phylogenetic study of amiid fishes (Amiidae) based on comparative skeletal anatomy. An empirical search for interconnected patterns of natural history. *J. Vertebr. Paleontol. Mem.* **4,** 1–690
73. Debiais-Thibaud, M., Borday-Birraux, V., Germon, I., Bourrat, F., Metcalfe, C. J., Casane, D., and Laurenti, P. (2007) Development of oral and pharyngeal teeth in the medaka (*Oryzias latipes*): comparison of morphology and expression of *eve1* gene. *J. Exp. Zool.* **308,** 693–708
74. Portz, D., and Tyus, H. (2004) Fish humps in two Colorado River fishes: a morphological response to cyprinid predation? *Env. Biol. Fishes* **71,** 233–245
75. Britz, R., Conway, K. W., and Rüber, L. (2009) Spectacular morphological novelty in a miniature cyprinid fish, *Danionella dracula* n. sp. *Proc. R. Soc. Lond. B Biol. Sci.* **276,** 2179–2186

Received for publication April 3, 2012.
Accepted for publication August 20, 2012.

2. Altered retinoic acid signalling underpins dentition evolution

Ce travail est en cours de soumission pour publication dans le journal *Proceedings of the Royal Society B.*



CONFIDENTIEL

CONFIDENTIEL

CONFIDENTIEL

CONFIDENTIEL

CONFIDENTIEL

CONFIDENTIEL

CONFIDENTIEL

CONFIDENTIEL

CONFIDENTIEL

CONFIDENTIEL

CONFIDENTIEL

IV. Discussion - Perspectives

1. La voie de l'AR au cœur de l'évolution.

1.1. Le potentiel d'une voie de signalisation pléiotrope

La nature pléiotrope de la voie de l'AR fait de cette signalisation un candidat privilégié dans la mise en place de diversités morphologiques et phénotypiques. En effet, nous avons largement illustré dans l'introduction que l'AR contrôle la développement et la spécification de nombreux axes et organes de l'embryon, ainsi la modulation de son activité a un fort potentiel de modification sur de nombreuses structures. De plus, l'AR intervient à différents moments au cours de l'embryogénèse et son rôle peut être multiple sur une même structure au cours du temps. Ainsi, l'éventail des conséquences d'une modulation de son activité est extrêmement large selon le moment et l'endroit où elle s'applique. Par exemple, concernant le rôle de l'AR pour la mise en place de la denture chez les poissons téléostéens, nos résultats illustrent que des modifications de la voie de l'AR à différents moments au cours du développement vont mener à des phénotypes distincts (Figure 40). En effet, une augmentation du niveau d'AR très précocement après la neurulation influence la migration et la maintenance des cellules de crêtes neurales (NCCs) qui formeront le mésenchyme dentaire (Gibert et al., 2010) alors qu'un traitement à l'AR d'embryons à partir de 24 hpf étend antérieurement la denture pharyngienne chez le

poisson-zèbre (Seritrakul et al., 2012), page 166). Ce type de traitement induit également un changement de la forme des dents observées à 5,5 jours comme détaillé précédemment (Samarut et al., soumis, page 177). De plus, certains de mes travaux non décrits ici montrent qu'un traitement à l'AR à partir de 36 hpf augmente la vitesse de mise en place de la denture chez le poisson-zèbre. Enfin, un traitement à l'AR à partir de 48 hpf module le nombre de dents et induit une 6ème dent sur le 5ème arc branchial chez les cypriniformes (Figure 40). Nos résultats illustrent donc que la modulation de la voie de l'AR à différents moments au cours du développement conduit à divers phénotypes. Ainsi, le potentiel de la voie de l'AR dans la genèse de diversité réside dans sa nature pléiotrope à la fois spatialement et temporellement.

1.2. La force des gènes paralogues

L'apparition de nouveaux gènes paralogues à l'issu d'évènements de duplication de génome est fortement propice à la genèse de nouvelles fonctions géniques par néofonctionnalisation et ainsi à la mise en place de nouveautés morphologiques (section B-III-1.2, page 40). Toutefois, nos résultats suggèrent que la sous-fonctionnalisation des gènes paralogues l'est également, plus particulièrement la « sous-localisation d'expression » des différents gènes paralogues. En effet, nos résultats montrent que les différents gènes codants pour les enzymes CYP26 chez le poisson-zèbre ont des patrons d'expression majoritairement exclusifs (Fig. 3, page 181). En particulier, l'enzyme CYP26B1 est spécifiquement exprimée dans la région de formation des dents. Ainsi, la modulation de l'expression temporelle de cette enzyme permet une modulation locale du métabolisme de l'AR, alors que l'expression des autres enzymes CYP26 est inchangée assurant le métabolisme général de l'AR au sein de l'embryon. Dans la plupart de nos expériences, les traitements à l'AR conduisent à des malformations létales des embryons en moins d'une semaine. Selon les concentrations et le moment du traitement, les embryons développent des œdèmes cardiaques ou encore une fusion des arcs branchiaux caractéristiques des effets tératogènes de l'AR. Ainsi, de fortes contraintes développementales vont à l'encontre de modulations drastiques du niveau d'AR dans l'embryon et il est probable que l'expression différentielle des gènes paralogues codants pour les enzymes CYP26 permette un échappatoire à ces contraintes via des actions locales. Ainsi il existe une balance sensible entre le degré et le temps de modulation de la voie de l'AR nécessaire pour le

développement de nouveaux phénotypes et les contraintes développementales directement liées à cette modulation.

2. Comment des modulations de la voie de l'AR se traduisent-elles au niveau des gènes ?

i. Différences d'expression

L'origine de la modulation de l'expression temporelle et/ou spatiale des gènes se trouve dans leurs séquences *cis*-régulatrices. En effet, la modification des séquences déjà présentes ou l'acquisition de nouvelles séquences *cis*-régulatrices peut permettre l'expression du gène dans un nouveau tissu ou à un moment différent au cours du développement. Dans le cas de la différence d'expression temporelle de *cyp26b1* entre le poisson-zèbre et *Tanichthys albonubes*, il est probable que des modifications subtiles au sein d'éléments *cis*-régulateurs du gène en soient responsables. La corrélation du décalage temporel d'expression avec une telle mutation dans la région promotrice du gène serait très intéressante. Cependant, le génome de *Tanichthys albonubes* n'étant pas séquencé, cela nous empêche toute étude comparative des séquences promotrices du gène *cyp26b1* entre ces espèces.

ii. Différences d'activité

La voie de l'AR peut également être modulée par des changements de l'activité d'un des membres de la voie. La voie de l'AR comporte de nombreux acteurs dont les enzymes de métabolisme, les RAR et de nombreux corégulateurs. La transduction du signal de l'AR à un endroit et à un temps donné nécessite la coordination de l'activité des différents acteurs (Fig.3, page 58). Ainsi, la modulation de l'activité d'un des acteurs peut mener à la modification du signal de l'AR. En particulier, les modifications post-traductionnelles dont les phosphorylations sont susceptibles de contrôler l'activité d'une protéine en agissant comme un interrupteur réversible et dynamique. Nos études sur l'évolution des sites de phosphorylations de RARα montrent que l'acquisition d'un site phosphorylé dans la séquence codante de la protéine permet une régulation plus fine de son activité (chapitre 2, page 139). L'expression spécifique des kinases qui phosphorylent les RAR pourrait ainsi conduire à une régulation locale et à un temps donné de l'activité du récepteur et ainsi moduler finement la voie de l'AR. Ainsi, il serait intéressant d'étudier les

phénotypes induits par l'expression de RAR phospho-mutants chez des embryons de poisson-zèbre pour inférer un rôle évolutif aux phosphorylations des RAR (projet de financement, page 160).

En résumé, la nature pléiotrope de la voie de l'AR, ainsi que les nombreux acteurs qui régulent son activité la positionne comme cible privilégiée au cours de l'évolution pour la genèse de nouveaux phénotypes.

3. Une diversité de denture, mais pourquoi ?

La diversité de la denture pharyngienne des Cypriniformes, en particulier au niveau de la forme des dents, pose la question du rôle adaptatif d'une telle diversité. Chez les Cypriniformes, l'absence de dents orales oblige le poisson à gober sa nourriture mais il y a en revanche trituration au niveau des dents pharyngiennes (Pasco-Viel et al., 2010). Ainsi, l'hypothèse principale est que la forme des dents pharyngiennes chez ces poissons pourrait être corrélée au régime alimentaire. Une étude récente impliquant le laboratoire de Vincent Laudet a recensé les différences de nombre et de forme des dents pharyngiennes chez deux tribus de Cyprininés (une des six familles de cypriniformes): les Labeonini et les Poropuntiini. L'étude de la denture pharyngienne de plus d'une centaine d'espèces par microtomographie a permis de mettre en évidence une différence essentielle entre ces deux tribus. Alors que le nombre et la forme des dents sont extrêmement conservés dans l'ensemble des Labeonini, ces deux caractères sont au contraire très variables chez les Poropuntiini (Pasco-viel et al., en préparation). Les données sur le régime alimentaire, bien que très limitées, tendent à montrer que les Labeonini sont plutôt des espèces se nourrissant essentiellement de plancton et de détritus alors que les Poropuntiini ont des régimes alimentaires variables (insectivores, végétivores, etc.). Il semble donc qu'il existe une corrélation entre morphologie dentaire et régime alimentaire chez ces groupes. Par ailleurs, plusieurs autres caractères morpho-anatomiques vont dans le même sens : position de la bouche différente entre Labeonini (bouche inférieure) et Poropuntiini (bouche terminale), ainsi que l'espace entre les dents et le coussinet kératinisé contre lequel viennent s'écraser les dents. Au final, l'étude comparée de ces deux tribus de Cyprininés tend à démontrer une corrélation entre morphologie des dents pharyngiennes et régime alimentaire.

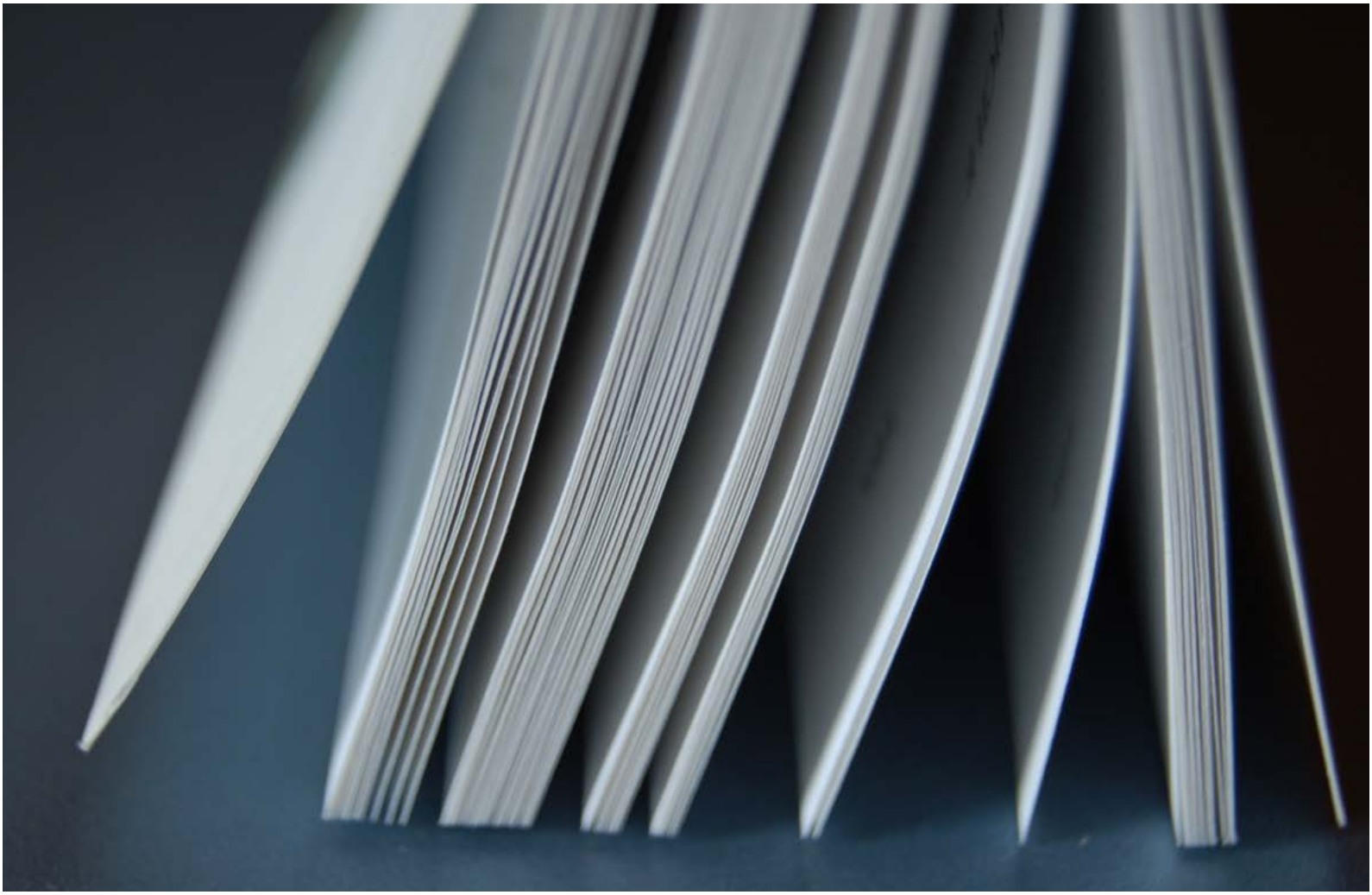
Conclusion générale

Mes travaux de thèse ont donc permis de générer de nouveaux outils pour l'étude des RAR chez le poisson-zèbre. De plus, ils participent à la compréhension des mécanismes moléculaires régissant l'activité transcriptionnelle des RAR *in vivo*. Ils ouvrent de nouvelles perspectives de recherche, visant à utiliser les dernières techniques de génomique fonctionnelle, qui font actuellement l'objet de demandes de financement et qui devraient permettre d'aller plus loin en caractérisant les cistromes des RAR endogènes *in vivo*.

L'étude des phosphorylations des RAR chez les vertébrés a également permis de mettre en avant un rôle de ces modifications post-traductionnelles dans l'évolution de la régulation de l'activité des récepteurs. Plus largement, nos résultats proposent de considérer les sites de phosphorylation comme des cibles majeures dans l'évolution de la fonction des protéines. En effet, l'étude expérimentale sur les RAR de poisson-zèbre a permis de proposer un scénario évolutif de l'acquisition des sites de phosphorylation des RAR et des conséquences sur la structure et la fonction du récepteur. Ces travaux consolident la base des recherches à venir sur l'étude du rôle joué par ces phosphorylations *in vivo* au cours du développement du poisson-zèbre.

Enfin, mes travaux ont mis en avant un rôle majeur de la voie de l'acide rétinoïque dans l'évolution, notamment pour la genèse de nouveautés morphologiques. En particulier, nous avons mis en avant que des changements locaux du métabolisme de l'AR dans le pharynx postérieur pouvait être à l'origine de la diversité de denture observée chez les cypriniformes, ou plus largement chez les poissons téléostéens. A plus large échelle, ces résultats positionnent la voie de l'acide rétinoïque au centre des questions d'évolution des phénotypes et ouvrent de nouvelles perspectives de recherche en Evo-Devo.

Annexes



The FASEB Journal • Research Communication

Vinexinβ, an atypical "sensor" of retinoic acid receptor γ signaling: union and sequestration, separation, and phosphorylation

Sébastien Lalevée, Gaétan Bour,[1] Marc Quinternet, Eric Samarut, Pascal Kessler, Marc Vitorino,[2] Nathalie Bruck, Marc-André Delsuc, Jean-Luc Vonesch, Bruno Kieffer, and Cécile Rochette-Egly[3]

Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de la Santé et de la Recherche Médicale (INSERM), U596, and Centre National de la Recherche Scientifique (CNRS), Unité Mixte de Recherche (UMR) 7104, Université de Strasbourg, Illkirch, France

ABSTRACT **The transcriptional activity of nuclear retinoic acid receptors (RARs) relies on the association/dissociation of coregulators at the ligand-binding domain. However, we determined that the N-terminal domain (NTD) also plays a role through its phosphorylation, and we isolated vinexinβ, a cytoskeleton protein with three SH3 domains, as a new partner of the RARγ NTD. Here we deciphered the mechanism of the interaction and its role in RARγ-mediated transcription. By combining molecular and biophysical (surface plasmon resonance, NMR, and fluorescence resonance energy transfer) approaches, we demonstrated that the third SH3 domain of vinexinβ interacts with a proline-rich domain (PRD) located in RARγ NTD and that phosphorylation at a serine located in the PRD abrogates the interaction. The affinity of the interaction was also evaluated. *In vivo*, vinexinβ represses RARγ-mediated transcription and we dissected the underlying mechanism in chromatin immunoprecipitation experiments performed with F9 cells expressing RARγ wild type or mutated at the phosphorylation site. In the absence of retinoic acid (RA), vinexinβ does not occupy RARγ target gene promoters and sequesters non-phosphorylated RARγ out of promoters. In response to RA, RARγ becomes phosphorylated and dissociates from vinexinβ. This separation allows RARγ to occupy promoters. This is the first report of an RAR corepressor association/dissociation out of promoters and regulated by phosphorylation.—Lalevée, S., Bour, G., Quinternet, M., Samarut, E., Kessler, P., Vitorino, M., Bruck, N., Delsuc, M.-A., Vonesch, J.-L., Kieffer, B., Rochette-Egly, C. Vinexinβ, an atypical "sensor" of retinoic acid receptor γ signaling: union and sequestration, separation, and phosphorylation. *FASEB J.* 24, 4523–4534 (2010). www.fasebj.org**

Key Words: nuclear receptors · coregulators · retinoic acid · transcription

RETINOIC ACID (RA) INFLUENCES the differentiation, proliferation, and apoptosis of a variety of cell types through modifications in the expression of target genes (1). The RA response is mediated by two classes of nuclear receptors, the retinoic acid receptors (RARs) (α, β, and γ) and the retinoid X receptors (RXRs) (α, β, and γ), which function as ligand-dependent heterodimeric RAR/RXR transcription activators (2, 3). Gene induction by RAR/RXR heterodimers relies on a complex network of dynamic interactions with coregulatory proteins (for review, see ref. 4). Indeed, after the RA signal, RAR/RXR heterodimers bound at RA response elements (RAREs) undergo major structural rearrangements, which initiate an ordered and coordinated dissociation and/or recruitment of a series of coregulator complexes with different enzymatic activities including histone acetyl and methyl transferases, and DNA-dependent ATPases. At the end, these events alter the chromatin structure surrounding the promoter of target genes and pave the way for the recruitment of the transcription machinery including RNA polymerase II and general transcription factors. This network, which is directed by the C-terminal AF-2 domain (helix 12) located in the ligand-binding domain (LBD), was typically described for the RARα subtype.

In addition to this scenario, it emerged that the N-terminal domain (NTD) of RARs, although of naturally disordered structure (5–7), also plays a role, adding more complexity to RAR-mediated transcription mechanisms. In this context, we recently highlighted a novel regulation mechanism for the RARγ subtype. Indeed, we found that the NTD of RARγ is phosphorylated (8, 9) and interacts with vinexinβ (10), a protein with three SH3 domains, which belongs to the vinexin/CAP/ponsin/ArgBP2 family of adaptor proteins (11, 12). Although generally associated as a scaffolding protein to complexes involved in cytoskeleton organization and signal transduction, vinexinβ is also one of a growing number of actin-binding proteins that

[1] Current address: IRCAD, 67000 Strasbourg, France.
[2] Current address: Astex Therapeutics, Cambridge CB4 0QA, UK.
[3] Correspondence: IGBMC, 1 rue Laurent Fries, BP 10142, 67404 Illkirch Cedex, France. E-mail: cegly@igbmc.fr
doi: 10.1096/fj.10-160572

are present in the nucleus and modulate transcription (13), and we found that vinexinβ represses RARγ-mediated transcription (10).

Here, we aimed to decipher the underlying mechanism of vinexinβ interaction with RARγ, by combining molecular, cellular, and biophysical approaches. We demonstrated that the interaction involves the third C-terminal SH3 domain of vinexinβ and a proline-rich domain (PRD), which is located in the NTD of RARγ and contains phosphorylation sites. The affinity of the interaction and the influence of phosphorylation were determined in surface plasmon resonance (SPR) and NMR experiments. We reveal a completely novel regulation mechanism that is distinct from the classic corepressor network. In this new model, vinexinβ represses transcription through sequestering RARγ out of gene promoters. Then the release of vinexinβ that occurs on RARγ phosphorylation, in response to RA, allows liganded and phosphorylated RARγ to occupy the promoters of target genes and to initiate transcription.

MATERIALS AND METHODS

Plasmids and peptides

The following plasmids have been described previously (8, 10, 14): hRARγ1WT, S79A, S77A, S77A/S79A, ΔNTD, and ΔH12 in pSG5; hRARγ1, hRARα1, and hRARβ2 in pGEX-2T; FLAG-vinexinβ in the pCX vector; and the RARβ2-CAT reporter gene. The vectors encoding vinexinβ deleted for each SH3 domain, the glutathione *S*-transferase (GST)-RARγS77E, S79E, and S77E/S79E chimera and the pSG5 RARγ1 mutants P78A, P80A, P82A, P84A, R85A, and R88A, were constructed by double PCR amplification. The pSG5 RARγS77E, RARγS79E, and RARγS77E/S79E vectors were obtained by subcloning each mutant from pGEX2T to pSG5. GST-vinexinβ and GST-V-SH3-3 were constructed by subcloning the corresponding fragments amplified by PCR into pGEX2T.

Cyan fluorescent protein (CFP)-RARγ and yellow fluorescent protein (YFP)-vinexinβ constructs were made by first introducing the fragment corresponding to the fluorochromes into the HA-FLAG-pCX vector. Then the SH3-3 domain of vinexinβ, RARγWT, RARγS79A, or RARγS79E or the AB regions of RARγS79A were amplified by PCR and ligated in-frame at the C-terminal end of the fluorochrome. RARγS79A and RARγ(AB)S79A were also introduced at the N-terminal end of CFP.

Phospho and nonphospho peptides were synthesized in solid phase using Fmoc chemistry (model 431A peptide synthesizer; Applied Biosystems, Foster City, CA, USA), purified by reversed-phase HPLC, and verified by electrospray ionization-tandem mass spectrometry.

Antibodies

Polyclonal antibodies recognizing RARγ [RPγ(F)] were described previously (8) and purified on SulfoLink gel columns (Pierce Chemical, Rockford, IL, USA) coupled to the corresponding immunizing peptide. Monoclonal antibodies against human vinexinβ and cyclin H have been described previously (10, 15). Polyclonal antibodies against β-actin and the cdk7 subunit of TFIIH (C-19) were from Santa Cruz Biotechnology (Santa Cruz, CA, USA). Anti-FLAG monoclonal antibodies were from Sigma-Aldrich (St. Louis, MO, USA). Monoclonal antibodies recognizing specifically RARγ phosphorylated at position S79 (RARγ-P-S79) were generated by immunization of BALB/c mice with synthetic phosphopeptides coupled to ovalbumin according to standard procedures.

Recombinant protein expression and purification

Transformed BL21-DE3 cells were grown in minimal medium supplemented with $^{15}NH_4Cl$, to OD_{600} ≈0.6, and protein expression was induced using 1 mM isopropyl β-D-thiogalactoside for 8 h. After centrifugation, cells were resuspended in a minimal volume of 20 mM Tris (pH 7.8), containing 200 mM NaCl and 1 mM DTT, and sonicated. The GST fusion proteins were isolated using a GSTrap HP affinity column (GE Healthcare, Little Chalfont, UK). Then, the GST part was cleaved with thrombin and removed by exclusion size chromatography on a HiLoad 16/60 Superdex 75 PG column (GE Healthcare) in 20 mM Tris (pH 7.0), containing 100 mM NaCl. Fractions containing purified ^{15}N-labeled V-SH3-3 were pooled and concentrated using a Vivaspin system (Sartorius Stedim Biotech, Aubagne, France).

NMR spectroscopy and chemical shift perturbation

NMR experiments were performed in 50-μl capillaries at 300K on a DRX600 spectrometer (Bruker, Newark, DE, USA) equipped with a cryoprobe. Several ^{15}N-^{1}H heteronuclear single quantum correlation (HSQC) spectra were recorded to follow the titration of V-SH3-3 with synthetic peptides corresponding to the PRD of RARγ or RARα. All peptides were first desalted and recovered in 100 mM phosphate buffer (pH 7.0) using NAP-5 columns (Pharmacia Biotech, Uppsala, Sweden). The final NMR samples contained 4 mM Tris, 20 mM NaCl, and 80 mM phosphate buffer. Titrations were done with 54 μM uniformly ^{15}N-labeled V-SH3-3 to which peptides were added. Composite ^{15}N and ^{1}H chemical shifts were used as described in ref. 16, using the 5 most perturbed ^{15}N-^{1}H correlations that were added for averaging of statistical errors in the K_d (dissociation constant) calculations. Two-parameter nonlinear curve fitting was performed using in-house Python scripts to get the K_d values. The statistical errors were estimated from Monte Carlo simulations.

SPR measurements

SPR experiments were performed with a BIAcore T100 instrument (Biacore AB, Uppsala, Sweden) and research grade CM5 sensor chips (17). The ligand (GST-V-SH3-3) was captured *via* anti-GST antibodies that were immobilized on the sensor surface using standard amine coupling procedures and following the manufacturer's instructions. Synthetic peptides corresponding to the PRD of RARs were diluted in running buffer [10 mM Hepes (pH 7.4), 150 mM NaCl, 3.4 mM EDTA, and 0.005% (v/v) P20] and injected over the surface in a continuous flow at 25°C.

Cell lines, immunoblotting, immunoprecipitation, and immunofluorescence

COS-1 cells and the different F9 cell lines were cultured and transiently transfected as in ref. 10. When 80–90% confluent, cells were treated with RA (10^{-7} M) after 24 h in a medium containing 1% dextran-charcoal-treated FCS. Extract preparation, immunoprecipitation, immunoblotting, and immunofluorescence analyses were described previously (10).

Fluorescence resonance energy transfer (FRET)

COS-1 cells coexpressing CFP-RARγ and YFP-V-SH3-3 were imaged on a confocal SP2 AOBS MP microscope (Leica, Wetzlar, Germany). FRET was measured by using the sensitized emission method. For each sample, three images were acquired: one for the donor (donor excitation and donor emission), one for the acceptor (acceptor excitation and acceptor emission), and one for the FRET channel (donor excitation and acceptor emission). The CFP and YFP images of cells were collected after excitation with 458 and 514 nm laser lights, respectively. Samples were analyzed using the Leica application wizard for FRET sensitized emission. For each sample, the efficiency (E) of FRET was measured using the formula $E = [B - A \times b - C \times (c - a \times b)]/C$ with A, B, and C corresponding to the intensities of the three channels (donor, FRET, and acceptor) and a, b, and c being the calibration factors generated by acceptor-only and donor-only references. For the donor-only sample, the correction of donor cross-talk in the FRET image was calculated as $b = B/A$. For the acceptor-only sample, the two correction factors a and c were calculated as $c = B/C$ and $a = A/C$.

GST pulldown assays

Equimolar GST, GST-RARγ1, GST-RARα1, and GST-RARβ2 fusion proteins were purified on glutathione-Sepharose 4B beads and incubated with COS-1 cell extracts expressing the FLAG-vinexinβ proteins as described previously (10).

Chromatin and nuclear extract preparation

The cytosolic, soluble nucleoplasmic, and insoluble native chromatin fractions were prepared according to the protocol described in ref. 18. In brief, cells were lysed in 10 mM Tris · HCl (pH 7.65) containing 3 mM $MgCl_2$, 10 mM KCl, and protease inhibitors, and intact nuclei were pelleted by centrifugation (10 min at 10,000 *g*). Nuclei were then lysed in 20 mM Tris HCl (pH 7.65) containing 25% glycerol, 3 mM $MgCl_2$, 0.2 mM EDTA, 0.9 M NaCl, and protease inhibitors. The nucleoplasmic fraction was cleared by centrifugation, and the chromatin-enriched pellet was resuspended in 20 mM Tris HCl (pH 7.65) containing 15 mM KCl, 60 mM NaCl, 0.34 M sucrose, 0.15 M spermine, 0.5 mM spermidine, and protease inhibitors. Then solubilized native chromatin proteins were obtained by DNA digestion with micrococcal nuclease (5 U/μl) and cleared by centrifugation.

RNA isolation and quantitative RT-PCR

These were performed as described previously (10).

Chromatin immunoprecipitation (ChIP) experiments

Subconfluent cells were treated with RA (10^{-7} M), and ChIP experiments were performed as described previously (15). Control ChIPs were performed without antibodies, and RARγ was immunoprecipitated with purified RPγ(F). Immunoprecipitated DNA was amplified by quantitative PCR with primers for RARβ2, 5′-CGATCCCAAGTTCTCCCTTC-3′ and 5′-CAGACTGGTTGGGTCATTTG-3′ and 36B4, 5′-TTTGCTGTACTGACTCGGTGA-3′ and 5′-CCTCCCACAACAAAACAACC-3′. Occupancy of the promoters was calculated by normalizing the PCR signals from the immunoprecipitation samples to the signals obtained from the input DNA.

Electromobility shift assays (EMSAs)

Nuclear extracts (2 μg) from COS-1 cells transfected with RARγ were incubated with purified RXRα overexpressed in *Escherichia coli* (50 ng) and then with a biotin-labeled oligonucleotide probe containing a direct repeat of a RARE in which the motifs are interspaced by 5 bp (DR5) as described previously (19). Complexes were resolved on a 5% native acrylamide gel and revealed by chemiluminescence using a streptavidin-horseradish peroxidase conjugate, according to the manufacturer's instructions (Thermo Fisher Scientific, Waltham, MA, USA).

RESULTS

Third C-terminal SH3 domain of vinexinβ interacts with the PRD of RARγ

Vinexinβ being characterized by the presence of three SH3 domains (11, 12), we first investigated whether one of these domains is involved in the interaction with RARγ. FLAG-tagged vinexinβ full-length or deleted for each SH3 domain (**Fig. 1*A***) was analyzed for its ability to interact with GST-RARγ1. We found that the interaction involves the third C-terminal SH3 domain of vinexinβ. Indeed, deletion of this domain (in vinexinβ-ΔSH3-3) abrogated the interaction (Fig. 1*B*, lane 9), whereas deletion of the first or second N-terminal SH3 domain (in vinexinβ-ΔSH3-1 and ΔSH3-2) had no effect (Fig. 1*B*, lanes 7 and 8). These results were confirmed in coimmunoprecipitation experiments performed with extracts from COS-1 cells overexpressing RARγ1 and the same FLAG-vinexinβ deletion mutants (Fig. 1*C*). In addition, we found that the isolated third C-terminal domain of vinexinβ (V-SH3-3) fused to GST interacted with highly purified bacterially expressed RARγ, indicating that the interaction is direct (Fig. 1*D*, lane 3). This interaction was abrogated on deletion of the N-terminal RARγ regions A and B (in RARγΔNTD), corroborating that it involves the NTD (Fig. 1*D*, lane 7). It was not affected on RA addition to RARγ, indicating that it is not modulated by RA-induced conformational changes of the receptor (Fig. 1*D*, compare lanes 3 and 4).

Given that SH3 domains are known to interact with PRDs (20, 21), we also investigated which residues within this domain located in the NTD of RARγ1 are involved in the interaction. Interestingly, RARα that differs from RARγ by the substitution of one proline residue (P83) within the pentaproline segment by a leucine residue (Fig. 1*E*) did not interact with vinexinβ in GST pulldown experiments (Fig. 1*F*). In contrast, RARβ, in which P83 is conserved but P81 substituted by a leucine, retained the ability to interact with vinexinβ (Fig. 1*F*). This result suggests that P83 is required for the binding of RARγ to vinexinβ. Then we substituted individually alanine residues for all of the other proline residues (from P78 to P84) and the flanking basic R85 and K88 residues. The corresponding RARγ mutants were tested for their ability to bind vinexinβ. Substitution of P80 and R85 abrogated the interaction (Fig. 1*G*, lanes 4 and 7). Taken together, these results indicate that vinexinβ interacts with RARγ *via* a PxxPxR motif

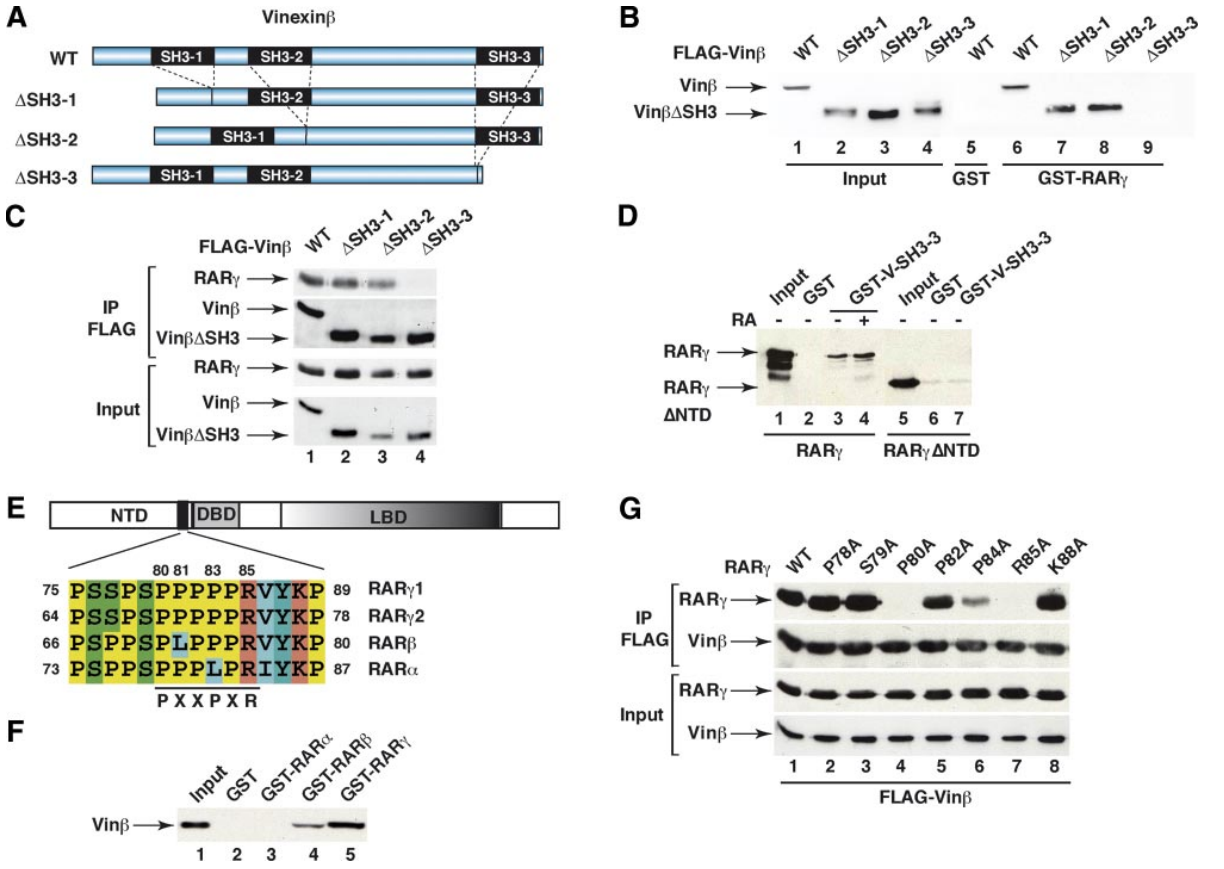


Figure 1. The third C-terminal SH3 domain of vinexinβ (Vinβ) interacts with the PRD of RARγ. *A*) Schematic representation of vinexinβ and the mutants lacking the SH3 domains. *B*) Extracts (5 μg) from COS-1 cells overexpressing FLAG-vinexinβ WT or lacking the first, second, or third SH3 domain were incubated with GST or GST-RARγ1 immobilized on glutathione-Sepharose beads. Bound vinexinβ was analyzed by immunoblotting with FLAG antibodies. *C*) COS-1 cells were transfected with the RARγ1 vector along with the different FLAG-vinexinβ deletion mutants, and nuclear extracts were incubated with FLAG antibodies and protein G-Sepharose beads. Immunocomplexes were resolved by SDS-PAGE and immunoblotted with RPγ (*F*) and FLAG antibodies. Bottom 2 panels correspond to 5% of the amount of immunoprecipitated extracts. *D*) Immobilized GST and GST-V-SH3-3 was incubated with highly purified RARγ, either WT or deleted for the NTD (5 μg), in the absence or presence of RA, as indicated. Bound RARγ was analyzed by immunoblotting as in *B*. *E*) Amino acid sequence of the PRD of RARγ1, RARγ2, RARα1, and RARβ2. *F*) Immobilized GST, GST-RARα1, GST-RARβ2, or GST-RARγ1 was incubated with extracts (5 μg) from COS-1 cells overexpressing FLAG-vinexinβ. Bound vinexinβ was analyzed by immunoblotting as in *B*. *G*) COS-1 cells were transfected with the FLAG-vinexinβ vector along with RARγ, either WT or with each residue of the PRD substituted by an alanine. Nuclear extracts were immunoprecipitated with FLAG antibodies and analyzed as in *C*.

(Fig. 1*E*), which corresponds to a typical class II ligand for SH3 domains (20). Note that vinexinβ interacted as efficiently with the RARγ2 isoform (data not shown), in which the PRD is conserved (Fig. 1*E*).

Phosphorylation of the RARγ PRD at serine 79 prevents vinexinβ binding

The RARγ1 PxxPxR motif involved in the interaction with vinexinβ is flanked at the N-terminal end by two serine residues (S77 and S79) (**Fig. 2*A***), which can be phosphorylated *in vivo* and *in vitro* (8, 9). Therefore, we hypothesized that phosphorylation of these residues might affect the capacity of RARγ to interact with vinexinβ.

GST-RARγ1 mutants with S77 and/or S79 substituted by a glutamic acid that mimics a phosphorylated residue (Fig. 2*A*) were constructed and compared with GST-RARγ1WT for their interaction with vinexinβ. When expressed in *E. coli*, GST-RARγ1WT was not phosphorylated and interacted with vinexinβ (Fig. 2*B*, lane 3). GST-RARγS77E interacted as efficiently as the wild-type (WT) receptor (Fig. 2*B*, lane 4). However, the S79E mutant did not interact efficiently with vinexinβ (Fig. 2*B*, lane 5). The double S77E/S79E mutant did not interact either (Fig. 2*B*, lane 6). These results indicate that phosphorylation of S79 prevents the interaction of RARγ with vinexinβ. They were confirmed in coimmunoprecipitation experiments performed with extracts from COS-1 cells overexpressing vinexinβ and the same RARγ1S77E, S79E, or S77E/S79E mutants (Fig. 2*C*). Interestingly, RARγ1 with S79 substituted by an alanine residue, which mimics a nonphosphorylated form, interacted more efficiently with vinexinβ than

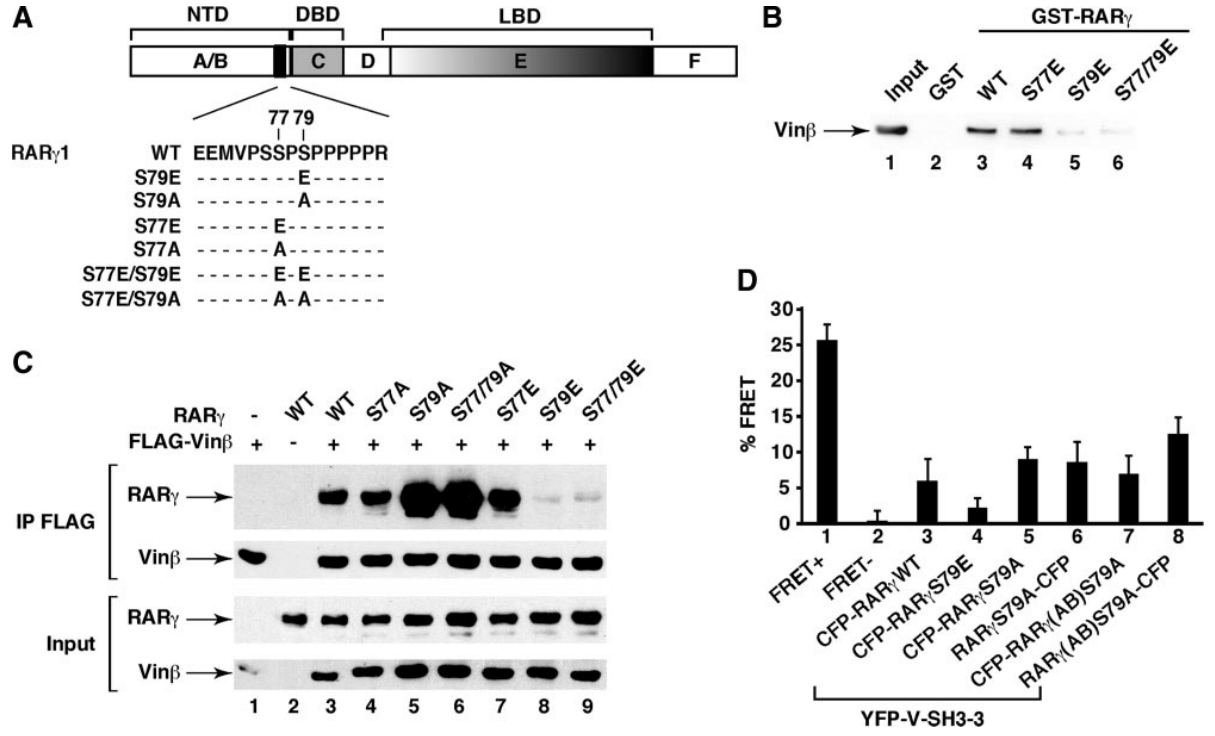


Figure 2. RARγ phosphorylation at S79 located in the PRD prevents vinexinβ (Vinβ) binding. *A*) Schematic representation of the RARγ PRD with S77 and/or S79 substituted by alanine or glutamic acid residues. *B*) Extracts (5 μg) from COS-1 cells overexpressing FLAG-vinexinβ were incubated with immobilized GST or GST-RARγ1WT, S77E, S79E, or S77E/S79E. Bound vinexinβ was analyzed by immunoblotting with FLAG antibodies. *C*) Coimmunoprecipitation experiments showing that in transfected COS-1 cells, substitution of serine 79 by a glutamic acid or an alanine residue decreases and increases, respectively, the interaction with vinexinβ. *D*) FRET analysis of interaction between YFP-V-SH3-3 and CFP-RARγ in COS-1 cells. Highest and lowest FRET efficiencies were observed with RARγS79A and RARγS79E respectively. Similar results were obtained with CFP fused at the C-terminal end of RARγ. Lanes 1 and 2 correspond to positive and negative FRET controls. Values are means ± SE of ≥3 analyses.

RARγ1WT (Fig. 2*C*, lanes 5 and 6), suggesting that in transfected cells, a fraction of the RARγ1 pool might be already phosphorylated at this residue.

Then we used imaging FRET to examine direct RARγ-vinexinβ interactions in living cells. As a positive control, COS-1 cells were transfected with a CFP-YFP construct in which CFP and YFP were linked *via* 10 amino acids (Fig. 2*D*, lane 1). Negative FRET controls were established with cells cotransfected independently with CFP and YFP (Fig. 2*D*, lane 2). COS-1 cells were transfected with a CFP-RARγ vector along with the SH3-3 domain of vinexinβ (V-SH3-3), fused at the C-terminal end of the YFP fluorochrome (YFP-V-SH3-3). A significant FRET signal was produced in nuclei coexpressing CFP-RARγWT and YFP-V-SH3-3 (Fig. 2*D*, lane 3). The FRET signal decreased significantly when S79 within the PRD of RARγ was substituted by a glutamic acid (Fig. 2*D*, lane 4) but was higher when the same serine was substituted by an alanine (Fig. 2*D*, lane 5). Similar results were obtained when the CFP fluorochrome was fused at the C-terminal end of RARγ (Fig. 2*D*, lane 6). Finally, a positive FRET signal was also produced when the NTD alone of RARγ was fused to CFP (Fig. 2*D*, lanes 7 and 8), corroborating the importance of this domain in the interaction. Collectively these data indicate that RARγ interacts directly with vinexinβ and that S79 phosphorylation impedes the interaction.

Affinity of the interaction between the SH3-3 domain of vinexinβ and the PRD of RARγ

Two biophysical techniques (SPR and NMR) were then used to determine the affinity of the interaction between the SH3-3 domain of vinexinβ (V-SH3-3) and synthetic peptides corresponding to the PRD of RARγ. Such peptides were synthesized with S77 and/or S79 phosphorylated or not (**Table 1**) to get more accurate insights into the influence of phosphorylation on the interaction.

SPR was first used to measure the equilibrium affinity and kinetic parameters of the interaction of V-SH3-3 with the PRD of RARγ (**Fig. 3*A***). Affinity between V-SH3-3 and peptide PI121 that corresponds to the nonphosphorylated PRD was calculated with the simple Langmuir 1:1 model and Biacore T100 evaluation software (version 1.1.1) and found to be 80 ± 16μM. Affinity for peptide PI120 in which S77 is phosphorylated was 5 times weaker. The weakest affinity was found for peptides PI119 (phosphorylated at S79), PI118 (phosphorylated at both S77 and S79), and PI80 (PRD of RARα).

The affinity of the PRD peptides targeting the

TABLE 1. *Amino acid sequence of the synthetic peptides corresponding to the phosphorylated or nonphosphorylated proline-rich domains of RARγ and RARα*

Peptide	Amino acid sequence	PRD
	77 79	
PI121	EEMVPSSPSPPPPPRVYK	Nonphosphorylated PRD of RARγ
PI120	EEMVPSSPSPPPPPRVYK P	PRD of RARγ phosphorylated at S77
PI119	EEMVPSSPSPPPPPRVYK P	PRD of RARγ phosphorylated at S79
PI118	EEMVPSSPSPPPPPRVYK P P	PRD of RARγ phosphorylated at S77 and S79
	77	
PI80	EEIVPSPPSPPPLPRIYK	Nonphosphorylated PRD of RARα

Amino acid sequences are given in 1-letter codes.

SH3-3 domain of vinexinβ was further investigated in solution using ^{1}H-^{15}N heteronuclear NMR. The titration of the ^{15}N labeled V-SH3-3 domain with increasing amounts of peptides led to a progressive shift of a set of correlation peaks in the ^{1}H-^{15}N HSQC spectra (Fig. 3*B*), as expected for weak affinity ligands. In this situation, the resonance frequencies are weighted averages of the frequencies characterizing the free and peptide-bound forms of the protein, allowing the determination of an accurate K_d value from the analysis of the frequency shifts as a function of peptide concentrations (Fig. 3*C*). The K_d for the interaction between V-SH3-3 and the nonphosphorylated PI121 peptide was found to be 37 ± 9 μM, a

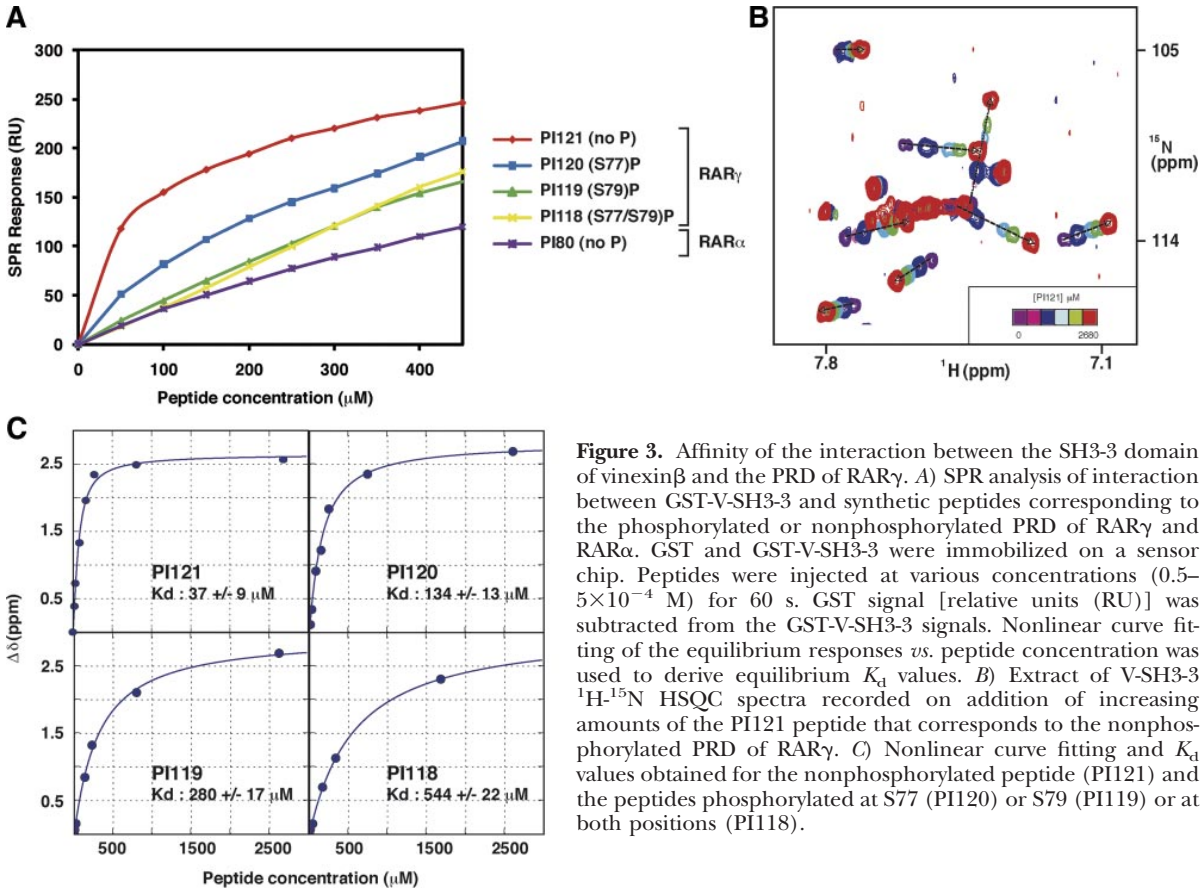


Figure 3. Affinity of the interaction between the SH3-3 domain of vinexinβ and the PRD of RARγ. *A*) SPR analysis of interaction between GST-V-SH3-3 and synthetic peptides corresponding to the phosphorylated or nonphosphorylated PRD of RARγ and RARα. GST and GST-V-SH3-3 were immobilized on a sensor chip. Peptides were injected at various concentrations (0.5–5×10^{-4} M) for 60 s. GST signal [relative units (RU)] was subtracted from the GST-V-SH3-3 signals. Nonlinear curve fitting of the equilibrium responses *vs.* peptide concentration was used to derive equilibrium K_d values. *B*) Extract of V-SH3-3 ^{1}H-^{15}N HSQC spectra recorded on addition of increasing amounts of the PI121 peptide that corresponds to the nonphosphorylated PRD of RARγ. *C*) Nonlinear curve fitting and K_d values obtained for the nonphosphorylated peptide (PI121) and the peptides phosphorylated at S77 (PI120) or S79 (PI119) or at both positions (PI118).

value that is compatible with SPR experiments, the slight difference being due to the different state of the V-SH3-3 protein in both measurements (surface-immobilized *vs.* in solution). Such a K_d value is in the range generally observed for SH3 domains and proline-rich peptides (20). The affinity for the peptide phosphorylated at S77 (PI120) is 4 times weaker, whereas peptides phosphorylated at S79 (PI119) and at both sites (PI118) displayed a significant decrease in affinity (by a factor of 8 and 15, respectively) (Fig. 3*C*). Taken together, these results indicate that the interaction between the PRD of RARγ and V-SH3-3 is weak and confirm that it is down-regulated by phosphorylation of S79 in the PRD.

RA induces RARγ phosphorylation and the subsequent dissociation of vinexinβ

RA has been shown to induce the rapid phosphorylation of RARα at a serine residue located in the PRD (15). Given that phosphorylation affects the affinity of the RARγ PRD for vinexinβ, we investigated whether the phosphorylation of RARγ and therefore its interaction with vinexinβ are affected in response to RA.

In transfected COS-1 cells, RA induced the rapid phosphorylation of RARγ1 at S79, as assessed by immunoblotting after immunoprecipitation of RARγ with a monoclonal antibody recognizing specifically RARγ phosphorylated at this residue (**Fig. 4*A*, *C*,** lanes 1–5). No signal was observed with the RARγS79A mutant, confirming the specificity of the antibody (Fig. 4*C*, lanes 6–10).

Interestingly, RARγ from untreated cell extracts, but not RARγ from RA-treated cells, interacted significantly with GST-V-SH3-3 (Fig. 4*B*, lanes 7 and 10). In contrast, the RARγS79A mutant, which cannot be phosphorylated, interacted more efficiently than RARγWT with the SH3 domain of vinexinβ, and this interaction was not significantly affected by RA (Fig. 4*B*, lanes 8 and 11). Such results indicate that the RA-induced phosphorylation of the serine residue located in the NTD of RARγ abrogates the interaction with vinexinβ. Corroborating this, deletion of the N-terminal domain of RARγ completely abrogated the interaction (Fig. 4*B*, lanes 9 and 12). Similar results were obtained with GST-vinexinβ full-length (data not shown)

Then in coimmunoprecipitation experiments, we found that the amount of RARγ bound to vinexinβ decreased rapidly after RA addition and was completely abrogated at 30–60 min at the peak of RARγ phosphorylation (Fig. 4*C*, *D*, lanes 1–5). This finding indicates that vinexinβ is released in response to RA. A similar dissociation was observed with RARγS77A (data not shown). In contrast, the RARγS79A mutant, which cannot be phosphorylated in response to RA, dissociated only slightly from vinexinβ at 5 min and interacted even more efficiently at 60 min (Fig. 4*C*, lanes 6–10). Such results indicate that the full RA-induced dissociation of vinexinβ requires the phosphorylation of RARγ at S79 but that *in vivo* additional rapid RA-induced regulatory mechanisms are likely to be also involved.

Therefore, we investigated whether the LBD and more precisely the C-terminal helix 12, the conformation of which changes in response to RA (4), plays a

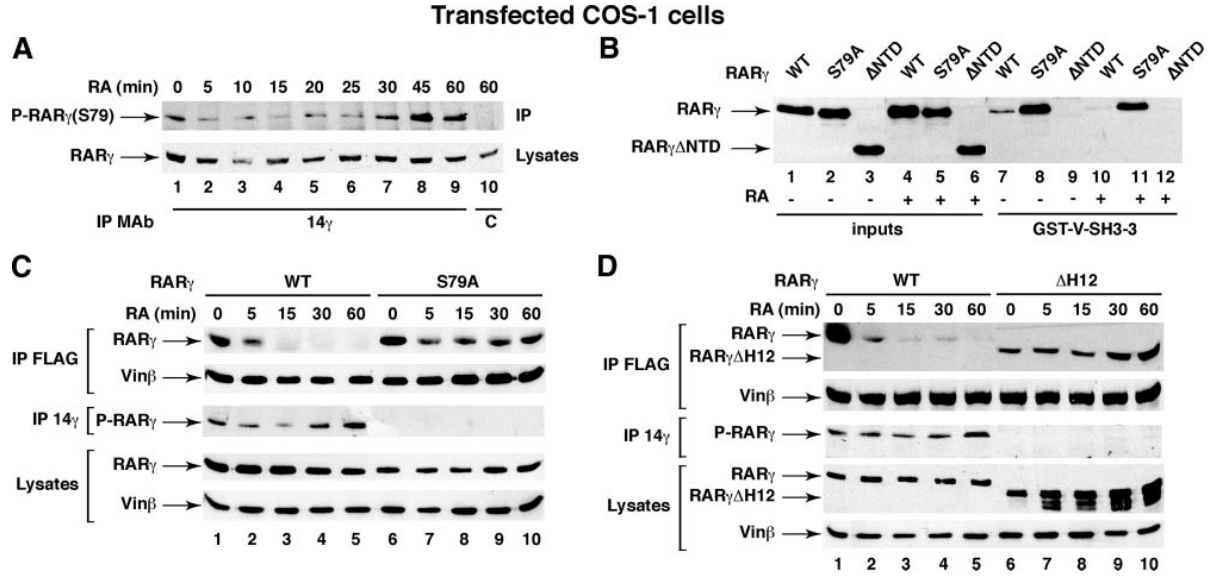


Figure 4. RA induces RARγ phosphorylation and vinexinβ (Vinβ) dissociation. *A*) In transfected COS-1 cells, RA induces the rapid phosphorylation of RARγ1 at S79. Phosphorylated RARγ was immunoprecipitated (IP) with monoclonal antibodies (MAb) recognizing specifically RARγ phosphorylated at S79 and immunoblotted with RPγ(F). Bottom panel corresponds to the inputs. *B*) Extracts (5 μg) from COS-1 cells overexpressing RARγ, either WT, S79A, or ΔNTD, and treated or not with RA for 45 min were incubated with immobilized GST-V-SH3-3. Bound RARγ was analyzed by immunoblotting. *C*) Coimmunoprecipitation experiments showing that after RA addition, vinexinβ dissociates from RARγWT and less efficiently from RARγS79A. *D*) Vinexinβ interacts less efficiently with RARγΔH12 and does not dissociate from this mutant in coimmunoprecipitation experiments.

role in the association-dissociation of vinexinβ. RARγ deleted for helix 12 (RARγΔH12) was analyzed in coimmunoprecipitation experiments for its ability to interact with vinexinβ. In the absence of RA, less vinexinβ interacted with RARγΔH12 compared with RARγWT (Fig. 4*C*, compare lanes 1 and 6), suggesting that helix 12 would contribute, in addition to the NTD, to vinexinβ binding. Given that vinexinβ is a scaffolding protein with three SH3 domains, which have been shown to interact with several proteins (22, 23), one can propose that it interacts not only with the NTD of RARγ *via* its third SH3 domain but also with the LBD through intermediary proteins recruited *via* the two other SH3 domains. Such a model would be in agreement with the rapid and slight RA-induced dissociation of vinexinβ that occurs before and independently of RARγ phosphorylation at the NTD.

Of note is that, in response to RA, the interaction between vinexinβ and RARγΔH12 was not disrupted but rather increased (Fig. 4*C*, lanes 6–10), most probably because this mutant was not phosphorylated at S79 either in the absence or in the presence of RA (Fig. 4*C*, lanes 6–10) for reasons that remain to be determined.

Then the question was whether phosphorylation also controls the association/dissociation of endogenous RARγ and vinexinβ proteins. We used mouse embryo carcinoma cells (F9 cell line), which constitute a well-established model system for investigating RA signaling (1). In these cells, RARγ (the RARγ2 isoform) was found exclusively in nuclei as assessed in epifluorescence and confocal experiments (**Fig. 5*A***, panels *1*, *3*, and *7*). Vinexinβ was present in both nuclei and cytoplasm (Fig. 5*A*, panels *4*, *6*, and *9*) and colocalized with RARγ in nuclei (Fig. 5*A*, panel *10*).

Highly purified and intact nuclei were also prepared from F9 cells, solubilized and analyzed in coimmunoprecipitation experiments. In the absence of RA, endogenous vinexinβ and RARγ did interact (Fig. 5*B*, lane 2). Interestingly, the amount of RARγ phosphorylated at S79 (S68 in RARγ2) increased rapidly after RA addition (Fig. 5*C*). Consequently, the amount of vinexinβ bound to RARγ decreased in the same time slot (Fig. 5*B*, lanes 2–5). Taken together, these results indicate that the interaction of vinexinβ with RARγ is disrupted by phosphorylation of the RARγ partner both *in vitro* and *in vivo*.

Vinexinβ is a repressor that does not bind RARγ at the promoters of RA target genes

Vinexinβ has been shown previously to inhibit the transcription of RARγ target genes (10). Therefore, we aimed to investigate the mechanism of this repressive activity. In F9 cells, several genes such as the *RARβ2* gene are rapidly induced in response to RA (1). In F9 cells reexpressing in a RARγ-null background, RARγ with the serine residues of the PRD substituted by alanines (RARγm), the expression of the *RARβ2* gene was less efficient than that in F9 cells reexpressing RARγWT (24) (**Fig. 6*A***, lane 4). Therefore, we hypothesized that this decrease might be due to the interaction of vinexinβ with the nonphosphorylated form of RARγ at the promoter of the *RARβ2* gene, according to the classic model of corepressors.

ChIP experiments were performed with F9 cells expressing RARγWT or RARγm to assess the vinexinβ and RARγ occupancy of the *RARβ2* gene promoter, which contains one DR5 RARE at position −60 (15). Antibodies directed against RARγ or vinexinβ were used to immunoprecipitate RARγ- or vinexinβ-bound DNA fragments that were further analyzed by quantitative PCR using specific pairs of primers spanning the RARE. The specificity of our experimental conditions

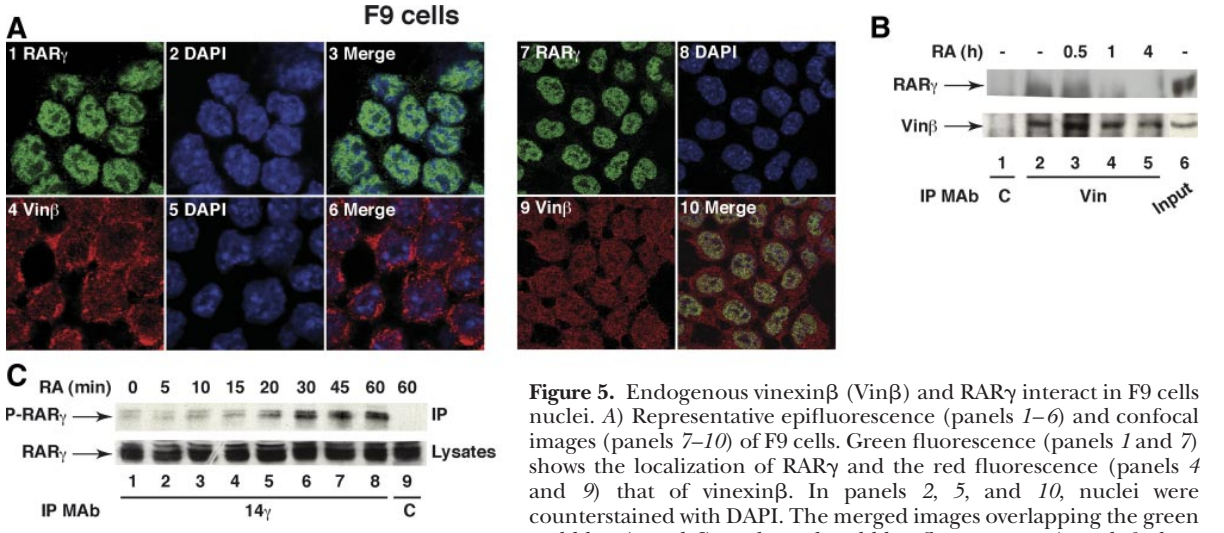


Figure 5. Endogenous vinexinβ (Vinβ) and RARγ interact in F9 cells nuclei. *A*) Representative epifluorescence (panels *1–6*) and confocal images (panels *7–10*) of F9 cells. Green fluorescence (panels *1* and *7*) shows the localization of RARγ and the red fluorescence (panels *4* and *9*) that of vinexinβ. In panels *2*, *5*, and *10*, nuclei were counterstained with DAPI. The merged images overlapping the green and blue (panel *3*) or the red and blue fluorescence (panel *6*) show the localization of RARγ and vinexinβ, respectively, in nuclei. In panel *10*, the merged image overlapping the red, green, and blue fluorescence shows the colocalization of RARγ and vinexinβ in nuclei. *B*) Coimunoprecipitation (IP) experiments performed with F9 cell nuclear extracts showing that vinexinβ interacts with RARγ and dissociates subsequently to RA addition. *C*) In F9 cells, RA induces the rapid phosphorylation of RARγ as shown by immunoprecipitation with a monoclonal antibody (MAb) recognizing specifically RARγ phosphorylated at S79.

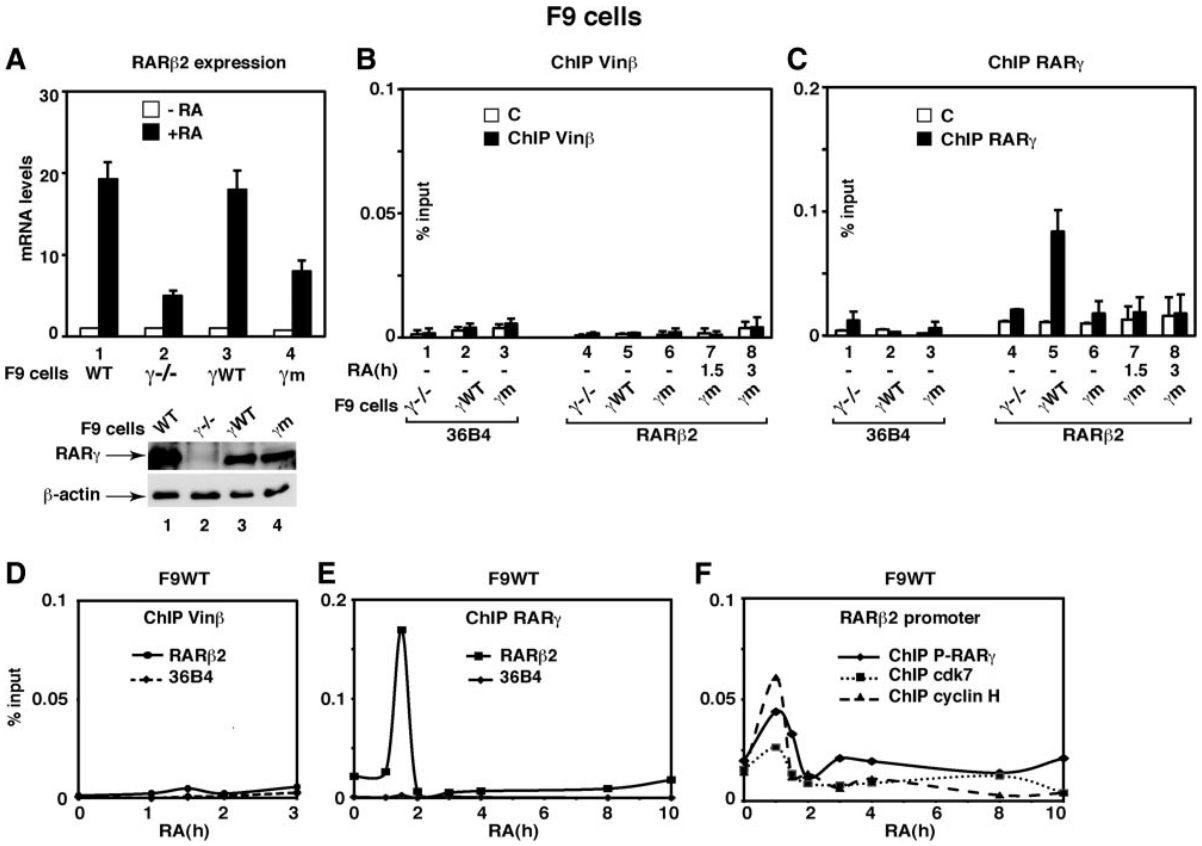


Figure 6. Vinexinβ (Vinβ) does not occupy the promoter of RA target genes. *A*) F9 cells (WT, RARγ$^{-/-}$, or reexpressing in a RARγ$^{-/-}$ background RARγWT or RARγm) were RA-treated for 8 h, and transcripts for *RARβ2* were analyzed by quantitative RT-PCR. Results are means ± SE of 3 independent experiments and correspond to the fold induction relative to the amount of transcripts present in vehicle-treated cells. Expression of RARγ in the different F9 cell lines was checked by immunoblotting. *B, C*) ChIP analysis of the basal (in the absence of RA) occupancy of the *RARβ2* gene promoter by vinexinβ (*B*) and RARγ (*C*) in the different F9 cell lines. ChIP experiments performed with RA-treated F9 cells expressing RARγm are also shown. Specificity of the experimental conditions was checked in the absence of antibodies and with the promoter of the control *36B4* gene, which does not contain any RARE. Values (% of inputs) are means ± SE of triplicate experiments performed on 3 separate chromatin preparations. *D–F*) Kinetic ChIP experiments performed with F9WT cells determining the recruitment of vinexinβ (*D*), RARγ, (*E*) and phosphorylated RARγ, cdk7, and cyclin H to the RARβ2 promoter (*F*).

was checked in the absence of antibodies, with the promoter of the control *36B4* gene, which does not contain any RARE, and with RARγ$^{-/-}$ cells.

Unexpectedly, in F9 cells expressing RARγm, vinexinβ was hardly detected at the RARβ2 promoter region, either in the absence or presence of RA (Fig. 6*B*, lanes 6–8). Then we investigated whether RARγm occupied the promoter. Most important, RARγm was also hardly detected at the promoter in the absence or in the presence of RA (Fig. 6*C*, lanes 6–8). Such results indicate that in F9 cells the repression of transcription observed on mutation of the RARγ phosphorylation sites does not reflect the binding of vinexinβ to RARγ at the promoter.

Interestingly, vinexinβ could not be detected either at the RARβ2 promoter in F9 cells expressing RARγWT (Fig. 6*B*, lane 5; *D*). However, in these cells, RARγWT significantly occupied the promoter region in the absence of RA (Fig. 6*C*, lane 5) in line with other studies (25, 26), with a rapid and transient enrichment at 1.5 h after RA addition (Fig. 6*E*). Additional ChIP experiments performed with our antibodies recognizing specifically the phosphorylated form of RARγ indicated that DNA-bound RARγ (either in the absence or the presence of RA) was phosphorylated (Fig. 6*F*). Moreover, the cdk7 and cyclin H subunits of TFIIH, which are involved in RAR phosphorylation (15), were also recruited to the promoter concomitantly with phosphorylated RARγ (Fig. 6*F*). Taken together, these results indicate that the recruitment of RARγ at the promoter is controlled by phosphorylation of the PRD, as previously reported for RARα (15). Because the phosphorylated form of RARγ is unable to interact with vinexinβ, they are in agreement with the absence of vinexinβ at the promoter, even in the absence of RA.

Vinexinβ sequesters RARγ in the nucleoplasm, out of chromatin

To further investigate how vinexinβ represses RARγ-mediated transcription, additional ChIP experiments

were performed with transfected COS-1 cells, in which overexpression of vinexinβ inhibits the RA-induced expression of a CAT reporter gene under the control of the natural mRARβ2 promoter (10). In COS-1 cells, endogenous vinexinβ was hardly detected at the promoter, corroborating the results obtained with F9 cells, but no enrichment was observed on vinexinβ overexpression either in the absence (**Fig. 7*A***, lanes 6 and 7) or in the presence of RA (Fig. 7*B*). This pattern was not significantly affected on overexpression of RARγWT or S79A (Fig. 7*A*, lanes 8 and 9), which did and did not occupy the promoter, respectively (Figs. 7*A*, lanes 2 and 4, and 1*C*, lanes 1–3).

However, on vinexinβ overexpression, the RARγWT occupancy of the promoter observed in the absence of RA was decreased (Fig. 7*A*, compare lanes 3 and 5). The RA-induced recruitment of RARγ was also decreased (Fig. 7*C*, lanes 4–6). In contrast, vinexinβ with the SH3-3 domain deleted had no effect (Fig. 7*C*, lanes 7–9). Collectively these results confirm that vinexinβ does not occupy the promoters of RARγ target genes and does not interact with DNA-bound RARγ. They also suggest that vinexinβ impedes RARγ recruitment to the promoters. Therefore, we aimed to determine whether vinexinβ would affect the ability of RARγ/RXRα heterodimers to bind DNA. EMSAs were performed with RARγWT expressed in COS-1 cells, a biotin-labeled double-stranded oligonucleotide containing a DR5 RARE and increasing concentrations of vinexinβ. Such assays demonstrated that the amount of bound RARγ decreased in the presence of an excess of vinexinβ (Fig. 7*D*), corroborating the inhibitory role of vinexinβ.

Then we investigated whether vinexinβ could be detected in chromatin. Highly purified and intact nuclei were prepared from F9 cells and chromatin was separated from nucleoplasm. As shown in Fig. 7*E*, vinexinβ was detected only in the nucleoplasm and not in chromatin, whereas RARγ was present in both. This result suggests that vinexinβ would repress transcription *via* the sequestration of RARγ in the nucleoplasm, out of chromatin and of the promoters of RA target genes.

DISCUSSION

In this study, we outlined a new paradigm in which, in the absence of RA, the nonphosphorylated form of RARγ is associated with vinexinβ, which sequesters RARγ out of the promoters of target genes. Then in response to RA, RARγ becomes phosphorylated at a serine residue close to the interaction motif, leading to the dissociation of vinexinβ and finally to promoter recruitment of RARγ. To our knowledge, this is the first

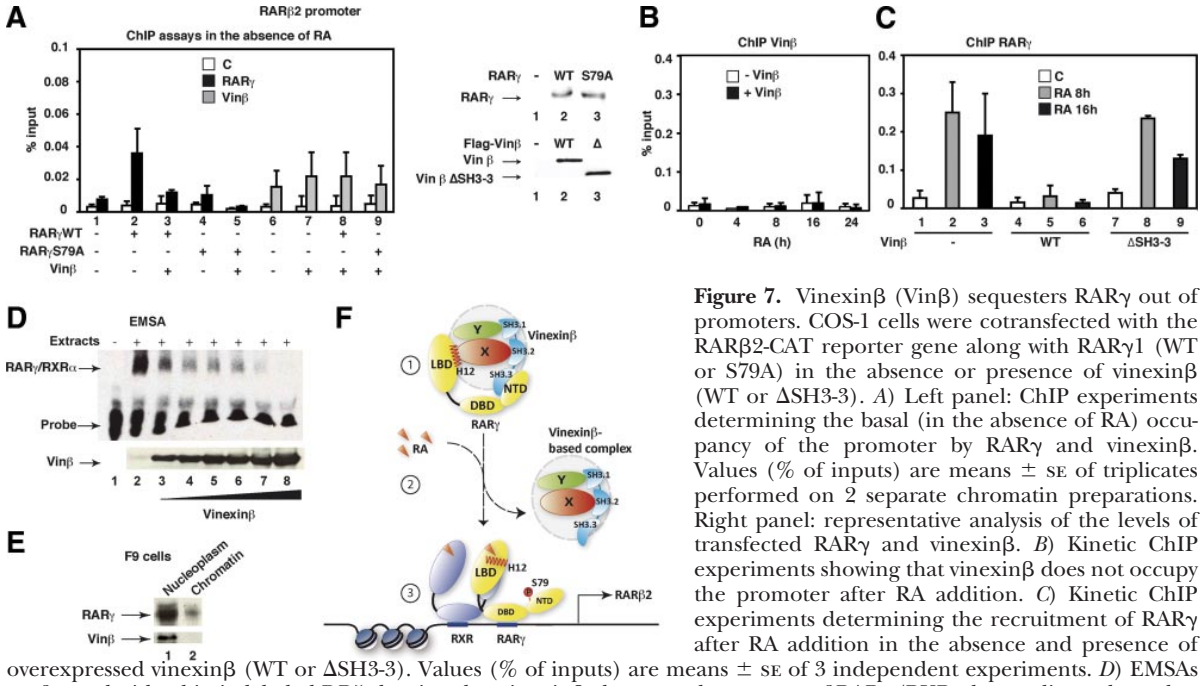


Figure 7. Vinexinβ (Vinβ) sequesters RARγ out of promoters. COS-1 cells were cotransfected with the RARβ2-CAT reporter gene along with RARγ1 (WT or S79A) in the absence or presence of vinexinβ (WT or ΔSH3-3). *A*) Left panel: ChIP experiments determining the basal (in the absence of RA) occupancy of the promoter by RARγ and vinexinβ. Values (% of inputs) are means ± SE of triplicates performed on 2 separate chromatin preparations. Right panel: representative analysis of the levels of transfected RARγ and vinexinβ. *B*) Kinetic ChIP experiments showing that vinexinβ does not occupy the promoter after RA addition. *C*) Kinetic ChIP experiments determining the recruitment of RARγ after RA addition in the absence and presence of overexpressed vinexinβ (WT or ΔSH3-3). Values (% of inputs) are means ± SE of 3 independent experiments. *D*) EMSAs performed with a biotin-labeled DR5 showing that vinexinβ abrogates the amount of RARγ/RXRα heterodimers bound to DNA. Vinexinβ overexpression was checked by immunoblotting. Data correspond to a representative experiment among 3. *E*) Nuclei were isolated from F9 cells, and nucleoplasm was separated from chromatin. The presence of RARγ and vinexinβ was analyzed by immunoblotting. *F*) Model for the role of vinexinβ in the control of RARγ transcriptional activity. *1*) In the absence of RA, the third C-terminal domain of vinexinβ interacts with the nonphosphorylated N-terminal PRD of RARγ. Through its two other SH3 domains, vinexinβ might also act as a scaffold linking the NTD to the LBD of RARγ *via* protein complexes (exemplified as X and Y). Such a complex sequesters RARγ out of promoters, thereby impeding transcription. *2*) After RA addition, phosphorylation of the RARγ PRD and the conformational changes of the LBD cooperate to induce the dissociation of the vinexinβ-based complex. *3*) Once separated from vinexinβ, phosphorylated RARγ can dimerize with RXR and occupy the promoters of target genes to initiate the transcription of RA-target genes such as the *RARβ2* gene.

report of a RAR corepressor association/dissociation regulated by phosphorylation processes and occurring out of promoters. It was an unexpected finding and sheds light on a new level of complexity in the fine-tuning of RARγ-mediated transcription.

Vinexinβ is a protein with 3 SH3 domains. The unique property of such domains is to recognize PRDs without requiring high affinities (20). Such a PRD is present in the NTD of the RARγ subtype and contains phosphorylation sites. Here we combined classic molecular and cellular approaches with biophysical techniques to decipher the parameters of the interaction. We have shown that the third C-terminal domain of vinexinβ (SH3-3) interacts with a consensus PxxPxR motif (a classic class II ligand) located in the PRD of RARγ. Interestingly, although well conserved between RARs, the PRD of the RARα subtype differs from that of RARγ at one residue within the PxxPxR motif. Such a discrepancy might be the basis of the specificity of the interaction of vinexinβ with RARγ. Interestingly, this interaction fulfills the classic criteria of interactions between PRDs and their binding partners. Indeed, the SH3-3 domain of vinexinβ interacts with the target PRD with typical K_d values of ~40–80 μM, in line with the idea that rather weak interactions are required for efficient signaling processes.

Most important, we found that phosphorylation of RARγ at a serine residue (S79) adjacent to the PxxPxR motif significantly reduces the binding of vinexinβ. Moreover, *in vivo* the amount of RARγ phosphorylated at S79 increases rapidly in response to RA, and subsequently vinexinβ is released. Given that S79 is located in close proximity to the SH3-3 binding surface, it can be proposed that phosphorylation of this residue propagates a signal to this surface, promoting the dissociation of vinexinβ.

An important clue in the present study is that vinexinβ represses RARγ-mediated transcription, in line with its ability to interact with the nonphosphorylated form of RARγ, which is transcriptionally deficient. The first hypothesis was that this repression might result from the binding of vinexinβ to RARγ at the promoter of target genes, according to the model of classic corepressors. Unexpectedly, no significant amounts of vinexinβ could be detected at promoters, either in absence or presence of RA. Moreover, vinexinβ was found to inhibit RARγ binding to DNA. Finally, taking advantage of F9 cells expressing RARγWT or mutated at the phosphorylation sites and of antibodies recognizing specifically the phosphorylated form of RARγ, DNA-bound RARγ was phosphorylated. Such data are indeed not compatible with the presence of vinexinβ at the promoter and lead to the conclusion that the repressive activity of vinexinβ does not occur through its binding to RARγ at the promoters of target genes and thus differs from that of the classic corepressors.

In line with this result, we found that, in contrast to RARγ, vinexinβ could not be detected in chromatin either in the absence or the presence of RA. In fact, vinexinβ colocalized with RARγ in the nucleoplasm. This finding suggests that vinexinβ would sequester the nonphosphorylated form of RARγ in an inactive state out of chromatin. The release of vinexinβ subsequent to RA binding and to the phosphorylation of the PRD would increase the amount of RARγ present in chromatin and competent to bind DNA and to activate the expression of the target genes. Accordingly, the RA-induced increase in the RARγ occupancy of promoters occurred in the same time slot (1 h) as the phosphorylation of RARγ and the release of vinexinβ. Thus, vinexinβ would act as a "sensor" controlling the amount of transcriptionally active RARγ. Moreover, phosphorylation of the PRD, which is in close proximity to the DNA-binding domain (DBD) appears to play an important role in the recruitment of RARγ to DNA, as described previously for RARα (15). Given the close proximity of the phosphorylation site to the DBD, one cannot exclude the possibility that phosphorylation increases the affinity of RARγ for RAREs by an allosteric effect that has yet to be understood. Experiments are in progress to address this hypothesis.

Then the question is, how does vinexinβ sequester RARγ in the nucleoplasm? Interestingly, vinexinβ is a scaffolding protein with 3 SH3 domains, which have been shown to interact with several proteins such as actin and several actin-binding proteins (22, 23), which have an increasingly important role in the regulation of transcription, some enhancing and others repressing the transcriptional activity of nuclear receptors (13). Moreover, a direct interaction between vinexinβ and SAFB2 a novel nuclear receptor corepressor, has been described (27). All of these proteins might be interesting candidates for maintaining RARγ in an inactive state. Because, according to our results, vinexinβ also seems to interact weakly with the LBD of RARγ, one can propose that vinexinβ and RARγ belong to a multiprotein complex in which vinexinβ interacts directly with the NTD of RARγ through its third SH3 domain and indirectly with the LBD of RARγ *via* its two other SH3 domains and intermediary proteins (see model in Fig. 7*F*). Such a hypothesis is in agreement with the weak affinity of the direct interaction between RARγ and the SH3-3 domain of vinexinβ and with our observation that *in vivo* the dissociation of vinexinβ that occurs in response to RA requires not only the phosphorylation of RARγ but also additional rapid RA-induced events that may be either conformational changes or additional phosphorylation processes affecting the RARγ LBD and/or protein partners.

A final supporting result for the proposed model is the fact that helix 12 of RARγ has been shown to adopt, even in the absence of ligand, a constitutively closed conformation that approximates the conformation of liganded RARα but is not compatible with the binding of the classic corepressors and coactivators (28–30). Thus, vinexinβ-based complexes might be an alternative for controlling negatively the transcriptional activity of RARγ. The characterization of the proteins associated with the RARγ/vinexinβ complex is in progress and will shed light on the mechanism of regulation of RARγ transcriptional activity. FJ

The authors thank Dr. N. Kioka (Kyoto University, Kyoto, Japan) for the gift of the vinexinβ vector, M. Oulad Abdelghani (IGBMC) for the mouse monoclonal antibodies, and P. Eberling (IGBMC) for the synthetic peptides. The authors are

also grateful to D. Altschuh and the Biacore platform (UMR 7100 Illkirch). The authors warmly acknowledge Régis Lutzing, Julie Goepp, Hussein Raad, Xiaojun Gao, Anne Chapelle, and members of the cell culture facilities for help. We give special thanks to Christine Ferry and Ali Hamiche for fruitful discussions and suggestions. This work was supported by funds from Centre National de la Recherche Scientifique, INSERM, the Agence Nationale pour la Recherche (ANR-05-BLAN-0390-02 and ANR-09-BLAN-0297-01), the Association pour la Recherche sur le Cancer (ARC-07-1-3169), the Fondation pour la Recherche Médicale (FRM; DEQ 20090515423), and the Institut National du Cancer (INCa-PL06-095 and PL07-96099). S.L. was supported by the Ministère de l'Enseignement Supérieur et de la Recherche and ARC.

REFERENCES

1. Bour, G., Taneja, R., and Rochette-Egly, C. (2006) Mouse embryocarcinoma F9 cells and retinoic acid. A model to study the molecular mechanisms of endodermal differentiation. In *Nuclear Receptors in Development*, Vol. 16 (Taneja, R., ed) pp. 211–253, Elsevier, New York
2. Germain, P., Chambon, P., Eichele, G., Evans, R. M., Lazar, M. A., Leid, M., De Lera, A. R., Lotan, R., Mangelsdorf, D. J., and Gronemeyer, H. (2006) International Union of Pharmacology. LXIII. Retinoid X receptors. *Pharmacol. Rev.* **58,** 760–772
3. Germain, P., Chambon, P., Eichele, G., Evans, R. M., Lazar, M. A., Leid, M., De Lera, A. R., Lotan, R., Mangelsdorf, D. J., and Gronemeyer, H. (2006) International Union of Pharmacology. LX. Retinoic acid receptors. *Pharmacol. Rev.* **58,** 712–725
4. Rochette-Egly, C., and Germain, P. (2009) Dynamic and combinatorial control of gene expression by nuclear retinoic acid receptors. *Nucl. Recept. Signal.* **7,** e005
5. Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6,** 197–208
6. Lavery, D. N., and McEwan, I. J. (2005) Structure and function of steroid receptor AF1 transactivation domains: induction of active conformations. *Biochem. J.* **391,** 449–464
7. Liu, J., Perumal, N. B., Oldfield, C. J., Su, E. W., Uversky, V. N., and Dunker, A. K. (2006) Intrinsic disorder in transcription factors. *Biochemistry* **45,** 6873–6888
8. Bastien, J., Adam-Stitah, S., Riedl, T., Egly, J. M., Chambon, P., and Rochette-Egly, C. (2000) TFIIH interacts with the retinoic acid receptor γ and phosphorylates its AF-1-activating domain through cdk7. *J. Biol. Chem.* **275,** 21896–21904
9. Gianni, M., Bauer, A., Garattini, E., Chambon, P., and Rochette-Egly, C. (2002) Phosphorylation by p38MAPK and recruitment of SUG-1 are required for RA-induced RARγ degradation and transactivation. *EMBO J.* **21,** 3760–3769
10. Bour, G., Plassat, J. L., Bauer, A., Lalevee, S., and Rochette-Egly, C. (2005) Vinexin β interacts with the non-phosphorylated AF-1 domain of retinoid receptor γ (RARγ) and represses RARγ-mediated transcription. *J. Biol. Chem.* **280,** 17027–17037
11. Kioka, N., Ueda, K., and Amachi, T. (2002) Vinexin, CAP/ponsin, ArgBP2: a novel adaptor protein family regulating cytoskeletal organization and signal transduction. *Cell Struct. Funct.* **27,** 1–7
12. Zhang, J., Li, X., Yao, B., Shen, W., Sun, H., Xu, C., Wu, J., and Shi, Y. (2007) Solution structure of the first SH3 domain of human vinexin and its interaction with vinculin peptides. *Biochem. Biophys. Res. Commun.* **357,** 931–937
13. Zheng, B., Han, M., Bernier, M., and Wen, J. K. (2009) Nuclear actin and actin-binding proteins in the regulation of transcription and gene expression. *FEBS J.* **276,** 2669–2685
14. Plassat, J., Penna, L., Chambon, P., and Rochette-Egly, C. (2000) The conserved amphipatic α-helical core motif of RARγ and RARα activating domains is indispensable for RA-induced differentiation of F9 cells. *J. Cell Sci.* **113,** 2887–2895
15. Bruck, N., Vitoux, D., Ferry, C., Duong, V., Bauer, A., de The, H., and Rochette-Egly, C. (2009) A coordinated phosphorylation cascade initiated by p38MAPK/MSK1 directs RARα to target promoters. *EMBO J.* **28,** 34–47
16. Williamson, R. A., Carr, M. D., Frenkiel, T. A., Feeney, J., and Freedman, R. B. (1997) Mapping the binding site for matrix metalloproteinase on the N-terminal domain of the tissue inhibitor of metalloproteinases-2 by NMR chemical shift perturbation. *Biochemistry* **36,** 13882–13889
17. Altschuh, D., Oncul, S., and Demchenko, A. P. (2006) Fluorescence sensing of intermolecular interactions and development of direct molecular biosensors. *J. Mol. Recognit.* **19,** 459–477
18. Ouararhni, K., Hadj-Slimane, R., Ait-Si-Ali, S., Robin, P., Mietton, F., Harel-Bellan, A., Dimitrov, S., and Hamiche, A. (2006) The histone variant mH2A1.1 interferes with transcription by down-regulating PARP-1 enzymatic activity. *Genes Dev.* **20,** 3324–3336
19. Gaillard, E., Bruck, N., Brelivet, Y., Bour, G., Lalevee, S., Bauer, A., Poch, O., Moras, D., and Rochette-Egly, C. (2006) Phosphorylation by protein kinase A potentiates retinoic acid receptor α activity by means of increasing interaction with and phosphorylation by cyclin H/cdk7. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 9548–9553
20. Ball, L. J., Kuhne, R., Schneider-Mergener, J., and Oschkinat, H. (2005) Recognition of proline-rich motifs by protein-protein-interaction domains. *Angew. Chem. Int. Ed. Engl.* **44,** 2852–2869
21. Kay, B. K., Williamson, M. P., and Sudol, M. (2000) The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* **14,** 231–241
22. Mitsushima, M., Sezaki, T., Akahane, R., Ueda, K., Suetsugu, S., Takenawa, T., and Kioka, N. (2006) Protein kinase A-dependent increase in WAVE2 expression induced by the focal adhesion protein vinexin. *Genes Cells* **11,** 281–292
23. Mitsushima, M., Takahashi, H., Shishido, T., Ueda, K., and Kioka, N. (2006) Abl kinase interacts with and phosphorylates vinexin. *FEBS Lett.* **580,** 4288–4295
24. Taneja, R., Rochette-Egly, C., Plassat, J. L., Penna, L., Gaub, M. P., and Chambon, P. (1997) Phosphorylation of activation functions AF-1 and AF-2 of RARα and RARγ is indispensable for differentiation of F9 cells upon retinoic acid and cAMP treatment. *EMBO J.* **16,** 6452–6465
25. Gillespie, R. F., and Gudas, L. J. (2007) Retinoic acid receptor isotype specificity in F9 teratocarcinoma stem cells results from the differential recruitment of coregulators to retinoic response elements. *J. Biol. Chem.* **282,** 33421–33434
26. Gillespie, R. F., and Gudas, L. J. (2007) Retinoid regulated association of transcriptional co-regulators and the polycomb group protein SUZ12 with the retinoic acid response elements of *Hoxa1*, *RARβ*$_2$, and *Cyp26A1* in F9 embryonal carcinoma cells. *J. Mol. Biol.* **372,** 298–316
27. Townson, S. M., Dobrzycka, K. M., Lee, A. V., Air, M., Deng, W., Kang, K., Jiang, S., Kioka, N., Michaelis, K., and Oesterreich, S. (2003) SAFB2, a new scaffold attachment factor homolog and estrogen receptor corepressor. *J. Biol. Chem.* **278,** 20059–20068
28. Farboud, B., Hauksdottir, H., Wu, Y., and Privalsky, M. L. (2003) Isotype-restricted corepressor recruitment: a constitutively closed helix 12 conformation in retinoic acid receptors β and γ interferes with corepressor recruitment and prevents transcriptional repression. *Mol. Cell. Biol.* **23,** 2844–2858
29. Privalsky, M. L. (2004) The role of corepressors in transcriptional regulation by nuclear hormone receptors. *Annu. Rev. Physiol.* **66,** 315–360
30. Hauksdottir, H., Farboud, B., and Privalsky, M. L. (2003) Retinoic acid receptors β and γ do not repress, but instead activate target gene transcription in both the absence and presence of hormone ligand. *Mol. Endocrinol.* **17,** 373–385

Received for publication March 29, 2010.
Accepted for publication July 1, 2010.

Supplemental Material can be found at:
http://www.jbc.org/content/suppl/2011/07/29/M111.263681.DC1.html

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 38, pp. 33322–33334, September 23, 2011

Genome-wide *in Silico* Identification of New Conserved and Functional Retinoic Acid Receptor Response Elements (Direct Repeats Separated by 5 bp)*[S]

Received for publication, May 22, 2011, and in revised form, July 28, 2011 Published, JBC Papers in Press, July 29, 2011, DOI 10.1074/jbc.M111.263681

Sébastien Lalevée[‡¶1,2], Yannick N. Anno[§2], Amandine Chatagnon[¶‖], Eric Samarut[‡], Olivier Poch[§], Vincent Laudet[**], Gerard Benoit[¶‖], Odile Lecompte[§], and Cécile Rochette-Egly[‡3]

*From the [‡]Department of Functional Genomics and Cancer and [§]Department of Integrated Structural Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire, INSERM U596 and CNRS UMR7104, Université de Strasbourg, F_67404 Illkirch Cedex, [¶]CNRS UMR5534, F-69622 Villeurbanne, [‖]Université Lyon 1, UMR5534, F-69622 Villeurbanne, and **Institut de Génomique Fonctionelle de Lyon, UMR 5242, Institut National de la Recherche Agronomique, Université de Lyon, Ecole Normale Supérieure de Lyon, 69364 Lyon Cedex 07, France*

Background: Retinoic acid (RA) receptors regulate gene expression through binding-specific response elements (RAREs).
Results: A collection of new DR5 RAREs located ±10 kb from TSSs and conserved among 6 vertebrates species or more has been amassed.
Conclusion: We provide a wider knowledge base for analyzing RA target genes.
Significance: The RA response of the conserved target genes differs between species and tissues.

The nuclear retinoic acid receptors interact with specific retinoic acid (RA) response elements (RAREs) located in the promoters of target genes to orchestrate transcriptional networks involved in cell growth and differentiation. Here we describe a genome-wide *in silico* analysis of consensus DR5 RAREs based on the recurrent RGKTSA motifs. More than 15,000 DR5 RAREs were identified and analyzed for their localization and conservation in vertebrates. We selected 138 elements located ±10 kb from transcription start sites and gene ends and conserved across more than 6 species. We also validated the functionality of these RAREs by analyzing their ability to bind retinoic acid receptors (ChIP sequencing experiments) as well as the RA regulation of the corresponding genes (RNA sequencing and quantitative real time PCR experiments). Such a strategy provided a global set of high confidence RAREs expanding the known experimentally validated RAREs repertoire associated to a series of new genes involved in cell signaling, development, and tumor suppression. Finally, the present work provides a valuable knowledge base for the analysis of a wider range of RA-target genes in different species.

Retinoic acid (RA)[4] is an active derivative of vitamin A that influences a range of essential biological processes such as development and homeostasis (1–4). RA exerts its action through nuclear RA receptors (RARs), which are typical ligand-dependent regulators of transcription with a central DNA binding domain linked to a ligand binding domain (for review, see Refs. 5 and 6). In response to RA signaling, RARs heterodimerize with retinoid X receptors (RXRs) and occupy characteristic RA response elements (RAREs) located in the promoter of target genes involved in cell proliferation and differentiation. RXR/RAR heterodimer occupancy at cognate response elements is commonly a determinant of transcriptional responsiveness. Within a given cell type, binding of RXR/RAR heterodimers to RAREs can either up- or down-regulate transcription in a gene-specific manner. RAREs are composed of two direct repeats of a core hexameric motif (A/G)G(G/T)TCA. The classical RARE is a 5-bp-spaced direct repeat (referred to as a DR5), but RXR/RAR heterodimers can also bind to direct repeats separated by 2 bp (DR2) or 1 bp (DR1) (6, 7).

The development of high throughput technologies such as DNA microarrays revealed that within a given cell type or tissue, the RA response is composed of a huge and complex network of responsive genes (8–10). However, such techniques could not discriminate between direct primary and secondary target genes (which are modulated by the product of a primary target gene rather than by RXR/RAR heterodimers), and only a few of the RA target genes contained identified RAREs. More recently, chromatin immunoprecipitation coupled with array hybridization (ChIP-chip) allowed the identification of new RAR binding loci (11, 12). However, whether such loci bind RARs directly or indirectly through other bound factors could not be easily discriminated. Moreover, the identified loci do not correspond to the full repertoire, as the arrays do not represent all possible regions in a genome. The nascent genome-wide ChIP-seq (chromatin immunoprecipitation coupled with deep sequencing) technology should expand the repertoire of potential high affinity response elements (13, 14). Nevertheless, although powerful, such ChIP-based approaches are highly cell context-specific.

* This work was supported by funds from CNRS, INSERM, the Association pour la Recherche sur le Cancer (ARC 3169), the Agence Nationale pour la Recherche (ANR-05-BLAN-0390-02 and ANR-09-BLAN-0127-01), the Fondation pour la Recherche Médicale (DEQ20090515423), and the Institut National du Cancer (INCa-PLO7-96099 and PL09-194).
[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables S1–S5.
[1] Supported by the Association pour la Recherche sur le Cancer (ARC). Present address: Laboratory for Prenatal Medicine, Dept. of Biomedicine, University Hospital Basel, 4031 Basel, Switzerland.
[2] Supported by the Ministère de l'Enseignement Supérieur et de la Recherche.
[3] To whom correspondence should be addressed: 1 rue Laurent Fries, BP 10142, 67404 Illkirch Cedex, France. Tel.: 33-3-88-65-34-59; Fax: 33-3-88-65-32-01; E-mail: cegly@igbmc.fr.
[4] The abbreviations used are: RA, retinoic acid; RAR, RA receptor; RXR, retinoid X receptor; DR, direct repeat; RARE, RA response element; TSS, transcription start site; qPCR, quantitative PCR; seq, sequencing.

Now, with the availability of an increasing number of genome sequences, *in silico* analysis of RAREs can be also performed. The advantage of computational techniques is that it overcomes the chromatin structure and, thus, the cellular context and provides a direct glance on the whole repertoire of possible RAREs.

Here we conducted a genome-wide *in silico* study of RA response elements. Although RXR-RAR heterodimers can bind to DR5, DR2, or DR1 response elements, the significance and the specificity of the DR2 and DR1 is still unclear. Therefore, we focused on DR5 RAREs. Computational techniques were developed for the genome-wide identification of DR5 RAREs and for the characterization of their genomic and phylogenetic context. In this way we amassed a collection of DR5 RAREs that is conserved across vertebrate species and that was validated for its occupancy and functionally analyzed for the RA-responsiveness of the associated genes. Such a strategy allowed us to characterize a new set of high confidence conserved DR5 RAREs associated to a series of new potential RA-target genes, thus providing a wider knowledge base for the analysis of the RA response in different species.

EXPERIMENTAL PROCEDURES

Bioinformatics—In silico analyses were performed using the Genomic Context data base (GeCo).[5] This data warehouse, which was already exploited in a genome-wide study of the Staf transcription factor binding sites (15), aggregates genomic, phylogenetic, and epigenomic data from different sources, allowing the high-throughput contextual characterization of a given set of genetic elements. The underlying data base of annotated genes was built by computing refGene (proteins), rnaGene (snRNA, snoRNA, tRNA, rRNA, scaRNA), kgXref tables from the University of Santa Cruz California, mirna, mirna_literature_references, mirna_mature, and literature_references tables from the Sanger Institute and piRNA file from the piRNA Database. The data base also includes sequence conservation data extracted from the University of Santa Cruz California blastZ alignments. The data base is implemented in high speed DB2 architecture called Biological Integration and Retrieval of Data (BIRD), which can quickly address the whole set of sequences, genomic features, and alignments (15). RGKTSA DR5 motifs were searched in the human (NCBI build 36, hg18)- and mouse (NCBI build 37, mm9)-masked genomes (RepeatMasker) using an in-house tool dedicated to the automatic search of short motifs and implemented in the GeCo system. The obtained motifs were subsequently characterized. For each motif, we retrieved the nearest gene and its localization as well as, if applicable, the position of the motif regarding the gene elements, exon, intron, transcription start site (TSS), and gene end (end of the last exon). Motif conservation was then analyzed on the basis of University of Santa Cruz California blastZ alignments between the human or mouse and 13 other vertebrate genomic sequences. The considered species were selected for the confidence of their sequencing, the quality of their annotation, and for their repartition through the vertebrate phylogenetic tree: zebrafish (danRer5), fugu (fr2), xenopus (xenTro2), lizard (anoCar1), chicken (galGal3), platypus (ornAna1), opossum (monDom4), dog (canFam2), horse (equCab1), cow (bosTau4), rat (rn4), rhesus (rheMac2), and chimpanzee (panTro2). We considered a motif as conserved in a given species if the region encompassing the motif in human or mouse is aligned with a genomic region of the species also containing a RGKTSA DR5 motif.

Cell Culture, RNA Extraction, and qRT-PCR—F9 and P19 mouse embryocarcinoma cells, human MCF7 cells, and zebrafish PAC2 cells were cultured according to standard conditions as previously described (16–19). RNAs were extracted and subjected to qRT-PCR as previously described (20). Transcripts were normalized according to the ribosomal protein gene *RPLP0*. All mouse primers are listed in supplemental Table S1. The others are available upon request.

RNA Sequencing—After isolation of total RNA, a library of template molecules suitable for high throughput DNA sequencing was created according to the instructions of Illumina. Briefly, the poly(A)-containing mRNAs were isolated from total RNA (4 mg) by two runs of purification on Sera-Mag Oligo-dT Beads (Thermoscientific) and fragmented using divalent cations and heat-catalyzed hydrolysis. Fragmented mRNAs were used as a template to synthesize single-stranded cDNA with Superscript II reverse transcriptase and random primers. After second-strand synthesis, the cDNAs went through end-repair and ligation reactions using paired-end adapter oligos from Illumina and were electrophoresed on an agarose gel. A slice containing fragments in the 300-bp range was excised, and after elution and purification, the library was amplified with 15 cycles of PCR with Illumina sequencing primers and purified using Agencourt AMPure XP beads from Beckman.

The library was then used to build clusters on the Illumina flow cell according to protocol. Image analysis and base calling was performed using the Illumina pipeline. Reads were then mapped onto the mm9/NCBI37 assembly of the mouse genome using Tophat (21). Quantification of gene expression was done using Cufflinks (22) and annotations from Ensembl release 57. For each transcript the number of FPKM (fragments/kb of transcript/million fragments mapped) was converted into raw read counts, which were added for each gene locus by using an R script that we implemented. Then data normalization and identification of significantly differentially expressed genes were performed with the method proposed by Anders and Huber (23) and implemented into the DESeq Bioconductor package. The final *p* values were adjusted for multiple testing according to the method proposed by Benjamini and Hochberg (24), and a cutoff *p* value of 0.05 was applied for finding significant responsive genes.

RESULTS

Bioinformatic Genome-wide Research of DR5 RAREs Corresponding to the RGKTSA Motif—Only a few RAREs have been identified to date and associated to RA-target genes. Most of them are represented by two direct repeats of the hexameric motif (A/G)G(G/T)TCA, separated by five nucleotides (DR5) (25, 26). Such DR5s have been found in the promoters of human and mouse genes involved in RA metabolism (*Cyp26A1*) (27), in

[5] Y. N. Anno, O. Poch, and O. Lecompte, manuscript in preparation.

Conserved and Functional DR5 RAREs

RA signaling (*RARα2*, *RARβ2*, *RARγ2*) (28–30), or in development (*Hoxa1*, *Hoxa4*, *Hoxb1*) (31–33). Alignment of these RAREs (Fig. 1) clearly delineates a recurrent motif RGKTSA (coding is according to the IUPAC convention: R = AG; K = GT; S = CG), which differs from the classical consensus motif RGKTCA at position 5, with a G instead of a C (in *RARγ2* and *Hoxa4*). Therefore, with the aim of identifying novel RA-driven primary target genes, we screened the masked human and mouse genomes for DR5 corresponding to two direct repeats of the RGKTSA motif at the genome-wide scale (see "Experimental Procedures"). Such *in silico* screens have the potential of identifying target genes independently of their tissue of expression. We identified 15,925 DR5s corresponding to two direct repeats of the RGKTSA motif in the mouse genome and 14,571 in the human genome (supplemental Tables S2 and S3).

Conservation of the DR5 RAREs during Evolution—A way to assess the potential relevance of response elements is to determine whether they are conserved between species (phylogenetic footprinting). Indeed, highly *in vivo* relevant DR5 RAREs are expected to be conserved and, thus, to be under an ancient strong selective constraint. Therefore, to delineate functional RAREs, we analyzed the conservation of the human and mouse RAREs across 13 additional vertebrate organisms (see "Experimental Procedures") by using the BlastZ alignment of the University of Santa Cruz California genome browser. Due to the shortness and the divergence of the RGKTSA sequence, the criterion of conservation was deduced from the presence/absence of the complete DR5 motif RGKTSANNNNNRGKTSA in all considered genomes. We considered that a motif was conserved in a given species if the region encompassing the motif in human or mouse is aligned with a genomic region of the species also containing a RGKTSA DR5 motif (see "Experimental Procedures").

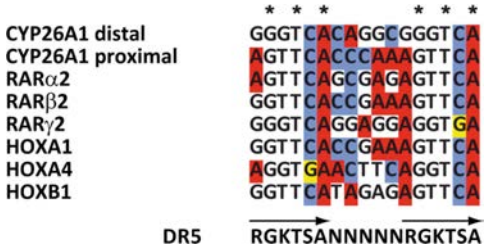


FIGURE 1. **Alignments of known DR5 RARE motifs in the promoters of the Cyp26A1, RARα2, *RARβ2*, RARγ2, Hoxa1, Hoxa4, and Hoxb1 genes and definition of a RGKTSA motif**.

In a phylogeny of vertebrates, we visualized the number of human or mouse RAREs that are conserved in each studied species (Fig. 2). We also calculated for each relevant clade of vertebrates the number of RAREs that are conserved in all the members of these clades. Although these data can be influenced by the coverage of the studied genome (34), this analysis raised three interesting conclusions. (i) Overall, human RAREs are less conserved in rodents than in other mammals. As an example, about 900 human RAREs are conserved in the mouse genome, whereas more than 1500 are conserved in the cow genome. This is in accordance with the known increased evolutionary rates in rodents (35) but questions the use of mouse as a unique *in vivo* experimental system for studying RA signaling in mammals. (ii) There is a striking difference between the number of RAREs conserved in placental mammals (Eutherians) and in all mammals or in eutherians + marsupials. Indeed around 309 human RAREs (or 319 mouse RAREs) are conserved in placental mammals, whereas only half (162 human RAREs, 170 mouse RAREs) is conserved in eutherians + marsupials. The decrease is even higher if we consider the number of RAREs conserved in all mammals (101 from human and 107 from mouse genomes). This suggests that a specific elaboration of the RA regulatory network occurred in eutherians and highlights the importance of studying the corresponding RA target genes. (iii) Only six RAREs are conserved in all jawed verte-

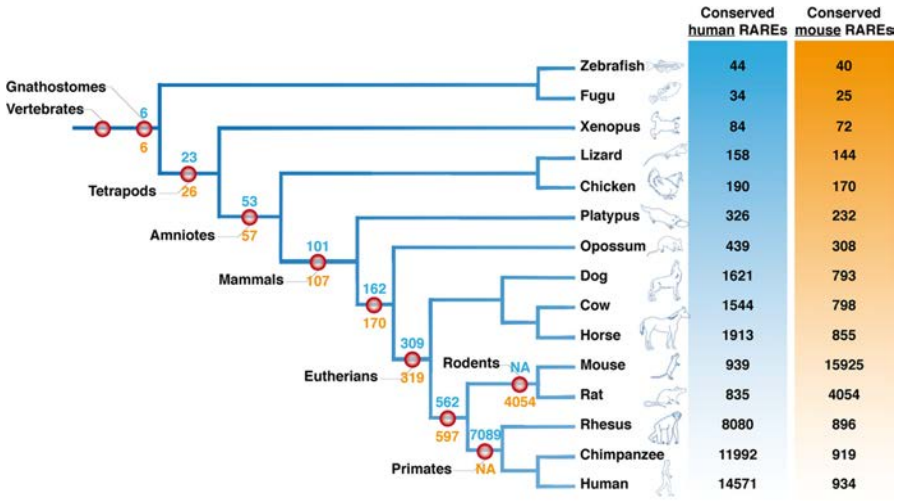


FIGURE 2. **Phylogenetic tree of jawed vertebrates showing the phylogenetic conservation of the DR5 RAREs.** On the *right*, the number of human RAREs (*blue*) or mouse RAREs (*orange*) conserved in each species is indicated. For example, in chimpanzee, there are 11,992 RAREs conserved from the 14,571 found in human. In contrast there are 919 RAREs conserved from the 15,925 found in the mouse genome. At each relevant node of the tree, the number of RAREs conserved in the species of the relevant node is indicated in *red*. The *blue numbers* represent the human RAREs, and the *orange numbers* represent the mouse RAREs. For example, in the rodent primate clade we found 562 human RAREs and 597 mouse RAREs conserved in the 5 relevant species (mouse, rat, rhesus, chimpanzee, and human). *NA*, not applicable.

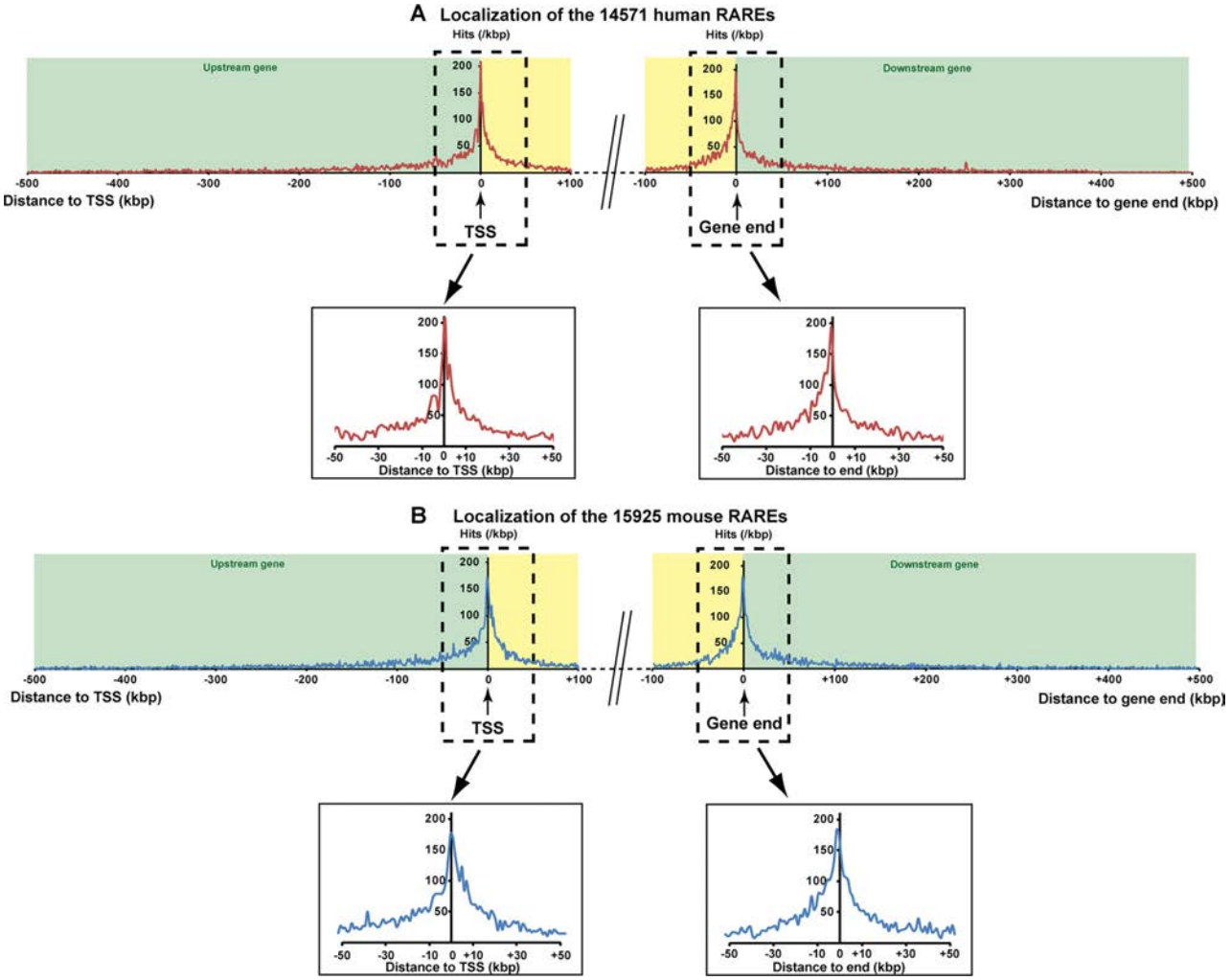


FIGURE 3. **Genome-wide distribution of the identified human (*A*) and mouse (*B*) DR5 RAREs.** The distance between the RARE and the corresponding gene was calculated by identifying its proximity to both boundaries of the genes, the TSS, and the end (end of the last exon). Following this rule for a RARE located upstream of the TSS, the distance was calculated from the TSS and was negative. For a RARE located downstream of a gene, the distance was calculated from the end of the gene and was positive. In the case of a RARE present inside the gene, both distances to the TSS and to the end of the gene were calculated, and the minimal distance in absolute value, called dTSS*, was considered (*cf.* Fig. 5). The genome-wide mapping of the RARE *versus* a canonical gene was calculated by cumulating the number of hits present in a 1-kb sliding window from each side from the TSS or from the end of the gene (the first point being attributed to 500 bp before and after TSS or gene end). These calculations were applied to a distance of 500 kb outside of the gene and 100 kb inside of the gene.

brates (gnathostomes). Three of these RAREs are associated to developmental genes, *Dach1* (Daschung homolog 1) (36), *Meis1* (Meis homeobox 1) (37), and *TSHZ3* (Teashirt 3) (38). The three others are associated to the *Gria2* (glutamate receptor 2) (39), *Lphn2* (latrophilin 2) (40), and *Paqr3* (an adiponectin receptor) (41) genes. It is interesting to note that, except Meis1 (8, 10, 37), these gene are not known RAR target genes. Nevertheless, they are likely to be RA-regulated in virtually any vertebrate species and thus might be considered as new interesting models.

Genome-wide Analysis of the Location of the Identified DR5 RAREs—The identified mouse and human RAREs were also annotated by analyzing genome-wide their locations using the GeCo system (see "Experimental Procedures"), which allows users to retrieve the genes in the neighborhood of factor binding sites with respect to annotated Refseq genes.[5] Then, in both the human and mouse genomes, the RAREs were localized relative to the nearest matched gene boundary: upstream and downstream distance from the TSS and from the end of genes. As shown in Fig. 3, *A* and *B*, the regions flanking TSSs and the ends of genes depict the highest concentration of RAREs compared with the further regions (±500 kb). This suggests that the RAREs located in the vicinity of TSSs and gene ends would be more relevant than the others, as described for most nuclear receptors and transcription factors (14, 42–46). Therefore, we selected the RAREs located between −10 and +10 kb, *i.e.* the RAREs ±10 kb from the TSSs and ±10 kb from gene ends. According to this criterion, 3862 RAREs were selected in the mouse genome and 3429 in the human one (supplemental Tables S2 and S3).

Selection of a List of RAREs Located ±10 kb from Gene Limits and Conserved in Six Organisms or More—Considering the low number (6) of highly conserved RAREs and the overall reparti-

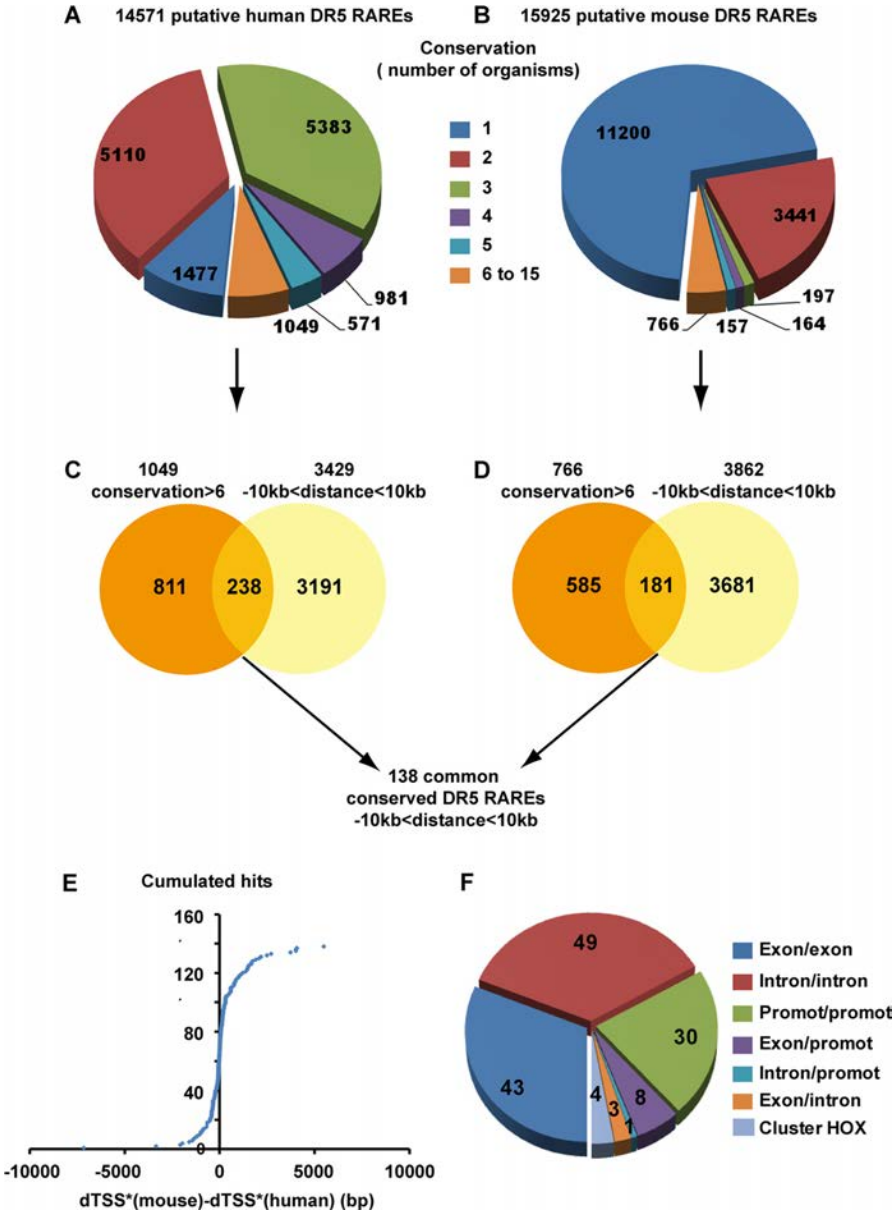


FIGURE 4. **Selection of 138 RAREs located ±10 kb from TSSs and gene ends and conserved in more than 6 organisms.** *A* and *B*, shown is conservation of the human and mouse RAREs among the 15 organisms tested. *C* and *D*, for both the human and mouse genomes, the RAREs conserved in more than 6 organisms were crossed with those located ±10 kb from gene boundaries. Crossing the resulting mouse and human RAREs led to a list of 138 RAREs with highly confident conservation and located ±10kb. *E*, the differences between mouse and human dTSS* were calculated and plotted into cumulative hits. *F*, conservation of the RARE positions (intron, exon, and promoter) between mouse and human is shown.

tion of the 15 organisms among the vertebrate tree, we arbitrary selected a criterion of conservation in 6 organisms. Only 7% of the human RAREs (1049 sites) (Fig. 4*A*) and 5% of the mouse RAREs (766 sites) (Fig. 4*B*) were found to be conserved in 6 organisms and more.

Then these human and mouse RAREs conserved across six and more organisms were further analyzed for their localization relative to the matched gene annotations. Among these RAREs, 238 human RAREs and 181 mouse RAREs were found to be located at the proximity of genes in the ±10 kb regions that we defined above (Fig. 4, *C* and *D*, and supplemental Tables 2 and 3). By using these two criteria of restriction, we obtained a list of 138 RAREs that are common to both mouse and human organisms and that are reliable in terms of genome annotation, name of the corresponding genes, localization, and conservation in more than 6 organisms (Fig. 5).

The orientation and localization of each conserved RARE listed in Fig. 5 were compared. Most interestingly, 100% of these RAREs showed the same orientation in the human and mouse genomes, and ~70% showed less than a 500-bp difference in their distances to the nearest associated genes (Fig. 4*E*), suggesting that these RAREs are good candidates for being func-

MOUSE												Mouse-Human
Official gene name	Position	BS strand	Gene strand	Phylogenetic Conservation	dTSS*	RARE Sequence	Position	BS strand	Gene strand	Phylogenetic Conservation	dTSS*	dTSS* difference
2310007D09Rik	Exon	+	+	8	-3436	GGTTCACGAAAAGTTCA	Exon	+	+	8	-5116	1680
4121402D02Rik	Intron	-	+	8	-5908	AGGTGAGTAAAAGTTGA	Intron	-	+	8	-5567	-341
4632412N22Rik		-	+	6	1259	GGGTCACAATCAGGTGA		-	+	6	1282	-23
4930452B06Rik	Exon	+	-	7	-459	AGTTCAGACAAGGGTCA	Exon	+	-	7	-495	36
4930562D19Rik	Exon	-	-	8	8578	GGGTGACCCCGAGTTCA	Exon	+	+	8	9552	-974
6430527G18Rik		-	-	8	-3426	GGGTGACGCGAGGTTCA		-	-	8	-4278	852
Abhd2	Exon	-	+	7	-1367	AGGTCAGATCCAGGTCA	Exon	-	+	7	-6849	5482
Aco2	Intron	+	+	9	404	AGGTGAGGACAAGGTGA	Intron	+	+	9	383	21
Actn4	Exon	-	-	6	-1391	GGGTGAGGCCGAGTTCA	Exon	+	+	6	-1535	144
Agap1	Intron	+	+	9	-9441	GGGTCACTCTGGGGTCA	Intron	+	+	8	-7485	-1956
Aipl1	Intron	-	-	6	-2338	GGGTCAAGGTCAGGTCA	Intron	-	-	6	-3811	1473
Alk	Intron	-	-	8	5054	AGGTGAGATGTGGGTCA	Intron	-	-	8	5420	-366
Ankrd50	Exon	+	-	9	-1989	AGTTGAAGCTAGGGTCA	Exon	+	-	9	-4700	2711
Ar	Exon	-	+	8	-1258	AGTTCATTCGAAGTTCA	Exon	-	+	8	-1413	155
Arpp21	Intron	-	-	7	-6279	GGTTCATTCATAGGTCA	Intron	+	+	7	-5225	-1054
Atxn2		+	+	6	-561	GGTTCAAATGAAGGTCA		-	-	6	-495	-66
BC052040	Exon	-	+	12	-1662	AGTTCAGCTGAAGTTCA	Exon	-	+	12	-1701	39
Bhlhbe40		-	+	12	1556	AGGTCAGCGCTGGGTGA		-	+	12	1775	-219
Bmf	Exon	+	-	8	-864	AGGTCACTGGGAGGTGA	Exon	+	-	8	-876	12
Bmp7	Intron	+	-	12	4301	AGTTCAAAGCTGGGTCA	Intron	+	-	12	5173	-872
Brs3	Exon	-	+	6	115	AGTTCAAATGCAGTTGA	Exon	-	+	6	95	20
Cacna1g	Intron	+	-	9	6105	GGGTCAGCCAGGGGTCA	Intron	-	+	9	6019	86
Camk2b	Exon	+	-	6	-3030	AGGTCAGCCCTGGGTCA	Exon	+	-	6	-3536	506
Ccdc132	Exon	-	+	8	-8558	AGTTCACTGGCGGTTGA	Exon	-	+	8	-9145	587
Ccno		-	+	7	-5719	AGTTCACTAAAAGTTCA		+	-	6	-8211	2492
Cdcp2		+	+	7	663	GGGTGACACACAGTTCA		-	-	7	1092	-429
Cited2		+	+	8	574	GGTTGACTCAAAGGTCA		-	-	8	632	-58
Clpb	Intron	-	+	7	5382	AGGTGATTCCAGGGTCA	Intron	+	-	7	5499	-117
Col24a1	Exon	-	+	7	292	AGTTCACTTGCGGTTCA		+	-	7	-296	588
Crygn		+	-	6	3244	GGGTGAGGGGGAGGTCA		+	-	6	1061	2183
Ctsk	Exon	-	+	7	2154	GGTTCATAGCCAGTTCA	Exon	+	-	7	2220	-66
Cugbp1	Exon	-	+	11	-379	AGGTCAACACAAGGTCA	Exon	+	-	11	-375	-4
Cugbp2	Exon	+	-	14	-5337	AGTTCAGCAAAAGGTGA	Exon	-	+	14	-5629	292
Cyp26a1		-	+	14	-31	AGTTCACCCAAAGTTCA	Exon	-	+	14	329	-360
Cyp26a1		+	+	8	-1898	GGGTCACAGGCGGGTCA		+	+	8	-1557	-341
Diap3	Exon	-	-	6	-229	AGTTGACAGAAGGGTCA	Exon	-	-	6	-1227	998
E130309F12Rik	Intron	+	+	9	745	GGGTGAGGCTCGGGTCA	Intron	+	+	9	958	-213
Ebf3		-	-	11	-7615	AGGTGAAACCGAGTTCA		-	-	11	-8385	770
Elf4	Exon	+	-	7	-7145	GGGTCAGTGACAGGTGA	Exon	+	-	7	-6455	-690
Enpp2		+	-	7	2163	AGTTCAACCTGAGGTGA		+	-	7	2866	-703
Ephb3	Intron	-	+	8	-6596	GGGTCACCTGGAGTTCA	Intron	-	+	8	-7178	582
Esrrg	Exon	-	+	10	-3113	GGTTCACAATAAGTTCA	Exon	+	-	10	-3061	-52
Ext2	Intron	-	-	9	-7409	AGTTCAAGCCCAGGTCA	Intron	+	+	9	-8911	1502
Fbxo30	Exon	+	+	6	-5674	AGGTCATTTGGAGTTCA	Exon	-	-	6	-6502	828
Fit4	Intron	+	+	8	3918	AGGTCAGCAGGGGTTCA	Intron	-	-	8	4073	-155
Foxa2		+	-	9	8392	AGGTCAGGGGGAGGTCA		+	-	9	8252	140
Frap1	Exon	-	+	9	-7491	AGTTCAGCAAGGGGTCA	Exon	+	-	9	-7802	311
Gabarapl2	Intron	-	+	7	522	AGGTGACAGCGAGGTGA	Intron	-	+	7	579	-57
Gpr149	Intron	+	-	6	-8572	AGGTGAATCCAGGGTGA	Intron	+	-	6	-8030	-542
Gria2	Intron	-	-	15	-6381	AGGTGATTCCAAGGTCA	Intron	+	+	15	-4962	-1419
Gzmk		+	-	8	-2144	GGGTGAGATTTAGTTCA		-	+	7	-2337	193
Hey1	Exon	+	-	7	-49	AGGTCAAACCCAGTTCA	Exon	+	-	7	-49	0
Hic1	Intron	-	-	9	1444	GGTTCACGGCGGGGTCA	Intron	+	+	9	322	1122
Hip1r	Exon	+	+	7	-2808	AGGTGAACGAGAGGTGA	Exon	+	+	8	-3142	334

FIGURE 5. **List of the 138 conserved RAREs located ±10 kb from TSSs and genes ends.** For the 138 conserved RAREs located ±10 kb, gene names and orthology were analyzed manually in both the human and mouse genomes. The sequence, localization, and dTSS* of each RARE are shown as well as the name of the associated gene. *BS*, binding sequence.

tional. Moreover, 43 RAREs located in exons, 49 in introns, and 30 in promoters correlated well between the two genomes (Fig. 4*F*). The other RAREs, although associated with a same gene in both genomes, depicted different localizations, most probably due to differences in genes annotations between the two genomes. Note that three RAREs associated with Hox genes differed between both genomes, most probably due to the complex organization and evolution of the Hox clusters.

RAR Binding to the Selected DR5 RAREs and Analysis of the Associated Genes—Then the key question to address was whether the DR5 RAREs that we selected *in silico* reflect biological significance *in vivo*; in other words, whether they are able to bind RAR/RXR heterodimers. To address this, we first crossed the list of 15,925 DR5 RAREs found in the mouse genome with the RAR and RXR binding sites mapped in ChIP-seq experiments[6] performed with a mouse embryocarcinoma cell line (F9 cell lines), which is well known to respond to RA (1). In these cells, 4% of these RAREs were occupied by RAR/RXR heterodimers in the absence of RA (Table 1). This percentage increased up to 9% after 48 h of RA treatment. In fact, taking into account that some sites become occupied whereas others dissociate from RAR/RXR heterodimers in response to RA, 11% of the RAREs were found to be able to bind RAR/RXR heterodimers (Table 1). As a control, a random list of 15,925 17-bp sequences extracted from the mouse-masked genome (supplemental Table S4) was crossed with the same binding sites (Table 1). Most interestingly, the percentage of occupied RAREs increased up to 42% when the same crossing was applied to the list of 181 conserved mouse RAREs and to our final in silico list of 138 RAREs (Table 1), thus validating our strategy.

Among these RAREs (Table 2), 39 were occupied in the absence of RA, among which 17 depicted an important increase in their occupancy in response to RA. In addition, 19 RAREs, although unoccupied in the absence of RA, became occupied after RA addition, raising to 58 the number of RAREs that can be occupied in F9 cells. Note that the increase in occupancy started rapidly (within 2 h) or later (24–48 h) depending on the RARE.

Some of these RAREs have been already reported to be direct RAR targets in EMSA, ChIP, or ChIP-chip experiments. It is the case for the canonical RAREs associated to the well known RA target genes involved in transcription regulation such as *RARβ2*, *Hoxa1*, *Hoxb1*, and *Wt1* (19, 32, 47–50) or in RA metabolism (*Cyp26A1*, *Rbp1*). Most inter-

[6] A. Chatagnon and G. Benoit, manuscript in preparation.

Conserved and Functional DR5 RAREs

Hmbox1		+	-	9	48	AGGTGATCACAGGGTGA	Exon	-	+	9	-133	181
Hoxa1		-	-	12	1940	GGTTCACCGAAAGTTCA		-	-	12	1749	191
Hoxa3	Intron	-	-	14	-6472	AGGTGAACTTCAGGTCA		-	-	14	-3124	-3348
Hoxa3	Intron	-	-	14	6064	GGTTCAAGAAGAGTTCA	Intron	-	-	14	6071	-7
Hoxb1		+	+	13	5100	GGTTCATAGAGAGTTCA		-	-	14	7170	-2070
Hoxb3	Intron	+	+	14	-6694	GGTTCAAGAAGAGTTCA	Intron	-	-	14	-6561	-133
Hoxc4		+	+	14	-5034	AGGTGAAATGCAGGTCA	Intron	+	+	14	-6699	1665
Hoxd3	Intron	+	+	13	-6731	GGTTCAAGCAGAGTTCA	Intron	+	+	12	-7185	454
Hoxd4	Intron	+	+	14	2778	AGGTGAAATGCAGGTCA		+	+	14	-1240	4018
Hoxd4	Intron	+	+	11	2696	AGTTGAGCGGGAGGTGA		+	+	11	-1322	4018
Ivns1abp		-	+	9	-187	AGGTCAGCGCAAGGTGA		+	-	9	-139	-48
Jmjd3	Intron	-	-	7	4801	GGGTCACATCGGGGTCA	Intron	+	+	7	4877	-76
Klk13	Intron	+	+	6	435	AGTTCACAGTGGGTTCA		-	-	6	-778	1213
Lgals2	Intron	-	-	7	-394	AGGTCAAGGTGAGGTCA	Intron	-	-	7	-322	-72
Lman2l	Intron	+	-	6	6032	AGGTCAAATCAGGTTGA	Intron	+	-	6	6093	-61
Lrrc29	Exon	+	-	6	-3339	GGTTGAAGCTGAGGTCA	Intron	+	-	6	-3496	157
Meis2	Intron	+	-	14	-7698	GGTTCACTCAAAGGTCA	Intron	+	-	14	-7710	12
Meis2	Intron	+	-	12	-6268	GGGTCATTCAGAGGTCA	Intron	+	-	12	-6210	-58
Meis2		-	-	8	-6129	AGGTCAAGAATAGTTCA		-	-	9	-4855	-1274
Mll1	Intron	-	-	7	1810	AGTTGAGTTCAGGTTCA	Intron	+	+	7	1873	-63
Mmp24	Exon	+	+	6	-1581	AGGTGAACTAGAGGTGA	Exon	+	+	6	-1658	77
mmu-let-7c-2		+	+	8	-3954	AGGTGACTTCGGGGTGA		+	+	8	-5394	1440
mmu-mir-10a		+	+	13	-2566	AGGTGAACCGCAGGTCA		-	-	14	-2623	57
Myf6	Intron	+	-	6	680	GGGTCACTGGGGGTTCA	Intron	-	+	6	705	-25
Myo3b	Intron	-	+	6	-7559	GGTTCAACTAGAGTTCA	Intron	-	+	6	-6568	-991
Nkapl	Exon	-	-	6	-386	AGGTGAAATTGGGTTGA	Exon	+	+	6	-624	238
Onecut2	Intron	+	+	12	9911	AGTTCAGCTATGGTTCA	Intron	+	+	12	9964	-53
Osr1		+	+	8	-988	GGGTCAGCCGGAGGTCA		-	-	8	-1004	16
Otp		+	+	7	-415	AGGTCACGCCAGGGTCA		-	-	6	-401	-14
Parp8		-	-	8	9175	AGTTCATTTACAGGTCA		+	+	7	8462	713
Pcbp2	Intron	-	+	9	-6347	GGTTGATGCTGAGGTGA	Intron	-	+	10	-6242	-105
Pcbp4	Intron	+	+	8	3290	AGGTGAGCTGGAGGTCA	Intron	-	-	8	3438	-148
Pld2	Intron	-	+	7	-7034	AGGTGACAGCAAGGTCA	Intron	-	+	7	-9119	2085
Polb	Exon	-	-	7	-8806	GGGTGAGACAAAGTTCA	Exon	+	+	7	-9115	309
Pou3f1		-	+	6	3357	AGGTGATGCTGGGTTCA		+	-	6	4039	-682
Prnpip1		+	+	10	-3868	AGTTCAGATTCAGGTCA		-	-	10	-3986	118
Prss27		+	+	8	-148	GGTTCAGGGCTAGGTCA	Exon	-	-	8	195	-343
Ptch1	Intron	+	-	11	6913	GGGTCACAGGCAGGTCA	Intron	+	-	9	7203	-290
Ptchd1	Exon	-	-	11	-4895	AGGTCAGTGAAGGGTCA	Exon	+	+	11	-3721	-1174
Ptprj	Exon	-	-	7	-917	AGGTCACACGAAGGTCA	Exon	+	+	8	-963	46
Ptprj	Exon	+	-	6	-7803	AGTTGATGAGCAGGTCA	Exon	-	+	6	-7307	-496
Qk		+	-	14	608	AGGTCATCATTGGGTGA	Exon	-	+	14	-3124	3732
Rab11fip2	Exon	-	-	8	5995	GGGTGAGATAAAGGTCA	Exon	-	-	8	6099	-104
Rab39b	Exon	-	-	7	410	GGGTCAAGAGAGGTTCA	Exon	-	-	7	492	-82
Raph1	Intron	-	-	7	-8272	GGGTCAAGTTAAGGTCA	Intron	-	-	7	-1126	-7146
Rarb		-	-	12	-303	GGTTCACCGAAAGTTCA		+	+	13	-52	-251
Rarb	Exon	+	-	9	-4293	AGGTGAACACAAGGTCA	Exon	-	+	9	-4316	23
Rbbp7	Exon	-	+	8	1280	GGGTCATAACCAGGTCA	Exon	+	-	9	1283	-3
Rbm35b	Exon	+	-	8	-943	AGTTCAGGTAGAGTTGA	Exon	+	-	8	-2314	1371
Rbp1	Intron	-	+	8	7034	GGGTCATCCTAAGTTCA	Intron	+	-	8	5998	1036
Rnf10	Exon	+	-	8	-2326	GGGTGACAGAGGGGTCA	Exon	-	+	8	-6394	4068
Rnf214	Exon	-	-	6	6752	GGTTCATCAGCAGGTGA	Exon	+	+	6	5981	771
Ror1	Intron	+	+	12	-6573	AGGTCAAGCAGAGGTCA	Intron	+	+	12	-6197	-376
Rxrb	Intron	-	+	8	-1559	AGGTCACTCAAAGGTCA	Intron	+	-	8	-1884	325
Sat1	Exon	+	-	8	-460	AGTTGATGGATGGTTCA	Exon	-	+	8	-471	11
Sema3e	Exon	-	+	8	63	AGTTCAGGCAGGGTTGA		+	-	8	-117	180
Spk2	Exon	-	+	8	-1190	GGGTGAACTCTGGGTCA	Exon	-	+	8	-755	-435
Shank3	Intron	-	+	7	-620	GGGTCAGCGCCAGGTGA	Intron	-	+	7	-2523	1903
Shf	Intron	-	-	11	-6710	AGTTGACATTAAGGTCA	Exon	-	-	11	-6489	-221
Slc22a5	Intron	-	-	9	-7623	GGGTGAAGCTCAGGTCA	Intron	+	+	8	-7207	-416
Slc22a5	Intron	+	-	8	-7756	AGTTCAACAAAAGGTCA	Intron	-	+	7	-7343	-413
Slc25a23	Exon	-	-	8	-1767	GGGTCACGTCCAGGTGA	Exon	-	-	8	-1912	145
Slc9a3	Intron	+	+	7	682	AGGTGACAGGAAGGTCA	Intron	-	-	8	931	-249
Smyd5	Exon	-	+	7	-1309	AGGTCACATTGGGTTCA	Exon	-	+	7	-1380	71
Sp7	Exon	-	-	8	-384	GGGTGACCCCAGGGTCA	Exon	-	-	8	-972	588
Srp68	Exon	+	-	9	-633	GGTTGAGGGCCAGGTCA	Exon	+	-	9	-692	59
Sspn		+	+	8	3592	AGTTCATCTATAGGTCA		+	+	8	5189	-1597
Tcf7l2	Intron	-	+	9	-2852	AGGTCAGAATCAGGTGA	Intron	-	+	9	-1478	-1374
Tcfap2c		+	+	6	2676	GGGTCATGGGTGGGTGA		+	+	6	3016	-340
Tfg	Intron	-	-	9	-1690	GGTTCAAGGTGAGTTCA	Intron	+	+	9	-1666	-24
Tnks1bp1	Exon	+	+	6	-2692	AGGTCAGTGGAGGGTGA	Intron	-	-	6	-2816	124
Top2b	Intron	+	+	8	1145	AGTTCAGACCAGGGTCA	Intron	-	-	9	907	238
Trpc1		-	-	6	304	AGTTCACATATAGTTCA	Exon	+	+	6	-1225	1529
Ttc27		-	+	8	3495	AGTTGACTCAGAGGTCA		-	+	8	4339	-844
Wnt1		-	+	7	153	GGGTCATCCAAGGGTCA		-	+	6	85	68
Wnt5a	Intron	+	+	10	1493	GGGTCACTGAGAGTTCA	Intron	-	-	9	1239	254
Wt1		+	+	13	-4329	AGTTGAATCTGGGGTCA		-	-	13	-4441	112
Wt1	Intron	-	+	10	7818	AGTTCACCCAAAGTTGA	Intron	+	-	10	8308	-490
Yes1	Intron	+	+	8	3362	AGGTCACCTTCAGTTCA	Intron	-	-	8	3781	-419
Ypel5	Exon	+	+	9	-2592	AGTTCAAGATCGGGTCA	Exon	+	+	9	-1891	-701
Zbtb5	Exon	+	-	9	-3813	AGTTCAAATGCAGGTGA	Exon	+	-	9	-4120	307
Zdhhc3	Intron	-	-	7	676	AGGTGAAGCTGGGGTCA	Intron	-	-	7	948	-272
Zfp503	Exon	+	-	13	1645	AGGTGATTGGAGGGTCA		+	-	12	-91	1736
Zfp598	Intron	-	+	6	6048	GGGTCATGCCAAGGTCA	Intron	+	-	6	5149	899

FIGURE 5—*continued*

estingly, this analysis revealed that two RAREs are associated to the *RARβ2*, *Wt1*, and *Cyp26A1* genes. However, only one of the RAREs associated to the *Wt1* and *Cyp26A1* genes was occupied in F9 cells, whereas both RAREs associated to the *RARβ2* gene were occupied. Of note is that for *RARβ2*, the RARE located in the promoter was more efficiently occupied than the other one, located in an exon, increasing the complexity of the transcriptional regulation of this gene in F9 cells.

Other occupied RAREs were associated to genes that are already known as RA-responsive genes but for which no RAREs had been identified yet. Among these genes are the "stimulated by RA" (*Stra*) genes such as *Bhlhb40* (*Stra13*) (51), *Tcfap2C* (*Stra2*) (52), and *Meis2* (*Stra10*) (53), and zinc finger proteins (*Zfp598* and *Zfp503*) (8, 10). Note that three RAREs are associated to the *Meis2* gene but that only two were occupied by RAR/RXR heterodimers in F9 cells. The analysis also revealed occupied RAREs associated to gene regulatory regions, which were recently found to be occupied by RARs in ChIP-chip and ChIP-qPCR experiments performed with other cell lines but without any indication whether this occupancy was direct or indirect through other bound factors (11). This is exemplified by the RAREs associated to the *Atxn2*, *Top2b*, *Wnt1*, and *Wnt5* genes.

TABLE 1

Numbers of occupied mouse DR5 RAREs in the initial list (15,925 predicted sites), the list of conserved 181 sites located ±10 kb from TSSs, and in the final list of 138 RAREs

For finding significant occupied sites, a cutoff *p* value of 0,00001 was applied. As a control, the same strategy was applied to a random list of 15,925 sequences of 17 kb.

	RAREs			
F9 cells	Random list	Predicted (159,25)	Conserved (181)	Final list (138)
Untreated	32 (0.2%)	691 (4%)	48 (26%)	39 (28%)
RA 2 h	62	1,109	49	45
RA 24 h	39	964	46	33
RA 48 h	85	1,523 (9%)	71	55
Total	112	1,791 (11%)	76 (42%)	58 (42%)

Most interestingly, a new repertoire of occupied RAREs was found to be associated to new potential RA target genes encoding transcription regulators (*RXRβ*, *Jmjd3*, *Foxa2*), several Homeobox genes belonging to clusters (*Hoxa3*, *Hoxb3*, *Hoxd3*), galectins (*Lgals2*), membrane-associated proteins (*Sema3e*, *Abhd2*, *Crygn*), RNA-binding proteins (*Cugbp1*, *Qk*, *Srp68*, *Pcbp4*), ATPases (*Clpb*), and proteins involved in cell death (*Sspn*), neuronal functions (*Agap1*), developmental processes (*Otp*), cell signaling (*Raph1*, *Arpp21*, *Zdhhc3*, *Cacna1g*, *Camk2b*, *Ephb3*, and *Pld2*), and cytoskeleton organization (*Ivns1abp*). RAREs were also found associated to the tumor suppressor *HIC1* gene, the kallikrein-related peptidase 13 (*KLK13*) gene, the *Myf6* gene, which belongs to the family of muscle regulatory factors, and the *Prss27* gene, which encodes a membrane-anchored protease. Note that in the two latter cases, the occupation of the sites decreased after RA addition.

The other RAREs of the bioinformatics list were not occupied by RARs in F9 cells either in the absence or presence of RA. However, as RAR binding relies on the cellular and/or physiological context, one cannot exclude that these RAREs would be occupied in other cell lines or tissues or in other species.

RA Regulation of the Genes Associated to the Selected RAREs—Next we assessed whether the genes associated to the selected RAREs are RA-regulated. Our *in silico* screen identified 138 DR5 RAREs, but taking into account that several RAREs were found associated to a same gene, there are 129 potential RA-regulated genes. First, the set of RA-regulated genes was analyzed by high throughput qPCR sequencing (RNA-seq) using F9 cells for which we already had a list of 58 occupied DR5 RAREs. A list of 167 genes that were either induced or repressed after a 4-h treatment with RA was generated after data normalization and identification of the significant differentially expressed genes (supplemental Table S5). This list was finally reduced to 164 distinct genes after removal of the duplicated genes.

Then this list of 164 RA-regulated genes was crossed with the list of 129 RARE-associated genes raised *in silico*, resulting in the selection of 9 RA-responsive genes common to the two lists (Fig. 6). This list includes indeed the canonical RAR target genes (*Cyp26A1*, *RARβ2*, *Rbp1*, *Hoxa1*, and *Hoxb1*). It also includes two new *Hox* genes (*Hoxa3*, *Hoxb3*) as well as two new *Stra* genes, *Tcfap2C and BHBLH4.* In F9 cells, for all these nine RA-responsive genes, the associated DR5 RAREs were occupied by RAR/RXR heterodimers, and this occupancy was increased in response to RA (see Table 2).

Note that several other genes that are not in our bioinformatics list are activated in RA-treated F9 cells. However, some rely on other DR elements (*cdx1*) and/or reflect the complexity of Hox clustering (*Hoxa5*, *Hoxb5*, *Hoxa4*, and *Hoxb2*) (54–57). Others are known RA-responsive genes (1, 58–60) with DR5 RAREs (see supplemental Table S2) but are not conserved in several species (*Cyp26B1*, *Stra6*, *Stra8*, *Foxa1*, *Gbx2/Stra7*) or are located out of the ±10-kb limits (*Gata6*).

The RA-responsiveness of the genes we selected *in silico* was also analyzed in qRT-PCR experiments performed with F9 cells after RA treatment for different times up to 8 h. According to the confidence of the quality of their annotation and sequencing, 49 genes among the 129 genes (supplemental Table 1) were analyzed. This approach confirmed the RA inducibility of the nine genes selected above (Fig. 7, *A–D*). Interestingly, it also revealed that the inducibility of these genes increases with time (Fig. 7, *A–D*). Moreover, some additional RAR-bound genes, such as *Meis2*, *KLK13*, and *HIC1*, can be also activated in response to RA but with a low efficiency and at later times (8 h) (Fig. 7, *D* and *E*), raising to 12 the list of RA-responsive genes controlled by conserved DR5 RAREs and located ±10 kb from TSSs in F9 cells (Fig. 6).

Given that the RA response of target genes is well known to be cell type-specific, the same qRT-PCR experiments were performed with another RA-responsive mouse embryocarcinoma cell line, the P19 cell line. As shown in Fig. 7, *F–J*, the same genes were activated in response to RA, although with different intensities and kinetics. As an example, the *Hoxa1* and *Meis2* genes were more efficiently activated in P19 cells than in F9 cells. Note that the *Myf6* gene, which was not activated in F9 cells, responded to RA in P19 cells (Fig. 7*I*), raising the number of RA-responsive genes to 13.

Finally, as the RAREs controlling these 13 genes are highly conserved between species (Fig. 5), we analyzed whether they also responded to RA in other cell lines from other species such as a human breast cancer cell line (MCF7 cells) (Fig. 8*A*) and a zebrafish cell line (PAC2) (Fig. 8*B*). The *Bhlhe40* gene was significantly activated in MCF7 cells but not in zebrafish PAC2 cells. In contrast, *Meis2* was strongly activated in PAC2 cells and not in MCF7 cells. These results are summarized in Fig. 9 and point out that the RA response of the new RARE-associated genes we identified may vary from one cell type to the other and from one species to the other.

DISCUSSION

Here we describe a genome-wide *in silico* analysis of consensus DR5 RAREs with recurrent RGKTSA motifs. The advantage of such a computational approach is that it overcomes the chromatin and cellular context and thus provides a direct glance on the whole repertoire of possible RAREs. Moreover, the choice of recurrent RGKTSA motifs was expected to expand this repertoire of RAREs.

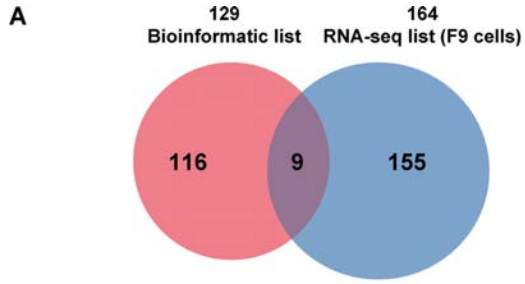
This computational study revealed around 15,000 DR5 RAREs in the human and mouse genomes. Among these RAREs, 24% are concentrated in regions located ±10 kb from the TSSs and the gene ends, and 5–7% are conserved in 6 organisms or more. It also revealed that the degree of conservation of the overall RAREs is not linear with time in the various verte-

Conserved and Functional DR5 RAREs

TABLE 2
RAR/RXR occupancy of the selected RAREs in F9 cells treated or not with RA (10^{-7} M) for the indicated times

Official gene name	dTSS*	Occupied RAREs in F9 cells			
		Control	RA 2h	RA 24h	RA 48h
2310007D09Rik	-3436				
4121402D02Rik	-5908				
4632412N22Rik	1259				
4930452B06Rik	-459				
4930562D19Rik	8578				
6430527G18Rik	-3426	+	+	+	+
Abhd2	-1367	+	+	+	+
Aco2	404				
Actn4	-1391	-	+	-	+
Agap1	-9441	-	-	++	++
Aipl1	-2338				
Alk	5054				
Ankrd50	-1989				
Ar	-1258				
Arpp21	-6279	-	+	-	+
Atxn2	-561	+	++	++	++
BC052040	-1662	-	-	-	+
Bhlhe40	1556	+	++	++	++
Bmf	-864	-	-	-	+
Bmp7	4301				
Brs3	115				
Cacna1g	6105	+	+	++	++
Camk2b	-3030	+	+	+	+
Ccdc132	-8558				
Ccno	-5719				
Cdcp2	663	-	-	-	+
Cited2	574	-	+	-	+
Clpb	5382	+	+	+	+
Col24a1	292	-	-	-	+
Crygn	3244	++	++	++	++
Ctsk	2154				
Cugbp1	-379	+	+	-	+
Cugbp2	-5337				
Cyp26a1	-31				
Cyp26a1	-1898	++	++++	+++	+++
Diap3	-229				
E130309F12Rik	745				
Ebf3	-7615				
Elf4	-7145				
Enpp2	2163				
Ephb3	-6596	+	+	-	+
Esrrg	-3113				
Ext2	-7409				
Fbxo30	-5674				
Flt4	3918				
Foxa2	8392	-	+	-	+
Frap1	-7491				
Gabarapl2	522				
Gpr149	-8572				
Gria2	-6381				
Gzmk	-2144				
Hey1	-49	-	-	-	+
Hic1	1444	+	++	++	++
Hip1r	-2808				
Hmbox1	48				
Hoxa1	1940	+	+++	++++	++
Hoxa3	-6472	+	+	++	+
Hoxa3	6064	+	+++	+++	++
Hoxb1	5100	+	+	++	++
Hoxb3	-6694	+	-	++	++
Hoxc4	-5034				
Hoxd3	-6731	-	-	+	+
Hoxd4	2778				
Hoxd4	2696				
Ivns1abp	-187	-	+	+	+
Jmjd3	4801	++	++	++	++
Klk13	435	-	+	-	+
Lgals2	-394	+	++	++	++
Lman2l	6032				

Official gene name	dTSS*	Occupied RAREs in F9 cells			
		Control	RA 2h	RA 24h	RA 48h
Lrrc29	-3339				
Meis2	-7698	++	++	++	++
Meis2	-6268	+	++	++	+
Meis2	-6129				
Mll1	1810				
Mmp24	-1581				
mmu-let-7c-2	-3954				
mmu-mir-10a	-2566				
Myf6	680	+	-	-	-
Myo3b	-7559				
Nkapl	-386				
Onecut2	9911				
Osr1	-988				
Otp	-415	+	++	++	++
Parp8	9175				
Pcbp2	-6347				
Pcbp4	3290	+	+	+	+
Pld2	-7034	+	+	-	+
Polb	-8806				
Pou3f1	3357				
Prnpip1	-3868				
Prss27	-148	+	+	-	-
Ptch1	6913				
Ptchd1	-4895				
Ptprj	-917				
Ptprj	-7803				
Qk	608	+	+++	++	++
Rab11fip2	5995				
Rab39b	410				
Raph1	-8272	+	+	-	+
Rarb	-303	+	++	-	+
Rarb	-4293	+++	++++	++++	++++
Rbbp7	1280	+	+	-	+
Rbm35b	-943				
Rbp1	7034	++	++	+++	++
Rnf10	-2326				
Rnf214	6752				
Ror1	-6573				
Rxrb	-1559	++	++	++	++
Sat1	-460				
Sema3e	63	+	+	+	+
Sgk2	-1190				
Shank3	-620				
Shf	-6710				
Slc22a5	-7623				
Slc22a5	-7756	-	-	-	+
Slc25a23	-1767				
Slc9a3	682				
Smyd5	-1309				
Sp7	-384				
Srp68	-633	+	+	-	+
Sspn	3592	-	+	+	+
Tcf7l2	-2852				
Tcfap2c	2676	-	-	-	+
Tfg	-1690				
Tnks1bp1	-2692				
Top2b	1145	-	+	++	++
Trpc1	304				
Ttc27	3495				
Wnt1	153	+	-	+	+
Wnt5a	1493	+	+	-	+
Wt1	-4329	-	+	+	-
Wt1	7818				
Yes1	3362				
Ypel5	-2592				
Zbtb5	-3813				
Zdhhc3	676	-	+	-	+
Zfp503	1645	++	++	++	++
Zfp598	6048	+	-	-	+



B Genes from the bioiformatic list that are RA regulated in F9 cells

Cyp26A1
Hoxa1
Hoxa3
Hoxb1
Hoxb3
RARβ2
Rbp1
Tcfap2c
Bhlhe40

C Other genes from the bioinformatic list that are RA-regulated later and/or in other cell types

Hic1
KLK13
Meis2
Myf6

FIGURE 6. **RA-responsiveness of the conserved DR5 RARE-associated genes identified *in silico* as assessed by RNA-seq in F9 cells.** *A*, shown are Venn diagrams. *B* and *C*, shown is a summary of the RA-regulated genes.

brates and that the RA gene regulatory network is specifically elaborated in specific groups. Surprisingly, this occurred specifically in placental mammals (eutherians) *versus* all mammals. Indeed, 3-fold more RAREs are conserved in the former than in the latter. As no major events of genomic reorganization are known to have occurred at the base of placental mammals, this elaboration might be specific to RA signaling.

Finally it provided a list of 138 RAREs located ±10 kb from TSSs and gene ends and conserved in 6 organisms or more. This list includes the majority of known RAREs, validating the restrictive criteria of our analysis. It also includes RAREs associated to "stimulated by RA" (*Stra*) genes for which no RAREs had been identified yet. The interesting point is that it provided a newly expanded set of high confidence conserved DR5 RAREs associated to a series of new genes involved in transcription, cell signaling, development, neuronal functions, and tumor suppression. The other interesting point is that in some cases, two to three RAREs were found to be associated to a same gene (*e.g. Cyp26A1*, *RARβ2*, and *Meis2*), increasing the complexity of the transcriptional regulation of these genes.

However, *in silico* identification of RAREs does not assure their functionality. Therefore, we combined the present computational analysis to experimental biology to determine whether the selected RAREs can bind RARs (ChIP-seq) and respond to RA (RNA-seq and qRT-PCR). Such an integrated strategy performed with mouse embryocarcinoma cells (F9 cell line) revealed that 11% of the 15,925 mouse RAREs present in the starting list were occupied by RAR/RXR heterodimers. Interestingly, this percentage increased to 40% in the final list of

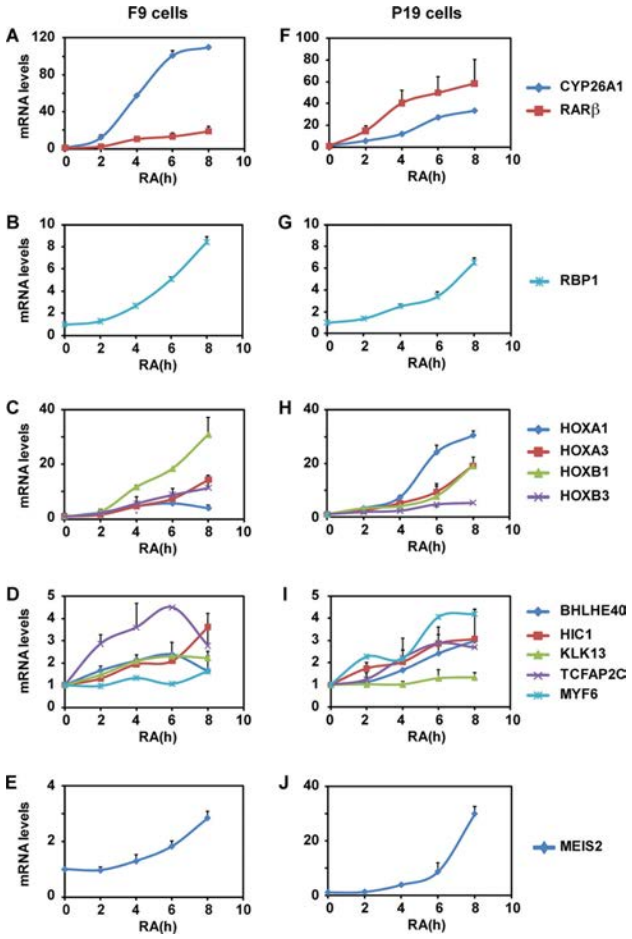


FIGURE 7. **Real time RT-PCR analysis of the RA regulation of the genes associated to the conserved DR5 RAREs identified *in silico* in F9 (*A–E*) and P19 (*F–J*) mouse embryocarcinoma cells.** The results correspond to the mean ± S.D. of three independent experiments.

conserved RAREs located ±10 kb from TSSs, validating our selection strategy.

Of note is that, in F9 cells, among the 58 occupied RAREs of our final list, only 12 of the corresponding genes were rapidly activated in response to RA. These genes include indeed the canonical RA target genes (*Cyp26A1*, *RARβ2*, *Rbp1*, *Hoxa1*, *Hoxb1*) as well as new *Hox* genes (*Hoxa3* and *Hoxb3*), *Stra* genes (*Tcfap2c*, *Bhlhe40*, *Meis2*), *HIC1*, and *KLK13*. These 12 genes were also activated in another mouse embryocarcinoma cell line (P19). However, some of them (exemplified by the *Bhlhe40* and *Meis2* genes) did not respond to RA in human MCF7 cells or in zebrafish PAC2 cells. In contrast, another gene, *Myf6*, which was occupied but not RA-responsive in F9 cells, was significantly induced in P19 cells. This corroborates that the RA regulation of target genes differs from one cell type to the other (Fig. 9), most probably in line with their chromatin context and final feature (differentiation or proliferation). In fact, the majority of the genes associated to occupied RAREs were not RA-regulated in F9 cells. This lack of RA response may be due to the fact that the genes are already expressed (and thus cannot be further stimulated). However, one cannot exclude that RA regulation requires longer times as exemplified for

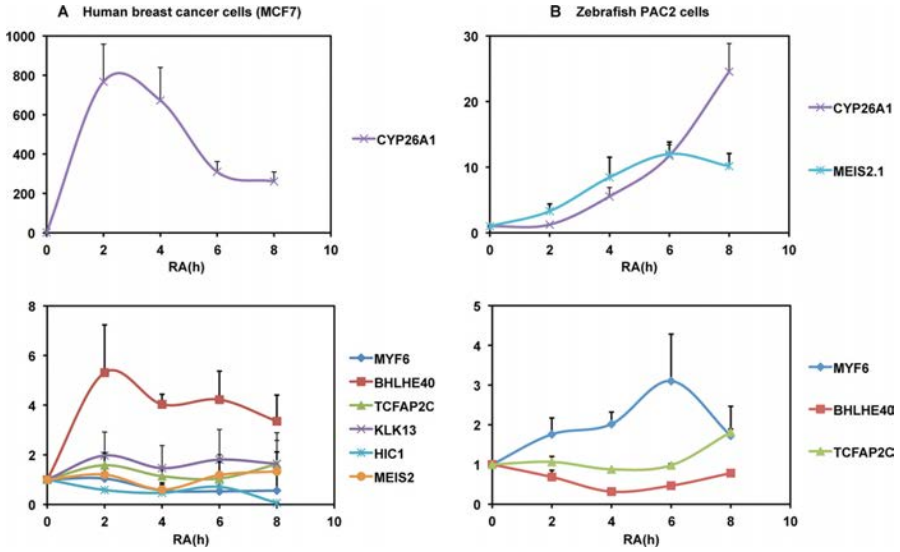


FIGURE 8. **Real time RT-PCR analysis of the RA regulation of the genes associated to the conserved DR5 RAREs in human MCF7 (*A*) and zebrafish PAC2 (*B*) cells.** The results correspond to the mean ± S.D. of three independent experiments.

RA-target gene	qRT-PCR validation in different cell lines				Associated RARE(s)		
	Mouse F9	Mouse P19	Human MCF7	Zebrafish PAC2	Number	Conservation	Occupancy in F9 cells
Canonical DR5 RARE-associated genes							
Cyp26a1	+	+	+	+	1	14	=
					2	8	+++
Hoxa1	+	+	+	ND	1	12	+++
Hoxb1	+	+	+	ND	1	13	+
Rarb	+	+	±	Gene loss	1	12	++++
					2	9	++
Rbp1	+	+	ND	ND	1	8	++
New DR5 RARE-associated genes							
Hoxa3	+	+	ND	ND	1	14	+
Hoxb3	+	+	ND	ND	1	14	+
Bhlhbe40	+	+	+	-	1	12	++
Tcfap2c	+	+	-	-	1	6	+
Meis2	+	+	-	+	1	14	++
					2	12	++
					3	8	-
Hic1	+	+	-	ND	1	9	+
Klk13	+	-	-	Gene loss	1	6	+
Myf6	-	+	-	+	1	6	±

FIGURE 9. **Recapitulation of the conserved DR5 RAREs that are RA-activated in mouse embryocarcinoma cells (F9 and P19 cell lines), human breast cancer cells (MCF7 cells), and a zebrafish cell line (PAC2 cells).**

Zfp503 (10), specific RARE-mediated conformational changes of the bound RAR (61), and/or cross-talks with other signaling pathways (12, 62), emphasizing the complexity of the RAR-mediated regulation of gene expression.

Remarkably, the majority of the RAREs present in our *in silico* list were not occupied *in vivo* in F9 cells. This is not surprising, as RAR binding relies on the cellular and physiological context and/or may require other cell specific transcription factors (12). Thus, one can predict that the other RAREs present in the *in silico* list would be occupied in other appropriate cell types or tissues with the corresponding genes being RA regulated under specific conditions.

The final interesting point of this study is the identification of 6 RAREs that are conserved in all the 15 species studied. However, except the RARE associated to the *Gria2* gene, all these RAREs are located out of the ±10-kb limits we defined. Moreover, none of the corresponding genes were RA-regulated in F9 cells, as assessed in RNA-seq experiments, except *Meis1*, which was activated 24 h after the RA addition to F9 cells (10). Nevertheless, these genes are mostly developmental genes (36–41) that are expressed in specific cell types and tissues and at specific developmental stages. Therefore, they might be new markers of the RA response, valid at specific times, in specific tissues from any jawed vertebrate species, opening new avenues for the study of RA signaling during development.

In conclusion, the novelty of the present study resides in an integrated strategy combining genome-wide biocomputing analysis and biological experiments for discovering and characterizing new RAR target genes and response elements. In addition to providing a wider valuable knowledge base for the anal-

ysis of robust RA-responsive genes, such a strategy also brought significant biological information. Indeed, it revealed (i) low conservation of RAREs between human and mouse (6%) and significant differences in the RA regulation of the highly conserved RAR target genes between species. Thus, it suggests that the RA response will differ from one species to the other as well as from one tissue to the other and under different situations. Finally, one can predict that the small set of conserved RAR direct target genes would act as key effectors of evolutionary steps.

Acknowledgments—We warmly acknowledge all members of the team and of the cell culture facilities for help. Special thanks to Regis Lutzing for qRT-PCR and to Bernard Jost and Celine Keime for RNA-seq.

REFERENCES

1. Bour, G., Taneja, R., and Rochette-Egly, C. (2006) in *Nuclear Receptors in Development* (Taneja, R., ed.) pp. 211–253, Elsevier Science Publishing Co., Inc., New York
2. Duong, V., and Rochette-Egly, C. (2011) *Biochim. Biophys. Acta* **1812,** 1023–1031
3. Mark, M., Ghyselinck, N. B., and Chambon, P. (2009) *Nucl. Recept. Signal.* **7,** e002
4. Samarut, E., and Rochette-Egly, C. (2011) *Mol. Cell. Endocrinol.*, in press
5. Rochette-Egly, C., and Germain, P. (2009) *Nucl. Recept. Signal.* **7,** e005
6. Bastien, J., and Rochette-Egly, C. (2004) *Gene* **328,** 1–16
7. Cotnoir-White, D., Laperrière, D., and Mader, S. (2011) *Mol. Cell. Endocrinol.* **334,** 76–82
8. Eifert, C., Sangster-Guity, N., Yu, L. M., Chittur, S. V., Perez, A. V., Tine, J. A., and McCormick, P. J. (2006) *Mol. Reprod. Dev.* **73,** 796–824
9. Harris, T. M., and Childs, G. (2002) *Funct. Integr. Genomics* **2,** 105–119
10. Su, D., and Gudas, L. J. (2008) *Biochem. Pharmacol.* **75,** 1129–1160
11. Delacroix, L., Moutier, E., Altobelli, G., Legras, S., Poch, O., Choukrallah, M. A., Bertin, I., Jost, B., and Davidson, I. (2010) *Mol. Cell. Biol.* **30,** 231–244
12. Hua, S., Kittler, R., and White, K. P. (2009) *Cell* **137,** 1259–1271
13. Hoffman, B. G., and Jones, S. J. (2009) *J. Endocrinol.* **201,** 1–13
14. Reddy, T. E., Pauli, F., Sprouse, R. O., Neff, N. F., Newberry, K. M., Garabedian, M. J., and Myers, R. M. (2009) *Genome Res.* **19,** 2163–2171
15. Anno, Y. N., Myslinski, E., Ngondo-Mbongo, R. P., Krol, A., Poch, O., Lecompte, O., and Carbon, P. (2011) *Nucleic Acids Res.* **39,** 3116–3127
16. Rochette-Egly, C., Gaub, M. P., Lutz, Y., Ali, S., Scheuer, I., and Chambon, P. (1992) *Mol. Endocrinol.* **6,** 2197–2209
17. Taneja, R., Rochette-Egly, C., Plassat, J. L., Penna, L., Gaub, M. P., and Chambon, P. (1997) *EMBO J.* **16,** 6452–6465
18. Samarut, E., Amal, I., Markov, G., Stote, R., Dejaegere, A., Laudet, V., and Rochette-Egly, C. (2011) *Mol. Biol. Evol.* **28,** 2135–2137
19. Bruck, N., Vitoux, D., Ferry, C., Duong, V., Bauer, A., de Thé, H., and Rochette-Egly, C. (2009) *EMBO J.* **28,** 34–47
20. Bour, G., Plassat, J. L., Bauer, A., Lalevée, S., and Rochette-Egly, C. (2005) *J. Biol. Chem.* **280,** 17027–17037
21. Trapnell, C., Pachter, L., and Salzberg, S. L. (2009) *Bioinformatics* **25,** 1105–1111
22. Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J., and Pachter, L. (2010) *Nat. Biotechnol.* **28,** 511–515
23. Anders, S., and Huber, W. (2010) *Genome Biol.* **11,** R106
24. Benjamini, Y., and Hochberg, Y. (1995) *J. R. Stat. Soc. Series B Stat. Methodol.* **57,** 289–300
25. Umesono, K., Murakami, K. K., Thompson, C. C., and Evans, R. M. (1991) *Cell* **65,** 1255–1266
26. Leid, M., Kastner, P., and Chambon, P. (1992) *Trends Biochem. Sci* **17,** 427–433
27. Loudig, O., Maclean, G. A., Dore, N. L., Luu, L., and Petkovich, M. (2005) *Biochem. J.* **392,** 241–248
28. de Thé, H., Vivanco-Ruiz, M. M., Tiollais, P., Stunnenberg, H., and Dejean, A. (1990) *Nature* **343,** 177–180
29. Leroy, P., Nakshatri, H., and Chambon, P. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88,** 10138–10142
30. Lehmann, J. M., Zhang, X. K., and Pfahl, M. (1992) *Mol. Cell. Biol.* **12,** 2976–2985
31. Langston, A. W., and Gudas, L. J. (1992) *Mech. Dev.* **38,** 217–227
32. Huang, D., Chen, S. W., Langston, A. W., and Gudas, L. J. (1998) *Development* **125,** 3235–3246
33. Doerksen, L. F., Bhattacharya, A., Kannan, P., Pratt, D., and Tainsky, M. A. (1996) *Nucleic Acids Res.* **24,** 2849–2856
34. Milinkovitch, M. C., Helaers, R., Depiereux, E., Tzika, A. C., and Gabaldón, T. (2010) *Genome Biol* **11,** R16
35. Bromham, L., and Penny, D. (2003) *Nat. Rev. Genet.* **4,** 216–224
36. Jing, Y., Machon, O., Hampl, A., Dvorak, P., Xing, Y., and Krauss, S. (2011) *Cell. Mol. Neurobiol.* **31,** 715–727
37. Mercader, N., Leonardo, E., Piedra, M. E., Martínez-A, C., Ros, M. A., and Torres, M. (2000) *Development* **127,** 3961–3970
38. Faralli, H., Martin, E., Coré, N., Liu, Q. C., Filippi, P., Dilworth, F. J., Caubit, X., and Fasano, L. (2011) *J. Biol. Chem.* **286,** 23498–23510
39. Mead, A. N., and Stephens, D. N. (2003) *J. Neurosci.* **23,** 9500–9507
40. Xing, Y., Nakamura, Y., and Rainey, W. E. (2009) *Mol. Cell. Endocrinol.* **300,** 43–50
41. Garitaonandia, I., Smith, J. L., Kupchak, B. R., and Lyons, T. J. (2009) *J. Recept. Signal. Transduct. Res.* **29,** 67–73
42. Lin, C. Y., Ström, A., Vega, V. B., Kong, S. L., Yeo, A. L., Thomsen, J. S., Chan, W. C., Doray, B., Bangarusamy, D. K., Ramasamy, A., Vergara, L. A., Tang, S., Chong, A., Bajic, V. B., Miller, L. D., Gustafsson, J. A., and Liu, E. T. (2004) *Genome Biol.* **5,** R66
43. Smeenk, L., van Heeringen, S. J., Koeppel, M., van Driel, M. A., Bartels, S. J., Akkers, R. C., Denissov, S., Stunnenberg, H. G., and Lohrum, M. (2008) *Nucleic Acids Res.* **36,** 3639–3654
44. Carroll, J. S., Meyer, C. A., Song, J., Li, W., Geistlinger, T. R., Eeckhoute, J., Brodsky, A. S., Keeton, E. K., Fertuck, K. C., Hall, G. F., Wang, Q., Bekiranov, S., Sementchenko, V., Fox, E. A., Silver, P. A., Gingeras, T. R., Liu, X. S., and Brown, M. (2006) *Nat. Genet.* **38,** 1289–1297
45. Carroll, J. S., and Brown, M. (2006) *Mol. Endocrinol.* **20,** 1707–1714
46. Fullwood, M. J., Liu, M. H., Pan, Y. F., Liu, J., Xu, H., Mohamed, Y. B., Orlov, Y. L., Velkov, S., Ho, A., Mei, P. H., Chew, E. G., Huang, P. Y., Welboren, W. J., Han, Y., Ooi, H. S., Ariyaratne, P. N., Vega, V. B., Luo, Y., Tan, P. Y., Choy, P. Y., Wansa, K. D., Zhao, B., Lim, K. S., Leow, S. C., Yow, J. S., Joseph, R., Li, H., Desai, K. V., Thomsen, J. S., Lee, Y. K., Karuturi, R. K., Herve, T., Bourque, G., Stunnenberg, H. G., Ruan, X., Cacheux-Rataboul, V., Sung, W. K., Liu, E. T., Wei, C. L., Cheung, E., and Ruan, Y. (2009) *Nature* **462,** 58–64
47. Bollig, F., Perner, B., Besenbeck, B., Köthe, S., Ebert, C., Taudien, S., and Englert, C. (2009) *Development* **136,** 2883–2892
48. Gillespie, R. F., and Gudas, L. J. (2007) *J. Biol. Chem.* **282,** 33421–33434
49. Gillespie, R. F., and Gudas, L. J. (2007) *J. Mol. Biol.* **372,** 298–316
50. Lalevée, S., Bour, G., Quinternet, M., Samarut, E., Kessler, P., Vitorino, M., Bruck, N., Delsuc, M. A., Vonesch, J. L., Kieffer, B., and Rochette-Egly, C. (2010) *FASEB J.* **24,** 4523–4534
51. Boudjelal, M., Taneja, R., Matsubara, S., Bouillet, P., Dolle, P., and Chambon, P. (1997) *Genes Dev.* **11,** 2052–2065
52. Oulad-Abdelghani, M., Bouillet, P., Chazaud, C., Dollé, P., and Chambon, P. (1996) *Exp. Cell Res.* **225,** 338–347
53. Oulad-Abdelghani, M., Chazaud, C., Bouillet, P., Sapin, V., Chambon, P., and Dollé, P. (1997) *Dev. Dyn.* **210,** 173–183
54. Lickert, H., and Kemler, R. (2002) *Dev. Dyn.* **225,** 216–220
55. Tabariès, S., Lapointe, J., Besch, T., Carter, M., Woollard, J., Tuggle, C. K., and Jeannotte, L. (2005) *Mol. Cell. Biol.* **25,** 1389–1401
56. Coulombe, Y., Lemieux, M., Moreau, J., Aubin, J., Joksimovic, M., Bérubé-Simard, F. A., Tabariès, S., Boucherat, O., Guillou, F., Larochelle, C., Tuggle, C. K., and Jeannotte, L. (2010) *PLoS One* **5,** e10600
57. Balmer, J. E., and Blomhoff, R. (2005) *J. Steroid Biochem. Mol. Biol.* **96,** 347–354

58. Bouillet, P., Sapin, V., Chazaud, C., Messaddeq, N., Décimo, D., Dollé, P., and Chambon, P. (1997) *Mech. Dev.* **63,** 173–186
59. Bouillet, P., Chazaud, C., Oulad-Abdelghani, M., Dollé, P., and Chambon, P. (1995) *Dev. Dyn.* **204,** 372–382
60. Oulad-Abdelghani, M., Bouillet, P., Décimo, D., Gansmuller, A., Heyberger, S., Dollé, P., Bronner, S., Lutz, Y., and Chambon, P. (1996) *J. Cell Biol.* **135,** 469–477
61. Meijsing, S. H., Pufall, M. A., So, A. Y., Bates, D. L., Chen, L., and Yamamoto, K. R. (2009) *Science* **324,** 407–410
62. Ross-Innes, C. S., Stark, R., Holmes, K. A., Schmidt, D., Spyrou, C., Russell, R., Massie, C. E., Vowler, S. L., Eldridge, M., and Carroll, J. S. (2010) *Genes Dev.* **24,** 171–182

Cullin 3 mediates SRC-3 ubiquitination and degradation to control the retinoic acid response

Christine Ferry[a], Samia Gaouar[a], Benoit Fischer[b], Marcel Boeglin[c], Nicodeme Paul[d], Eric Samarut[a], Aleksandr Piskunov[a], Gabriella Pankotai-Bodo[a], Laurent Brino[b], and Cecile Rochette-Egly[a,1]

[a]Department of Functional Genomics and Cancer, [b]High Throughput Screening Facility, [d]Bioinformatics Platform; and [c]Imaging Platform, Institut de Génétique et de Biologie Moléculaire et Cellulaire, Institut National de la Santé et de la Recherche Médicale U964, Centre National de la Recherche Scientifique, Unité Mixte de Recherche 7104, Université de Strasbourg, BP 10142, 67404 Illkirch Cedex, France

Edited by Johan Auwerx, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland, and accepted by the Editorial Board November 1, 2011 (received for review February 16, 2011)

SRC-3 is an important coactivator of nuclear receptors including the retinoic acid (RA) receptor α. Most of SRC-3 functions are facilitated by changes in the posttranslational code of the protein that involves mainly phosphorylation and ubiquitination. We recently reported that SRC-3 is degraded by the proteasome in response to RA. Here, by using an RNAi E3-ubiquitin ligase entry screen, we identified CUL-3 and RBX1 as components of the E3 ubiquitin ligase involved in the RA-induced ubiquitination and subsequent degradation of SRC-3. We also show that the RA-induced ubiquitination of SRC-3 depends on its prior phosphorylation at serine 860 that promotes binding of the CUL-3–based E3 ligase in the nucleus. Finally, phosphorylation, ubiquitination, and degradation of SRC-3 cooperate to control the dynamics of transcription. In all, this process participates to the antiproliferative effect of RA.

Retinoic acid (RA) influences cell differentiation, proliferation, and apoptosis through modifications in the expression of target genes. The transcription of RA target genes is a highly coordinated process that requires a well-defined cross-talk among RA nuclear receptors (RARs), basal transcription machinery, and several transcriptional coregulators including the p160 family of coactivators (SRC-1, SRC-2, and SRC-3) (1). For each transcriptional component, there is a fine-tuned code of posttranslational modifications that control their activity, partners' association/dissociation, localization, and turnover (2, 3). This regulation is especially true for the coactivator SRC-3, which is a key regulator of nuclear receptors, metabolic homeostasis, and cell proliferation. Indeed, much of its function is facilitated through changes in the posttranslational code of the protein including phosphorylation and several types of posttranslational modifications (2, 4, 5).

In response to RA, SRC-3 binds to RARs and then recruits a battery of coregulatory proteins such as chromatin remodelers and modifiers that act in a coordinated and combinatorial manner to decompact chromatin and direct the transcriptional machinery to the promoter. Recently, we demonstrated that, in response to RA, SRC-3 is degraded by the proteasome (6, 7). However, the underlying mechanism of SRC-3 degradation and its link with the transcription of RA target genes was still unclear. Here, in a high-throughput screen based on the use of a siRNA thematic library and chemical transfection to create transient gene knockdown in MCF7 cells, we identified cullin 3 (CUL-3) and the Ring protein RBX1 as components of the E3 ligase complex involved in SRC-3 ubiquitination and degradation. We also show that SRC-3 degradation is involved in the transcription of RAR target genes and in the antiproliferative action of RA, through a phosphorylation-dependent ubiquitination "code."

Results

CUL-3–Based E3 Ligase Controls RA-Induced Degradation of SRC-3. Given that in human MCF7 breast cancer cells, SRC-3 is degraded in response to RA by the 26S proteasome (Fig. 1*A*) (6), we addressed whether this process is regulated by ubiquitination. In immunoprecipitation experiments, SRC-3 was constitutively ubiquitinated in agreement with other reports (4) and ubiquitinated SRC-3 accumulated in the presence of the proteasome inhibitor MG132 (Fig. 1*B*). Ubiquitination was also enhanced in response to RA, either in the absence or presence of MG132 (Fig. 1*B*).

Then we aimed at investigating which E3-ubiquitin ligase is involved in the RA-induced ubiquitination and degradation of SRC-3. We performed a high-throughput screen based on the use of a siRNA thematic library to create transient gene knockdown in MCF7 cells. The screen was based on the immunofluorescence analysis of SRC-3 with specific antibodies. Through combining the imaging of cells in microtiter plates with powerful image analysis algorithms, the screen determines whether silencing of a specific E3 ligase reverses the RA-induced degradation of SRC-3.

First, the technique was validated by checking that the signal disappears upon knockdown of SRC-3 with specific siRNAs (Fig. 1*C*). Then kinetic experiments performed after RA addition indicated that SRC-3 degradation occurs within 3–5 h (Fig. 1*D*). This degradation process was reversed by siRNAs targeting subunits of the 20S core proteasome (PSMB1 and PSMB2) or the SUG-1 subunit of the 19S subcomplex (Fig. 1*E*) corroborating that it involves the 26S proteasome.

For the screen, we used a library of 111 siRNAs with four different siRNAs per target (Dataset S1). Upon statistical analysis of SRC-3 nuclear intensities displayed in the transfected cells, we determined two lists of candidate genes, differing by the level of selection stringency ($\alpha = 1.5$ and $\alpha = 2$ for maximum stringency). Seven potential hits validated by at least two siRNAs ($\alpha = 2$) were found, among which CUL-3 and RBX1 were highly significant (*P* values) and validated by 3 and 4 siRNAs, respectively (Dataset S1 and Fig. 1*F*). However, the screen did not identify with high confidence any of the other E3 ligases (CUL-1-skp1-Fbw7α and E6AP) previously reported to regulate SRC-3 degradation (4, 8) (Fig. 1*F*). Finally, CUL-3 silencing also reversed the increase in SRC-3 ubiquitination observed at 2.5 h after RA addition (before degradation) (Fig. 1*G*). In conclusion, our screen indicates that the RA-induced ubiquitination and degradation of SRC-3 involves a cullin-RING Ligase (CRL) assembled with CUL-3 (CRL3).

SRC-3 Is Phosphorylated at S860 Before Degradation. CRLs consist of three core components: a cullin scaffold protein, a RING domain protein (RBX1) that recruits the E2 conjugating enzyme, and a substrate-adaptor protein that binds target proteins (9). Given that CRL3s recognize substrates containing serine-rich

Author contributions: C.R.-E. designed research; C.F., S.G., M.B., E.S., A.P., and G.P.-B. performed research; C.F., B.F., M.B., N.P., and A.P. contributed new reagents/analytic tools; L.B. analyzed data; and C.R.-E. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. J.A. is a guest editor invited by the Editorial Board.

[1]To whom correspondence should be addressed. E-mail: cegly@igbmc.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1102572108/-/DCSupplemental.

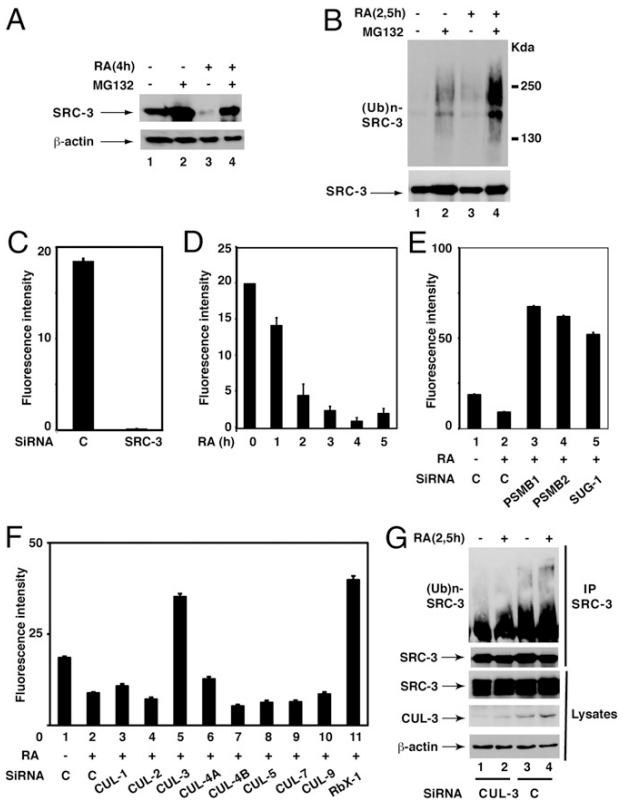


Fig. 1. Screening of the E3 ligase involved in the RA-induced degradation and ubiquitination of SRC-3. (*A* and *B*) Extracts from MCF7 cells treated or not with RA (0.1 μM) and MG132 (4 μM) were analyzed by immunoblotting for SRC-3 degradation and for SRC-3 ubiquitination after immunoprecipitation. (*C*) Silencing of SRC-3 abrogates the immunofluorescence signal obtained with SRC-3 antibodies. (*D*) RA induces the degradation of SRC-3 as assessed by the disappearance of the fluorescence signal. (*E*) The RA-induced degradation of SRC-3 is reversed with siRNAs targeting proteasome subunits. (*F*) In the high-throughput screen, siRNAs against CUL-3 and RBX1 reverse the degradation of SRC-3. Values are the mean ± SD of at least three different experiments. (*G*) Analysis of SRC-3 ubiquitination as in *B*, after CUL-3 silencing with specific siRNAs (50 nM).

domains that can be phosphorylated (10–12), we investigated whether the RA-induced ubiquitination and degradation of SRC-3 is controlled by phosphorylation.

SRC-3 depicts serine-rich motifs (Fig. 2*A*), and the amount of endogenous phosphorylated SRC-3 markedly increased 2 h after RA addition to MCF7 cells or transfected COS-1 cells before SRC-3 degradation (Fig. 2*B*). According to our previous studies (6, 13), p38MAPK is rapidly activated in response to RA through nongenomic effects (14) and phosphorylates SRC-3 (6) and several other targets. SRC-3 depicts 4 p38MAPK consensus phosphorylation sites: S505, S543, S860, and S867 (Fig. 2*A*) (5). Therefore, we investigated whether one of these residues was phosphorylated in response to RA. FLAG-tagged SRC-3 mutants with S505, S543, S860, or S867 substituted with alanines were constructed and overexpressed in COS-1 cells. The S505A, S543A, and S867A mutants were phosphorylated in response to RA as efficiently as WT SRC-3 (Fig. 2*C*). However, the S860A mutant was not phosphorylated (Fig. 2*C*, lanes 9 and 10), indicating that S860 is a target for RA signaling.

Next, antibodies recognizing specifically SRC-3 phosphorylated at S860 were generated and used in immunoblotting experiments after phosphoprotein affinity purification: A signal at the right position was detected in MCF7 cells and in COS-1 cells overexpressing SRC-3 WT (Fig. 2*D*). No signal was obtained with SRC-3 (S860A) (Fig. 2*D*, lanes 4–6), validating the specificity of the phosphospecific antibodies. Collectively these results indicate that SRC-3 becomes phosphorylated at S860 in response to RA.

Phosphorylation at S860 Is the Signal for SRC-3 Ubiquitination/Degradation and for CUL-3 Binding. Next, we analyzed whether S860 phosphorylation controls the ability of SRC-3 to be ubiquitinated and degraded in response to RA. Immunoprecipitation experiments were performed with MCF7 cells and our antibodies recognizing specifically SRC-3 phosphorylated at S860. The amount of phosphorylated SRC-3 increased in response to RA as well as the amount of ubiquitinated SRC-3 (Fig. 2*E*), and this effect was inhibited upon knockdown of CUL-3 (Fig. 2*F*). Finally, in contrast to SRC-3 WT, the S860A mutant was not degraded after RA addition (Fig. 2*G*). Thus, ubiquitination by CRL3 and degradation concern SRC-3 phosphorylated at S860.

Then we compared the ability of FLAG-SRC-3 to coimmunoprecipitate with HA-CUL-3 in transfected COS cells. SRC-3WT but not SRC-3 (S860A) was pulled down with CUL-3 in response to RA (Fig. 2*H*). Altogether these results indicate that the RA-induced ubiquitination and degradation of SRC-3 require a priming phosphorylation at S860 that controls the binding of CUL-3 complexes.

CUL-3 Migrates to the Nucleus in Response to RA. Then we analyzed the intracellular distribution of CUL-3 and SRC-3 in MCF7 cells by immunofluorescence and confocal analysis (Fig. 3*A*). In the absence of RA, SRC-3 was present essentially in nuclei (Fig. 3*A*, *a*) and CUL-3 in the cytoplasm (Fig. 3*A*, *c*). After RA treatment, CUL-3 accumulated at the perinuclear surface and in nuclei where it colocalized with SRC-3 (Fig. 3*A*, *d*, *f*, and *h*). That CUL-3 migrates to nuclei was corroborated by immunoblotting of nuclear extracts (Fig. 3*B*). Most interestingly, the amount of SRC-3 phosphorylated at S860 also increased in nuclei (Fig. 4*A*).

Next, to explore further the interaction between CUL-3 and SRC-3 phosphorylated at S860, we used a proximity ligation assay (PLA) (14, 15), which allows the in situ detection of interacting endogenous proteins. Rabbit anti–CUL-3 and mouse anti–phospho-SRC-3 antibodies were used followed by species-specific secondary antibodies, called PLA probes, each attached with a unique short DNA strand. When in close proximity, the DNA strands can be joined by a circle-forming DNA oligonucleotide. After amplification and revelation with labeled complementary oligonucleotide probes, the complexes are easily visible as bright red spots under a fluorescence microscope.

A few CUL-3/P-SRC-3 complexes were seen in the cytosol of control cells (Fig. 4*B*, *b* and *c*), in line with some constitutive phosphorylation of SRC-3. After RA treatment, the number of complexes increased in nuclei (Fig. 4*B*, *e* and *f*), suggesting that SRC-3 phosphorylation that occurs in nuclei would target CUL-3 to nuclei (16).

In contrast, the other cullins (CUL-1 and CUL-2) were present essentially in nuclei and RA did not affect their localization (Fig. S1), corroborating the specific role of CUL-3 in the RA response. RBX1 was nuclear either in the absence or presence of RA, in line with its participation to all CRL complexes (Fig. S2).

CUL-3 and SRC-3 Phosphorylation Are Required for the Transcription of RA-Target Genes. SRC-3 contributes to the transcription of RA receptor α (RARα)-target genes in several cell types (6, 7). Because the ubiquitin/proteasome pathway is implicated in the transcriptional functions of nuclear receptors and their coactivators (1, 4, 6), we investigated whether CUL-3 complexes are involved in the transcription of endogenous RARα target genes exemplified by the *Cyp26A1, Hoxa1,* and *Btg2* genes. In MCF7 cells, knockdown of CUL-3 or RBX1 decreased the RA-induced activation of the three genes (Fig. 5 *A–D*) as assessed by quantitative RT-PCR. Knockdown of the other cullins, CUL-1 and CUL-2, had no significant effects (Fig. 5 *A–D*). Overexpression

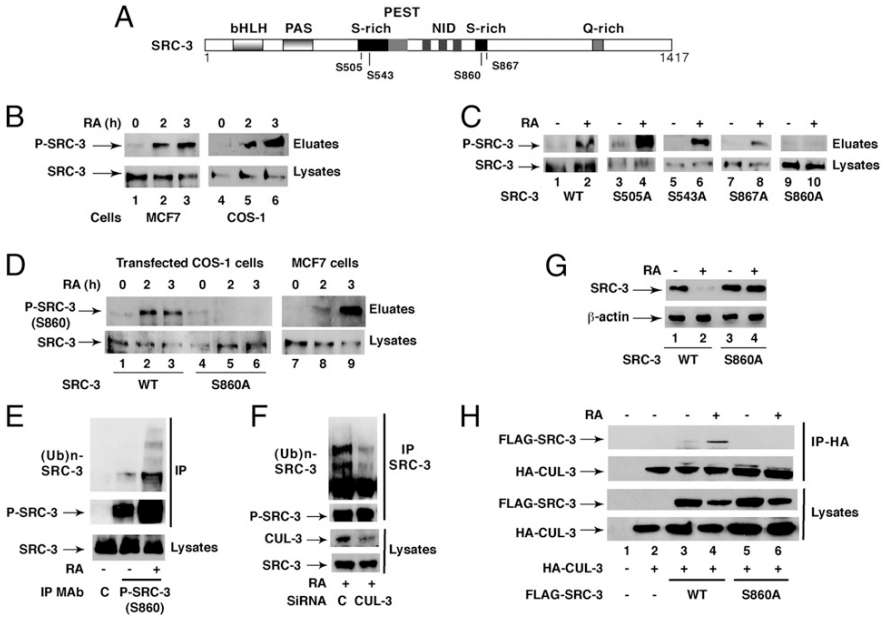


Fig. 2. Phosphorylation at S860 is the signal for SRC-3 ubiquitination and degradation in response to RA. (*A*) Schematic representation of SRC-3 with the main phosphorylation sites. (*B*) Kinetics of SRC-3 phosphorylation in RA-treated MCF7 and transfected COS-1 cells, after phosphoprotein affinity purification and immunoblotting. (*C*) Analysis as in *B* of the phosphorylation of the SRC-3 mutants in transfected COS-1 cells. (*D*) Kinetics of SRC-3 phosphorylation at S860 after phosphoprotein affinity purification and immunoblotting with antibodies recognizing specifically SRC-3 phosphorylated at this residue. (*E*) S860 phosphorylation controls SRC-3 ubiquitination, as assessed by immunoprecipitation of MCF7 cells extracts with the phospho-antibodies. (*F*) Knockdown of CUL-3 diminishes the ubiquitination of phosphorylated SRC-3 immunoprecipitated as in *E*. (*G*) Immunoblots showing that in transfected COS-1 cells, SRC-3WT but not SRC-3 (S860A) is degraded. (*H*) In transfected COS-1 cells, HA-CUL-3 coimmunoprecipitates with FLAG-SRC-3 WT but not with the S860A mutant.

of CUL-3 also had no effect, indicating that CUL-3 is not in limiting amounts in MCF7 cells (Fig. S3). Altogether these results corroborate the importance of a CRL assembled with CUL-3 in the expression of RARα-target genes.

Next, because SRC-3 ubiquitination by the CUL-3 complex depends on the prior phosphorylation of SRC-3, the relevance of SRC-3 phosphorylation at S860 for the transcription of RA target genes was investigated in COS-1 cells transfected with RARα and a DR5-tk-CAT reporter gene. CAT activity increased in response to RA and overexpression of SRC-3 WT but not of SRC-3(S860A) enhanced this effect (Fig. 5*E*), corroborating that SRC-3 phosphorylation facilitates transcription (6).

SRC-3 Ubiquitination and Degradation Occur out of Chromatin. To substantiate the role of the phosphorylation-dependent ubiquitination of SRC-3 in transcription, we correlated the phosphorylation and ubiquitination state of SRC-3 to the recruitment of the coactivator to RARα-target promoters in chromatin immunoprecipitation (ChIP) experiments. In MCF7 cells, SRC-3 was rapidly recruited concomitantly with RARα, to the RA response element (RARE) located in the promoter of the *Btg2* gene, and to both the proximal and distal RAREs located in the promoter of the *Cyp26A1* gene, with a peak at 1 h after RA addition (Fig. 6 *A* and *B*) and a decrease at 2 h. The decrease occurred when active phosphorylated p38MAPK was recruited (Fig. 6 *B* and *D*) and when SRC-3 was phosphorylated (Fig. 6*C*), ubiquitinated, and degraded (Fig. 1), raising the hypothesis that these latter processes might play a role in clearing out SRC-3 from the promoters.

Therefore, we investigated whether phosphorylation, ubiquitination, and degradation concern SRC-3 associated to chromatin or out of chromatin. After preparation of highly purified intact nuclei from MCF7 cells, insoluble chromatin and soluble nucleoplasm were separated and subjected to immunoprecipitation with our phosphospecific antibodies. In chromatin, SRC-3 was only transiently phosphorylated 2 h after RA addition (Fig. 6*E*) when its promoter occupancy decreased. However, no ubiquitination (Fig. 6*E*) and no degradation could be detected (Fig. 6*F*). In contrast, in soluble nucleoplasm, SRC-3 was more abundant and became markedly phosphorylated, ubiquitinated, and degraded in response to RA (Fig. 6 *E* and *F*). Collectively, these results suggest that ubiquitination and degradation of phosphorylated SRC-3 occur out of chromatin. Indeed, we demonstrated that S860 phosphorylation induces the dissociation of SRC-3 from RARα (6). Moreover, CUL-3 and RBX1 were detected in nucleoplasmic extracts (Fig. 6*G*) and not in chromatin.

SRC-3 Is Not Phosphorylated nor Degraded in erbB-2 Positive Cells. Given the importance of SRC-3 phosphorylation/degradation in the dynamics of RARα target genes transcription, we examined whether this process is affected in erbB-2 positive breast cancer cells that are characterized by aberrant kinase pathways downstream of erbB-2 (17). In these cells, exemplified by the BT474 cell line, SRC-3 is overexpressed (Fig. 7*A*) (18) and the nongenomic effects of RA, i.e., the activation of the p38MAPK pathway, were abrogated (Fig. 7*B*) (14). Consequently, SRC-3 was not phosphorylated at S860 (Fig. 7*D*, *c* and *d*) nor degraded

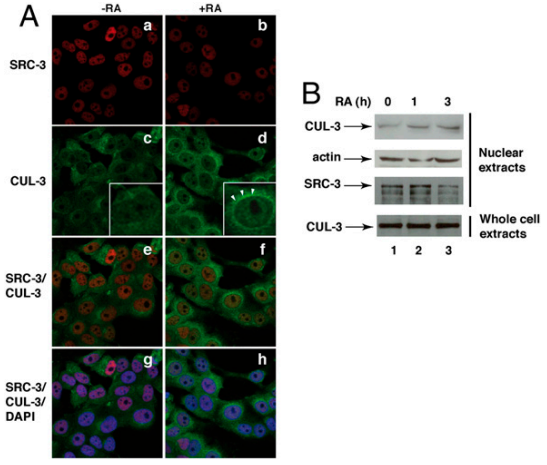


Fig. 3. CUL-3 migrates to nuclei in response to RA. (*A*) MCF7 cells treated (*Right*) or not (*Left*) with RA for 2.5 h were triple stained with DAPI (blue), SRC-3 (red; *a* and *b*), and CUL-3 (green; *c* and *d*) antibodies and examined by confocal microscopy. The merge images overlapping the red and green (*e* and *f*) or the red, green, and blue fluorescence (*g* and *h*) are shown. (*B*) Immunoblots showing that CUL-3 levels increase in the nuclei of RA-treated cells. The arrows in *d* show the accumulation of CUL-3 at the perinuclear surface.

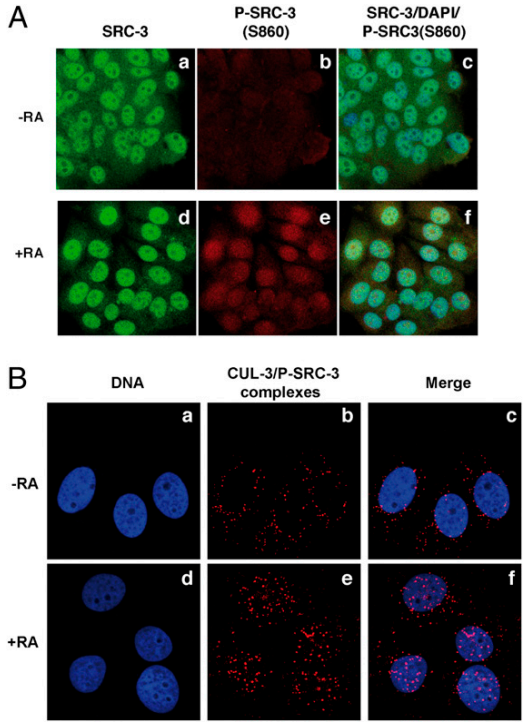


Fig. 4. SRC-3 phosphorylated at S860 interacts with CUL-3 in nuclei. (*A*) MCF7 cells treated (*d–f*) or not (*a–c*) with RA for 2 h were triple-stained with DAPI (blue), and antibodies recognizing SRC-3 (green) or its phosphorylated form (red) and examined by confocal microscopy. The merge images overlapping the red, green, and blue fluorescence are shown (*c* and *f*). (*B*) Proximity ligation assay showing the CUL-3/P-SRC-3(S860) complexes (red; *b* and *e*) in MCF7 cells. The merge between blue and red is shown (*c* and *f*).

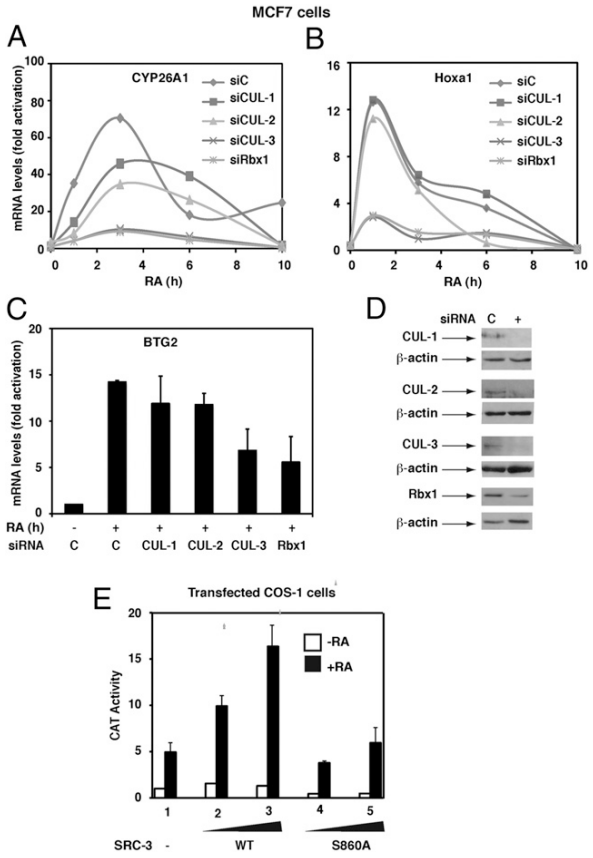


Fig. 5. CUL-3 participates in the transcription of RA-target genes. (*A–C*) Silencing of CUL-3 or RBX1 decreases the RA-induced expression of the *Cyp26A1*, *Hoxa1*, and *Btg2* genes as monitored by quantitative RT-PCR. Values, expressed as fold induction relative to untreated cells, correspond to a representative experiment among three or are the mean ± SD of three different experiments. (*D*) Knockdown efficiency was controlled by immunoblotting. (*E*) CAT activity in COS-1 cells transfected with the SRC-3WT or SRC-3(S860A) vectors along with RARα and the DR5-tk-CAT reporter gene and RA-treated for 6 h. Results are the mean ± SD of three experiments.

(Fig. 7 *C* and *D*, *a* and *b*) and the classical RA target genes (*Cyp26A1*, *Btg2,* and *Hoxa1*) were not regulated by RA (Fig. S4). Similar observations were made with another erbB-2 positive cell line, MDA-MB361 (Fig. S5 *A–C*). Collectively, these results corroborate the importance of the p38MAPK pathway and the subsequent phosphorylation, ubiquitination, and degradation of SRC-3 in the RA response.

SRC-3 Degradation by a CUL-3–Based E3 Ligase Contributes to the Antiproliferative Effect of RA. We next asked whether SRC-3 levels and/or SRC-3 degradation directly influence cell growth. MCF7 cells respond to RA through a decrease in their proliferation rate (Fig. 7*E*, *Left*). Knockdown of CUL-3 did not affect MCF7 cells growth but abrogated the decrease observed after RA addition (Fig. 7*F*, *Left*). In contrast, the erbB-2 positive BT474 and MDA-MB361 cell lines were both resistant to the antiproliferative action of RA (Fig. 7*E*, *Right*, and Fig. S5*D*) (17) and knockdown of CUL-3 had no effect, either in the presence or absence of RA (Fig. 7*F*, *Right*, and Fig. S5*D*). Collectively, these observations highlight the importance of SRC-3 phosphorylation and turnover in the antiproliferative effect of RA. Of note is that knockdown of SRC-3 decreased significantly the proliferation rate of the three cell lines (Fig. 7*F* and Fig. S5*D*), suggesting that a reduction in SRC-3 levels inhibits cell growth. However, knockdown of SRC-3 did not restore the RA sensitivity of the BT474 and MDA-MB361 cells, in terms of cell proliferation (Fig. 7*F* and Fig. S5*D*).

Discussion

SRC-3 is a model coactivator of nuclear receptors for studying the influence of posttraductional modifications. Indeed, SRC-3 has been shown to be phosphorylated at several residues by different kinases in response to different signaling pathways (2, 5). Moreover, our laboratory demonstrated that in response to RA, SRC-3 is phosphorylated by p38MAPK and subsequently degraded by the proteasome (6). Here we identified S860 as the residue that is phosphorylated in response to RA and that promotes SRC-3 ubiquitination and degradation. We also expanded the repertoire of SRC-3 E3 ligases by characterizing a CUL-3–based complex as the E3 ubiquitin ligase involved in the ubiquitination/degradation of SRC-3 in a RA and phospho-dependent manner.

Classically, CUL-3–based complexes recognize their substrate through an adaptor containing a Bric-a Brac/Tramtrack/Broad (BTB) domain (19). The human genome encodes 190 BTB proteins (20), which were not included in our screen, but one can speculate that SRC-3 ubiquitination involves a CRL3 with a BTB protein that recognizes phosphorylated motifs (10, 11), as described for CUL-1 complexes (4, 21), or a nearby domain created by S860 phosphorylation through an allosteric mechanism (22). However, one cannot exclude that CUL-3 interacts directly with the serine-rich domain of SRC-3 containing phosphorylated S860, independently of any adaptor (12).

It is worth noting that SRC-3 can be ubiquitinated by other CRLs such as CUL-1-skp1-Fbw7α, in response to other signal kinase pathways and to estrogens via other phosphorylated domains (4). When this manuscript was submitted, another study

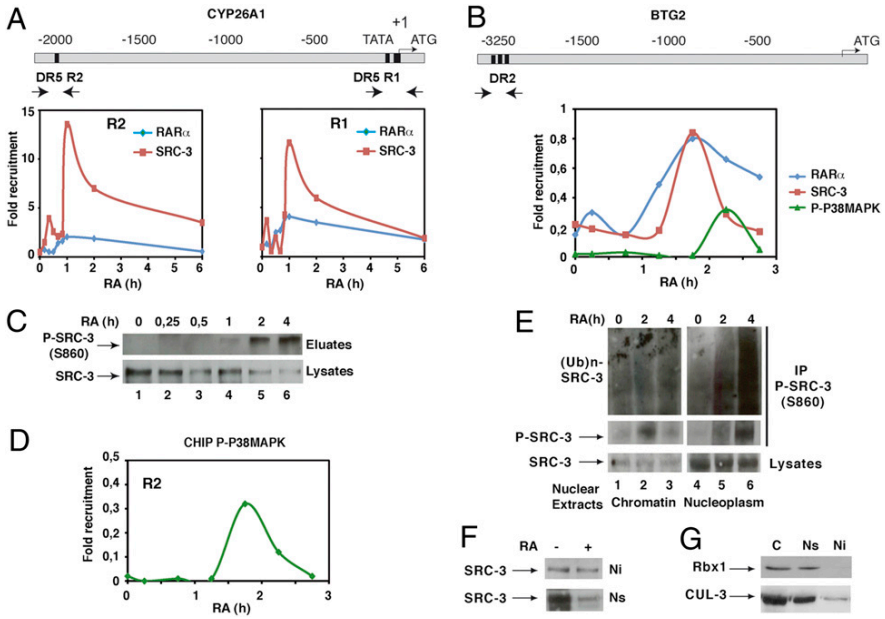


Fig. 6. RA-induced SRC-3 ubiquitination and degradation occur out of chromatin. (*A*) Kinetic ChIP experiments performed with RA-treated MCF7 cells and showing the recruitment of SRC-3 and RARα to the R1 and R2 regions of the *Cyp26A1* gene promoter. Values are expressed as fold enrichment relative to untreated cells and are the mean of three distinct experiments. (*B*) Recruitment of RARα, SRC-3, and active p38MAPK to the *Btg2* gene promoter. (*C*) Kinetics of SRC-3 phosphorylation at S860 after phosphoprotein affinity purification and immunoblotting with the phospho-antibodies. (*D*) Recruitment of active phosphor-p38MAPK to the R2 region of the Cyp26A1 promoter. (*E*) Soluble nucleoplasmic (Ns) and insoluble chromatin (Ni) extracts were immunoprecipitated with the phospho-SRC-3 antibodies and immunoblotted with SRC-3 or ubiquitin antibodies. (*F*) Ni and Ns extracts were analyzed for SRC-3 degradation by immunoblotting. (*G*) Cytoplasmic (C), Ns, and Ni extracts were compared for the presence of RBX1 and CUL-3 by immunoblotting.

was issued, reporting also a role for a CRL3 complex in SRC-3 ubiquitination and proteolysis (23). However, in this study, CUL-3 was recruited through the BTB protein SPOP (speckle-typePOZ protein) to a different motif that was phosphorylated by a different kinase. Thus, different signals and/or phosphorylation of distinct serine residues can select distinct E3 ligases to regulate SRC-3 ubiquitination/degradation and, thus, SRC-3 levels.

The other interesting point of this study is that ubiquitination and degradation of SRC-3 occur out of chromatin and that these processes, together with CUL-3, are required for the expression of RA target genes. Therefore, we proposed a model (Fig. 8) in which the RA-induced phosphorylation of DNA-bound SRC-3 at S860 is the signal that promotes the dissociation of SRC-3 from chromatin (6) and SRC-3 interaction with CUL-3 complexes out of chromatin. Then SRC-3 is ubiquitinated and degraded. Thus, one can suggest that ubiquitination and degradation cooperate with phosphorylation to clear SRC-3 out of the promoters according to the model proposed for another CUL-3 target (10) so that other coregulators can come and participate to transcription (3), in line with the dynamics of transcription. Corroborating this model, in erbB-2 positive breast cancer cell lines, where the p38MAPK pathway is not activated by RA, SRC-3 is not phosphorylated nor degraded and most of the classical RA target genes are not regulated.

Finally, our knockdown experiments indicate that SRC-3 participates to cell growth and that SRC-3 degradation via a CUL-3 complex is involved in the antiproliferative action of RA. Such results highlight the importance of SRC-3 phosphorylation

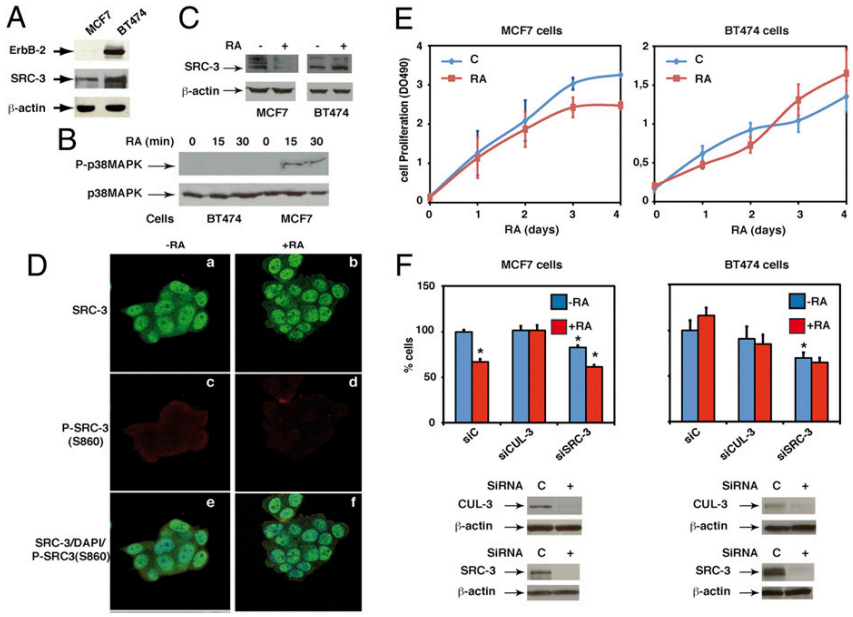


Fig. 7. SRC-3 is not phosphorylated nor degraded in erbB-2 positive breast cancer cells. (*A–C*) Immunoblots showing that in BT474 cells, erbB-2, and SRC-3 are overexpressed, p38MAPK is not activated, and SRC-3 not degraded. (*D*) In BT474 cells, SRC-3 is not phosphorylated at S860 in confocal microscopy experiments performed as in Fig. 4*A*. (*E*) RA decreases the proliferation rate of MCF7 cells but not of BT-474 cells. (*F*) MCF7 (*Left*) and BT474 (*Right*) cells, RA-treated or not, were analyzed for proliferation after knockdown of CUL-3 and SCRC-3. Knockdown efficiency was checked by immunoblotting. Results are the mean ± SD of two distinct experiments performed in quadruplate. Statistically significant differences are indicated (*$P <$ 0.05, control versus RA or siRNA).

CELL BIOLOGY

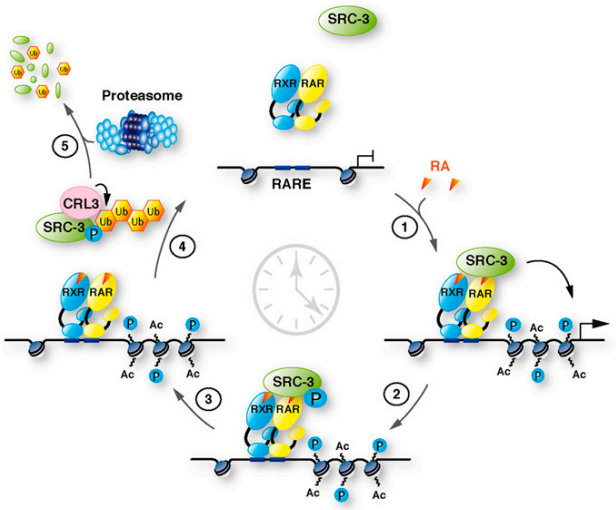


Fig. 8. Working model for the role of SRC-3 phosphorylation, ubiquitination, and degradation in RA target genes transcription. In response to RA, RARα and SRC-3 are recruited to target genes promoters to initiate transcription (1). Then SRC-3 becomes phosphorylated at S860, dissociates from RARα and DNA (2), and interacts with CUL-3 complexes that promote its ubiquitination (3) and degradation by the proteasome (4).

and turnover in the RA response. Thus, one can predict that cancers characterized by aberrant signaling pathways (24) would be RA resistant. In line with this hypothesis, erbB-2 positive breast cancer cell lines, in which SRC-3 is not degraded in response to RA, are resistant to the antiproliferative effect of RA. Note that in such cells, overexpression of CUL-3 markedly inhibited cell growth (Fig. S5*E*), most probably through the ability of the CUL-3–based ubiquitin ligases, which are known to function as breast cancer tumor suppressors, to target several signaling proteins for ubiquitination and degradation (25).

Remarkably, SRC-3 is a coactivator not only for RARs, but also for several nuclear receptors such as the estrogen receptor (3). Because several RAR regulated genes cross-talk with estrogen signaling (26, 27), one can speculate that the RA-induced degradation of SRC-3 would influence RA target genes at the cost of estrogen-ER function. With estrogen being the predominant hormone involved in the proliferation of breast cancer cells, SRC-3 segradation might be part of the RA rationale in the treatment of breast cancer with functional signaling pathways (28). Reciprocally, according to our results, one cannot exclude that the degradation of SRC-3 that occurs in response to estrogens (4) might potentiate the antiproliferative action of RA.

In conclusion, our work highlights the importance of phosphorylation and ubiquitination processes in the regulation of RA target genes through the control of SRC-3 turnover. It also reveals that RA resistance may be correlated, at least in part, to the deregulation of these processes.

Materials and Methods

Plasmids, reagents, antibodies, and cell lines are described in *SI Experimental Procedures*.

Complete details for the RNAi E3 Ubiquitine Ligase screen, immunoblotting, immunoprecipitation, chromatin immunoprecipitation, qRT-PCR, immunofluorescence analysis, and proximity ligation assay (PLA) are also described in the *SI Experimental Procedures*.

ACKNOWLEDGMENTS. We thank Dr. W. Krek for the cullin vectors; Dr. B. W. O'Malley for the phosphorylation defective SRC-3 mutants; M. Oulad Abdelghani (Institut de Génétique et de Biologie Moléculaire et Cellulaire; IGBMC) for the mouse monoclonal antibodies; J. M. Garnier (IGBMC) for constructs; Amélie Weiss and Laure Froidevaux from the high-throughput screening facility (IGBMC), the cell culture facilities; and Regis Lutzing for help. This work was supported by funds from Centre National de la Recherche Scientifique, Institut National de la Santé et de la Recherche Médicale, Agence Nationale pour la Recherche Grants ANR-05-BLAN-0390-02 and ANR-09-BLAN-0297-01, Association pour la Recherche sur le Cancer Grant ARC-07-1-3169, Fondation pour la Recherche Médicale (FRM) Grant DEQ20090515423), and Institut National du Cancer Grants INCa-PL09-194 and PL07-96099. C.F. and E.S. were supported by the Ministère de l'Enseignement Supérieur et de la Recherche and A.P. by FRM and the Lady TATA Memorial Trust. Association pour la Recherche sur le Cancer supported G.P.-B. and C.F.

1. Rochette-Egly C, Germain P (2009) Dynamic and combinatorial control of gene expression by nuclear retinoic acid receptors. *Nucl Recept Signal* 7:e005.
2. Han SJ, Lonard DM, O'Malley BW (2009) Multi-modulation of nuclear receptor coactivators through posttranslational modifications. *Trends Endocrinol Metab* 20:8–15.
3. O'Malley BW, Qin J, Lanz RB (2008) Cracking the coregulator codes. *Curr Opin Cell Biol* 20:310–315.
4. Wu RC, Feng Q, Lonard DM, O'Malley BW (2007) SRC-3 coactivator functional lifetime is regulated by a phospho-dependent ubiquitin time clock. *Cell* 129:1125–1140.
5. Wu RC, et al. (2004) Selective phosphorylations of the SRC-3/AIB1 coactivator integrate genomic reponses to multiple cellular signaling pathways. *Mol Cell* 15: 937–949.
6. Gianni M, et al. (2006) P38MAPK-dependent phosphorylation and degradation of SRC-3/AIB1 and RARalpha-mediated transcription. *EMBO J* 25:739–751.
7. Ferry C, et al. (2009) SUG-1 plays proteolytic and non-proteolytic roles in the control of retinoic acid target genes via its interaction with SRC-3. *J Biol Chem* 284:8127–8135.
8. Mani A, et al. (2006) E6AP mediates regulated proteasomal degradation of the nuclear receptor coactivator amplified in breast cancer 1 in immortalized cells. *Cancer Res* 66:8680–8686.
9. Petroski MD, Deshaies RJ (2005) Function and regulation of cullin-RING ubiquitin ligases. *Nat Rev Mol Cell Biol* 6:9–20.
10. Spoel SH, et al. (2009) Proteasome-mediated turnover of the transcription coactivator NPR1 plays dual roles in regulating plant immunity. *Cell* 137:860–872.
11. Zhang Q, et al. (2009) Multiple Ser/Thr-rich degrons mediate the degradation of Ci/Gli by the Cul3-HIB/SPOP E3 ubiquitin ligase. *Proc Natl Acad Sci USA* 106:21191–21196.
12. Xing H, Hong Y, Sarge KD (2010) PEST sequences mediate heat shock factor 2 turnover by interacting with the Cul3 subunit of the Cul3-RING ubiquitin ligase. *Cell Stress Chaperones* 15:301–308.
13. Bruck N, et al. (2009) A coordinated phosphorylation cascade initiated by p38MAPK/MSK1 directs RARalpha to target promoters. *EMBO J* 28:34–47.
14. Piskunov A, Rochette-Egly C (2011) A retinoic acid receptor RARα pool present in membrane lipid rafts forms complexes with G protein αQ to activate p38MAPK. *Oncogene*, in press.
15. Söderberg O, et al. (2006) Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat Methods* 3:995–1000.
16. Welcker M, Larimore EA, Frappier L, Clurman BE (2011) Nucleolar targeting of the fbw7 ubiquitin ligase by a pseudosubstrate and glycogen synthase kinase 3. *Mol Cell Biol* 31:1214–1224.
17. Tari AM, Lim SJ, Hung MC, Esteva FJ, Lopez-Berestein G (2002) Her2/neu induces all-trans retinoic acid (ATRA) resistance in breast cancer cells. *Oncogene* 21:5224–5232.
18. Xu J, Wu RC, O'Malley BW (2009) Normal and cancer-related functions of the p160 steroid receptor co-activator (SRC) family. *Nat Rev Cancer* 9:615–630.
19. van den Heuvel S (2004) Protein degradation: CUL-3 and BTB—partners in proteolysis. *Curr Biol* 14:R59–R61.
20. Stogios PJ, Downs GS, Jauhal JJ, Nandra SK, Privé GG (2005) Sequence and structural analysis of BTB domain proteins. *Genome Biol* 6:R82.
21. Hao B, Oehlmann S, Sowa ME, Harper JW, Pavletich NP (2007) Structure of a Fbw7-Skp1-cyclin E complex: Multisite-phosphorylated substrate recognition by SCF ubiquitin ligases. *Mol Cell* 26:131–143.
22. Samarut E, et al. (2011) Evolution of nuclear retinoic acid receptor alpha (RARα) phosphorylation sites. Serine gain provides fine-tuned regulation. *Mol Biol Evol* 28: 2125–2137.
23. Li C, et al. (2011) Tumor-suppressor role for the SPOP ubiquitin ligase in signal-dependent proteolysis of the oncogenic co-activator SRC-3/AIB1. *Oncogene* 30: 4350–4364.
24. Blume-Jensen P, Hunter T (2001) Oncogenic kinase signalling. *Nature* 411:355–365.
25. Emanuele MJ, et al. (2011) Global identification of modular cullin-RING ligase substrates. *Cell* 147:459–474.
26. Hua S, Kittler R, White KP (2009) Genomic antagonism between retinoic acid and estrogen signaling in breast cancer. *Cell* 137:1259–1271.
27. Ross-Innes CS, et al. (2010) Cooperative interaction between retinoic acid receptor-alpha and estrogen receptor in breast cancer. *Genes Dev* 24:171–182.
28. Darro F, et al. (1998) Growth inhibition of human in vitro and mouse in vitro and in vivo mammary tumor models by retinoids in comparison with tamoxifen and the RU-486 anti-progestagen. *Breast Cancer Res Treat* 51:39–55.

Supporting Information

Ferry et al. 10.1073/pnas.1102572108

SI Materials and Methods

Plasmids and Reagents. The B10-HA-FLAG–tagged SRC-3 and RARα vectors and the DR5-tk-CAT reporter gene were described (1, 2). The expression vectors for SRC-3 S505A, S543A, S860A, and S867A were constructed by cloning the cDNA of the corresponding mutants provided by B. W. O'Malley (Baylor College of Medicine, Houston, TX) into a pTL2 vector containing a FLAG-tag. The pCDNA-based expression vectors for CUL-1, CUL-2, and CUL-3 with an HA tag were kindly provided by W. Krek (Swiss Federal Institute of Technology, Zurich, Switzerland).

Antibodies. The rabbit polyclonal antibodies used were against RARα, SRC-3, β-actin, CUL-1, and CUL-2 (all from Santa Cruz Biotechnologies), CUL-3 (Sigma-Aldrich), and phospho-p38 MAPK (Thr180/Tyr182) (Cell Signaling). Goat polyclonal antibodies against RBX1 were from Santa Cruz Biotechnologies. Mouse monoclonal antibodies were against FLAG-tag (Sigma-Aldrich), HA-Tag, SRC-3 (BD-Bioscience), and mono- and polyubiquitinated conjugates (Enzo Life Sciences). Mouse monoclonal antibodies recognizing SRC-3 phosphorylated at S860 were generated by immunization of BALB/c mice with synthetic phosphopeptides (3).

RNAi E3-Ubiquitin Ligase Screen. Screening was performed at the high-throughput screening facility of the Institut de Génétique et de Biologie Moléculaire et Cellulaire, using a siRNA library (4 individual siRNA per protein) targeting 111 E3 Ubiquitin Ligases and associated proteins (Flexiplate siRNA sets; Qiagen) (Dataset S1). Controls were performed with smartpool siRNAs from Dharmacon (Thermo Scientific) [siRNAs against human SRC-3 (M-003759-02) and SUG-1 (M-009484-02) and scramble siRNA] or from Qiagen [siRNAs against CUL-1 (Hs_Cul1_5), CUL-2 (Hs_Cul2_3), CUL-3 (Hs_Cul3_5), RBX1 (Hs_RBX1_5), PSMB1 (Hs_PSMB1_2), and PSMB2 (Hs_PSMB2_2)]. For each target, 3.5 pmol siRNA was reverse transfected in 5,000 MCF7 cells per 0.3 cm^2 by using Interferin (Polyplus).

All screens were performed in 96-well cell culture microplates with a particular focus on avoiding microplate edge effects. SRC-3 degradation was analyzed by immunofluorescence 3 d after siRNA transfection and 3 h after RA addition. Cells were washed, fixed with 3% paraformaldehyde, permeabilized with 0.1% Triton X-100, blocked with 2% BSA, and incubated with SRC-3 antibodies followed by ALEXAFluor 488-conjugated second antibodies (Invitrogen). The screens were achieved owing to a TECAN robotic station (for cell transfection, staining, and immunocytochemistry) and to a Caliper Twister II robotic arm coupling microplate stacks to the InCELL1000 analyzer microscope (GE LifeSciences).

Statistical Analysis of the siRNA Screens. After background-correction and interplate normalization, internal positive and negative plate controls were used to determine a threshold for cell classification. The distance from negative controls was quantified by determining the percent of control (i.e., the percentage of positive cells in the test experiment versus the negative control). After multiple testing correction, P values were determined. Then gene selection was based on the one-sided tests procedure with H0: $Xp \leq \alpha$ and H1: $Xp > \alpha$, where Xp is the percent of control estimator and α is a sensibility parameter ($\alpha \geq 1$ and $\alpha = 2$ for high stringency).

Cell Lines, Cell Proliferation, Transfection, Immunoprecipitation, and Immunoblotting. COS-1, MCF7, BT474, and MDA-MB361 cells were cultured and transiently transfected under standard conditions. After RA (10^{-7} M) addition (4), whole-cell extracts or subcellular fractions (cytosol, nucleoplasm, and chromatin) were prepared, immunoblotted, and immunoprecipitated (2, 3). Cell proliferation was analyzed by using the XTT (2,3-bis-(2-methoxy-4-nitro-5-sulfophenyl)-^{2}H-tetrazolium-5-carboxanilide) assay kit (Roche Diagnostics).

Detection of in Vivo Phosphorylated SRC-3. Cell extracts were applied to phosphoprotein purification columns (Qiagen). Column eluates containing protein peaks were concentrated and analyzed by immunoblotting (4).

Immunofluorescence Analysis and Proximity Ligation Assay (PLA). Cells were seeded onto labtek chamberslides (Fisher Scientific), proceeded as described for the RNAi screen, and incubated with the primary antibodies followed by ALEXAFluor 448- or 555-conjugated secondary antibodies or by the PLA probes (Duolink II; Eurogentec) as described (5). Images were edited by using ImageJ.

RNA isolation, quantitative RT-PCR, and chromatin immunoprecipitations (ChIP) were performed as described (2, 4). Primer sequences are available upon request.

1. Gianni M, et al. (2006) P38MAPK-dependent phosphorylation and degradation of SRC-3/AIB1 and RARalpha-mediated transcription. *EMBO J* 25:739–751.
2. Ferry C, et al. (2009) SUG-1 plays proteolytic and non-proteolytic roles in the control of retinoic acid target genes via its interaction with SRC-3. *J Biol Chem* 284:8127–8135.
3. Lalevée S, et al. (2010) Vinexinß, an atypical "sensor" of retinoic acid receptor gamma signaling: Union and sequestration, separation, and phosphorylation. *FASEB J* 24: 4523–4534.
4. Bruck N, et al. (2009) A coordinated phosphorylation cascade initiated by p38MAPK/MSK1 directs RARalpha to target promoters. *EMBO J* 28:34–47.
5. Piskunov A, Rochette-Egly C (2011) A retinoic acid receptor RARα pool present in membrane lipid rafts forms complexes with G protein αQ to activate p38MAPK. *Oncogene*, in press.

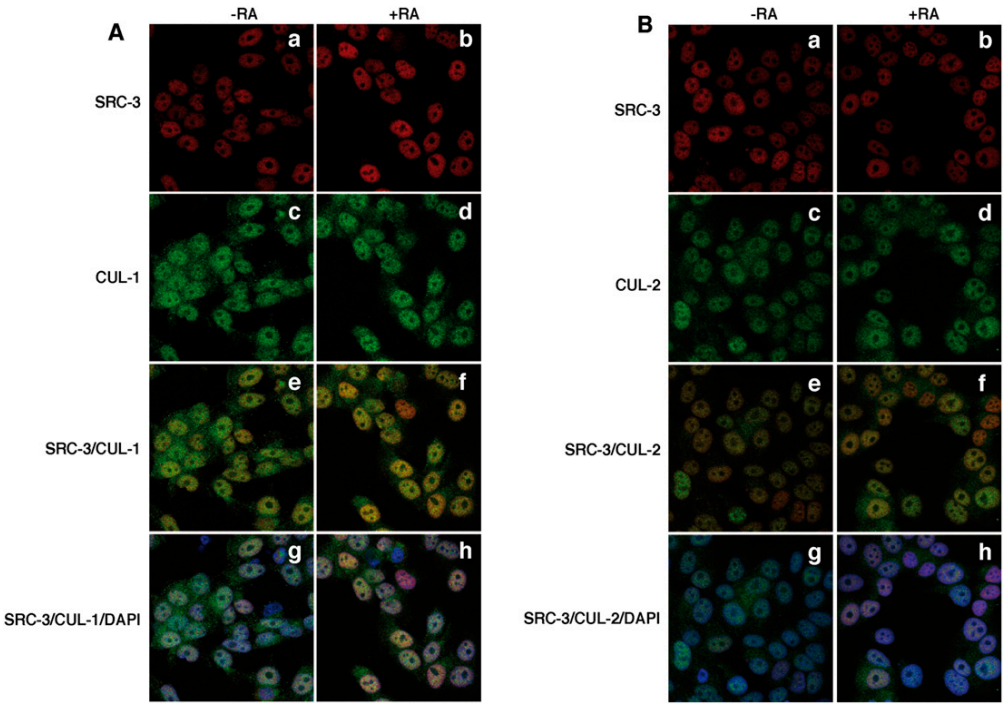


Fig. S1. The nuclear localization of CUL-1 and CUL-2 is not affected by RA. (*A*) MCF7 cells were RA-treated (*b, d, f,* and *h*) or not (*a, c, e,* and *g*) for 2.5 h, fixed, triple stained with DAPI (blue), SRC-3 antibodies (red), and CUL-1 antibodies (green), and examined by confocal microscopy. Merges overlapping the red and green (*e* and *f*) and the red, green, and blue (*g* and *h*) are shown. (*B*) Same as in *A* with CUL-2 antibodies.

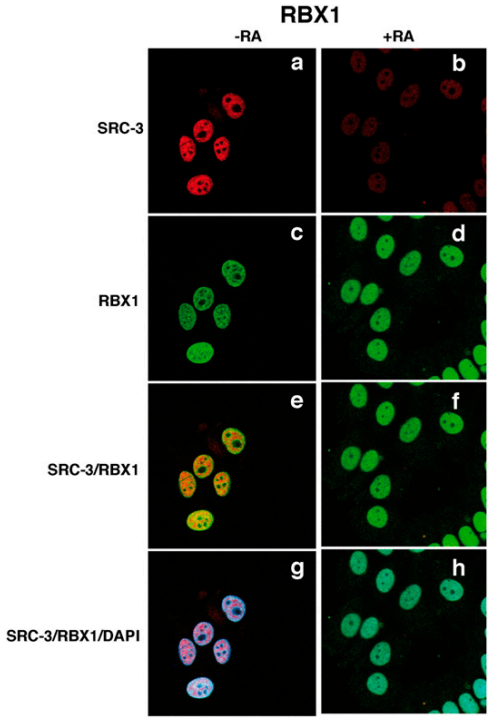


Fig. S2. The nuclear localization of Rbx1 is not affected by RA. MCF7 cells were RA-treated (*b, d, f,* and *h*) or not (*a, c, e,* and *g*) for 2.5 h, fixed, triple stained with DAPI (blue), SRC-3 antibodies (red), and Rbx1 antibodies (green), and examined by confocal microscopy. Merges overlapping the red and green (*e* and *f*) and the red, green, and blue (*g* and *h*) are shown.

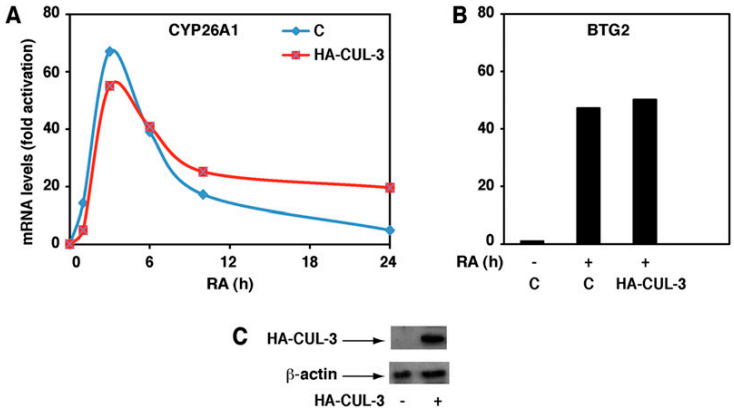


Fig. S3. The RA-induced expression of the *Cyp26A1* (*A*) and *Btg2* (*B*) genes is not affected by CUL-3 overexpression. Values are expressed as fold induction relative to untreated cells and correspond to a representative experiment among three. Overexpression efficiency was controlled by immunoblotting (*C*).

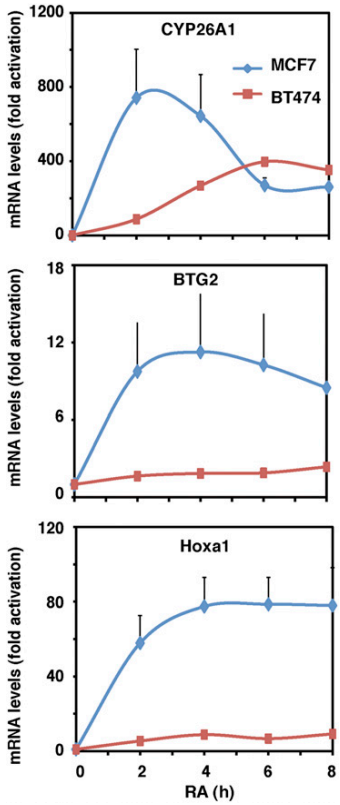


Fig. S4. Real-time RT-PCR experiments showing that in BT474 human breast cancer cells, the *Cyp26A1* (*Top*), *Btg2* (*Middle*), and *Hoxa1* (*Bottom*) genes are not induced, compared with MCF7 cells.

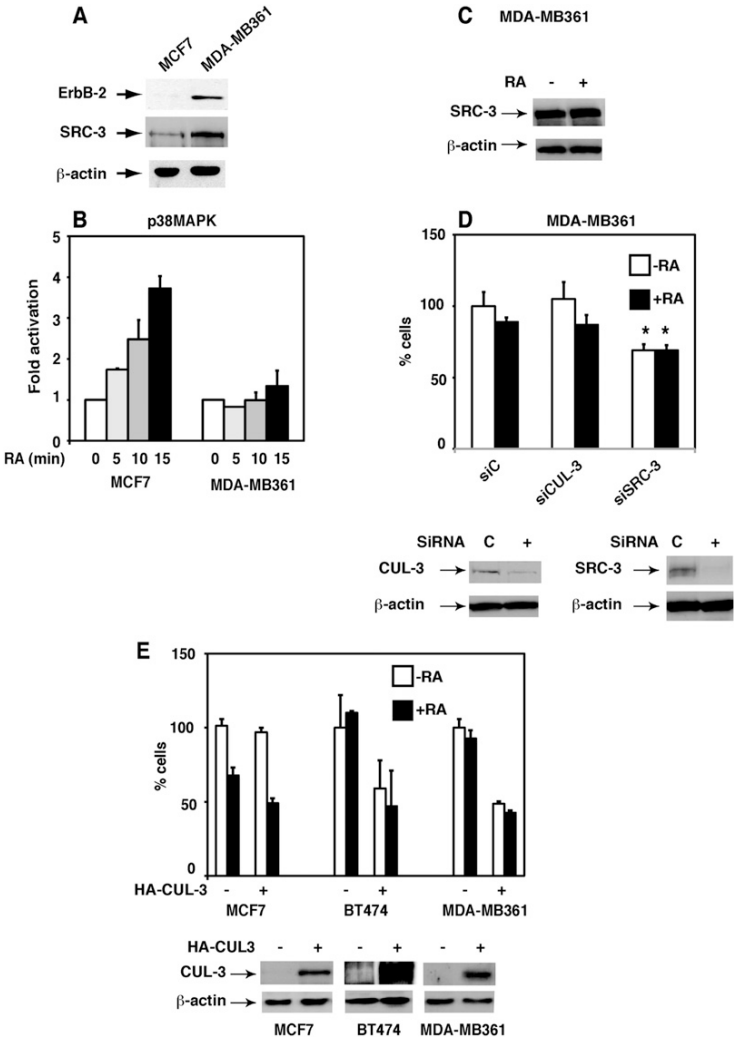


Fig. S5. In RA-resistant MDA-MB361 cells, SRC-3 is not degraded. (*A*) Immunoblots showing the overexpression of erbB-2 and SRC-3 in MDA-MB361 cells, compared with MCF7 cells. (*B* and *C*) In MDA-MB361 cells, p38MAPK is not activated and SRC-3 is not degraded. (*D*) The growth of MDA-MB361 cells is not affected by RA nor by knockdown of CUL-3 but is decreased upon knockdown of SRC-3 (*Upper*). The efficiency of knockdown was checked by immunoblotting (*Lower*). The results are the mean ± SD of two distinct experiments performed in quadruplate. Statistically significant differences were indicated (*$P < 0.05$, control versus siRNA). (*E*) The growth of BT474 and MDA-MB361 cells is markedly decreased upon overexpression of CUL-3 (*Upper*). Overexpression efficiency was checked by immunoblotting (*Lower*). The results are the mean ± SD of two distinct experiments performed in quadruplate.

Other Supporting Information Files

Dataset S1 (XLSX)

Bibliographie

Abu-Abed, S., Dolle, P., Metzger, D., Beckett, B., Chambon, P., and Petkovich, M. (2001). The retinoic acid-metabolizing enzyme, CYP26A1, is essential for normal hindbrain patterning, vertebral identity, and development of posterior structures. Genes Dev *15*, 226-240.

Acampora, D., Merlo, G.R., Paleari, L., Zerega, B., Postiglione, M.P., Mantero, S., Bober, E., Barbieri, O., Simeone, A., and Levi, G. (1999). Craniofacial, vestibular and bone defects in mice lacking the Distal-less-related gene Dlx5. Development *126*, 3795-3809.

Adam-Stitah, S., Penna, L., Chambon, P., and Rochette-Egly, C. (1999). Hyperphosphorylation of the retinoid X receptor alpha by activated c-Jun NH2-terminal kinases. J Biol Chem *274*, 18932-18941.

Al Tanoury, Z., Piskunov, A., and Rochette-Egly, C. (2013). Vitamin A and retinoid signaling: genomic and nongenomic effects. J Lipid Res *54*, 1761-1775.

Albalat, R. (2009). The retinoic acid machinery in invertebrates: ancestral elements and vertebrate innovations. Mol Cell Endocrinol *313*, 23-35.

Alsayed, Y., Uddin, S., Mahmud, N., Lekmine, F., Kalvakolanu, D.V., Minucci, S., Bokoch, G., and Platanias, L.C. (2001). Activation of Rac1 and the p38 mitogen-activated protein kinase pathway in response to all-trans-retinoic acid. J Biol Chem *276*, 4012-4019.

Aoto, J., Nam, C.I., Poon, M.M., Ting, P., and Chen, L. (2008). Synaptic signaling by all-trans retinoic acid in homeostatic synaptic plasticity. Neuron *60*, 308-320.

Baker, R.E., Schnell, S., and Maini, P.K. (2006). A clock and wavefront mechanism for somite formation. Dev Biol *293*, 116-126.

Bastien, J., Adam-Stitah, S., Plassat, J.L., Chambon, P., and Rochette-Egly, C. (2002). The phosphorylation site located in the A region of retinoic X receptor alpha is required for the antiproliferative effect of retinoic acid (RA) and the activation of RA target genes in F9 cells. Journal of Biological Chemistry *277*, 28683-28689.

Bastien, J., Adam-Stitah, S., Riedl, T., Egly, J.M., Chambon, P., and Rochette-Egly, C. (2000). TFIIH interacts with the retinoic acid receptor gamma and phosphorylates its AF-1-activating domain through cdk7. Journal of Biological Chemistry *275*, 21896-21904.

Bastien, J., Plassat, J.L., Payrastre, B., and C., R.-E. (2006). The phosphoinositide 3-kinase/Akt pathway is essential for the retinoic acid-induced differentiation of F9 cells. Oncogene *25*, 2040-2047.

Begemann, G., and Ingham, P.W. (2000). Developmental regulation of Tbx5 in zebrafish embryogenesis. Mech Dev *90*, 299-304.

Bentzinger, C.F., Wang, Y.X., and Rudnicki, M.A. (2012). Building muscle: molecular regulation of myogenesis. Cold Spring Harb Perspect Biol *4*.

Bertrand, S., Thisse, B., Tavares, R., Sachs, L., Chaumot, A., Bardet, P.L., Escriva, H., Duffraisse, M., Marchand, O., Safi, R., *et al.* (2007). Unexpected novel relational links uncovered by extensive developmental profiling of nuclear receptor expression. PLoS Genet *3*, e188.

Beverdam, A., Merlo, G.R., Paleari, L., Mantero, S., Genova, F., Barbieri, O., Janvier, P., and Levi, G. (2002). Jaw transformation with gain of symmetry after Dlx5/Dlx6 inactivation: mirror of the past? Genesis *34*, 221-227.

Bibel, M., Richter, J., Lacroix, E., and Barde, Y.A. (2007). Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. Nat Protoc *2*, 1034-1043.

Biddie, S.C., John, S., and Hager, G.L. (2010). Genome-wide mechanisms of nuclear receptor action. Trends Endocrinol Metab *21*, 3-9.

Bielska, A.A., and Zondlo, N.J. (2006). Hyperphosphorylation of tau induces local polyproline II helix. Biochemistry *45*, 5527-5537.

Blentic, A., Gale, E., and Maden, M. (2003). Retinoic acid signalling centres in the avian embryo identified by sites of expression of synthesising and catabolising enzymes. Dev Dyn *227*, 114-127.

Bloom, S., Ledon-Rettig, C., Infante, C., Everly, A., Hanken, J., and Nascone-Yoder, N. (2013). Developmental origins of a novel gut morphology in frogs. Evol Dev *15*, 213-223.

Boch, J. (2011). TALEs of genome targeting. Nat Biotechnol *29*, 135-136.

Bour, G., Gaillard, E., Bruck, N., Lalevee, S., Plassat, J.L., Busso, D., Samama, J.P., and Rochette-Egly, C. (2005a). Cyclin H binding to the RAR{alpha} activation function (AF)-2 domain directs phosphorylation of the AF-1 domain by cyclin-dependent kinase 7. Proc Natl Acad Sci U S A *102*, 16608-16613.

Bour, G., Plassat, J.L., Bauer, A., Lalevee, S., and Rochette-Egly, C. (2005b). Vinexin beta interacts with the non-phosphorylated AF-1 domain of retinoid receptor gamma (RARgamma) and represses RARgamma-mediated transcription. J Biol Chem *280*, 17027-17037.

Bourguet, W., Germain, P., and Gronemeyer, H. (2000). Nuclear receptor ligand-binding domains: three-dimensional structures, molecular interactions and pharmacological implications. Trends Pharmacol Sci *21*, 381-388.

Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H., and Moras, D. (1995). Crystal structure of the ligand-binding domain of the human nuclear receptor RXR-alpha. Nature *375*, 377-382.

Brelivet, Y., Kammerer, S., Rochel, N., Poch, O., and Moras, D. (2004). Signature of the oligomeric behaviour of nuclear receptors at the sequence and structural level. Embo Reports *5*, 423-429.

Britz, R., Conway, K.W., and Ruber, L. (2009). Spectacular morphological novelty in a miniature cyprinid fish, Danionella dracula n. sp. Proc Biol Sci *276*, 2179-2186.

Bruck, N., Bastien, J., Bour, G., Tarrade, A., Plassat, J.L., Bauer, A., Adam-Stitah, S., and Rochette-Egly, C. (2005). Phosphorylation of the Retinoid X Receptor at the Omega loop, modulates the expression of retinoic-acid-target genes with a promoter context specificity. Cellular Signalling *17*, 1229-1239.

Bruck, N., Vitoux, D., Ferry, C., Duong, V., Bauer, A., de The, H., and Rochette-Egly, C. (2009). A coordinated phosphorylation cascade initiated by p38MAPK/MSK1 directs RARalpha to target promoters. Embo J *28*, 34-47.

Campo-Paysaa, F., Marletaz, F., Laudet, V., and Schubert, M. (2008). Retinoic acid signaling in development: tissue-specific functions and evolutionary origins. Genesis *46*, 640-656.

Cañestro, C., Postlethwait, J.H., Gonzàlez-Duarte, R., and Albalat, R. (2006). Is retinoic acid genetic machinery a chordate innovation? Evol Dev *8*, 394-406.

Carroll, J.S., Liu, X.S., Brodsky, A.S., Li, W., Meyer, C.A., Szary, A.J., Eeckhoute, J., Shao, W., Hestermann, E.V., Geistlinger, T.R., *et al.* (2005). Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. Cell *122*, 33-43.

Carroll, J.S., Meyer, C.A., Song, J., Li, W., Geistlinger, T.R., Eeckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C., Hall, G.F., *et al.* (2006). Genome-wide analysis of estrogen receptor binding sites. Nat Genet *38*, 1289-1297.

Carroll, S.B. (2005). Endless forms most beautiful: The new science of evo devo and the making of the animal kingdom. (Weidenfeld).

Carroll, S.B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell *134*, 25-36.

Carroll, S.B., Grenier, J., and Weatherbee, S.D. (2004). From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design.

Chakravarti, D. (2009). Volume 87: Regulatory Mechanisms in Transcriptional Signaling, 1st Edition, Elsevier edn.

Chakravarti, D., LaMorte, V.J., Nelson, M.C., Nakajima, T., Schulman, I.G., Juguilon, H., Montminy, M., and Evans, R.M. (1996). Role of CBP/P300 in nuclear receptor signalling. Nature *383*, 99-103.

Chambers, D., Wilson, L., Maden, M., and Lumsden, A. (2007). RALDH-independent generation of retinoic acid during vertebrate embryogenesis by CYP1B1. Development *134*, 1369-1383.

Chambon, P. (1996). A decade of molecular biology ot retinoic acid receptors. FASEBJ *10*, 940-954.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Jr., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J., *et al.* (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. Science *327*, 302-305.

Chandra, V., Huang, P., Hamuro, Y., Raghuram, S., Wang, Y., Burris, T.P., and Rastinejad, F. (2008). Structure of the intact PPAR-gamma-RXR- nuclear receptor complex on DNA. Nature *456*, 350-356.

Chang, C.Y., Norris, J.D., Gron, H., Paige, L.A., Hamilton, P.T., Kenan, D.J., Fowlkes, D., and McDonnell, D.P. (1999). Dissection of the LXXLL nuclear receptor-coactivator interaction motif using combinatorial peptide libraries: Discovery of peptide antagonists of estrogen receptors alpha and beta. Molecular and Cellular Biology *19*, 8226-8239.

Charite, J., McFadden, D.G., Merlo, G., Levi, G., Clouthier, D.E., Yanagisawa, M., Richardson, J.A., and Olson, E.N. (2001). Role of Dlx6 in regulation of an endothelin-1-dependent, dHAND branchial arch enhancer. Genes Dev *15*, 3039-3049.

Chatonnet, F., Guyot, R., Benoit, G., and Flamant, F. (2013). Genome-wide analysis of thyroid hormone receptors shared and specific functions in neural cells. Proc Natl Acad Sci U S A *110*, E766-775.

Chebaro, Y., Amal, I., Rochel, N., Rochette-Egly, C., Stote, R.H., and Dejaegere, A. (2013). Phosphorylation of the retinoic acid receptor alpha induces a mechanical allosteric regulation and changes in internal dynamics. PLoS Comput Biol 9, e1003012.

Chen, H., Howald, W.N., and Juchau, M.R. (2000). Biosynthesis of all-trans-retinoic acid from all-trans-retinol: catalysis of all-trans-retinol oxidation by human P-450 cytochromes. Drug Metab Dispos *28*, 315-322.

Chen, N., and Napoli, J.L. (2008). All-trans-retinoic acid stimulates translation and induces spine formation in hippocampal neurons through a membrane-associated RARalpha. Faseb J *22*, 236-245.

Chen, N., Onisko, B., and Napoli, J.L. (2008). The nuclear transcription factor RARalpha associates with neuronal RNA granules and suppresses translation. J Biol Chem *283*, 20841-20847.

Cohen, R.N., Brzostek, S., Kim, B., Chorev, M., Wondisford, F.E., and Hollenberg, A.N. (2001). The specificity of interactions between nuclear hormone receptors and corepressors is mediated by distinct amino acid sequences within the interacting domains. Molecular Endocrinology *15*, 1049-1061.

Collop, A.H., Broomfield, J.A.S., Chandraratna, R.A.S., Yong, Z., Deimling, S.J., Kolker, S.J., Weeks, D.L., and Drysdale, T.A. (2006). Retinoic acid signaling is essential for formation of the heart tube in Xenopus. Developmental Biology *291*, 96-109.

Cooke, J., and Zeeman, E.C. (1976). A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. J Theor Biol *58*, 455-476.

Corcoran, J., Shroot, B., Pizzey, J., and Maden, M. (2000). The role of retinoic acid receptors in neurite outgrowth from different populations of embryonic mouse dorsal root ganglia. J Cell Sci *113 (Pt 14)*, 2567-2574.

Darimont, B.D., Wagner, R.L., Apriletti, J.W., Stallcup, M.R., Kushner, P.J., Baxter, J.D., Fletterick, R.J., and Yamamoto, K.R. (1998). Structure and specificity of nuclear receptor-coactivator interactions. Genes Dev *12*, 3343-3356.

de The, H., Vivanco-Ruiz, M.M., Tiollais, P., Stunnenberg, H., and Dejean, A. (1990). Identification of a retinoic acid responsive element in the retinoic acid receptor beta gene. Nature *343*, 177-180.

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

Delacroix, L., Moutier, E., Altobelli, G., Legras, S., Poch, O., Choukrallah, M.A., Bertin, I., Jost, B., and Davidson, I. (2010). Cell-specific interaction of retinoic acid receptors with target genes in mouse embryonic fibroblasts and embryonic stem cells. Mol Cell Biol *30*, 231-244.

Demarest, S.J., Martinez-Yamout, M., Chung, J., Chen, H., Xu, W., Dyson, H.J., Evans, R.M., and Wright, P.E. (2002). Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. Nature *415*, 549-553.

Denuce, J.M. (1991). [Teratogenic and metamorphosis inhibiting activity of retinoic acid in Ciona intestinalis]. Z Naturforsch C *46*, 1094-1100.

Depew, M.J., Simpson, C.A., Morasso, M., and Rubenstein, J.L. (2005). Reassessing the Dlx code: the genetic regulation of branchial arch skeletal pattern and development. J Anat *207*, 501-561.

Diez del Corral, R., Olivera-Martinez, I., Goriely, A., Gale, E., Maden, M., and Storey, K. (2003). Opposing FGF and retinoid pathways control ventral neural pattern, neuronal differentiation, and segmentation during body axis extension. Neuron *40*, 65-79.

Dilworth, F.J., and Chambon, P. (2001). Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription. Oncogene *20*, 3047-3054.

Dmetrichuk, J.M., Carlone, R.L., and Spencer, G.E. (2006). Retinoic acid induces neurite outgrowth and growth cone turning in invertebrate neurons. Dev Biol *294*, 39-49.

Duboc, V., and Logan, M.P. (2011). Regulation of limb bud initiation and limb-type morphology. Dev Dyn *240*, 1017-1027.

Duboc, V., Röttinger, E., Lapraz, F., Besnardeau, L., and Lepage, T. (2005). Left-right asymmetry in the sea urchin embryo is regulated by nodal signaling on the right side. Dev Cell *9*, 147-158.

Duboule, D., and Dolle, P. (1989). The Structural and Functional-Organization of the Murine Hox Gene Family Resembles That of Drosophila Homeotic Genes. Embo Journal *8*, 1497-1505.

Dubrulle, J., McGrew, M.J., and Pourquie, O. (2001). FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. Cell *106*, 219-232.

Dubrulle, J., and Pourquie, O. (2004). Coupling segmentation to axis formation. Development *131*, 5783-5793.

Dupe, V., Davenne, M., Brocard, J., Dolle, P., Mark, M., Dierich, A., Chambon, P., and Rijli, F.M. (1997). In vivo functional analysis of the Hoxa-1 3' retinoic acid response element (3'RARE). Development *124*, 399-410.

Dupe, V., and Lumsden, A. (2001). Hindbrain patterning involves graded responses to retinoic acid signalling. Development *128*, 2199-2208.

Durand, B., Saunders, M., Leroy, P., Leid, M., and Chambon, P. (1992). All-trans and 9-cis retinoic acid induction of CRABPII transcription is mediated by RAR-RXR heterodimers bound to DR1 and DR2 repeated motifs. Cell *71*, 73-85.

Durston, A.J., Timmermans, J.P., Hage, W.J., Hendriks, H.F., de Vries, N.J., Heideveld, M., and Nieuwkoop, P.D. (1989). Retinoic acid causes an anteroposterior transformation in the developing central nervous system. Nature *340*, 140-144.

Dyson, H.J., and Wright, P.E. (2005). Intrinsically unstructured proteins and their functions. Nature Reviews Molecular Cell Biology *6*, 197-208.

Eeckhoute, J., Carroll, J.S., Geistlinger, T.R., Torres-Arzayus, M.I., and Brown, M. (2006). A cell-type-specific transcriptional network required for estrogen regulation of cyclin D1 and cell cycle progression in breast cancer. Genes Dev *20*, 2513-2526.

Emoto, Y., Wada, H., Okamoto, H., Kudo, A., and Imai, Y. (2005). Retinoic acid-metabolizing enzyme Cyp26a1 is essential for determining territories of hindbrain and spinal cord in zebrafish. Dev Biol *278*, 415-427.

Epping, M.T., Wang, L.M., Edel, M.J., Carlee, L., Hernandez, M., and Bernards, R. (2005). The human tumor antigen repressor of retinoic acid PRAME is a dominant receptor signaling. Cell *122*, 835-847.

Erickson, T., Pillay, L.M., and Waskiewicz, A.J. (2011). Zebrafish Tshz3b Negatively Regulates Hox Function in the Developing Hindbrain. Genesis *49*, 725-742.

Escriva, H., Bertrand, S., Germain, P., Robinson-Rechavi, M., Umbhauer, M., Cartry, J., Duffraisse, M., Holland, L., Gronemeyer, H., and Laudet, V. (2006). Neofunctionalization in vertebrates: the example of retinoic acid receptors. PLoS Genet *2*, e102.

Escriva, H., Holland, N.D., Gronemeyer, H., Laudet, V., and Holland, L.Z. (2002). The retinoic acid signaling pathway regulates anterior/posterior patterning in the nerve cord and pharynx of amphioxus, a chordate lacking neural crest. Development *129*, 2905-2916.

Farboud, B., and Privalsky, M.L. (2004). Retinoic acid receptor-alpha is stabilized in a repressive state by its C-terminal, isotype-specific F domain. Molecular Endocrinology *18*, 2839-2853.

Farooqui, M., Franco, P.J., Thompson, J., Kagechika, H., Chandraratna, R.A., Banaszak, L., and Wei, L.N. (2004). Effects of retinoid ligands on RIP140: molecular interaction with retinoid receptors and biological activity.
. Biochemistry *42*, 971-979.

Feng, L., Hernandez, R.E., Waxman, J.S., Yelon, D., and Moens, C. (2010). Dhrs3a regulates retinoic acid biosynthesis through a feedback inhibition mechanism. Dev Biol *338*, 1-14.

Fernandes, I., Bastien, Y., Wai, T., Nygard, K., Lin, R., Cormier, O., Lee, H.S., Eng, F., Bertos, N.R., Pelletier, N., *et al.* (2003). Ligand-dependent nuclear receptor corepressor LCoR functions by histone deacetylase-dependent and -independent mechanisms. Mol Cell *11*, 139-150.

Ferry, C., Gaouar, S., Fischer, B., Boeglin, M., Paul, N., Samarut, E., Piskunov, A., Pankotai-Bodo, G., Brino, L., and Rochette-Egly, C. (2011). Cullin 3 mediates SRC-3 ubiquitination and degradation to control the retinoic acid response. Proc Natl Acad Sci U S A *108*, 20603-20608.

Ferry, C., Gianni, M., Lalevee, S., Bruck, N., Plassat, J.L., Raska, I., Jr., Garattini, E., and Rochette-Egly, C. (2009). SUG-1 plays proteolytic and non-proteolytic roles in the control of retinoic acid target genes via its interaction with SRC-3. J Biol Chem *284*, 8127-8135.

Flajollet, S., Lefebvre, B., Rachez, C., and Lefebvre, P. (2006). Distinct roles of the steroid receptor coactivator 1 and of MED1 in retinoid-induced transcription and cellular differentiation. J Biol Chem *281*, 20338-20348.

Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L., and Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. Genetics *151*, 1531-1545.

Freedman, L.P., Luisi, B.F., Korszun, Z.R., Basavappa, R., Sigler, P.B., and Yamamoto, K.R. (1988). The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain. Nature *334*, 543-546.

Gaillard, E., Bruck, N., Brelivet, Y., Bour, G., Lalevee, S., Bauer, A., Poch, O., Moras, D., and Rochette-Egly, C. (2006). Phosphorylation by PKA potentiates retinoic acid receptor alpha activity by means of increasing interaction with and phosphorylation by cyclin H/cdk7. Proc Natl Acad Sci U S A *103*, 9548-9553.

Garcia-Fernandez, J., and Holland, P.W. (1994). Archetypal organization of the amphioxus Hox gene cluster. Nature *370*, 563-566.

Germain, P., Altucci, L., Bourguet, W., Rochette-Egly, C., and Gronemeyer, H. (2003). Nuclear receptor superfamily: Principles of signaling. In Pure Appl Chem,, J. Miyamoto, and J. Burger, eds., pp. 1619-1664.

Germain, P., Chambon, P., Eichele, G., Evans, R.M., Lazar, M.A., Leid, M., De Lera, A.R., Lotan, R., Mangelsdorf, D.J., and Gronemeyer, H. (2006a). International Union of Pharmacology. LX. Retinoic acid receptors. Pharmacol Rev *58*, 712-725.

Germain, P., Chambon, P., Eichele, G., Evans, R.M., Lazar, M.A., Leid, M., De Lera, A.R., Lotan, R., Mangelsdorf, D.J., and Gronemeyer, H. (2006b). International Union of Pharmacology. LXIII. Retinoid X receptors. Pharmacol Rev *58*, 760-772.

Gettemans, J., Van Impe, K., Delanote, V., Hubert, T., Vandekerckhove, J., and De Corte, V. (2005). Nuclear actin-binding proteins as modulators of gene transcription. Traffic 6, 847-857.

Gianni, M., Bauer, A., Garattini, E., Chambon, P., and Rochette-Egly, C. (2002). Phosphorylation by p38MAPK and recruitment of SUG-1 are required for RA-indced RARγ degradation and transactivation. EMBO J *21*, 3760-3769.

Gianni, M., Tarrade, A., Nigro, E.A., Garattini, E., and Rochette-Egly, C. (2003). The AF-1 and AF-2 domains of RAR gamma 2 and RXR alpha cooperate for triggering the transactivation and the degradation of RAR gamma 2/RXR alpha heterodimers. J Biol Chem *278*, 34458-34466.

Gibert, Y., Bernard, L., Debiais-Thibaud, M., Bourrat, F., Joly, J.S., Pottin, K., Meyer, A., Retaux, S., Stock, D.W., Jackman, W.R., *et al.* (2010). Formation of oral and pharyngeal dentition in teleosts depends on differential recruitment of retinoic acid signaling. FASEB J *24*, 3298-3309.

Glover, J.C., Renaud, J.S., and Rijli, F.M. (2006). Retinoic acid and hindbrain patterning. J Neurobiol *66*, 705-725.

Godsave, S.F., Koster, C.H., Getahun, A., Mathu, M., Hooiveld, M., van der Wees, J., Hendriks, J., and Durston, A.J. (1998). Graded retinoid responses in the developing hindbrain. Dev Dyn *213*, 39-49.

Goldbeter, A., and Pourquie, O. (2008). Modeling the segmentation clock as a network of coupled oscillations in the Notch, Wnt and FGF signaling pathways. J Theor Biol *252*, 574-585.

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

Graham, A., Papalopulu, N., and Krumlauf, R. (1989). The Murine and Drosophila Homeobox Gene Complexes Have Common Features of Organization and Expression. Cell *57*, 367-378.

Grandel, H., Lun, K., Rauch, G.J., Rhinn, M., Piotrowski, T., Houart, C., Sordino, P., Kuchler, A.M., Schulte-Merker, S., Geisler, R., *et al.* (2002). Retinoic acid signalling in the zebrafish embryo is necessary during pre-segmentation stages to pattern the anterior-posterior axis of the CNS and to induce a pectoral fin bud. Development *129*, 2851-2865.

Grandel, H., and Schulte-Merker, S. (1998). The development of the paired fins in the zebrafish (Danio rerio). Mech Dev *79*, 99-120.

Grant, P.R., and Grant, B.R. (2011). How and Why Species Multiply: The Radiation of Darwin's Finches.

Gronemeyer, H., Gustafsson, J.A., and Laudet, V. (2004). Principles for modulation of the nuclear receptor superfamily. Nat Rev Drug Discov *3*, 950-964.

Gu, P., LeMenuet, D., Chung, A.C., Mancini, M., Wheeler, D.A., and Cooney, A.J. (2005). Orphan nuclear receptor GCNF is required for the repression of pluripotency genes during retinoic acid-induced embryonic stem cell differentiation. Mol Cell Biol *25*, 8507-8519.

Guenther, M.G., Barak, O., and Lazar, M.A. (2001). The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3. Mol Cell Biol *21*, 6091-6101.

Guenther, M.G., Lane, W.S., Fischle, W., Verdin, E., Lazar, M.A., and Shiekhattar, R. (2000). A core SMRT corepressor complex containing HDAC3 and TBL1, a WD40-repeat protein linked to deafness. Genes Dev *14*, 1048-1057.

Gurevich, I., Flores, A.M., and Aneskievich, B.J. (2007). Corepressors of agonist-bound nuclear receptors. Toxicol Appl Pharmacol *223*, 288-298.

Gutierrez-Mazariegos, J., Theodosiou, M., Campo-Paysaa, F., and Schubert, M. (2011). Vitamin A: A mutifonctional tool for development. Seminars in Cell and Developmental Biology *22*, 603-610.

Guyot, R., Vincent, S., Bertin, J., Samarut, J., and Ravel-Chapuis, P. (2010). The transforming acidic coiled coil (TACC1) protein modulates the transcriptional activity of the nuclear receptors TR and RAR. BMC Mol Biol *11*, 3.

Hale, L.A., Tallafuss, A., Yan, Y.L., Dudley, L., Eisen, J.S., and Postlethwait, J.H. (2006). Characterization of the retinoic acid receptor genes raraa, rarab and rarg during zebrafish development. Gene Expr Patterns *6*, 546-555.

Hamy, F., Verwaerde, P., Helbecque, N., Formstecher, P., and Henichart, J.P. (1991). Nuclear targeting of a viral-cointernalized protein by a short signal sequence from human retinoic acid receptors. Bioconjug Chem *2*, 375-378.

Heery, D.M., Hoare, S., Hussain, S., Parker, M.G., and Sheppard, H. (2001). Core LXXLL motif sequences in CREB-binding protein, SRC1, and RIP140 define affinity and selectivity for steroid and retinoid receptors. J Biol Chem *276*, 6695-6702.

Heery, D.M., Kalkhoven, E., Hoare, S., and Parker, M.G. (1997). A signature motif in transcriptional co-activators mediates binding to nuclear receptors. Nature *387*, 733-736.

Hernandez, R.E., Putzke, A.P., Myers, J.P., Margaretha, L., and Moens, C.B. (2007). Cyp26 enzymes generate the retinoic acid response pattern necessary for hindbrain development. Development *134*, 177-187.

Hinman, V.F., and Degnan, B.M. (1998). Retinoic acid disrupts anterior ectodermal and endodermal development in ascidian larvae and postlarvae. Dev Genes Evol *208*, 336-345.

Hirokawa, N., Tanaka, Y., and Okada, Y. (2009). Left-right determination: involvement of molecular motor KIF3, cilia, and nodal flow. Cold Spring Harb Perspect Biol *1*, a000802.

Holland, L.Z., Albalat, R., Azumi, K., Benito-Gutierrez, E., Blow, M.J., Bronner-Fraser, M., Brunet, F., Butts, T., Candiani, S., Dishaw, L.J., *et al.* (2008). The amphioxus genome illuminates vertebrate origins and cephalochordate biology. Genome Res *18*, 1100-1111.

Holland, L.Z., and Holland, N.D. (1996). Expression of AmphiHox-1 and AmphiPax-1 in amphioxus embryos treated with retinoic acid: Insights into evolution and patterning of the chordate nerve cord and pharynx. Development *122*, 1829-1838.

Howard-Ashby, M., Materna, S.C., Brown, C.T., Chen, L., Cameron, R.A., and Davidson, E.H. (2006). Gene families encoding transcription factors expressed in early development of Strongylocentrotus purpuratus. Dev Biol *300*, 90-107.

Hu, X., and Lazar, M.A. (1999). The CoRNR motif controls the recruitment of corepressors by nuclear hormone receptors. Nature *402*, 93-96.

Hua, S., Kittler, R., and White, K.P. (2009). Genomic antagonism between retinoic acid and estrogen signaling in breast cancer. Cell *137*, 1259-1271.

Huang, S.M., Huang, C.J., Wang, W.M., Kang, J.C., and Hsu, W.C. (2004). The enhancement of nuclear receptor transcriptional activation by a mouse actin-binding protein, alpha actinin 2. J Mol Endocrinol *32*, 481-496.

Jackman, W.R., Draper, B.W., and Stock, D.W. (2004). Fgf signaling is required for zebrafish tooth development. Developmental Biology *274*, 139-157.

Jaillon, O., Aury, J.M., Brunet, F., Petit, J.L., Stange-Thomann, N., Mauceli, E., Bouneau, L., Fischer, C., Ozouf-Costaz, C., Bernot, A., *et al.* (2004). Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. Nature *431*, 946-957.

Jernvall, J., and Salazar-Ciudad, I. (2007). The economy of tinkering mammalian teeth. Novartis Found Symp *284*, 207-216; discussion 216-224.

John, S., Sabo, P.J., Johnson, T.A., Sung, M.H., Biddie, S.C., Lightman, S.L., Voss, T.C., Davis, S.R., Meltzer, P.S., Stamatoyannopoulos, J.A., *et al.* (2008). Interaction of the glucocorticoid receptor with the chromatin landscape. Mol Cell *29*, 611-624.

Jones, P.L., and Shi, Y.B. (2003). N-CoR-HDAC Corepressor Complexes: Roles in Transcriptional Regulation by Nuclear Hormone Receptors. In Protein Complexes that Modify Chromatin (Springer Berlin Heidelberg), pp. 237-268.

Kambhampati, S., Li, Y., Verma, A., Sassano, A., Majchrzak, B., Deb, D.K., Parmar, S., Giafis, N., Kalvakolanu, D.V., Rahman, A., *et al.* (2003). Activation of protein kinase C delta by all-trans-retinoic acid. J Biol Chem *278*, 32544-32551.

Kane, M.A., Folias, A.E., Pingitore, A., Perri, M., Obrochta, K.M., Krois, C.R., Cione, E., Ryu, J.Y., and Napoli, J.L. (2010). Identification of 9-cis-retinoic acid as a pancreas-specific autacoid that attenuates glucose-stimulated insulin secretion. 107 *50*, 21884-21889.

Katsuyama, Y., and Saiga, H. (1998). Retinoic acid affects patterning along the anterior-posterior axis of the ascidian embryo. Dev Growth Differ *40*, 413-422.

Kawakami, Y., Raya, A., Raya, R.M., Rodriguez-Esteban, C., and Izpisua Belmonte, J.C. (2005). Retinoic acid signalling links left-right asymmetric patterning and bilaterally symmetric somitogenesis in the zebrafish embryo. Nature *435*, 165-171.

Keidel, S., LeMotte, P., and Apfel, C. (1994). Different agonist- and antagonist-induced conformational changes in retinoic acid receptors analyzed by protease mapping. Mol Cell Biol *14*, 287-298.

Keriel, A., Stary, A., Sarasin, A., Rochette-Egly, C., and Egly, J.M. (2002). XPD mutations prevent TFIIH-dependent transactivation by nuclear receptors and phosphorylation of RARalpha. Cell *109*, 125-135.

Kerszberg, M., and Wolpert, L. (2000). A clock and trail model for somite formation, specialization and polarization. J Theor Biol *205*, 505-510.

King, M.C., and Wilson, A.C. (1975). Evolution at two levels in humans and chimpanzees. Science *188*, 107-116.

Kininis, M., Chen, B.S., Diehl, A.G., Isaacs, G.D., Zhang, T., Siepel, A.C., Clark, A.G., and Kraus, W.L. (2007). Genomic analyses of transcription factor binding, histone acetylation, and gene expression reveal mechanistically distinct classes of estrogen-regulated promoters. Mol Cell Biol *27*, 5090-5104.

Klaholz, B.P., Mitschler, A., and Moras, D. (2000). Structural basis for isotype selectivity of the human retinoic acid nuclear receptor. J Mol Biol *302*, 155-170.

Kopf, E., Plassat, J.L., Vivat, V., de The, H., Chambon, P., and Rochette-Egly, C. (2000). Dimerization with retinoid X receptors and phosphorylation modulate the retinoic acid-induced degradation of retinoic acid receptors alpha and gamma through the ubiquitin-proteasome pathway. J Biol Chem *275*, 33280-33288.

Kudoh, T., Wilson, S.W., and Dawid, I.B. (2002). Distinct roles for Fgf, Wnt and retinoic acid in posteriorizing the neural ectoderm. Development *129*, 4335-4346.

Kumar, M., Jordan, N., Melton, D., and Grapin-Botton, A. (2003). Signals from lateral plate mesoderm instruct endoderm toward a pancreatic fate. Dev Biol *259*, 109-122.

Kumar, M., and Melton, D. (2003). Pancreas specification: a budding question. Curr Opin Genet Dev *13*, 401-407.

Kurokawa, R., Soderstrom, M., Horlein, A., Halachmi, S., Brown, M., Rosenfeld, M.G., and Glass, C.K. (1995). Polarity-specific activities of retinoic acid receptors determined by a co-repressor. Nature *377*, 451-454.

Lalevee, S., Anno, Y.N., Chatagnon, A., Samarut, E., Poch, O., Laudet, V., Benoit, G., Lecompte, O., and Rochette-Egly, C. (2011). Genome-wide in silico identification of new conserved and functional retinoic acid receptor response elements (direct repeats separated by 5 bp). J Biol Chem *286*, 33322-33334.

Lalevee, S., Bour, G., Quinternet, M., Samarut, E., Kessler, P., Vitorino, M., Bruck, N., Delsuc, M.A., Vonesch, J.L., Kieffer, B., *et al.* (2010). Vinexin{beta}, an atypical "sensor" of

retinoic acid receptor {gamma} signaling: union and sequestration, separation, and phosphorylation. FASEB J.

Lamason, R.L., Mohideen, M.A., Mest, J.R., Wong, A.C., Norton, H.L., Aros, M.C., Jurynec, M.J., Mao, X., Humphreville, V.R., Humbert, J.E., *et al.* (2005). SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. Science *310*, 1782-1786.

Lammer, E.J., Chen, D.T., Hoar, R.M., Agnish, N.D., Benke, P.J., Braun, J.T., Curry, C.J., Fernhoff, P.M., Grix, A.W., Jr., Lott, I.T., *et al.* (1985). Retinoic acid embryopathy. N Engl J Med *313*, 837-841.

Laudet, V., and Gronemeyer, H. (2002). Nuclear Receptor Factsbook (London: Acedemic Press).

Lavery, D.N., and McEwan, I.J. (2005). Structure and function of steroid receptor AF1 transactivation domains: induction of active conformations. Biochem J *391*, 449-464.

Le Romancer, M., Poulard, C., Cohen, P., Sentis, S., Renoir, J.M., and Corbo, L. (2011). Cracking the estrogen receptor's posttranslational code in breast tumors. Endocr Rev *32*, 597-622.

Lee, H.Y., Suh, Y.A., Robinson, M.J., Clifford, J.L., Hong, W.K., Woodgett, J.R., Cobb, M.H., Mangelsdorf, D.J., and Kurie, J.M. (2000). Stress pathway activation induces phosphorylation of retinoid X receptor. J Biol Chem *275*, 32193-32199.

Lee, M.S., Kliewer, S.A., Provencal, J., Wright, P.E., and Evans, R.M. (1993). Structure of the retinoid X receptor alpha DNA binding domain: a helix required for homodimeric DNA binding. Science *260*, 1117-1121.

Lee, Y.H., Campbell, H.D., and Stallcup, M.R. (2004). Developmentally essential protein flightless I is a nuclear receptor coactivator with actin binding activity. Molecular and Cellular Biology *24*, 2103-2117.

Lefebvre, B., Ozato, K., and Lefebvre, P. (2002). Phosphorylation of histone H3 is functionally linked to retinoic acid receptor beta promoter activation. Embo Reports *3*, 335-340.

Leid, M., Kastner, P., and Chambon, P. (1992). Multiplicity generates diversity in the retinoic acid signalling pathways. Trends Biochem Sci *17*, 427-433.

Linville, A., Radtke, K., Waxman, J.S., Yelon, D., and Schilling, T.F. (2009). Combinatorial roles for zebrafish retinoic acid receptors in the hindbrain, limbs and pharyngeal arches. Dev Biol *325*, 60-70.

Lohnes, D., Mark, M., Mendelsohn, C., Dolle, P., Dierich, A., Gorry, P., Gansmuller, A., and Chambon, P. (1994). Function of the retinoic acid receptors (RARs) during development (I). Craniofacial and skeletal abnormalities in RAR double mutants. Development *120*, 2723-2748.

Loudig, O., Babichuk, C., White, J., Abu-Abed, S., Mueller, C., and Petkovich, M. (2000). Cytochrome P450RAI(CYP26) promoter: a distinct composite retinoic acid response element underlies the complex regulation of retinoic acid metabolism. Mol Endocrinol *14*, 1483-1497.

Lumsden, A., and Krumlauf, R. (1996). Patterning the vertebrate neuraxis. Science *274*, 1109-1115.

Lupien, M., Eeckhoute, J., Meyer, C.A., Krum, S.A., Rhodes, D.R., Liu, X.S., and Brown, M. (2009). Coactivator function defines the active estrogen receptor alpha cistrome. Mol Cell Biol *29*, 3413-3423.

Maden, M. (2002). Retinoid signalling in the development of the central nervous system. Nat Rev Neurosci *3*, 843-853.

Maden, M. (2007). Retinoic acid in the development, regeneration and maintenance of the nervous system. Nat Rev Neurosci *8*, 755-765.

Maden, M., Gale, E., Kostetskii, I., and Zile, M. (1996). Vitamin A-deficient quail embryos have half a hindbrain and other neural defects. Current Biology *6*, 417-426.

Maghsoodi, B., Poon, M.M., Nam, C.I., Aoto, J., Ting, P., and Chen, L. (2008). Retinoic acid regulates RARalpha-mediated control of translation in dendritic RNA granules during homeostatic synaptic plasticity. Proc Natl Acad Sci U S A *105*, 16015-16020.

Mallo, M. (1997). Retinoic acid disturbs mouse middle ear development in a stage-dependent fashion. Dev Biol *184*, 175-186.

Mangelsdorf, D.J., and Evans, R.M. (1995). The RXR heterodimers and orphan receptors. Cell *83*, 841-850.

Mangelsdorf, D.J., Umesono, K., Kliewer, S.A., Borgmeyer, U., Ong, E.S., and Evans, R.M. (1991). A direct repeat in the cellular retinol-binding protein type II gene confers differential regulation by RXR and RAR. Cell *66*, 555-561.

Mark, M., Ghyselinck, N.B., and Chambon, P. (2004). Retinoic acid signalling in the development of branchial arches. Curr Opin Genet Dev *14*, 591-598.

Mark, M., Ghyselinck, N.B., and Chambon, P. (2009). Function of retinoic acid receptors during embryonic development. Nucl Recept Signal *7*, e002.

Marquez, D.C., Chen, H.W., Curran, E.M., Welshons, W.V., and Pietras, R.J. (2006). Estrogen receptors in membrane lipid rafts and signal transduction in breast cancer. Mol Cell Endocrinol *246*, 91-100.

Marshall, H., Nonchev, S., Sham, M.H., Muchamore, I., Lumsden, A., and Krumlauf, R. (1992). Retinoic Acid Alters Hindbrain Hox Code and Induces Transformation of Rhombomeres 2-3 into a 4-5 Identity. Nature *360*, 737-741.

Martens, J.H., Rao, N.A., and Stunnenberg, H.G. (2011). Genome-wide interplay of nuclear receptors with the epigenome. Biochim Biophys Acta *1812*, 818-823.

Masia, S., Alvarez, S., de Lera, A.R., and Barettino, D. (2007). Rapid, nongenomic actions of retinoic acid on phosphatidylinositol-3-kinase signaling pathway mediated by the retinoic acid receptor. Mol Endocrinol *21*, 2391-2402.

Matt, N., Ghyselinck, N.B., Wendling, O., Chambon, P., and Mark, M. (2003). Retinoic acid-induced developmental defects are mediated by RARbeta/RXR heterodimers in the pharyngeal endoderm. Development *130*, 2083-2093.

Matthews, L., Berry, A., Ohanian, V., Ohanian, J., Garside, H., and Ray, D. (2008). Caveolin mediates rapid glucocorticoid effects and couples glucocorticoid action to the antiproliferative program. Mol Endocrinol *22*, 1320-1330.

McEwan, I.J., Lavery, D., Fischer, K., and Watt, K. (2007). Natural disordered sequences in the amino terminal domain of nuclear receptors: lessons from the androgen and glucocorticoid receptors. Nucl Recept Signal *5*, e001.

McKenna, N.J., Xu, J., Nawaz, Z., Tsai, S.Y., Tsai, M.J., and O'Malley, B.W. (1999). Nuclear receptor coactivators: multiple enzymes, multiple complexes, multiple functions. J Steroid Biochem Mol Biol *69*, 3-12.

Meijsing, S.H., Pufall, M.A., So, A.Y., Bates, D.L., Chen, L., and Yamamoto, K.R. (2009). DNA binding site sequence directs glucocorticoid receptor structure and activity. Science *324*, 407-410.

Meyer, A., and Van de Peer, Y. (2005). From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). Bioessays *27*, 937-945.

Miloso, M., Villa, D., Crimi, M., Galbiati, S., Donzelli, E., Nicolini, G., and Tredici, G. (2004). Retinoic acid-induced neuritogenesis of human neuroblastoma SH-SY5Y cells is ERK independent and PKC dependent. J Neurosci Res *75*, 241-252.

Moens, C.B., and Prince, V.E. (2002). Constructing the hindbrain: insights from the zebrafish. Dev Dyn *224*, 1-17.

Mollard, R., Viville, S., Ward, S.J., Decimo, D., Chambon, P., and Dolle, P. (2000). Tissue-specific expression of retinoic acid receptor isoform transcripts in the mouse embryo. Mech Dev *94*, 223-232.

Molotkov, A., Molotkova, N., and Duester, G. (2005). Retinoic acid generated by Raldh2 in mesoderm is required for mouse dorsal endodermal pancreas development. Dev Dyn *232*, 950-957.

Moras, D., and Gronemeyer, H. (1998). The nuclear receptor ligand-binding domain: structure and function. Curr Opin Cell Biol *10*, 384-391.

Moutier, E., Ye, T., Choukrallah, M.A., Urban, S., Osz, J., Chatagnon, A., Delacroix, L., Langer, D., Rochel, N., Moras, D., *et al.* (2012). Retinoic acid receptors recognize the mouse genome through binding elements with diverse spacing and topology. J Biol Chem *287*, 26328-26341.

Murakami, Y., Pasqualetti, M., Takio, Y., Hirano, S., Rijli, F.M., and Kuratani, S. (2004). Segmental development of reticulospinal and branchiomotor neurons in lamprey: insights into the evolution of the vertebrate hindbrain. Development *131*, 983-995.

Nagatomo, K., and Fujiwara, S. (2003). Expression of Raldh2, Cyp26 and Hox-1 in normal and retinoic acid-treated Ciona intestinalis embryos. Gene Expr Patterns *3*, 273-277.

Nagpal, S., Friant, S., Nakshatri, H., and Chambon, P. (1993). RARs and RXRs: evidence for two autonomous transactivation functions (AF-1 and AF-2) and heterodimerization in vivo. EMBO J *12*, 2349-2360.

Nagpal, S., Saunders, M., Kastner, P., Durand, B., Nakshatri, H., and Chambon, P. (1992). Promoter context- and response element-dependent specificity of the transcriptional activation and modulating functions of retinoic acid receptors. Cell *70*, 1007-1019.

Nagy, L., Kao, H.Y., Love, J.D., Li, C., Banayo, E., Gooch, J.T., Krishna, V., Chatterjee, K., Evans, R.M., and Schwabe, J.W. (1999). Mechanism of corepressor binding and release from nuclear hormone receptors. Genes Dev *13*, 3209-3216.

Nelson, J.S. (2006). Fishes of the world, Vol 4th edition.

Niederreither, K., and Dolle, P. (2008). Retinoic acid in development: towards an integrated view. Nat Rev Genet 9, 541-553.

Niederreither, K., McCaffery, P., Drager, U.C., Chambon, P., and Dolle, P. (1997). Restricted expression and retinoic acid-induced downregulation of the retinaldehyde dehydrogenase type 2 (RALDH-2) gene during mouse development. Mech Dev *62*, 67-78.

Niederreither, K., Vermot, J., Schuhbaur, B., Chambon, P., and Dolle, P. (2000). Retinoic acid synthesis and hindbrain patterning in the mouse embryo. Development *127*, 75-85.

Nielsen, R., Pedersen, T.A., Hagenbeek, D., Moulos, P., Siersbaek, R., Megens, E., Denissov, S., Borgesen, M., Francoijs, K.J., Mandrup, S., *et al.* (2008). Genome-wide profiling of PPARgamma:RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. Genes Dev 22, 2953-2967.

Novitch, B.G., Wichterle, H., Jessell, T.M., and Sockanathan, S. (2003). A requirement for retinoic acid-mediated transcriptional activation in ventral neural patterning and motor neuron specification. Neuron *40*, 81-95.

Ohno, S. (1970). Evolution by gene duplication.

Oliveira, E., Casado, M., Raldua, D., Soares, A., Barata, C., and Pina, B. (2013). Retinoic acid receptors' expression and function during zebrafish early development. J Steroid Biochem Mol Biol *138C*, 143-151.

Olivera-Martinez, I., and Storey, K.G. (2007). Wnt signals provide a timing mechanism for the FGF-retinoid differentiation switch during vertebrate body axis extension. Development *134*, 2125-2135.

Ollikainen, N., Chandsawangbhuwana, C., and Baker, M.E. (2006). Evolution of the thyroid hormone, retinoic acid, ecdysone and liver X receptors. Integr Comp Biol *46*, 815-826.

Orlov, I., Rochel, N., Moras, D., and Klaholz, B.P. (2012). Structure of the full human RXR/VDR nuclear receptor heterodimer complex with its DR3 target DNA. EMBO J *31*, 291-300.

Ostrom, M., Loffler, K.A., Edfalk, S., Selander, L., Dahl, U., Ricordi, C., Jeon, J., Correa-Medina, M., Diez, J., and Edlund, H. (2008). Retinoic Acid Promotes the Generation of Pancreatic Endocrine Progenitor Cells and Their Further Differentiation into beta-Cells. Plos One *3*.

Pan, F.C., Chen, Y., Bayha, E., and Pieler, T. (2007). Retinoic acid-mediated patterning of the pre-pancreatic endoderm in Xenopus operates via direct and indirect mechanisms. Mech Dev *124*, 518-531.

Pasco-Viel, E., Charles, C., Chevret, P., Semon, M., Tafforeau, P., Viriot, L., and Laudet, V. (2010). Evolutionary trends of the pharyngeal dentition in Cypriniformes (Actinopterygii: Ostariophysi). PLoS One *5*, e11293.

Pavri, R., Lewis, B., Kim, T.K., Dilworth, F.J., Erdjument-Bromage, H., Tempst, P., de Murcia, G., Evans, R., Chambon, P., and Reinberg, D. (2005). PARP-1 determines specificity in a retinoid signaling pathway via direct modulation of mediator. Mol Cell *18*, 83-96.

Pedram, A., Razandi, M., Sainson, R.C., Kim, J.K., Hughes, C.C., and Levin, E.R. (2007). A conserved mechanism for steroid receptor translocation to the plasma membrane. J Biol Chem *282*, 22278-22288.

Percipalle, P., and Visa, N. (2006). Molecular functions of nuclear actin in transcription. J Cell Biol *172*, 967-971.

Perissi, V., Scafoglio, C., Zhang, J., Ohgi, K.A., Rose, D.W., Glass, C.K., and Rosenfeld, M.G. (2008). TBL1 and TBLR1 phosphorylation on regulated gene promoters overcomes dual CtBP and NCoR/SMRT transcriptional repression checkpoints. Mol Cell *29*, 755-766.

Perissi, V., Staszewski, L.M., McInerney, E.M., Kurokawa, R., Krones, A., Rose, D.W., Lambert, M.H., Milburn, M.V., Glass, C.K., and Rosenfeld, M.G. (1999). Molecular determinants of nuclear receptor-corepressor interaction. Genes Dev *13*, 3198-3208.

Pierani, A., Brenner-Morton, S., Chiang, C., and Jessell, T.M. (1999). A sonic hedgehog-independent, retinoid-activated pathway of neurogenesis in the ventral spinal cord. Cell *97*, 903-915.

Piskunov, A., and Rochette-Egly, C. (2011a). MSK1 and nuclear receptor signaling. In MSKs, S. Arthur, and L. Vermeulen, eds. (Austin: Landes Biosciences), p. in press.

Piskunov, A., and Rochette-Egly, C. (2011b). A retinoic acid receptor RAR α pool present in membrane lipid rafts forms complexes with G protein αQ to activate p38MAPK. Oncogene *7*.

Pogenberg, V., Guichou, J.F., Vivat-Hannah, V., Kammerer, S., Perez, E., Germain, P., de Lera, A.R., Gronemeyer, H., Royer, C.A., and Bourguet, W. (2005). Characterization of the interaction between retinoic acid receptor/retinoid X receptor (RAR/RXR) heterodimers and transcriptional coactivators through structural and fluorescence anisotropy studies. Journal of Biological Chemistry *280*, 1625-1633.

Poon, M.M., and Chen, L. (2008). Retinoic acid-gated sequence-specific translational control by RARalpha. Proc Natl Acad Sci U S A *105*, 20303-20308.

Prin, F., and Dhouailly, D. (2004). How and when the regional competence of chick epidermis is established: feathers vs. scutate and reticulate scales, a problem en route to a solution. Int J Dev Biol *48*, 137-148.

Privalsky, M.L. (2004). The role of corepressors in transcriptional regulation by nuclear hormone receptors. Annu Rev Physiol *66*, 315-360.

Protas, M.E., Hersey, C., Kochanek, D., Zhou, Y., Wilkens, H., Jeffery, W.R., Zon, L.I., Borowsky, R., and Tabin, C.J. (2006). Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. Nat Genet *38*, 107-111.

Quinlan, R., Gale, E., Maden, M., and Graham, A. (2002). Deficits in the posterior pharyngeal endoderm in the absence of retinoids. Dev Dyn *225*, 54-60.

Ramagopalan, S.V., Heger, A., Berlanga, A.J., Maugeri, N.J., Lincoln, M.R., Burrell, A., Handunnetthi, L., Handel, A.E., Disanto, G., Orton, S.M., *et al.* (2010). A ChIP-seq defined genome-wide map of vitamin D receptor binding: associations with disease and evolution. Genome Res *20*, 1352-1360.

Reijntjes, S., Blentic, A., Gale, E., and Maden, M. (2005). The control of morphogen signalling: regulation of the synthesis and catabolism of retinoic acid in the developing embryo. Dev Biol *285*, 224-237.

Reimand, J., Hui, S., Jain, S., Law, B., and Bader, G.D. (2012). Domain-mediated protein interaction prediction: From genome to network. FEBS Lett *586*, 2751-2763.

Ren, X., Li, Y., Ma, X., Zheng, L., Xu, Y., and Wang, J. (2007). Activation of p38/MEF2C pathway by all-trans retinoic acid in cardiac myoblasts. Life Sci *81*, 89-96.

Renaud, J.P., and Moras, D. (2000). Structural studies on nuclear receptors. Cell Mol Life Sci *57*, 1748-1769.

Renaud, J.P., Rochel, N., Ruff, M., Vivat, V., Chambon, P., Gronemeyer, H., and Moras, D. (1995). Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid. Nature *378*, 681-689.

Rhinn, M., and Dolle, P. (2012). Retinoic acid signalling during development. Development *139*, 843-858.

Robinson-Rechavi, M., Carpentier, A.S., Duffraisse, M., and Laudet, V. (2001). How many nuclear hormone receptors are there in the human genome? Trends Genet *17*, 554-556.

Rochette-Egly, C., Adam, S., Rossignol, M., Egly, J.M., and Chambon, P. (1997). Stimulation of RAR alpha activation function AF-1 through binding to the general transcription factor TFIIH and phosphorylation by CDK7. Cell *90*, 97-107.

Rochette-Egly, C., and Germain, P. (2009). Dynamic and combinatorial control of gene expression by nuclear retinoic acid receptors (RARs). Nucl Recept Signal *7*, e005.

Ruest, L.B., Xiang, X., Lim, K.C., Levi, G., and Clouthier, D.E. (2004). Endothelin-A receptor-dependent and -independent signaling pathways in establishing mandibular identity. Development *131*, 4413-4423.

Sadier, A., Viriot, L., Pantalacci, S., and Laudet, V. (2013). The ectodysplasin pathway: from diseases to adaptations. Trends Genet.

Sakai, Y., Meno, C., Fujii, H., Nishino, J., Shiratori, H., Saijoh, Y., Rossant, J., and Hamada, H. (2001). The retinoic acid-inactivating enzyme CYP26 is essential for establishing an uneven distribution of retinoic acid along the anterio-posterior axis within the mouse embryo. Genes & Development *15*, 213-225.

Samarut, E., and Rochette-Egly, C. (2012). Nuclear retinoic acid receptors: conductors of the retinoic acid symphony during development. Mol Cell Endocrinol *348*, 348-360.

Santagati, F., and Rijli, F.M. (2003). Cranial neural crest and the building of the vertebrate head. Nat Rev Neurosci *4*, 806-818.

Schilling, T.F. (2008). Anterior-posterior patterning and segmentation of the vertebrate head. Integr Comp Biol *48*, 658-667.

Schubert, M., Holland, N.D., Laudet, V., and Holland, L.Z. (2006). A retinoic acid-Hox hierarchy controls both anterior/posterior patterning and neuronal specification in the developing central nervous system of the cephalochordate amphioxus. Developmental Biology *296*, 190-202.

Schubert, M., Yu, J.K., Holland, N.D., Escriva, H., Laudet, V., and Holland, L.Z. (2005). Retinoic acid signaling acts via Hox1 to establish the posterior limit of the pharynx in the chordate amphioxus. Development *132*, 61-73.

Seritrakul, P., Samarut, E., Lama, T.T., Gibert, Y., Laudet, V., and Jackman, W.R. (2012). Retinoic acid expands the evolutionarily reduced dentition of zebrafish. FASEB J *26*, 5014-5024.

Sharpe, C., and Goldstone, K. (2000). Retinoid signalling acts during the gastrula stages to promote primary neurogenesis. Int J Dev Biol *44*, 463-470.

Shimozono, S., Iimura, T., Kitaguchi, T., Higashijima, S., and Miyawaki, A. (2013). Visualization of an endogenous retinoic acid gradient across embryonic development. Nature *496*, 363-366.

Simoes-Costa, M.S., Azambuja, A.P., and Xavier-Neto, J. (2008). The search for non-chordate retinoic acid signaling: lessons from chordates. J Exp Zool B Mol Dev Evol *310*, 54-72.

Sirbu, I.O., and Duester, G. (2006). Retinoic-acid signalling in node ectoderm and posterior neural plate directs left-right patterning of somitic mesoderm. Nat Cell Biol *8*, 271-277.

Smith, M.M., Fraser, G.J., and Mitsiadis, T.A. (2009). Dental Lamina as Source of Odontogenic Stem Cells: Evolutionary Origins and Developmental Control of Tooth Generation in Gnathostomes. Journal of Experimental Zoology Part B-Molecular and Developmental Evolution *312B*, 260-280.

Smith, W.C., Nakshatri, H., Leroy, P., Rees, J., and Chambon, P. (1991). A retinoic acid response element is present in the mouse cellular retinol binding protein I (mCRBPI) promoter. EMBO J *10*, 2223-2230.

Soprano, D.R., and Soprano, K.J. (1995). Retinoids as Teratogens. Annual Review of Nutrition *15*, 111-132.

Sporn, M.B., Roberts, A.B., and Goodman, D.S. (1994). The Retinoids. Biology, Chemistry, and Medecine. (New York: Raven Press).

Srinivas, H., Xia, D., Moore, N.L., Uray, I.P., Kim, H., Ma, L., Weigel, N.L., Brown, P.H., and Kurie, J.M. (2006). Akt phosphorylates and suppresses the transactivation of retinoic acid receptor alpha. Biochem J *395*, 653-662.

Stafford, D., and Prince, V.E. (2002). Retinoic acid signaling is required for a critical early step in zebrafish pancreatic development. Curr Biol *12*, 1215-1220.

Stavridis, M.P., Collins, B.J., and Storey, K.G. (2010). Retinoic acid orchestrates fibroblast growth factor signalling to drive embryonic stem cell differentiation. Development *137*, 881-890.

Stevens, C.E., and Hume, I.D. (2004). Comparative Physiology of the Vertebrate Digestive System.

Stock, D.W. (2007). Zebrafish dentition in comparative context. J Exp Zool B Mol Dev Evol *308*, 523-549.

Stock, D.W., Jackman, W.R., and Trapani, J. (2006). Developmental genetic mechanisms of evolutionary tooth loss in cypriniform fishes. Development *133*, 3127-3137.

Sun, K., Montana, V., Chellappa, K., Brelivet, Y., Moras, D., Maeda, Y., Parpura, V., Paschal, B.M., and Sladek, F.M. (2007). Phosphorylation of a conserved serine in the deoxyribonucleic acid binding domain of nuclear receptors alters intracellular localization. Mol Endocrinol *21*, 1297-1311.

Tahayato, A., Dolle, P., and Petkovich, M. (2003). Cyp26C1 encodes a novel retinoic acid-metabolizing enzyme expressed in the hindbrain, inner ear, first branchial arch and tooth buds during murine development. Gene Expr Patterns *3*, 449-454.

Takechi, M., Adachi, N., Hirai, T., Kuratani, S., and Kuraku, S. (2013). The Dlx genes as clues to vertebrate genomics and craniofacial evolution. Seminars in Cell & Developmental Biology *24*, 110-118.

Tallafuss, A., Hale, L.A., Yan, Y.L., Dudley, L., Eisen, J.S., and Postlethwait, J.H. (2006). Characterization of retinoid-X receptor genes rxra, rxrba, rxrbb and rxrg during zebrafish development. Gene Expr Patterns *6*, 556-565.

Tanaka, Y., Okada, Y., and Hirokawa, N. (2005). FGF-induced vesicular release of Sonic hedgehog and retinoic acid in leftward nodal flow is critical for left-right determination. Nature *435*, 172-177.

Taneja, R., Rochette-Egly, C., Plassat, J.L., Penna, L., Gaub, M.P., and Chambon, P. (1997). Phosphorylation of activation functions AF-1 and AF-2 of RAR alpha and RAR gamma is indispensable for differentiation of F9 cells upon retinoic acid and cAMP treatment. Embo J *16*, 6452-6465.

Taylor, J.S., and Raes, J. (2004). Duplication and divergence: the evolution of new genes and old ideas. Annu Rev Genet *38*, 615-643.

Theodosiou, M., Laudet, V., and Schubert, M. (2010). From carrot to clinic: an overview of the retinoic acid signaling pathway. Cell Mol Life Sci *67*, 1423-1445.

Ting, H.J., and Chang, C. (2008). Actin associated proteins function as androgen receptor coregulators: An implication of androgen receptor's roles in skeletal muscle. Journal of Steroid Biochemistry and Molecular Biology *111*, 157-163.

Tsukui, T., Capdevila, J., Tamura, K., Ruiz-Lozano, P., Rodriguez-Esteban, C., Yonei-Tamura, S., Magallón, J., Chandraratna, R.A., Chien, K., Blumberg, B., *et al.* (1999). Multiple left-right asymmetry defects in Shh(-/-) mutant mice unveil a convergence of the shh and retinoic acid pathways in the control of Lefty-1. Proc Natl Acad Sci U S A 96, 11376-11381.

Umesono, K., Murakami, K.K., Thompson, C.C., and Evans, R.M. (1991a). Direct repeats as selective reponse elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. Cell *65*, 1255-1266.

Umesono, K., Murakami, K.K., Thompson, C.C., and Evans, R.M. (1991b). Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. Cell *65*, 1255-1266.

Veitch, E., Begbie, J., Schilling, T.F., Smith, M.M., and Graham, A. (1999). Pharyngeal arch patterning in the absence of neural crest. Curr Biol 9, 1481-1484.

Vermeulen, M., Carrozza, M.J., Lasonder, E., Workman, J.L., Logie, C., and Stunnenberg, H.G. (2004). In vitro targeting reveals intrinsic histone tail specificity of the Sin3/histone deacetylase and N-CoR/SMRT corepressor complexes. Mol Cell Biol *24*, 2364-2372.

Vermot, J., Gallego Llamas, J., Fraulob, V., Niederreither, K., Chambon, P., and Dolle, P. (2005). Retinoic acid controls the bilateral symmetry of somite formation in the mouse embryo. Science *308*, 563-566.

Vermot, J., and Pourquie, O. (2005). Retinoic acid coordinates somitogenesis and left-right patterning in vertebrate embryos. Nature *435*, 215-220.

Vicent, G.P., Ballare, C., Nacht, A.S., Clausell, J., Subtil-Rodriguez, A., Quiles, I., Jordan, A., and Beato, M. (2006). Induction of progesterone target genes requires activation of Erk and Msk kinases and phosphorylation of histone H3. Mol Cell *24*, 367-381.

Vieux-Rochas, M., Bouhali, K., Baudry, S., Fontaine, A., Coen, L., and Levi, G. (2010). Irreversible Effects of Retinoic Acid Pulse on Xenopus Jaw Morphogenesis: New Insight Into Cranial Neural Crest Specification. Birth Defects Research Part B-Developmental and Reproductive Toxicology *89*, 493-503.

Vieux-Rochas, M., Coen, L., Sato, T., Kurihara, Y., Gitton, Y., Barbieri, O., Le Blay, K., Merlo, G., Ekker, M., Kurihara, H., *et al.* (2007). Molecular dynamics of retinoic acid-

induced craniofacial malformations: implications for the origin of gnathostome jaws. PLoS One *2*, e510.

Vucetic, Z., Zhang, Z.P., Zhao, J.H., Wang, F., Soprano, K.J., and Soprano, D.R. (2008). Acinus-S' represses retinoic acid receptor (RAR)-Regulated gene expression through interaction with the B domains of RARs. Molecular and Cellular Biology *28*, 2549-2558.

Wada, H., Escriva, H., Zhang, S., and Laudet, V. (2006). Conserved RARE localization in amphioxus Hox clusters and implications for Hox code evolution in the vertebrate neural crest. Dev Dyn *235*, 1522-1531.

Wada, H., Garcia-Fernandez, J., and Holland, P.W. (1999). Colinear and segmental expression of amphioxus Hox genes. Dev Biol *213*, 131-141.

Waxman, J.S., and Yelon, D. (2007). Comparison of the expression patterns of newly identified zebrafish retinoic acid and retinoid X receptors. Dev Dyn *236*, 587-595.

Waxman, J.S., and Yelon, D. (2011). Zebrafish retinoic acid receptors function as context-dependent transcriptional activators. Developmental Biology *352*, 128-140.

Welboren, W.J., van Driel, M.A., Janssen-Megens, E.M., van Heeringen, S.J., Sweep, F.C., Span, P.N., and Stunnenberg, H.G. (2009). ChIP-Seq of ERalpha and RNA polymerase II defines genes differentially responding to ligands. EMBO J *28*, 1418-1428.

Wellik, D.M. (2007). Hox patterning of the vertebrate axial skeleton. Dev Dyn *236*, 2454-2463.

Wendling, O., Dennefeld, C., Chambon, P., and Mark, M. (2000). Retinoid signaling is essential for patterning the endoderm of the third and fourth pharyngeal arches. Development *127*, 1553-1562.

White, J.C., Highland, M., Kaiser, M., and Clagett-Dame, M. (2000). Vitamin A deficiency results in the dose-dependent acquisition of anterior character and shortening of the caudal hindbrain of the rat embryo. Dev Biol *220*, 263-284.

White, J.H., Fernandes, I., Mader, S., and Yang, X.J. (2004). Corepressor recruitment by agonist-bound nuclear receptors. Vitam Horm *68*, 123-143.

White, R.J., Nie, Q., Lander, A.D., and Schilling, T.F. (2007). Complex regulation of cyp26a1 creates a robust retinoic acid gradient in the zebrafish embryo. PLoS Biol *5*, e304.

Wilson, L., Gale, E., Chambers, D., and Maden, M. (2004). Retinoic acid and the control of dorsoventral patterning in the avian spinal cord. Dev Biol *269*, 433-446.

Wingert, R.A., Selleck, R., Yu, J., Song, H.D., Chen, Z., Song, A., Zhou, Y., Thisse, B., Thisse, C., McMahon, A.P., *et al.* (2007). The cdx genes and retinoic acid control the positioning and segmentation of the zebrafish pronephros. PLoS Genet *3*, 1922-1938.

Wong, C.W., and Privalsky, M.L. (1998). Transcriptional silencing is defined by isoform- and heterodimer-specific interactions between nuclear hormone receptors and corepressors. Molecular and Cellular Biology *18*, 5724-5733.

Wood, H., Pall, G., and Morrisskay, G. (1994). Exposure to Retinoic Acid before or after the Onset of Somitogenesis Reveals Separate Effects on Rhombomeric Segmentation and 3' Hoxb Gene-Expression Domains. Development *120*, 2279-2285.

Wray, G.A. (2007). The evolutionary significance of cis-regulatory mutations. Nat Rev Genet *8*, 206-216.

Wright, P.E., and Dyson, H.J. (2009). Linking folding and binding. Curr Opin Struct Biol *19*, 31-38.

Wuarin, L., Sidell, N., and de Vellis, J. (1990). Retinoids increase perinatal spinal cord neuronal survival and astroglial differentiation. Int J Dev Neurosci *8*, 317-326.

Yu, J.K., Satou, Y., Holland, N.D., Shin, I.T., Kohara, Y., Satoh, N., Bronner-Fraser, M., and Holland, L.Z. (2007). Axial patterning in cephalochordates and the evolution of the organizer. Nature *445*, 613-617.

Zákány, J., Kmita, M., Alarcon, P., de la Pompa, J.L., and Duboule, D. (2001). Localized and transient transcription of Hox genes suggests a link between patterning and the segmentation clock. Cell *106*, 207-217.

Zamir, I., Zhang, J., and Lazar, M.A. (1997). Stoichiometric and steric principles governing repression by nuclear hormone receptors. Genes Dev *11*, 835-846.

Zhao, J., Zhang, Z., Vucetic, Z., Soprano, K.J., and Soprano, D.R. (2009). HACE1: A novel repressor of RAR transcriptional activity. J Cell Biochem *107*, 482-493.

Zheng, B., Han, M., Bernier, M., and Wen, J.K. (2009). Nuclear actin and actin-binding proteins in the regulation of transcription and gene expression. FEBS J *276*, 2669-2685.

Zimmerman, T.L., Thevananther, S., Ghose, R., Burns, A.R., and Karpen, S.J. (2006). Nuclear export of retinoid X receptor alpha in response to interleukin-1beta-mediated cell signaling: roles for JNK and SER260. J Biol Chem *281*, 15434-15440.