



UNIL | Université de Lausanne

Unicentre

CH-1015 Lausanne

<http://serval.unil.ch>

Year : 2019

CIRCADIAN MODULATION OF GENOME-WIDE RXR BINDING IN MOUSE LIVER TISSUE

Trang Khanh Bao

Trang Khanh Bao, 2019, CIRCADIAN MODULATION OF GENOME-WIDE RXR BINDING IN
MOUSE LIVER TISSUE

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive <http://serval.unil.ch>

Document URN : urn:nbn:ch:serval-BIB_C4CE7516215F0

Droits d'auteur

L'Université de Lausanne attire expressément l'attention des utilisateurs sur le fait que tous les documents publiés dans l'Archive SERVAL sont protégés par le droit d'auteur, conformément à la loi fédérale sur le droit d'auteur et les droits voisins (LDA). A ce titre, il est indispensable d'obtenir le consentement préalable de l'auteur et/ou de l'éditeur avant toute utilisation d'une oeuvre ou d'une partie d'une oeuvre ne relevant pas d'une utilisation à des fins personnelles au sens de la LDA (art. 19, al. 1 lettre a). A défaut, tout contrevenant s'expose aux sanctions prévues par cette loi. Nous déclinons toute responsabilité en la matière.

Copyright

The University of Lausanne expressly draws the attention of users to the fact that all documents published in the SERVAL Archive are protected by copyright in accordance with federal law on copyright and similar rights (LDA). Accordingly it is indispensable to obtain prior consent from the author and/or publisher before any use of a work or part of a work for purposes other than personal use within the meaning of LDA (art. 19, para. 1 letter a). Failure to do so will expose offenders to the sanctions laid down by this law. We accept no liability in this respect.



UNIL | Université de Lausanne

Faculté de biologie
et de médecine

Département of Genomics
CIG - Centre Intégréatif de Génomique

CIRCADIAN MODULATION OF GENOME-WIDE RXR BINDING IN MOUSE LIVER TISSUE

Thèse de doctorat ès sciences de la vie (PhD) programme
Integrative Experimental and Computational Biology

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Khanh Bao TRANG

M.D diplômé, l'Université de Pharmacie et Médecine d'Ho-Chi-Minh-Ville, Vietnam
Maîtrise en Santé Publique, Université de Debrecen, Hongrie

Jury

Prof. Nicolas SALAMIN, Président
Prof. Béatrice DESVERGNE, Directeur de thèse
Dr Nicolas GUEX, Co-directeur de thèse
Dre Mercedes RICOTE, expert
Dr Philipp BUCHER, expert
Prof. Ioannis Xenarios, expert

Lausanne 2019

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président·e	Monsieur	Prof.	Nicolas	Salamin
Directeur·trice de thèse	Madame	Prof.	Béatrice	Desvergne
Co-directeur·trice	Monsieur	Dr	Nicolas	Guex
Expert·e-s	Monsieur	Prof.	Ioannis	Xenarios
	Madame	Dre	Mercedes	Ricote
	Monsieur	Dr	Philipp	Bucher

le Conseil de Faculté autorise l'impression de la thèse de

Madame Khanh Bao Trang

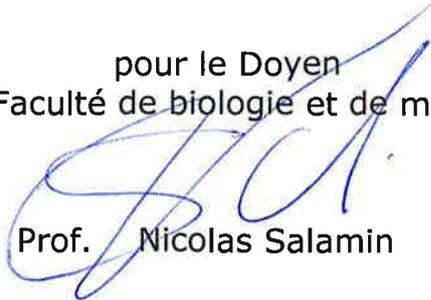
Master of Public Health, University of Debrecen, Hongrie

intitulée

**Circadian modulation of genome-wide RXR
binding in mouse liver tissue**

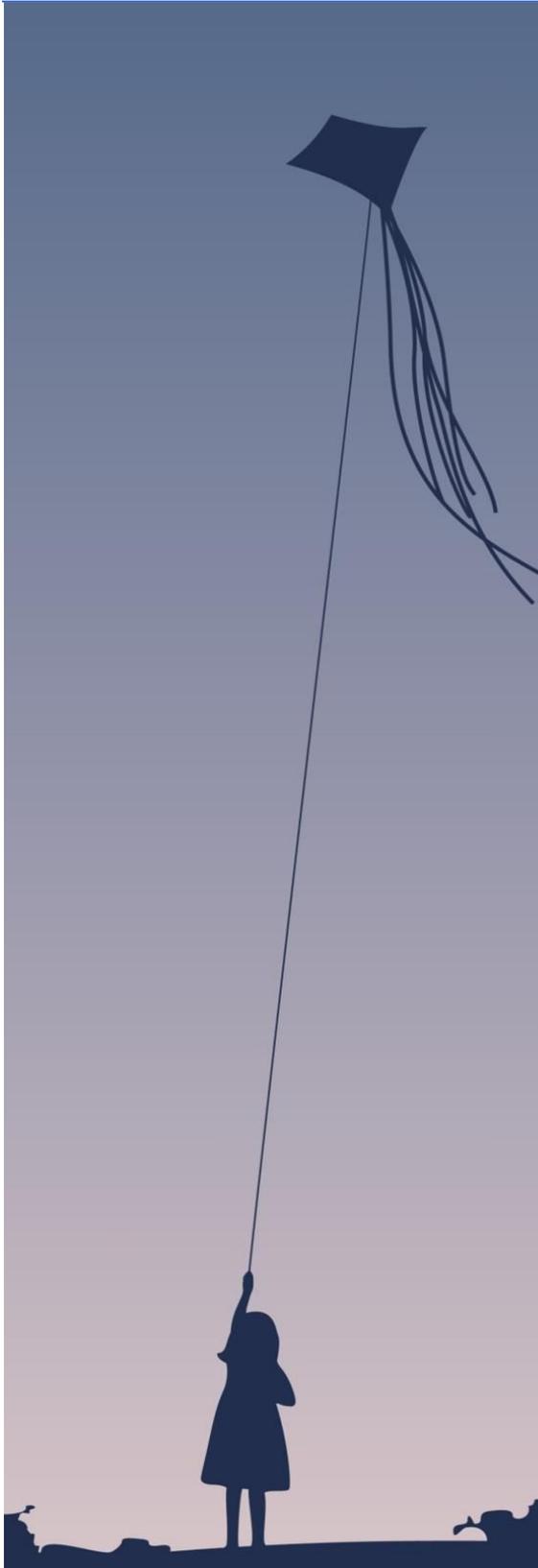
Lausanne, le 11 juillet 2019

pour le Doyen
de la Faculté de biologie et de médecine



Prof. Nicolas Salamin

ACKNOWLEDGEMENT



“Someone once told me that children are like kites. You struggle just to get them in the air; they crash; you add a longer tail. Then they get caught in a tree; you climb up and bring them down, and untangle the string; you run to get them aloft again. Finally, the kite is airborne, and it flies higher and higher, as you let out more string, until it’s so high in the sky, it looks like a bird. And if the string snaps, and you’ve done your job right, the kite will continue to soar in the wind, all by itself.” — Charmian Carr

For Nicolas – “the wind”, who lifted my bioinformatic passion,
and for Beatrice – “the string”, who kept my head earth-bound and steady.

You flew my kite high and strong.

Khanh – “the kite”

TABLE OF CONTENT

SUMMARY.....	1
RÉSUMÉ.....	2
ABBREVIATIONS.....	3
1. INTRODUCTION	5
1.1 RXR AS THE CENTRAL MEMBER OF NUCLEAR RECEPTORS SUPERFAMILY.....	6
1.1.1 <i>The discovery of Nuclear receptor and the birth of the Molecular Endocrinology field.....</i>	6
1.1.2 <i>Nuclear receptors genes and their evolution.....</i>	9
1.1.3 <i>The modular structure of RXR, and more broadly of all the NRs.....</i>	13
1.1.3.1 The A/B domain.....	13
1.1.3.2 The C domain.....	15
1.1.3.3 The D domain.....	15
1.1.3.4 The E domain.....	16
1.1.3.5 The F domain.....	18
1.1.4 <i>The dimerization property of NR and the specific role of RXR in heterodimerization.....</i>	19
1.1.5 <i>The controversial existence of endogenous ligands for RXR.....</i>	25
1.1.6 <i>RXR specific biological functions.....</i>	30
1.1.6.1 Total ablation and the roles of RXR isotypes in embryogenesis.....	30
1.1.6.2 Tissue specific ablation and diverse effects.....	34
1.1.6.3 Clinical relevance of subtypes.....	35
1.2 THE INVOLVEMENT OF RXR IN CIRCADIAN RHYTHM	36
1.2.1 <i>The CycliX project: interrogating the coordination of circadian, nutrient and cell cycles by transcription factors, including NRs.....</i>	36
1.2.1.1 The global landscape of epigenetics along the diurnal cycle.....	38
1.2.1.2 The integrating circadian cycle with feeding rhythm.....	39
1.2.1.3 Cell cycle - the third cog coming into focus.....	40
1.2.2 <i>The possible role of RXR in coordinating the three liver cycles.....</i>	40
1.2.3 <i>Aim of the project.....</i>	43
2. EXPERIMENTAL MATERIALS AND METHODS.....	44
2.1 ANIMALS AND TREATMENT	45
2.2 CHROMATIN IMMUNOPRECIPITATION FOLLOWED BY SEQUENCING PROTOCOL (CHIP-SEQ)	46
2.3 CHIP-SEQ LIBRARY CONTROL AND READ MAPPING	47
2.4 PEAK CALLING.....	49
2.4.1 <i>Identifying the enriched regions.....</i>	49
2.4.2 <i>Peak refinement within the regions.....</i>	51
2.5 MOTIF ANALYSIS.....	54
2.6 GENES ENRICHMENT ANALYSIS	54
2.7 BIOINFORMATICS TOOLS	55
2.8 LIST OF USED DATA SETS.....	55
3. RESULTS.....	56
3.1 THE PEAK-CALLER METHOD IS EFFICIENT TO IDENTIFY CHIP-SEQ ENRICHED REGIONS	57
3.2 PEAK REFINEMENTS.....	60
3.2.1 <i>Identifying single, double, and multiple peaks.....</i>	60
3.2.2 <i>RXR peaks validation by analyzing the overlaps with DNase hypersensitive sites.....</i>	63
3.2.3 <i>PhastCons analyses and peak shapes.....</i>	65

3.3	PEAK ANNOTATION	68
3.3.1	<i>RXR peaks prefer promoter regions, including transcription start sites</i>	68
3.3.2	<i>Relationship between peak annotation and gene expression levels</i>	70
3.3.3	<i>Clustered RXR peaks and gene expression</i>	72
3.4	THE PLATEAUS REVEAL A PROMISCUOUS RELATIONSHIP BETWEEN RXR AND POL-II.....	73
3.4.1	<i>The plateau peaks cover entire genes of small size</i>	73
3.4.2	<i>Colocalization of RNA Polymerase II and RXR on High plateaus and at transcription starting sites</i>	75
3.5	CIRCADIAN ANALYSES.....	78
3.5.1	<i>Clustering RXR peaks using peaks tag density values</i>	78
3.5.2	<i>The sinusoidal fit</i>	82
3.5.3	<i>Circadian functional analyses</i>	89
3.6	MOTIF SEARCH IDENTIFIES NUCLEAR RECEPTOR CONSENSUS BINDING SITES.....	93
3.6.1	<i>Unbiased motif searches</i>	93
3.6.2	<i>Searching for RXR heterodimer consensus-derived binding sites</i>	97
3.6.3	<i>Distribution of motif patterns across RXR peak groups, time points and annotations</i>	103
3.6.4	<i>Distribution of the different Nuclear receptor binding modes across target genes</i>	107
3.7	ENRICHED PATHWAYS ANALYSES	109
3.7.1	<i>The highly expressed – high plateau genes</i>	109
3.7.2	<i>Comparative pathways enrichment among RXR binding modes</i>	111
3.7.3	<i>The gating effect on pathway analysis</i>	112
4.	DISCUSSION	115
4.1	PEAK-CALLERS ARE NOT ONE-SIZE-FIT-ALL TOOLS	116
4.2	ASSOCIATION OF AN RXR PEAK WITH ITS TARGET GENE IS NOT PERFECT.....	118
4.3	CHROMATIN – REGULATORS ORGANIZATION FOR CO-EXPRESSED/CO-REGULATED GENES	120
4.4	POL-II: THE UNEXPECTED EXPECTED JOURNEY.....	122
4.5	WHAT IS THE FUNCTION OF HIGH PLATEAUS	125
4.6	CIRCADIAN RHYTHMS	127
4.7	THE HETERODIMER PARTNERS: PUBLIC PLAYGROUND (WITH RULES).....	130
	REFERENCES.....	134

SUMMARY

Nuclear receptors (NRs) are transcription factors that regulate genes involved in the development, homeostasis and metabolism of the organisms. Their expression is dynamic and, in the liver, follows a circadian pattern. Among the 48 members of this protein super-family, the retinoid receptor X (RXR) has a special status as an obligatory heterodimer partner for several other NRs. Unlike most of its partners, RXR is suitable for chromatin immunoprecipitation (ChIP) experiments, and is therefore a candidate of choice for genome-wide characterization of NR-dependent regulations. The aim of this thesis was to map the repertoire of RXR chromatin binding sites and explore the ability of RXR to regulate the rhythmic functions of the liver during the day. For that purpose, we used a pan-RXR antibody to perform a ChIP-seq experiment on mouse liver at 7 different time-points of the day over 24 hours.

Faced with the complexity of multiple comparisons between the 7 sets of ChIP-seq peaks obtained at the genome scale, we have developed an in-house method for analyzing the 7 samples together. Just over 30,000 binding sites have been identified, divided into 4 categories: solitary peaks, clustered peaks, high plateaus (wide regions with continuous high tag density), and low plateaus. RXR peaks were preferentially located in promoter regions, and the expression of three-quarters of the corresponding putative RXR target genes, representing almost half of mouse genes, was detectable by microarray. The abundance of RXR binding sites for each gene, but not the intensity of the signal observed at each binding site, correlated with gene expression levels. Intriguingly, the high plateaus most often cover the entire sequence of small and highly expressed genes, from promoter to polyadenylation site. The plateau signals were remarkably similar to the Pol-II signals, also obtained in ChIP-seq. Taken together with the fact that nearly 40% of the peaks located in promoters exactly overlap with the transcription initiation site (TSS), it suggests a possible direct contribution of RXR to Pol - II's dependent transcription machinery.

Analysis of the density variation of the RXR peaks during the circadian or ultradian rhythm reveals that only 10% of RXR binding sites are considered rhythmic, and only 10% of their putative target genes (a total of 219 genes) presented a rhythmic expression. This disproportion implies that the transcriptional response requires more than the binding rate of RXR.

After gating out the RXR peaks presenting high Pol-II co-occupancy, motif analyses of the sequences under each remaining peaks were performed to identify the associated heterodimeric partners. The binding sites of RXR coupled to a partner are made of two repeats of a canonical hexamer, separated by 1 to 5 base-pairs, and were identified in more than 90% of the remaining peaks. The direct repeat with one base pair spacing (DR1) was the dominant predicted binding mode, followed by the other direct repeat DR2-5, and finally by the IR1-2 inverted repeats. Half of the genes used in this analysis are adorned with multiple binding modes, and are therefore likely regulated by multiple NRs. Metabolic functions and circadian rhythm maintenance are particularly well represented in this subset of genes.

In summary, these results demonstrate that RXR and its partners interact directly with a significant number of liver genes, although RXR cannot be considered as the driving force behind the internal regulation of the liver clock. A key discovery is the close association of RXR with Pol-II, found on more than 8000 genes. Their co-occupancy, which could be observed on whole genes as well as on TSS, suggests a close involvement of RXR in the transcription machinery, particularly for genes highly expressed in the liver.

RÉSUMÉ

Les récepteurs nucléaires (NRs) sont des facteurs de transcription qui régulent des gènes impliqués dans le développement, l'homéostasie et le métabolisme de l'organisme. Leur expression est dynamique et, dans le foie, suit un schéma circadien. Parmi les 48 membres de cette super-famille de protéines, le récepteur rétinoïde X (RXR) occupe un statut particulier de partenaire obligatoire pour former un hétérodimère avec plusieurs autres NRs. Contrairement à la plupart de ses partenaires, RXR se prête aux expériences d'immunoprécipitation de la chromatine (ChIP), et est donc un candidat de choix pour caractériser à l'échelle du génome les régulations dépendantes des NRs. Dans le cadre de cette thèse, le but était de cartographier le répertoire des sites de liaison de RXR à la chromatine et explorer la capacité du RXR à réguler les fonctions rythmiques du foie pendant la journée. En conséquence, nous avons utilisé un anticorps pan-RXR pour réaliser, sur le foie de souris, une expérience dite de ChIP-seq, à 7 différentes heures de la journée sur 24 heures.

Face à la complexité des comparaisons multiples entre les 7 sets de pics ChIP obtenus à l'échelle du génome, nous avons développé une méthode permettant l'analyse des 7 échantillons ensemble. Un peu plus de 30'000 sites de liaison ont été identifiés, répartis en 4 catégories: pics solitaires, pics groupés, hauts plateaux (larges régions à forte densité continue de marquage), et bas plateaux. Les pics de RXR se regroupent préférentiellement au niveau des promoteurs, et l'expression des trois quarts des gènes cibles putatifs de RXR, représentant presque la moitié des gènes de souris, est détectable par microarray. L'abondance des sites de liaison RXR, mais non l'intensité du signal observé à chaque site de liaison, corrèle avec le niveau d'expression du gène. Curieusement, les hauts plateaux couvrent le plus souvent l'entier de gènes petits et fortement exprimés. Les plateaux partagent également une similarité remarquable avec les signaux Pol-II obtenus également en ChIP-seq. Enfin, près de 40 % des pics situés au niveau des promoteurs chevauchent exactement le site d'initiation de la transcription (TSS), suggérant une possible contribution directe de RXR à la machinerie de transcription dépendante de Pol - II.

L'analyse de la densité des pics de RXR au cours du rythme circadien ou ultradien révèle que seules 10 % des sites de liaison RXR sont considérés comme rythmiques. Et seulement 10 % des gènes cibles correspondants à ces sites (un total de 219 gènes) ont une expression rythmique. Cette disproportion implique que la réponse transcriptionnelle exige plus que le rythme de liaison de RXR.

Après avoir éliminé les pics RXR associés à une co-occupation élevée de Pol-II, les analyses de motifs sur les séquences des pics restants ont été effectuées dans le but d'identifier les partenaires hétérodimères correspondant à des éléments de réponse spécifiques. Les modes de liaison de RXR couplé à un partenaire sont formés de deux répétitions d'un hexamère canonique, et ont été trouvés dans plus de 90% des pics restants. Le tandem de répétitions directes avec un espacement d'une paire de base (DR1) était le mode le plus dominant, suivi par les autres répétitions directes DR2-5, et enfin par les répétitions inversées IR1-2. La moitié des gènes dans cette analyse présentent plusieurs modes de liaison, et sont donc probablement régulés par de multiples corégulateurs. Les fonctions métaboliques et le maintien du rythme circadien sont particulièrement bien représentés dans ce sous-ensemble de gènes.

En résumé, ces résultats démontrent que RXR et ses partenaires interagissent directement avec un nombre important de gènes hépatiques, bien que RXR ne puisse être considéré comme le moteur de la régulation interne de l'horloge hépatique. Une découverte clé est l'association étroite de RXR avec Pol-II, trouvée sur plus de 8000 gènes. Cette co-occupation, pouvant s'observer sur des gènes entiers comme au niveau des TSS, suggère une implication étroite de RXR dans les mécanismes de transcription, en particulier pour les gènes fortement exprimés dans le foie.

ABBREVIATIONS

3C	Chromosome conformation capture
3C-seq	chromosome conformation capture combined with sequencing
9CDHRA	9-cis-13,14-dihydroretinoic acid
ABAC	ATP-binding cassette
APL	Acute promyelocytic leukemia
BH	Benjamini–Hochberg
bp	base pair
CAR	Constitutive androstane receptor
CCL	Chemokine (C-C motif) ligand
ChIP-seq	chromatin immunoprecipitation followed by sequencing
Cryo-EM	Cryogenic electron microscopy
CTCF	CCCTC-binding factor / Transcriptional repressor
DBD	DNA-binding domain
DHA	Docosahexaenoic acid
DHSs	DNase I hypersensitive sites
DR	Direct repeat
ER	Everted repeat
ER α	Estrogen receptor alpha
FCHL	Familial combined hyperlipidemia
FFA	free fatty acid
FXR	Farnesoid X receptor
GR	Glucocorticoid receptor
H3K36me3	Tri-methylation of lysine 36 on histone H3
H3K4me3	Tri-methylation of lysine 4 on histone H3
HDAC	Histone deacetylase
HDL	High-density lipoprotein
HNF	Hepatic nuclear factor
IR	Inverted repeat
kb	kilobase - a length of double-stranded DNA containing one thousand nucleotides
LBD	Ligand-binding domain
LDL	Low-density lipoprotein
LPL	Lipoprotein lipase
LXR	Liver X receptor
MA	Mineralocorticoid receptor
MAPKs	Mitogen-activated protein kinases
NCoR	Nuclear receptor corepressor
NR	Nuclear receptor
NR-IL	Nuclear factor interleukin
nts	Nucleotides

PAS	Polyadenylation site
Pol-II	RNA polymerase II
Pol-III	RNA polymerase III
PPAR	Peroxisome proliferator-activated receptors
PR	Progesterone receptor
PWM	Position weight matrix
PXR	Pregnane X receptor
RA	Retinoic acid
RAR	Retinoid acid receptor
RE	Response element
ROR	RAR-related orphan receptor
RXR	Retinoid X receptor
SAP	Spike adjustment procedure
SCN	Suprachiasmatic nucleus
SMRT	Silencing mediator for retinoid and thyroid hormone receptor
Sp1	Specificity Protein 1
SREBP1	Sterol Regulatory Element Binding Protein 1
TF	Transcription factor
TGF	Tumor/Transforming growth factor
TR	Thyroid hormone receptor
TTS	Transcription starting site
VDR	Vitamin D receptor
VLDL	Very-low density lipoprotein
ZT	Zeitgeber time

1. INTRODUCTION

1.1 RXR as the central member of nuclear receptors superfamily

1.1.1 *The discovery of Nuclear receptor and the birth of the Molecular Endocrinology field*

The discovery of nuclear receptors has its historical roots in endocrinology with the studies of functional small lipophilic molecules that regulate the development, homeostasis, and metabolism of the organism. Examples of these molecules are steroid and thyroid hormones, retinoid (vitamin A) and vitamin D. Many of these molecules were associated with known human diseases, and glucocorticoids had gained popularity as a therapeutic agent very early on. The chemical characteristics of each of these small molecules are very distinct, but they share the unique property of being fat-soluble, which allowed them to pass through the lipid bilayer cell membrane. There was no presupposition of their common functional mechanism. However, and yet astonishingly, initial biochemical experiments revealed the presence of an intracellular receptor that, upon ligand binding could translocate from the cytoplasm into the nucleus (*Jensen et al., 1967*), and activate transcription of tissue-specific sets of target genes (*Ashburner et al., 1974*). Thus, the “Nuclear Receptor” (NR) is a transcription factor that is able to sense and translate simple chemical changes into distinct physiologic effects, by regulating the expression of specific genes, thereby controlling the physiologic states. The concept was early delineated, but the fundamental nature of this receptor, its means for recognizing specific chemical ligands, its mode of interaction with the genome, and its mechanism for control of gene transcription were beyond the limits of classic biochemical analysis (*Evans and Mangelsdorf, 2014*). The first biochemical, molecular, and genetic characterization of the encoding genes for steroid receptors was indeed a prerequisite for ultimately understanding the molecular basis of the nuclear receptors structure and

functions. The isolation of their encoding genes, the analysis of their structures and function revealed an evolutionarily conserved template that delineate the structural and functional features for nuclear receptors superfamily. Access to their coding genes cDNAs revealed three fundamental domains: the DNA binding, the ligand binding and the transactivation domains (*Giguere et al., 1986, Green et al., 1986, Miesfeld et al., 1986*). Importantly, this also enabled key experiments needed to test protein function, including alteration by mutagenesis of the receptor's primary structure to assess the importance of specific amino acids and characterization of the nucleotide code within the promoter sequences of target genes that allows gene-specific regulation (*Evans and Mangelsdorf, 2014*). These works highlighted the fundamental process of transcriptional regulation by hormone-receptor complexes, and their extracellular signal transduction by lipophilic endocrine mechanism. The discoveries of their ligands and their activation through ligand ignited great interest, especially for therapeutic medicine studies, in the attempt to control their signaling pathways.

The most revolutionary outcome sparked from the cloning of the first steroid receptors was the identification soon after of numerous evolutionarily related proteins along with their encoding genes. The explosion of characterization of these receptors created the new field of molecular endocrinology (*Evans and Mangelsdorf, 2014*). It grew with the suggestion of the existence of many previously unknown signaling pathways under the regulation of myriad of undiscovered ligands. The term “orphan” was thus given to the receptors for which the ligands were not (yet) identified, but these receptors were shown to be significantly phylogenetically conserved. Unlike in traditional endocrinology where a hormone was identified based on its predicted functions and confirmed by the use of its already known ligand(s), these receptors were found without precognitive functions. The development of cotransfection assay (*Giguere et al., 1986*) was a remarkable and innovative

technological achievement, that answered the conundrum of finding such ligands. In this assay, cells in culture were transfected with two distinct plasmids. The first one allows the expression of the receptor. The second one is carrying a promoter sequence which drives the expression of an easily quantifiable marker and that encompasses a binding site for the nuclear receptor. Upon this transfection, different combination of receptors or small lipophilic molecules could be tested to attest of the activation of the receptors. The cotransfection assay was versatile as a cell-based mean to study transcription, extremely sensitive to small-molecule, and provided quantitative results. It quickly became the mainstay of every molecular biology laboratory and pharmaceutical discovery/screening tool, and most importantly contributed greatly in the efforts of “de-orphaning” these orphan receptors (*Evans and Mangelsdorf, 2014*). The cotransfection assay was quickly adapted on the isolated cDNAs of the orphan receptors as an unbiased, high-throughput screening method that could be implemented for natural, synthetic, and xenobiotic, even when the receptor’s DNA binding site (or response element) was not precisely known (*Kliewer et al., 1999, Chawla et al., 2001*).

The retinoid X receptor (RXR) was the first established “adopted orphan” with the identification of the first (but still controversial) “endogenous” ligand – the 9-cis retinoic acid (RA), a metabolite of vitamin A (*Mangelsdorf et al., 1990, Heyman et al., 1992, Levin et al., 1992*). And with this “de-orphaning”, originated two major concepts in the nuclear receptor field: the discovery of many signaling pathways regulated by the orphan receptors; and a novel feature of multiple intertwined signaling pathways partly elucidated by the discovery of RXR heterodimerization, as presented later.

1.1.2 Nuclear receptors genes and their evolution

As of today, there are a total of 270 putative nuclear receptor genes, among which 48 in human and 49 in mouse genomes (*Evans and Mangelsdorf, 2014*). Their sequences obtained in various species show a high degree of conservation. Evolutionary genomics has extensively studied the phylogeny of the nuclear receptors. The conserved domain structures and the distribution in the various subfamilies suggest that nuclear receptors appeared very early during metazoan evolution and are present in all metazoan phyla (*Escriva et al., 1997, Escriva et al., 2000, Owen and Zelent, 2000*). No nuclear receptors have been found in fungi, plants or unicellular eukaryotes so far (*Szanto et al., 2004*). It is presumed that nuclear receptors emerged explosively after two waves of gene duplication events: the first wave during the emergence of metazoans, leading to the formation of members in six subfamilies, and the second wave later, mainly in vertebrates, leading to the divergence of the paralogues. A latter identified nuclear receptor in a sponge, *Suberites domuncula*, indicated that nuclear receptors were present at the base of metazoan evolution and RXR homologues might be indeed the most ancient members of the family (*Wiens et al., 2003*). From an evolutionary point of view, RXR seems to be an ancestral nuclear receptor from which many of the other receptor families emerged. The LBDs of the RXRs share a high similarity from jellyfish to humans. A homologue of the RXR, ultraspiracle (USP) was identified in advanced arthropods including the insects (*Henrich et al., 1990, Oro et al., 1990*), that bear different amino acid sequence but similar crystal structures with RXR, and does not bind 9-cis RA (*Billas et al., 2001, Clayton et al., 2001*). This suggested that RXR ligand binding may have appeared very early in metazoan development and was then lost mainly in arthropods. Since the appearance of RXR signaling happened much earlier than retinoic acid receptor (RAR)s signaling, we can assume the

independence of RXR signal pathways from RAR and that autonomous RXR signaling (independent of RAR and other partners) may still persist.

According to nomenclature provided by (*Germain et al., 2006*), the six subfamilies include 65 known nuclear receptors in vertebrates, arthropods, and nematodes. The first subfamily is a large family with its members regulated by lipophilic signaling molecules; including thyroid hormone receptors (TR), retinoic acid receptors (RAR), peroxisome proliferator activated receptors (PPAR, adopted), reverse-Erb receptors (REV-ERB), retinoic acid related receptors (ROR), farnesoid X receptors (FXR, adopted), liver X receptors (LXR, adopted), and vitamin D receptors (VDR). Second subfamily comprises of other “adopted” receptors that were shown to bind fatty acids via structural studies but with controversial in dynamic regulation; RXR, chicken ovalbumin upstream promoter transcription factors (COUP - TF), and hepatocyte nuclear Factor 4 (HNF4) are members. Third subfamily comprises the steroid receptors such as androgen receptor (AR), progesterone receptor (PR), glucocorticoid receptor (GR), mineralocorticoid receptor (MR), and estrogen receptors (ER- α and - β). Fourth subfamily contains the orphan NRs that involved in neuron development and maintenance like nerve growth Factor 1B (NGF1-B), nurr-related Factor-1 (NURR1), and neuron-derived orphan Receptor-1 (NOR-1). Fifth subfamily members, steroidogenic Factor 1 (SF-1)50 and liver receptor Homolog-1 (LRH-1), are regulated by phospholipids. The last subfamily contains only one receptor, germ cell nuclear factor (GCNF) that has critical difference in its structure and is known to drive gene silencing.

Among the 48 and 49 nuclear receptor genes in human and mouse, respectively, RXR stands for Retinoid X Receptor. The mysterious “X” was due to its discovery as a nuclear retinoid receptor that was activated by but did not bind to all-trans RA, (*Mangelsdorf et al., 1990, Heyman et al., 1992, Levin et al., 1992*) (hence, “retinoid X”, and later discovered as 9-cis RA), in contrast to the

first class of the RA receptors (RAR α , β and γ), which all bind all-trans RA with high affinity and share a high degree of structural conservation. Three isotypes RXR α (NR2B1), RXR β (NR2B2) and RXR γ (NR2B3) and their encoding genes *Rxra*, *Rxrb*, *Rxrg* in mice were isolated (*Mangelsdorf et al., 1992*). Although closely related and quickly identified as forming a separated group of the retinoic acid receptor, none of the *Rxr* loci co-segregated with each other or with the *Rar* loci previously mapped. *Rxra* gene locates on chromosome 2 near the centromere, forward strand, with 13 exons; *Rxrb* on chromosome 17 to the H-2 region, forward strand, with 11 exons, and *Rxrg* tightly linked to the *Pbx* gene on distal chromosome 1, forward strand with 11 exons (*Hoopes et al., 1992*). Alternative splicing can result in different isoforms of each isotype.

The three RXR genes were found to be abundantly expressed in many metabolic organs, such as the liver, kidney, and intestine. More specifically, their expression levels vary with cell types. RXR α was found to predominate in the epidermis, placenta, intestine, kidney, and liver; RXR β expression was ubiquitous; RXR γ has limited tissue expression, including brain, skeletal muscle, anterior pituitary gland and was weak in adipose tissue (*Dawson and Xia, 2012*).

Paradoxically from the great interest in their physiological importance, specific studies on the RXR genes regulatory models were rather limited. The early studies focused on measuring by northern blot the modulation of RXR genes expressions under different stimulators. Mano H reported the regulation effect of thyroid hormone positively on gene expression of RXR β at transcriptional level, and negatively on RXR γ gene expression possibly at a post-transcriptional level in intact rat (*Mano et al., 1994*). A study in osteoblastic MC3T3-E1 cells (*Chen et al., 1996*) showed that TGF- β 1 transcriptionally stimulated the expression of *Rara*, *Rarg*, and *Rxra* genes, but did not do so for *Rarb*, *Rxrb*, and *Rxrg* genes; and transcriptional factor AP-1 plays an important role in the signal pathway of these stimulations. Other study in rat hepatocytes (*Steineger et al.,*

1997) indicated that the RXR α gene expression is under distinct regulation by all-trans RA, fatty acids and dexamethasone acid which connected the lipid metabolizing system and the retinoid signaling pathway. Treatment of rat hepatocyte cultures with all-trans RA led to morphological differentiation and re-establishment of cell polarity, perhaps through activation of retinoid-responsive genes (Falasca et al., 1998). Lipopolysaccharide was reported to induce a rapid, dose-dependent reductions in mRNA and nuclear protein levels of all three RXR isoforms in hamster hepatocytes (Beigneux et al., 2000), and in mice cardiomyocytes (Feingold et al., 2004).

Sequencing and nuclease protection assay studies dwelled more in-depth to the regulatory elements on RXR genes loci. The cloning (Li et al., 2000a) and characterization of human RXRA gene and its mouse homolog *Rxra* (Li et al., 2006) showed highly conserved intron-exon positioning, surprisingly GC-rich promoter, upstream and downstream flanking regions and first exon, which resulted in a potentially stable folding pattern for transcription. These reminiscent hallmarks of housekeeping genes suggested regulation of RXR α at translational level. A part of the H2-Ke4 (of the major histocompatibility complex) genomic region, as well as the 250bp promoter, was reported to transcriptionally active as an mouse *Rxrb* promoter, on which TNF-alpha could repress the activity, and this repression is mediated by p38 MAP kinase independent of NF-kappaB (Sugawara et al., 1998). A study on cartilage cell culture and transgenic mouse showed an active 507-bp intergenic sequence from the *Col11a2* enhancer/silencer insulated the *Rxrb* promoter, probably associating with unknown factors that recognize a motif similar to CTCF (Murai et al., 2008). For *Rxrg*, a study isolated the mouse *Rxrg1* gene promoter region and identified the major start site, which is uniquely suppressed by 9-cis-RA in thyrotropes (McDermott et al., 2002). With the dawn of high-throughput era, many studies using chromatin immunoprecipitation followed by

sequencing (ChIP-seq) have put RXR genes into regulation target of many transcription factor including itself.

1.1.3 The modular structure of RXR, and more broadly of all the NRs

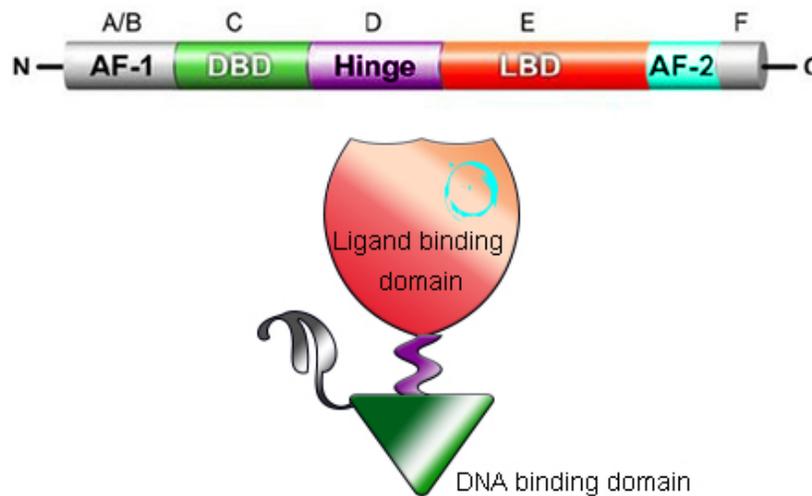


Figure 1: General nuclear receptor molecule with 6 domains and simplified tertiary structure

From the first three identified domains, molecular structure studies established six major functional/structural domains for nuclear receptors, including RXR. From the N-terminal to C-terminal, the order of domains are as follows: the A/B domain, the C or DNA-binding domain, the hinge D domain, the E or ligand-binding domain, and the F domain (**Fig.1**).

1.1.3.1 The A/B domain

This domain contains a ligand-independent activation function AF-1 to which coactivator proteins bind. The sequence and length are highly variable between receptors and among receptor subtypes (RXR α versus β). This domain is also frequently the site of alternative splicing and secondary start sites. It contains a variety of kinase recognition sequences, and is thus supposed to contribute for the receptor-, species-, and cell type-specific effects as well as promoter context-

dependent properties of NR transactivation. This N-terminal region is reported as an interaction surface for multiple transcriptional coregulatory proteins: steroid receptor coactivator-1 and -2, p300 and CBP. The functional synergism between AF-1 and AF-2 of the E domain (described later) cooperatively enhances transactivation function (*Bommer et al., 2002, Gianni et al., 2003, Bugge et al., 2009*).

The phosphorylation at various residues throughout the protein may modulate its stability and function (reviews (*Weigel, 1996, Shao and Lazar, 1999*)). Nuclear receptors are phosphoproteins, and multiple receptor functions can be affected by phosphorylation in response to various types of effector. (*Lefebvre et al., 1995*) have shown that dephosphorylation decreases the DNA binding of RXR α . The majority of the nuclear receptor phosphorylation sites lie within the N-terminal A/B region; they can be complex and comprise up to 13 residues, as in the case of the progesterone receptor. Most of the modified residues are serines surrounded by prolines and therefore correspond to consensus sites for proline-dependent kinases, which include cyclin-dependent kinases and mitogen-activated protein kinases (MAPKs). This was reported for RXR α function, although the ultimate phosphorylation sites and effects were subject for debate through several studies (*Adam-Stitah et al., 1999, Solomon et al., 1999, Lee et al., 2000, Matsushima-Nishiwaki et al., 2001, Adachi et al., 2002*). Studies have shown that RXR α is phosphorylated at serine 260 in human hepatocellular carcinoma, readily distinguishable from adjacent normal human tissue where this residue is not phosphorylated (*Adachi et al., 2002, Matsushima-Nishiwaki et al., 2001*). The phosphorylation of the serine/threonine-rich nuclear localization signal on MR N-terminus can influence receptor subcellular localization (*Walther et al., 2005*). An example of piggyback nuclear exclusion of RXR upon association with an AF-1 phosphorylated Nur77, resulting reduced transcription activity, provided interesting regulatory mechanism for RXR

(Katagiri *et al.*, 2000, Walther *et al.*, 2005). Similarly, PPAR γ nuclear localization inhibition was proposed to occur via MEK1-mediated phosphorylation of serine at position 84 (Burgermeister *et al.*, 2007), and alternatively via Pin1-mediated proteasomal degradation (Fujimoto *et al.*, 2010).

1.1.3.2 The C domain

DNA-binding domain (DBD) mediates NR binding to the response element (RE) in the vicinity of the promoter of target genes. DBD folds into a globular domain, is composed of two zinc fingers and is most conserved within NR superfamily. It is indeed a major hallmark of the nuclear receptor family. Each zinc atom is coordinated by four cysteine residues, and removal of one zinc ion results in protein unfolding and loss of DNA-binding activity (Freedman *et al.*, 1988). The N-terminal helix (helix 1) directly interacts with the major groove of binding DNA, making base-specific contacts, while the C-terminal helix (helix 2) overlays perpendicularly helix 1, and contributes to stabilization of the overall protein structure. In the dimerization formation (details presented later for RXR), the residues within dimer interface are located within the C-terminal zinc finger, called the D-box; the residues within sequence-specific DNA binding are located within helix 1, called the P-box. Interactions between the P-box and the DNA half-site coupled to conformational changes in the D-box are necessary for cooperative recruitment of the second monomer.

1.1.3.3 The D domain

The hinge, connects the DNA and ligand-binding domains; it contains the carboxy-terminal extension of the DBD, which may be involved in recognizing the extended 5' end of the RE. The D-domain appears to allow for conformational changes in the protein structure following ligand

binding. Also, this region may contain nuclear localization signals and protein-protein interaction sites. (*John P. Vanden Heuvel, nrresource.org*)

1.1.3.4 The E domain

For RXR, and for nuclear receptor in general, the ligand-binding domain (LBD) is most extensively studied. Crystallography has been used to investigate this domain in depth either alone (apo form), agonist/ligand bound (holo form), and in complexes with the LBDs of its dimeric NR partners and/or coactivator peptides. First high-resolution LBD structure of RXR was determined in 1995 (*Bourguet et al., 1995*). This domain varies substantially between NRs, but they mainly consist of 12 α -helices and a small β -sheet between helices H5 and H6. They typically form three antiparallel helical sheets that combine into the α -helical sandwich. The ligand-binding pocket for each receptor is located in the interior of the structure and is formed by a subset of the surrounding helices. The strength and specificity of LBD-ligand complexes are largely dependent on hydrophobic interactions, extensive hydrogen-bonding networks, and the steric size and shape of this binding pocket (*Li et al., 2003*). There is a correlation between receptor function with the overall size and shape of the binding pockets. To accommodate a number of different structures of diverse metabolic ligands, orphan receptors tend to have larger volume binding pockets (*Bain et al., 2007*).

The entrance to the binding pocket is guarded by the helix H12, which forms a movable lid. This lid forms with the LBD a crucial hydrophobic groove for the ligand-induced activation function, the AF-2 function, which capable of recruiting coactivators. Similar to the binding pocket conformation, the orientation of H12 is allosterically affected by the particular chemical structure of the specific binding ligand (*Gronemeyer et al., 2004*). In the absence of a ligand, all 12 helices of the LBD (apo form) position away from the LBD core structure, and thus are unable to complete

the formation of the hydrophobic groove. With the binding of ligand, LBD has a structural change (holo form): H2 unwound, providing a larger loop between helices H1 and H3 that permitted H3 to undergo a 13-Å tilt, H12 is stabilized and forms with H3 and H4 the hydrophobic groove that accommodates the binding of a coactivator (*Dawson and Xia, 2012*). Most coactivators that interact with nuclear receptors contain helical LxxLL motifs where L is leucine and x is any amino acid, called the nuclear-receptor box that is recognized by the groove. Interaction specificity is conferred by charge-clamp electrostatic interactions between the LBD and coactivator residues that cap each end of the two-turn LxxLL helix. For RXR which has the ability of forming tetramers, there exist another regulation mechanism at the quaternary level: the maintaining of a tetrameric state of RXR physically prevents coactivator binding. Upon agonist/ligand binding this tetramer is dissociated to dimers, resulting in the concomitant formation and exposure of the hydrophobic groove (*Gampe et al., 2000, Yasmin et al., 2010, Yasmin et al., 2004*).

In the absence of ligand, some nuclear receptors have the LBD bound to a set of transcriptional corepressors, such as nuclear receptor corepressor 1 (NCoR1) or silencing mediator for retinoid and thyroid hormone receptor (NCoR2, also known as SMRT). These corepressors recruit transcriptional complexes that contain specific histone deacetylases (HDACs) which generate condensed chromatin structure over the target promoter, resulting in gene repression (*Gronemeyer et al., 2004*). The corepressors can also prevent coactivator recruitment, such as for PPAR (*Chen and Evans, 1995*). Like coactivators, they contain LxxLL-binding motifs and thus are able to recognize the hydrophobic residues that make up AF-2. But in contrast to coactivators, the corepressor-binding motif forms a lengthy three-turn helix that sterically blocks H12 active conformation and thus prevents the formation of the hydrophobic binding groove (*Xu et al., 2002*). Antagonist ligands can inhibit coactivator binding by blocking H12 to approach the core LBD

structure (*Brzozowski et al., 1997, Xu et al., 2002*) or by inducing H12 binding to hydrophobic groove and thus unproductively mimic the coactivator (*Shiau et al., 1998*).

The allosteric ligand effect on the LBD conformation can also modulate the activity of the N-terminal AF1 – for example, through intramolecular crosstalk between N- and C-terminal domains, deletion of the PPAR γ N-terminal domain prevents corepressor binding (*Suzuki et al., 2010*).

1.1.3.5 The F domain

Whether any of the RXR isotypes has a functional F domain has yet to be established. For RXR structure, this domain is usually included with the LBD (*Mangelsdorf et al., 1992*). However, other NRs such as estrogen receptor (ER) α , Glucocorticoid receptor (GR), hepatic nuclear factor (HNF) 4 α , progesterone receptor (PR), and even RARs have F domains with specific sequences that modulate gene transcription that were described in many studies (*Patel and Skafar, 2015*). Several studies from (*Safe and Kim, 2008*) showed that the deletion of the entire F domain of the ER α eliminates its interaction with the transcription activator Specificity Protein 1 (Sp1). A recent study (*Arao and Korach, 2018*) analyzed the transcriptional activity of constructed mouse-human F domain-swapped ER α in the human hepatoma cell line HepG2, and deletion and point mutations experiments on a predicted β -strand region on F-domain; the results suggested that this region governs the species-specific 4-hydroxytamoxifen-mediated transcriptional activity of ER α . The F domain of the GR, despite having a relatively small number of residues, fundamentally contributes to protein dimerization (*Bledsoe et al., 2002*), ligand affinity (*Lanz and Rusconi, 1994, Kauppi et al., 2003*), and activation (*Charmandari et al., 2005*).

1.1.4 The dimerization property of NR and the specific role of RXR in heterodimerization

Early studies based on domains swapping and the characterization of DBD-DNA interactions suggested the autonomy of DBD (*Green and Chambon, 1987, Kumar et al., 1987, Umesono and Evans, 1989*). For more than two decades, structural efforts were focused on discrete domain/segment of nuclear receptor structure. However, as mentioned above, several studies have demonstrated high quaternary organization of NR domains, and the allosteric communications among them. The advances in crystallographic studies started to unveil how the entire receptor polypeptides act in a concerted fashion using its multiple domains to dimerize, to bind with ligands and DNA.

While the recognition two helix within DBD domain partially accounts for the selective site binding, this type of interaction alone did not fully account for response element selectivity in the NR family. With the paradox of conserved structures and diverse signaling pathways, nuclear receptors rely heavily on the ability of fine-tuned modulation by combination of ligand-binding and interaction with other receptors, coactivators, corepressors. Protein dimerization plays an important role in the functions of a wide variety of proteins, often altering as well as expanding the functions (*Jiang et al., 1997*). Particularly for transcription factors, dimerization influences nuclear localization, DNA binding affinity and specificity, and transactivation potentials (*Lee et al., 1992*). Studies using combinations of co-transfection assay and the electrophoretic mobility shift assay revealed that the steroid receptors (GR, PR, AR, and ER α) bind as homodimers to the sequence-specific bindings termed response elements (REs) configured as palindromes composed of two hexamer nucleotide sequences separated by three base pairs (*Beato, 1991*). Non-steroid receptors (RAR, VDR, and TR), in contrast bind preferentially to REs composed of two hexamer half-sites

arranged as direct tandem repeats (DR) (*Näär et al., 1991, Umesono et al., 1991*), with the nucleotide spacing between the two half-sites being specifically 3 for VDR, 4 for TR and 5 for RAR, hence the 3-4-5 rule (*Rochel et al., 2011*). The most striking difference between the steroid and non-steroid groups, however, is the ability of the later to heterodimer with RXR. Shortly after identified, RXR was shown to be the missing factor required for the high affinity binding of VDR, TR, and RAR to their cognate response elements (*Yu et al., 1991, Bugge et al., 1992, Leid et al., 1992b, Kliewer et al., 1992a, Zhang et al., 1992*). As for the orphan receptor, PPAR was the first then-orphan receptor shown to heterodimerize with RXR. While using cotransfection assays to identify interactions between an orphan receptor and compounds that promote peroxisome proliferation, PPAR was ultimately adopted as another first subfamily member of fatty acid receptors. And this heterodimer also confirmed the concept of crosstalk between hormone signaling and lipid metabolism (*Issemann and Green, 1990, Dreyer et al., 1992, Kliewer et al., 1992b, Wahli et al., 1995, Desvergne et al., 1998*). The paradigm of RXR heterodimerization quickly expanded into the universe of orphan receptors that now included the LXRs, FXR, PXR, and CAR (*Evans and Mangelsdorf, 2014*). Innate structure of the RXR ligand binding domain permits it to adopt multiple conformations and thereby dimerize with different nuclear receptors when bound to each of the binding half-sites (*Chandra et al., 2008, Lou et al., 2014*) as a ready anchor for heterodimer partners to come bind to the other half-site. This plasticity allowed the heterodimer rule expanded beyond the 3-4-5 rule.

The first polarized DR1 was reported for PPAR:RXR heterodimer (*Juge-Aubry et al., 1997, IJpenberg et al., 1997*) with the RXR positioned downstream from its partner. The full structural conformation for the PPAR γ :RXR α heterodimer was also the first to be reported by (*Chandra et al., 2008*) using crystallography, with the fold and dimerization of the DBDs and the LBDs are

similar to those of isolated DBDs and LBDs reported earlier, and each of the liganded LBDs was bound by one LxxLL coactivator peptide (*Gampe Jr et al., 2000, Connors et al., 2009*). The configuration of the PPAR γ :RXR α heterodimer on its DNA response element is largely asymmetrical (*Fig.2A*), with the RXR α being stretched out, enclosing the LBD of PPAR γ between its LBD and DBD, revealing an interface between PPAR γ LBD and RXR α DBD. Furthermore, the hinge of PPAR γ interacts with the 5'-upstream region of the DR1 determining the polarity of the PPAR γ -RXR α binding orientation (*Juge-Aubry et al., 1997, IJpenberg et al., 1997*). The hinge of RXR α , on the other hand, is involved in the DNA-dependent dimerization of the DBDs in a head-to-tail conformation on the DR1. In short, the PPAR γ LBD contacts both DBDs to stabilize DNA binding (*Chandra et al., 2008*). Moreover, phosphorylation of Ser273 in the DBD-LBD interface in mouse PPAR γ LBD results in changes in PPAR γ -dependent gene expression (*Choi et al., 2010*). This notion of communication between domains strongly suggests the possibility of selective genes expression regulation by alteration/modification of receptors structure conformations induced by ligands. Later, several studies using small-angle X-ray scattering, small-angle neutron scattering and fluorescence spectroscopy reported a different full size solution structures of PPAR:RXR heterodimer, which they termed “open” as oppose to “closed” conformation by crystallography (*Rochel et al., 2011, Osz et al., 2012*). Open conformation was reported where PPAR:RXR heterodimer formed an elongated asymmetric shape with the LBDs well separated from the DBDs by the hinges. The LBD dimer, however, is positioned above the DBD located at the 5' end of the element, with the two hinges in parallel. The hinge of PPAR is not contacting the DNA upstream of the 5 half-site which is in contrast with the crystal structure (*Helsen and Claessens, 2014*). While the position of the DBDs in these RXR heterodimers depends on the composition of the response

element, the position and the orientation of the dimerized LBDs is always very similar and positioned above the upstream DBD (Rochel et al., 2011).

The DR1 heterodimer structure with RXR of PPAR was later also predicted for RAR (Rastinejad et al., 2000), after the first identified pattern as DR5 and weakly at DR2 (Mader et al., 1993, Perlmann et al., 1993, Umesono et al., 1991), however with RXR positioned upstream (Fig.2B). The stabilization for the contacts between RXR and RAR in their complex takes advantage of the polar functional groups in the minor grooves of base pairs 8 and 9, and the well-ordered hydration spine associated with DNA structure. The downstream RXR uses exclusively its T-box to mediate protein-protein interactions with the upstream Zn-II region.

The DR1 structure was also proposed for RXR homodimer (Zhao et al., 2000) (Fig.2C). However, the subunit interactions in the homodimer rely to some extent on hydrogen bonding interactions between Glu74 from the T-box of the downstream subunit, and Gln49 and Arg52 from the Zn-II region of the upstream subunit, while this is not possible in the RXR-RAR interaction due to T49 in RAR disabling the salt-bridge at the dimerization surface. The bending of the DNA has also been suggested to be exploited to achieve cooperativity in the RXR-RAR. The RXR homodimer induces a 12° kink in DR1 compared with the RXR-RAR complex, which bends the DNA by 6° at the spacer as compared to 12° bend. The adoption to combination of tilt and roll

G. Structure of RXR-partner heterodimer complex bind to the DR4 (adapted from (Chandra et al., 2017). TR, LXR and CAR are the partners that heterodimer with RXR in this manner.

H. RXR and partners bind to the IR2 adapted from (Mohideen-Abdul et al., 2017, Suino et al., 2004)

I. RXR and FXR bind to the inverted repeat (IR) of two AGGTCA with one nucleotide spacer: IR1 (adapted from (Devarakonda et al., 2003)

J. DR0

K. IR0

L. Predicted everted repeats with various number of space nucleotide

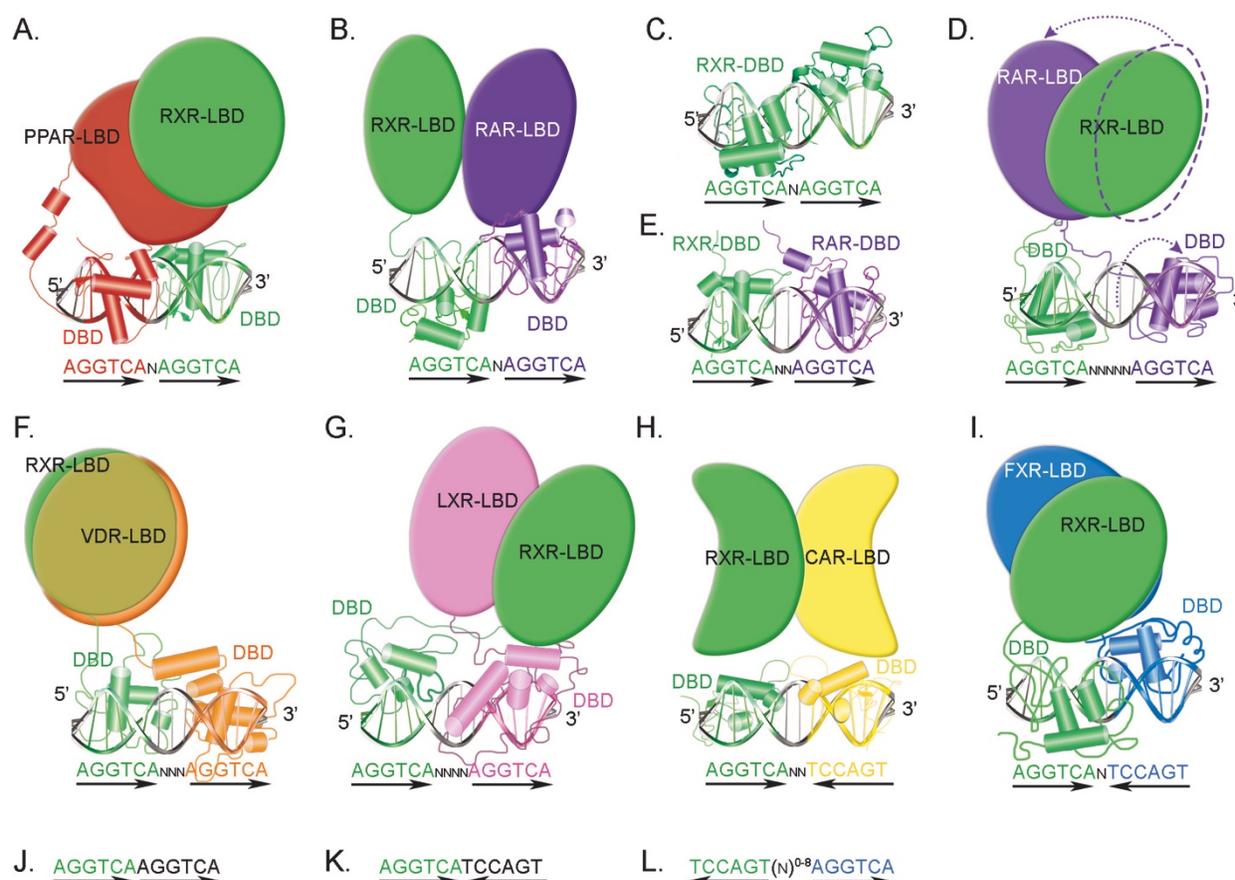


Figure 2: RXR heterodimer binding domain complexes structures and binding consensus

- A.** RXR and partners bind to the direct repeat (DR) of two AGGTCA with one nucleotide spacer – DR1. RXR-PPAR γ heterodimer structure arrangement with PPAR γ binds to the 5' upstream hexamer of the DR1 and RXR binds to the 3' downstream hexamer (adapted from (Huang et al., 2010))
- B.** RXR-RAR β heterodimer structure arrangement with RXR binds to the 5' upstream hexamer of the DR1 and RAR β binds to the 3' downstream hexamer (adapted from (Chandra et al., 2017))
- C.** The structure of two DNA binding domains of RXR-RXR homodimer in DR1 (adapted from (Osz et al., 2015))
- D.** Another conformation of RXR-RAR heterodimer structure (adapted from (Moutier et al., 2012)) where the DBD domain of RAR binds to the second hexamer 5 nucleotides downstream from the first (DR5) and on the same side with RXR-DBD domain on the DNA – as oppose to the arrangement in DR1
- E.** RXR and partners bind to the DR2 adapted from (Lefebvre, 2001)
- F.** Structure of RXR-VDR heterodimer complex bind to DR3: RXR binds to the upstream hexamer and VDR binds to the downstream hexamer with three nucleotide spacer (adapted from (Orlov et al., 2012))

angles of the DNA are a direct consequence of the dimerization arrangements, the distortion of helical axes of the RXR–RAR and RXR–RXR binding sites. Different from the tight connection in DR1, (*van Tilborg et al., 1999*) suggested the ability of the highly dynamic Zn-II region of RXR within the DR5 assembly of RAR-RXR heterodimer complex (**Fig.2D**). The differences between extrapolated structures were later confirmed by crystallographic study by (*Chandra et al., 2017*). While the DR2 were found along with DR5 in naturally occurring response elements of RAR-RXR heterodimer (*Mangelsdorf et al., 1991, Predki et al., 1994*), no structure of the DBDs complex has been established (**Fig.2E** extrapolated).

The hinge of the receptors seems to play an important role in adapting to differences in spacer lengths and hexamer orientations to still allow the regular DBD and LBD dimerization. The cryo-EM structure of DNA-bound RXR–VDR on a DR3 shows an extended RXR hinge region, while the hinge of VDR contains a rigid helix that determines the topology of the complex (**Fig.2F**). The localization of the short N-terminal domain of VDR interacts with the major groove of the DNA response element, next to the recognition helix of VDR, could point towards a modulating effect on DNA binding (*Orlov et al., 2012*). A naturally occurring polymorphism shortening the N-terminal domain by 3 amino acids, has been shown to influence the repressive function of the VDR in the absence of ligand. Both conformations could indeed support different aspects of receptor functioning. While the open conformation allows coregulators and other proteins to interact with the available surfaces on the DBD, the LBD and the hinge, the closed conformation could reinforce or stabilize binding of the receptor complex to the response element and/or conducting allosteric signals between these domains, the DNA and interacting complexes (*Helsen and Claessens, 2014*).

After TR, LXR was found to bind to DR4 in heterodimer with RXR, and their crystal structure were later reported in detail by (Lou et al., 2014)(Fig.2G). CAR and PXR were the two orphan receptor reported to heterodimer with RXR at DR4 and IR2 (Fig.2H) sites however only their LBD were explored (Echchgadda et al., 2007, Suino et al., 2004).

Using electrophoretic mobility shift assays and transient transfection assays, (Laffitte et al., 2000) reported the inverted repeat pattern for FXR/RXR heterodimer with a T or A as the one spacer nucleotide (Fig.2I), and for a stretch of T nucleotides flanking the downstream half-site (positions 8–11), referred to as the “extended half-site”. The structure was adapted from the homologs study (Devarakonda et al., 2003) and partially validated in recent study (Kim et al., 2015).

Aside from these traditional method of identifying binding consensus, high through put ChIP-seq also provided several computational predicted binding motif for RXR and partners heterodimer, such as tandem repeat with no spacer (DR0 – Fig.2J, IR0 – Fig.2K), or longer spacer (>5), and everted repeats (Fig.2L) (Moutier et al., 2012, Zhan et al., 2012).

1.1.5 The controversial existence of endogenous ligands for RXR

The autonomy of the DNA- and ligand-binding domains of the receptors inspired the swapping between these domains between receptors, and even other transcription factors, still retaining ligand-dependent transactivation (Giguere et al., 1987, Green and Chambon, 1987, Petkovich et al., 1987). Upon the ligand transactivation of each monomer, heterodimer partners of RXR have been characterized into three classes: non-permissive, permissive or conditionally permissive.

- **In non-permissive heterodimers**, RXR activity was classified as being subordinated compared to its partner, although RXR ability of binding its agonists and interacting with

coactivators are retained (*Cheskis and Freedman, 1996, Thompson et al., 1998, Germain et al., 2002*). The heterodimer would only be activated in response to the partner's ligand binding and induce transactivation response. The binding of RXR to a retinoid agonist would not enhance this response. The release of corepressor is also prompted in the same manner only by binding of a selective ligand. TRs and VDR belong to this category.

- ***In permissive heterodimers***, transactivation would occur through the binding of an agonist to either partner or to both partners. Binding by both agonists could have additive or synergistic effects. A corepressor protein would also be released when a retinoid agonist binds to RXR (*Lammi et al., 2008*). Within this classification are FXR (*Zheng et al., 2018*), LXR (*Willy and Mangelsdorf, 1997*), TR (*Castillo et al., 2004*) and PPAR (*DiRenzo et al., 1997*). PPAR was later found by (*Feige et al., 2005*) to be capable of readily heterodimerizing with RXR in living cells prior to the binding of ligand.

- The RAR heterodimers have a ***mixed*** status: while this heterodimer would not be activated by an RXR agonist, binding by an RAR agonist would induce both transactivation and permit RXR agonist to enhance the activation (*Shulman et al., 2004*).

Ironically, one of the most enigmatic and controversial areas of RXR research is the documentation of the presence of endogenous ligands activating RXR in vivo. Perlmann and co-workers created a ligand detector mouse line (*Solomin et al., 1998*). Transgenic lines were constructed using Gal-DBD RXR-LBD fusion constructs and a reporter gene containing Gal binding sites and a β -galactosidase reporter gene. X-gal staining of mouse embryos revealed sites of ligand production. Specific regions of the spinal cord lit up, suggesting that endogenous ligand production took place. Using a similar system with green fluorescent protein as a reporter, (*Luria and Furlow, 2004*) later have confirmed these results. RXR ligand also remains the only mean to

activate heterodimers with orphan receptors or RXR homodimers. Many molecules have been proposed as potential endogenous ligands. Here we will present the most representative ones:

- **9-cis-retinoic acid** was initially suggested as an endogenous ligand for RXR α (*Heyman et al., 1992, Levin et al., 1992, Mangelsdorf et al., 1992*) and quickly and widely accepted as “the” natural RXR ligand (*Evans and Mangelsdorf, 2014, Kane, 2012, Kane et al., 2010, Desvergne, 2007*). However, many groups failed to detect endogenous 9-cis RA in humans and other mammals (*Horton and Maden, 1995, Matt et al., 2005, Blomhoff and Blomhoff, 2006, Gundersen et al., 2007, KANE et al., 2005, Kane et al., 2008, Kane et al., 2010, Rühl, 2006, Schmidt et al., 2003, Wongsiriroj et al., 2014*) makes it doubtful to be RXR physiological ligand and partly questioned its endogenous existence as well as its nutritional significance. The identification of 9-cis RA was reported under supra-physiological conditions after administration of a high dosage of naturally-occurring retinoids or after ingestion of food with high content of vitamin A derivatives (*Arnhold et al., 1996, Ulven et al., 2001*), suggested its discrepancy in detectable versus efficient concentration in physiological conditions. Highly sensitive mass spectrometry-detectors can detect 9-cis RA levels as low as 0.03-0.003 nM (*Kane et al., 2008, de Lera et al., 2016*), however such levels are not sufficient to modulate RXR signaling. Multiple transactivation assays in diverse reporter cell lines suggested 10-100 nM is the minimal concentration of 9-cis RA that can facilitate RXR binding and transcriptional activation (*Heyman et al., 1992, Levin et al., 1992, Ruhl et al., 2015*), consequentially excluded the possibility of 9-cis RA function as endogenous and physiological ligand (*Ruhl et al., 2015, de Lera et al., 2016*).

- **Docosahexaenoic acid (DHA) and other free fatty acids (FFAs)** were proposed to be capable of inducing RXR transcriptional activation with different efficiencies in mouse brain by (*de Urquiza et al., 2000*). However, they reported a relatively high concentrations of DHA

necessary to induce transcription by RXR, which was confirmed by other studies (*Calderon and Kim, 2007, Goldstein et al., 2003*), although a much lower concentration (by one tenth) was applicable using a different method by (*Lengqvist et al., 2004*). Crystallography study of RXR α -LBD documented the direct interaction of DHA with the domain (*Egea et al., 2002*). In *in vitro* experiments, treatment with DHA (free acid form) increased the number of surviving neurons in primary culture cells, similarly to synthetic RXR agonists (LG100268 or BMS649), and prevented in the presence of the RXR agonist LG100268 (*Wallén-Mackenzie et al., 2003*). This effect was Nurr1-dependent indicating the driving of Nurr1-dependent signaling via ligand-dependent activation of RXR. DHA-dependent RXR activation was crucial for survival of cultured rat photoreceptors during early phases of primary cultures and in response to oxidative stress (*German et al., 2013*). Beside neuroprotection, RXR signaling is also critical for DHA cognitive and affective activities *in vivo*. In Rxrg-null mice and RXR antagonist pre-treated mice, the beneficial effect of DHA treatment on spatial working memory performance was abolished (*Wietrzych-Schindler et al., 2011*). They also implicated the relevance of RXR γ in antidepressant activities of DHA. DHA enhancement of antibody phagocytosis by microglial cells (*Hjorth et al., 2013*) could potentially be mediated by RXRs, and could be enhanced by bexarotene (LGD1069), a synthetic RXR ligand (see below), in an RXR α dependent manner (*Yamanaka et al., 2012*). The capability of vertebrates to synthesize DHA from α -linolenic acid as precursor of plant origin (*Rajaram, 2014*) supported the role of DHA as physiological ligand. However, the complication in measuring the administrative and effective concentration of DHA free form, especially hindered in neurology due to blood brain barrier, impeded the direct confirmation of DHA levels to function as a ligand-dependent switch, or the interaction between RXR and DHA in the nucleus. The relevance of DHA in modulating RXR signaling needs also to be considered in a context of other RXR endogenous

ligands due to potential competitions or cross-talks between metabolic pathways involved in the production of such ligands. Some retinol-binding proteins may bind with different efficiencies also to DHA (*Tachikawa et al., 2018*), suggesting that FFA and retinoid signaling may cooperate in controlling RXR signaling. Such control may involve transcriptional control and sharing of some proteins involved in ligand metabolism and traffic.

• **Phytanic acid** was identified as a natural agonist of RXRs by (*Lemotte et al., 1996, Kitareewan et al., 1996*) and its effective concentration confirmed by (*Zomer et al., 2000*). As an important metabolite of chlorophyll, phytanic acid can be produced in ruminating animals by their bacteria. Humans neither efficiently absorb nor metabolize chlorophyll, and entirely depend on meat and milk products. Thus, phytanic acid can hardly be considered endogenous ligand nor physiological relevance. Implication of RXR activation in mediating effects of high amounts of phytanic acid on glucose and lipid metabolism or adipocyte differentiation (reviewed in (*Roca-Saavedra et al., 2017*)) were not tested in competition with an RXR antagonist or in any RXR null mutations mice.

Recently, the endogenous presence of 9-cis-13,14-dihydroretinoic acid (9CDHRA) with its all-trans isomer was identified in several organs (liver, serum, brain) from mice through a combined liquid chromatography-tandem mass spectrometry using UV analytical set-up and comparison with synthetic standard samples (*Ruhl et al., 2015*). Crystallography analyses revealed that 9CDHRA employs a similar mode of binding to RXR α -LBD as 9-cis RA. The physiological relevance of 9CDHRA was documented by deficits of spatial working memory associated with reduced levels of 9CDHRA in Rbp1-null mice, that could be rescued by 9CDHRA supplementation. Classification of 9CDHRA is still debated, whether a new bioactive metabolite of ATROL (vitamin A1) and ATBC (pro-vitamin A1), or whether other dihydro-form(s) of retinol or carotenoids exist in vivo

and could play the role of such a precursor. Thus, if ATROL or ATBC act as precursors of 9CDHRA, the latter could be considered as a new metabolite of the vitamin A1 family. If 9CDHRA cannot be efficiently generated from those precursors then it could be considered as a founding member of a new class of vitamin A (as vitamin A5). More experiments are still needed to further identify relevant metabolic pathways and to understand whether this new-hope 9CDHRA may act as ligand fulfilling the food-to-ligand concept.

Beside the quest for endogenous ligand, pharmaceuticals also pressed for synthetic ligands as potential therapeutic methods. The design of synthetic ligands was much instructed by the chemical structures of endogenous ligands, the structure of RXR LBD and data on ligand – LBD interactions. The diverse modes of these interactions suggest a possibility of existence of alternative natural or endogenous ligands with such novel or yet unknown modes of interaction.

1.1.6 RXR specific biological functions

In order to characterize the mechanisms of how RXRs are implicated in regulatory pathways, many studies sought to generate genetically modified animals with mutations in the RXR genes.

1.1.6.1 Total ablation and the roles of RXR isotypes in embryogenesis

During development, RXR α and RXR β are present at early embryonic stages, while RXR γ appears a few days later in gestation (*Mangelsdorf et al., 1992*). RXR α was the first isotype that loss-of-function study revealed lethal effect between E13.5 and E16.5 on mouse embryo, due to hypoplastic development of the ventricular chambers of the heart (*Sucov et al., 1994*). An overexpression study 6 years later showed that overexpressing RXR α in cardiomyocytes causes

dilated cardiomyopathy but fails to rescue myocardial hypoplasia in RXR α -null fetuses (*Subbarayan et al., 2000*), indicating that the effect of RXR α deficiency is not cell-autonomous. Further studies supported the requirement of RXR α expression in epicardium for triggering paracrine signal that are necessary for myocardial growth (*Chen et al., 2002, Kang and Sucov, 2005, Merki et al., 2005*). Using a whole-mouse-embryo culture system, wild-type E11.5 embryos treated with TGFbeta2 protein for 24 hours, a study (*Kubalak et al., 2002*) could produce enhanced apoptosis in both the sinistroventralconal cushion and dextrodorsalconal cushion similar to that observed in the RXR α -null mice. This study also showed that RXR α -null embryos heterozygous for a null mutation in the TGFbeta2 allele exhibited a partial restoration of the elevated apoptosis and of the malformations. Other heart defects observed in RXR α -null mice were an enhanced rate of cell death in both the mesenchymal cells of the conotruncal ridges and the parietal conotruncal cardiomyocytes, which led to agenesis of the conotruncal septum (*Ghyselinck et al., 1998, Kastner et al., 1994*). This indicated that RXR α is required for the transduction of the RA signal that controls apoptosis in the conotruncal segment of the embryonic heart.

RXR α -null mice fetuses also displayed ocular characteristics such as persistent fetal vasculature, closer eyelid folds, a thickened ventral portion of the corneal stroma, a ventral rotation of the lens, an agenesis of the sclera and a shortening of the ventral (*Kastner et al., 1994*). Similar effects were found in RXR β/γ -null mice (*Ghyselinck et al., 1997*). However, RXR γ -null mice appeared normal and compound RXR $\alpha^{+/-}/RXR\beta^{-}/RXR\gamma^{-}$ triple mutants mice were viable (*Krezel et al., 1996*), indicated the important role of RXR α among other isotypes during morphogenesis of the mouse embryo.

To determine the transcriptional role of RXR α in vivo, particularly the transactivation function carried out by AF-1 and AF-2 domains, Mascrez *et al* engineered mouse mutants repressing truncated RXR α proteins that lack:

- the AF-2 activating domain core-containing H12 located at the C-terminus of the ligand binding domain: *Rxra*^{af2o} mutants (Mascrez *et al.*, 1998)
- the N-terminal activation function AF-1-containing A/B domain: *Rxra*^{af1o} mutants (Mascrez *et al.*, 2001)
- both AF-1 and AF-2: *Rxra*^{afo} mutants (Mascrez *et al.*, 2009)

The *Rxra*^{af2o} mutants display lower percentage of the myocardium hypoplasia and the ocular syndrome that are hallmarks of the RXR α -null (Kastner *et al.*, 1994). Additional inactivation of RXR β —which on its own, displayed no effect in (Krezel *et al.*, 1996) study, but increased the frequency of the myocardium hypoplasia from 5% to 50%, and the frequency of the ocular syndrome from ~15% to 100% in *Rxra*^{af2o} mutants (Mascrez *et al.*, 1998). This might reflect a functional compensation of RXR β in the absence of RXR α . The *Rxra*^{afo} displayed full ocular syndrome like the RXR α -null but normal heart histology in 80% of the fetuses (Mascrez *et al.*, 2009). These results suggested the “silent” role of RXR α in transcription activation during myocardial development. The *Rxra*^{af1o} mutants however displayed only occasional persistent fetal vasculature, and additional ablation of RXR β and RXR γ only increased the occurrence of this defect but displayed no additional defect (Mascrez *et al.*, 2001). Hence, both RXR α AF-1 and AF-2 are required for the involution of the primary vitreous body, while only AF-2 is required for the other RA-dependent ocular morphogenetic events. These findings indicated that the activation functions of RXR α are differentially required for eye morphogenesis.

While The AF-2 of RXR α may appear functionally more important during development, AF-1-containing A/B region may have unique role in the RA-dependent disappearance of the interdigital mesenchyme. In early studies, RXR α - and RXR γ -, RXR α - and RXR β -null showed various level of causing interdigital webbing (i.e., soft tissue syndactyly) as either heterozygous and or homozygous mutant (*Ghyselinck et al., 1997, Lohnes et al., 1993, Lufkin et al., 1993, Kastner et al., 1994*). Furthermore, RXR α -null embryos showed resistance to limb malformation usually caused by retinoid acid exposure (*Sucov et al., 1995*). Other study showed reduction levels of either RAR α or RXR α alone decreased cell sensitivity to growth inhibition by all-trans RA; and cellular maximum RA resistance was obtained when both RAR α and RXR α were reduced (*Wu et al., 1997*). From the engineered mutant studies of Mascrez, majority of *Rxra*^{af1o} mutants and all *Rxra*^{af1o}/RXR β/γ -null mutants display a soft tissue syndactyly (*Mascrez et al., 2001*), while *Rxra*^{af2o} and *Rxra*^{af2o}/RXR β/γ -null mutants never display this defect (*Mascrez et al., 1998*). Phosphorylation of RXR α at a specific serine residue located in the A domain was found necessary for the anti-proliferative response of F9 teratocarcinoma cells to RA (*Bastien et al., 2002, Rochette-Egly and Chambon, 2001*). The crucial turn of function by phosphorylation, consistent with other studies presented previously in the structure of A/B domain, may act in a functioning cascade of molecular events that, in vivo, leads to the normal disappearance of the interdigital mesenchyme.

Comparing the severity and penetrance of a given abnormality between various mutants also led to the indication of which heterodimers are preferentially involved in transducing RA signals in a given developmental process: RXR α /RAR α heterodimers are preferentially transducing the RA signal acting on myocardial growth, RXR α /RAR β and RXR α /RAR γ are the heterodimers instrument the ocular morphogenesis (*Mark et al., 2009*).

1.1.6.2 Tissue specific ablation and diverse effects

Using tamoxifen-inducible chimeric Cre recombinase (Cre-ERT2) to selectively ablate RXR α in adult mice adipocytes, (Imai *et al.*, 2001) showed that these mice were resistant to dietary and chemically induced obesity and impaired in fasting-induced lipolysis. Their results also indicated the involvement of RXR in adipocyte differentiation.

Pirinixic acid, or Wy14,643, is known to be a PPAR agonist with strong hypolipidemic effects. In hepatocyte RXR α -deficient mice, Wy14,643-induced hepatomegaly was partially inhibited, while Wy14,643-induced hepatocyte peroxisome proliferation was preserved in adipocytes (Wan *et al.*, 2000, Nagao *et al.*, 1998). This mutant also displayed reducing food intake due to leptin increase, increasing body weight, improving glucose tolerance through high serum IGF-1 level mechanism (Wan *et al.*, 2003). Hepatomegaly induced by 1,4-bis[2-(3,5-dichloropyridyloxy)]benzene (TCPOBOP - a CAR and PXR ligand), was also prevented in hepatocyte RXR α -deficient mice (Cai *et al.*, 2002), supposedly also through the cytochrome P450 (CYP) genes family.

Mice lacking RXR α in myeloid cells exhibit reduced levels of CCL6 and CCL9, impaired recruitment of leukocytes to sites of inflammation, and lower susceptibility to sepsis (Nunez *et al.*, 2010).

Using Cre-ERT recombinases to ablate RXR α selectively in adult mouse keratinocytes, (Li *et al.*, 2000b) generated hair loss and skin abnormalities, indicating that RXR α has key roles in epidermal keratinocyte proliferation and differentiation. In the epidermis of VDR $^{-/-}$ mice, alopecia was observed, but keratinocyte proliferation and differentiation were normal. This similarity in phenotype may reflect a major role of RXR/VDR heterodimers in skin hair follicle cycling.

Using a designed dominant negative RXR β mutated receptor, termed DBD-, lacking the DNA binding domain but retaining the ability to dimerize with partner receptors, (Minucci *et al.*, 1994) showed inhibition of retinoic acid-responsive gene regulation in embryonal carcinoma cells.

The same lab that designed the domain AF-mutant mentioned in the earlier section (1.1.6.1) also did specific mutant in Sertoli cell and showed that *Rxrb^{af20}* mutation does not display the spermatid release defects observed in RXR β -null mutants, indicating that the role of RXR β in spermatid release is ligand-independent (Mascrez *et al.*, 2004). But they accumulated cholesteryl esters in Sertoli cells (SCs) due to reduced ABCA1 transporter-mediated cholesterol efflux.

RXR γ was reported to not co-express with RAR β and RAR γ in various differentiating muscles including those of the face and limbs. However, in the striatum and in part of the ventral horns of the spinal cord there was a precise co-expression with RAR β , although with quantitative differences, which suggested a possible preferential heterodimerization between these two retinoic acid receptors in the developing central nervous system (Dolle *et al.*, 1994).

1.1.6.3 Clinical relevance of subtypes

Nohara and colleagues (Nohara *et al.*, 2009) reviewed the effects of RXR subtype polymorphisms: those of RXR α were not linked with any metabolic dysfunction, whereas the RXR β c.51C > T polymorphism was linked to higher body mass, gallstone risk, and bile duct cancer risk. The association of an RXR γ polymorphism with hyperlipidemia was noted. The most common form of hereditary hyperlipidemia is the familial combined type (FCHL), which has been associated with increased very-low density lipoprotein (VLDL) that could be accompanied by increased low-density lipoprotein (LDL). The locus of RXR γ gene has been termed the “FCHL” locus and linked with higher LDL-cholesterol and triglyceride levels in several families. The RXR γ p.Gly14Ser

variant was observed in hyperlipidemic patients, was more common in those with FCHL, and was higher in those with coronary stenosis. Interestingly, the Ser14 variant suppressed lipoprotein lipase (LPL) promoter activity by 60%, whereas wild-type Gly14 suppressed LPL by 40%. LPL plays a role in the hydrolysis of VLDL. RXR γ variants have also been linked to free FA and triglyceride levels in familial type 2 diabetes.

1.2 The involvement of RXR in circadian rhythm

1.2.1 The CycliX project: interrogating the coordination of circadian, nutrient and cell cycles by transcription factors, including NRs

As life on earth directly depends on the energy of the sun, most of the earth inhabitants have their biological activities synchronized with the cycle of day and night – the circadian cycle. Adapting to the light cycle on earth, this diurnal rhythm is embedded in the organ and cellular clock of living organisms. High-level multicellular organisms like mammals further developed a hierarchy of oscillators that keep track of time in specialized tissues and organs. The master pacemaker - central suprachiasmatic nucleus (SCN) of the hypothalamus - relays the base tempo in activity and rest, feeding, body temperature, and hormones throughout the body. The receiving organs tune their metabolism to the rhythmic signal, in adapting to their local different stimulations and entrain their own peripheral clock (*Mohawk et al., 2012*). A fair number of studies have established plausible models of autonomous circadian oscillators, detailing transcriptional (*Young et al., 2001, Rosbash, 2009*), non-transcriptional (*Roenneberg and Merrow, 2005, Merrow et al., 2006*) and other input mechanisms (*Bunger et al., 2000, Cho et al., 2012*) that play key roles in mammalian clocks. Among the earliest to be discovered and extensively studied is the feedback regulatory loop between core oscillators and metabolism, with many evidences indicating that

disruption of either the core clock or of the nutrient cycle would affect the other. Other models of transcription regulatory cycles are present in nature. As an example, in the same phase-adapted transcriptional manner, the cell-division cycle that corresponds to genome duplication and segregation during cell proliferation, allows cycle progression only if certain nutrition conditions have been met. On the other side, the nutrient-response cycle leads "fasting" cells to undergo a program of gene expression that returns to the fasting state once the nutrients have been exhausted. However, although each individual cycle has been studied extensively, we still know little about the global genomic responses to the cycles and their associated transcriptional regulatory programs. Intrigued by the lack of a study on the global view level, CycliX project emerged with the mission of seeking to understand three of these circuits and how they interact: the circadian rhythm, cell division and nutrient-response cycles, and to shed light on whether there is a causality-consequence connection between them. By integrating multiple information from different core transcription factors, using high throughput sequencing methods, CyliX project aims to identify "cyclic nodes" genes and their regulators to study the interaction networks connecting these cycles. Through their most important common cis-regulatory modules and the corresponding transcription factors, this project aims to enable the generation of interconnecting core regulatory networks. As a key metabolic tissue, the mouse liver has been widely reported to be one of the organs under the crucial control of the core clock and the feeding behavior. In addition, hepatic cells can regenerate upon hepatectomy (*Cho et al., 2012, Miller et al., 2007, Le Martelot et al., 2009*). Genome-wide profiles have shown that around 10% of genes exhibit cyclic mRNA levels in the liver (*Akhtar et al., 2002, Panda et al., 2002, Storch et al., 2002, Ueda et al., 2002*), and were found to be accompanied by a highly dynamic epigenetic landscape (*Vollmers et al., 2009*). Thus, mouse liver became the central

target for multiple projects, and for Cyclix in particular, in an attempt to extrapolate the full view of the global transcription system and how its regulators coordinately drive these circuits.

1.2.1.1 The global landscape of epigenetics along the diurnal cycle

From using ChIP-seq for RNA polymerase II and III (Pol-II, Pol-III) and several histone modification marks together with microarray on C57BL/6 mouse liver tissue, the first Cyclix studies have shed light on the kinetic of epigenetic change and the effect of such modulation on global transcription activity during diurnal cycle. They found that rhythmic Pol-II recruitment at promoters, rather than rhythmic transition from poised to active productive stage, preceded mRNA accumulation by 3 hours, consistent with mRNA half-lives. Promoters of transcribed genes had rhythmic tri-methylated H3K4 levels reached their peak 1 hour after Pol-II, rhythms tri-methylation of H3K36 lagged transcription by 3 hours. They also identified three classes of genes: one showing rhythmicity both in transcriptional and mRNA accumulation, second class with rhythmic transcription but flat mRNA levels, and a third with constant transcription but rhythmic mRNAs that emphasizes widespread temporally gated posttranscriptional regulation in the mouse liver (*Le Martelot et al., 2012*). The mouse liver Pol-III occupied loci were shown to include a conserved mammalian interspersed repeat (MIR) as a potential regulator of subunit-encoding gene. They indicated a synteny relationships between a number of human and mouse Pol-III genes, whose expression levels are significantly linked (*Canella et al., 2012*).

The exploration is followed by a study incorporate another layer of regulatory from a key transcription factor the Sterol Regulatory Element Binding Protein 1 (SREBP1) that connects the lipid homeostatic and nutrient cycle with the circadian cycle. They found that the recruitment of SREBP1 to the DNA showed a highly circadian behavior, with a maximum during the fed status, but the expressions of target genes were not always synchronized. A group of genes that harbor

bindings sites for both SREBP1 and Hepatocyte Nuclear Factor 4 (HNF4) showed a shift by about 8 hours in their rhythmicity compared to SREBP1 binding, hinting the cross-talk between hepatic HNF4 and SREBP1. In time-restricted feeding *Bmal1*^{-/-} mice, whose internal molecular clock was disrupted, the temporal expression profiles of these genes dramatically changed even though SREBP1 rhythmic binding remained. This suggested the existence of a second layer of modulation of SREBP1 transcriptional activity, strongly dependent from the circadian clock, besides the nutrient-driven regulation of SREBP1 nuclear translocation (*Gilardi et al., 2014*).

1.2.1.2 The integrating circadian cycle with feeding rhythm

Inspired from the results from one transcription factor SREBP1, more recent studies expanded the transcriptional regulatory interplay between the circadian clock and feeding rhythms by incorporating DNase I hypersensitive sites (DHSs) (*Sobel et al., 2017*) and Pol-III (*Mange et al., 2017*). They found that hypersensitivity cycled in phase with Pol-II loading and H3K27ac histone marks. For the clock disrupted phenotype *Bmal1*^{-/-} mice, they found that the total number of rhythmic genes was equal to that in the wild type, although the amplitudes of expressions were generally lower. They analyzed the DNase I cuts at nucleotide resolution and found that dynamically changing footprints consistent with dynamic binding of CLOCK:BMAL1 complexes, suggested a transient heterotetramer binding configuration driven at peak activity of these footprints (*Sobel et al., 2017*).

Concordant to previous study of (*Canella et al., 2012*), Pol-III occupancy of its target genes were found to rise before the onset of the night, to stay high during the night—mice feeding time and translation is known to be increased—, and then to decrease in daytime. MAF1 is a repressor of Pol-III in response to serum starvation or TORC1 inhibition in fasting phase, and thus Pol-III occupancy during the night reflects a MAF1-inactivated response to feeding. The rise of Pol-III

occupancy before the onset of the night however reflects a circadian clock-dependent response. With this bimodal rhythm, they concluded that Pol-III transcription during the diurnal cycle is regulated both in response to nutrients and by the circadian clock, which allows anticipatory Pol-III transcription (*Mange et al., 2017*).

1.2.1.3 Cell cycle - the third cog coming into focus

Aside from MAF1, Pol-III activity is also tightly regulated with cell growth and proliferation by factors such as MYC, RB1, TRP53. With Pol-III bridging the circadian and the nutrient response cycle, the cell proliferation cycle was brought into the interplay using the model of compensatory liver hyperplasia-or regeneration-induced two-thirds partial hepatectomy. Early activation of cell-division-cycle genes were found after 10-20 hours post-surgery and continued with a robust and coordinate cell-division-cycle gene-expression response before returning to the resting state after 1 week. The activation of Pol-II, H3K4me3, H3K36me3 during the post-surgery response revealed a general de novo promoter recruitment. Some unusual gene profiles with abundant Pol-II but little evidence of H3K4me3 or H3K36me3 modification, indicated that these modifications are neither universal nor essential partners to Pol-II transcription (*Rib et al., 2018*). The results also define two classes of genes in response to hepatectomy: one class included genes close to H3K4me3 and Pol-II peaks, whose Pol-III occupancy is high and stable, reminiscent of housekeeping genes; the other class, distant from Pol II peaks, whose Pol-III occupancy strongly increases, represented the ability of adaptation to increased demand (*Yeganeh et al., 2019*).

1.2.2 The possible role of RXR in coordinating the three liver cycles

The ability to regulate reproduction, development, and nutrient utilization have been long known to coincide with the evolution of nuclear receptors. Studies on nuclear receptor superfamily

has been shown dynamic and expressed in oscillations in key metabolic tissues in the frame of one or all of these cycles (Yang et al., 2006, Yang et al., 2007, Duez and Staels, 2010, Nakatsuka et al., 2013, Bookout et al., 2006). In particular, RXR was proposed to be tightly integrated in one or more of these three cycles, especially in the liver, where the rhythm of transcripts encoding rate-limiting enzymes is required and warranted by the temporal nature of metabolic processes and biological clock (De Cosmo and Mazzocchi, 2017, Bookout et al., 2006).

RXR α was shown by (McNamara et al., 2001) to ligand-dependently interact with NPAS2 and CLOCK, but not with BMAL1, and impede the transcriptional activity of CLOCK/NPAS2–BMAL1 heterodimers at the promoters of clock genes. They also showed that RA injection prompted small phase shifts of peripheral clocks in the mouse cardiovascular tissue. The expression of DEC1 and DEC2, who takes part in the molecular clockwork, were shown to influence and be affected by RXR in heterodimer with LXR (Cho et al., 2009). The involvement of RXR in the circadian clock was also indirectly exhibited through the effects of its putative ligand retinoid acid. All-trans RA, 9-cis RA, and 13-cis RA were found capable to entrain circadian rhythmicity in cultured fibroblasts expressing Per2-luciferase (Nakahata et al., 2006), and retinoic acid signaling was shown to reset central clock upon light sensing (Thompson et al., 2004, Fu et al., 2005). The nutritional vitamin A deficiency shifted BMAL1 phases and abolished PER1 circadian expression at both mRNA and protein levels in the hippocampus (Navigatore-Fonzo et al., 2013).

Within the cell cycle, RXR is bound and phosphorylated by CK1 α in an agonist-dependent manner (Zhao et al., 2004), and by GSK3 β with modulation of cell predisposition to RXR agonist-induced growth arrest and apoptosis and support of cell survival (Gao et al., 2013). (Yang et al., 2010) showed that in partially hepatectomy mouse, RXR α deficiency caused an approximately 20-hour delay in hepatocyte proliferation, impaired several pathways, including growth factors and the

circadian cell cycle. The expression patterns of hepatocyte growth factor, fibroblast growth factor 2, platelet-derived growth factor, and transforming growth factor alpha were also altered due to lack of RXR α . This indicated the central role of RXR in the interconnection between circadian and cell cycles. Moreover, several heterodimer partners of RXR showed oscillating expression in mouse liver, and PPAR, in particular, has been proposed to be the link between circadian clock and energy metabolism (*Kersten et al., 1999, Feige et al., 2006, Yang et al., 2006*). (*Nakamura et al., 2008*) described the interplay between the PPARs/RXR α -regulated system and the molecular clockwork driven by CLOCK/BMAL1 heterodimers.

Many studies have suggested that RXR binding has a stable nature and does not have a time-dependence occurrence (*Simicevic et al., 2013, Boergesen et al., 2012, Minucci et al., 1997b, Brazda et al., 2014, Shen et al., 2011b, Tzameli et al., 2003, Mangelsdorf and Evans, 1995*). Notwithstanding, due to the peculiar characteristic of heterodimerizing with many other nuclear hormone receptors in binding to chromatin with or without ligand presence (*Minucci et al., 1997a, Westin et al., 1998, Boergesen et al., 2012, Hong and Tontonoz, 2014*), RXR served as a center point for the study of astonishing cross-talk between the nuclear receptors signaling pathways. Many high throughput sequencing experiments like ChIP-seq have been conducted to study the interaction networks connecting different nuclear receptors and other transcription factors through their important common cis-regulatory modules with RXR. Biochemistry studies of the domains involved in the coupling between RXR and partners have opened up a whole new concept of molecular interactions and their regulatory functions. The most direct effect on ChIP-seq studies is the abundance of overlapping binding sites between different nuclear receptors. Henceforth, global binding sites of RXR and its partners, and the effect of ligands have been established in some mouse and human tissues (*Tzameli et al., 2003, Frank et al., 2005, Shen et al., 2011a, Brazda et al., 2014,*

Hosoda et al., 2015). Thus, we hypothesized that RXR could be an important node for cross-talk between the CycliX cycles, and expected to demonstrate in more detail the dynamic mechanism of RXR chromatin bindings through ChIP-seq, and how this affects its transcriptional regulation through gene expression as measured by microarray RNA-transcript levels.

1.2.3 Aim of the project

To functionally investigate RXR and its synergic partners within the circadian cycle, one may wish to perform multiple ChIP-seq assays using paired antibodies against RXR and its partners at multiple time points. Aside from being costly and effort-consuming, this approach is unlikely achievable due to the limited availability and questionable quality variation of the required antibodies. To overcome this hurdle, we performed ChIP-seq assays using a homemade pan RXR antibody and took advantage of the heterodimerization nature of RXR binding to map the chromatin binding profiles of RXR at 7 time points with 4-hour interval on circadian cycle. We then analyzed the pattern of their binding sites and target genes to eventually infer the RXR coupling partner, instead of using individual antibodies for each factor. Several partners of RXR have been reported to be expressed in distinct phases (*Yang et al., 2006, Bookout et al., 2006*). It was therefore conceivable to anticipate the phased activities in their respected binding motifs. By combining the available regulatory elements and transcription factors data of Pol-II, H3K36me3, H3K4me3, DHSs and microarray from the previous experiments of the CycliX project (mentioned above in (*Le Martelot et al., 2012, Canella et al., 2012, Sobel et al., 2017, Mange et al., 2017*)), in relation with genomic features, we hoped to further illuminate the effect on transcription activity of RXR binding.

2. EXPERIMENTAL MATERIALS AND METHODS

Laboratory experiments were performed by Federica Gilardi and Michael Baruchet.

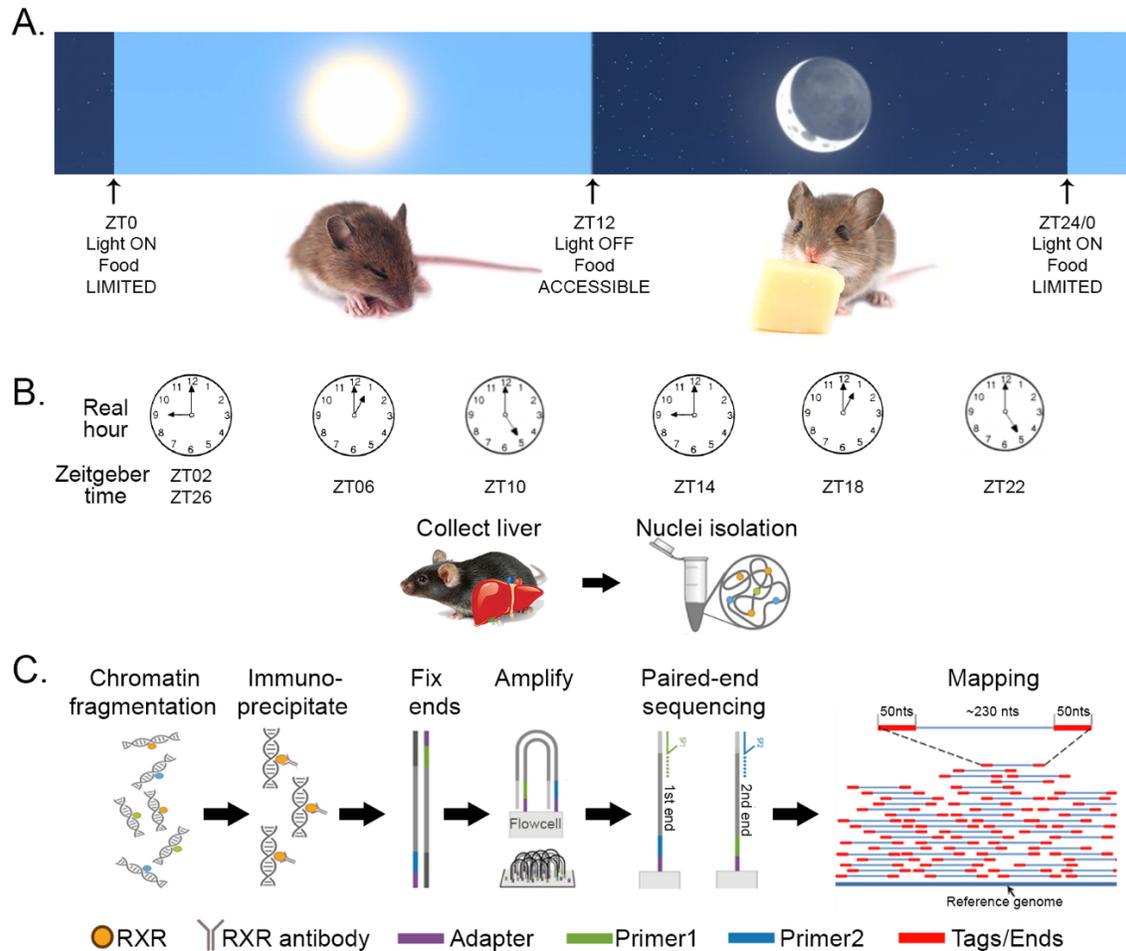


Figure 3: Study scheme

A. Experiment design through 7 days

B. 7 time points (in real hour and in ZT) when the mice were sacrificed, and their livers collected for nuclei extraction

C. Chromatin immunoprecipitation and paired-end sequencing steps.

For each precipitated chromatin fragment, the mapping was done on both ends

2.1 Animals and Treatment

C57/BL6 male 12-14-week-old mice were housed in the standard 12 hours light/12 hours dark (LD) cycle. For 7 days they had access to food only between ZT12 and ZT24 (Zeitgeber time;

ZT0 is the time when lights are turned on, ZT12 as time the lights turned off), but free access to water (**Fig.3A**). At 7 time points of 4 hours interval (ZT02, ZT06, ZT10, ZT14, ZT18, ZT22 and ZT26) 5 mice were anesthetized with isoflurane and sacrificed (**Fig.3B**). Their liver tissues were collected, pooled and immediately homogenized in PBS containing 1% formaldehyde for immunoprecipitation. All animal care and handling were performed based on the State of Lausanne's law for animal protection.

2.2 Chromatin Immunoprecipitation followed by sequencing protocol (ChIP-seq)

Livers of C57BL/6 mice were perfused with PBS before collection. The nuclei isolation was performed as described in (*Ripperger and Schibler, 2006*), except that the ultracentrifugation was performed for 1h at 24,000 rpm (100,000g) at 1°C in a Beckmann SW28 rotor. The nuclei were lysed with 20 mM Tris pH 7.5, 150 mM NaCl, 2 mM EDTA and 1% SDS. The sonication of the chromatin was performed with a Branson SLPe sonicator during 15 cycles of 10 sec at 50% amplitude. The sheared chromatin was then diluted 10x in 50 mM Hepes pH 7.9, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate, 0.5 mM PMSF, and protease inhibitor cocktail (Sigma, #P8340).

The inputs were isolated and the immunoprecipitation performed with a homemade anti-RXR antibody overnight at 4°C on a rotating platform. This antibody is a rabbit polyclonal antibody raised against the first 27 amino acids of RXR α A/B domain, generously provided by Cecile Rochette-Egly. Protein A-sepharose CL-4B beads (GE Healthcare, #17-0780-01) were added and the tubes were kept rotating for 2 h at 4°C. The beads were then washed 2x with 50 mM Hepes pH 7.9, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS, 0.5 mM PMSF, and protease inhibitor cocktail (Sigma, #P8340); 2x with 50 mM Hepes pH 7.9, 500 mM

NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS, 0.5 mM PMSF, and protease inhibitor cocktail (Sigma, #P8340); 2x with 20 mM Tris pH 8, 1 mM EDTA, 250mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate, 0.5 mM PMSF, and protease inhibitor cocktail (Sigma, #P8340); and 2x with 10 mM Tris pH 8 and 1 mM EDTA. The elution of the chromatin from the beads was done twice with 50 mM Tris pH 8, 1 mM EDTA, 1% SDS and 50 mM NaHCO₃ for 10 min at 65°C. NaCl was added to 200 mM and the chromatin was de-crosslinked at 65°C overnight. The samples were treated with 25 µg/ml of RNase A for 1 h at 37°C, then with 5 mM of EDTA and 50 µg/ml of proteinase K for 2 h at 42°C. The DNA was purified with the NucleoSpin® Gel and PCR Clean-up from Macherey-Nagel. Nuclei isolation method as described in *Nature Genetics* by (Ripperger and Schibler, 2006).

The DNA concentration was measured with the Qubit® Fluorometer (Invitrogen). The libraries were prepared starting from 2.1 ng to 4.2 ng of IP samples and 8 ng of inputs samples, using the Microplex Library Preparation Kit (Diagenode, #C05010010).

The immunoprecipitated chromatin fragments were pair-ended sequenced with Illumina HiSeq 2500 by 100 cycles each sample. The reads were aligned on mouse genome (Mus musculus NCBI m37 genome assembly (mm9; July 2007)) and quality controlled by Illumina pipeline Casava 1.82. (illustrated in [Fig.3C](#))

2.3 ChIP-seq library control and read mapping

For each time point, we processed two samples: one immunoprecipitated with pan-RXR antibody and one not immunoprecipitated as control (input). We used the mapping quality from the fastq files produced by Illumina pipeline to filter all the reads that did not pass the filter, reads that contain ambiguous match or more than 5 mismatches. Only the reads with both ends uniquely mapped were kept. All 14 samples (7 RXRs and 7 inputs) have a high percentage of well-aligned

uniquely mapping reads (over 80%). We also checked the redundancy level of all pair perfect reads. A read is considered redundant when it was mapped on the same genomic position multiple times; and contra wise the non-redundant read was mapped only once at a genomic position. All our samples have fraction of non-redundancy above 80% - which is endorsed by the ENCODE project (Landt et al., 2012). We used one copy of each redundant fragment in subsequent analyses.

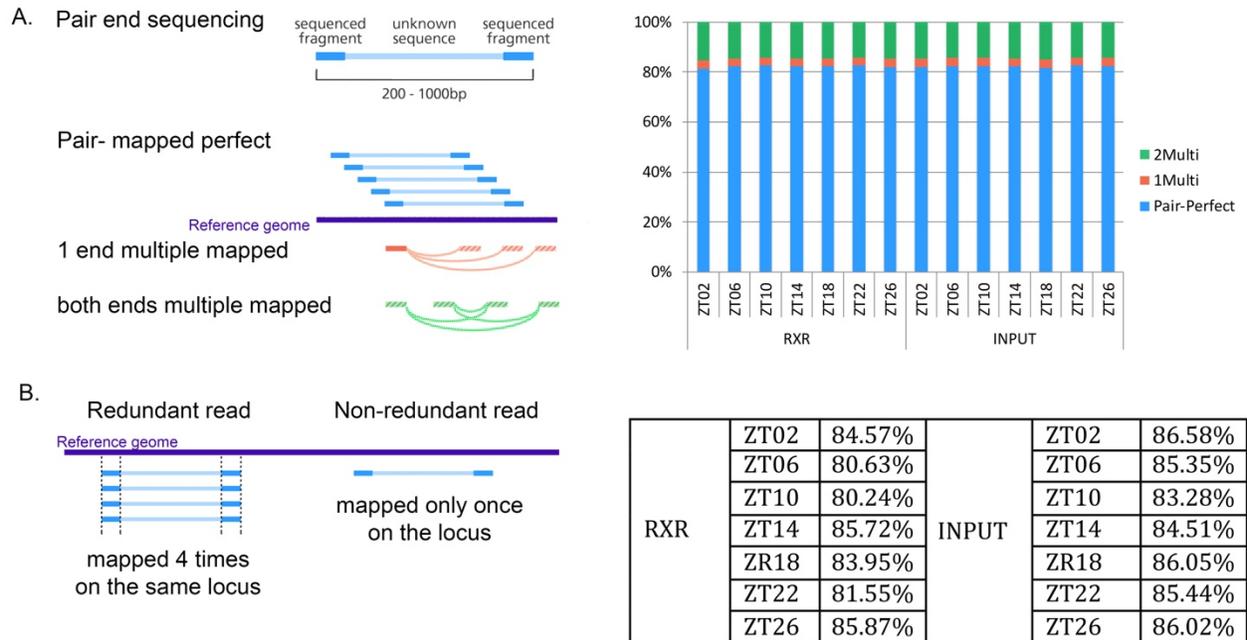


Figure 4: ChIP-seq library quality

- A.** Model of 3 mapping types of pair-end fragments onto the reference genome: pair-mapped perfect when both ends are uniquely mapped, 1end-mapped, and both ends mapped are when one or both ends map multiple times on the genome. The bar-plot shows the percentages of these 3 models in each sample
- B.** Threshold $T=1$ for non-redundant read and the table showing the non-redundant fractions of each sample

The mean fragment length is 214 bases long. Different from the normal ChIP-seq data with single-ended, where only information of a small portion at one end of precipitated fragments is known, here we have whole fragment span information. Thus the length of each fragment is known precisely, and the process of shifting the tags on the top and bottom strand to guess the location of each binding site is unnecessary. This facilitates the data processing and eliminates the effect

created by the overlap of different length reads. We used only the 50 bases located in the middle of each fragment for the enrichment step of the peak calling method.

Files generated:

- For analysis: BED files were generated from all the uniquely mapped reads remained after filter. The total reads for each sample refer to these numbers.
- For USCS genome viewer: bedGraph and bigWig files were generated from the BED files and scaled by the total tag number for each sample.

2.4 Peak calling

2.4.1 Identifying the enriched regions

The genome was divided into 500 bases long consecutive, non-overlapping bins to calculate the ChIP signal. For each bin, each sample has its own value of tag density which is the AUC (area under the curve) created by the tags piled-up. With the uniquely mapped pair-ended tags taken into account, the samples cover at least 3.3 fold genome size.

For each time point, the bin values of IP sample (RXR) and of the input sample were log 2 scaled, and then used to compute ratio-mean distribution. This distribution was sectioned into 200 step-wise proportions along the mean axis. Smoothing function Lowess with smoother-span 0.2 was applied for only the negative-ratio bin population of each step-wise proportion. From the predicted distribution of smooth function, the mirrored bins were selected on the positive-ratio population. Distribution function was applied for this positive-ratio population with the mean and standard deviation derived from smooth function. The p-values were adjusted for false discovery rate (Benjamini–Hochberg). All bins with adjusted p-value less than 0.05 were considered

significant. The second filter layer was the 4-fold ratio of RXR over input within the significant bins.

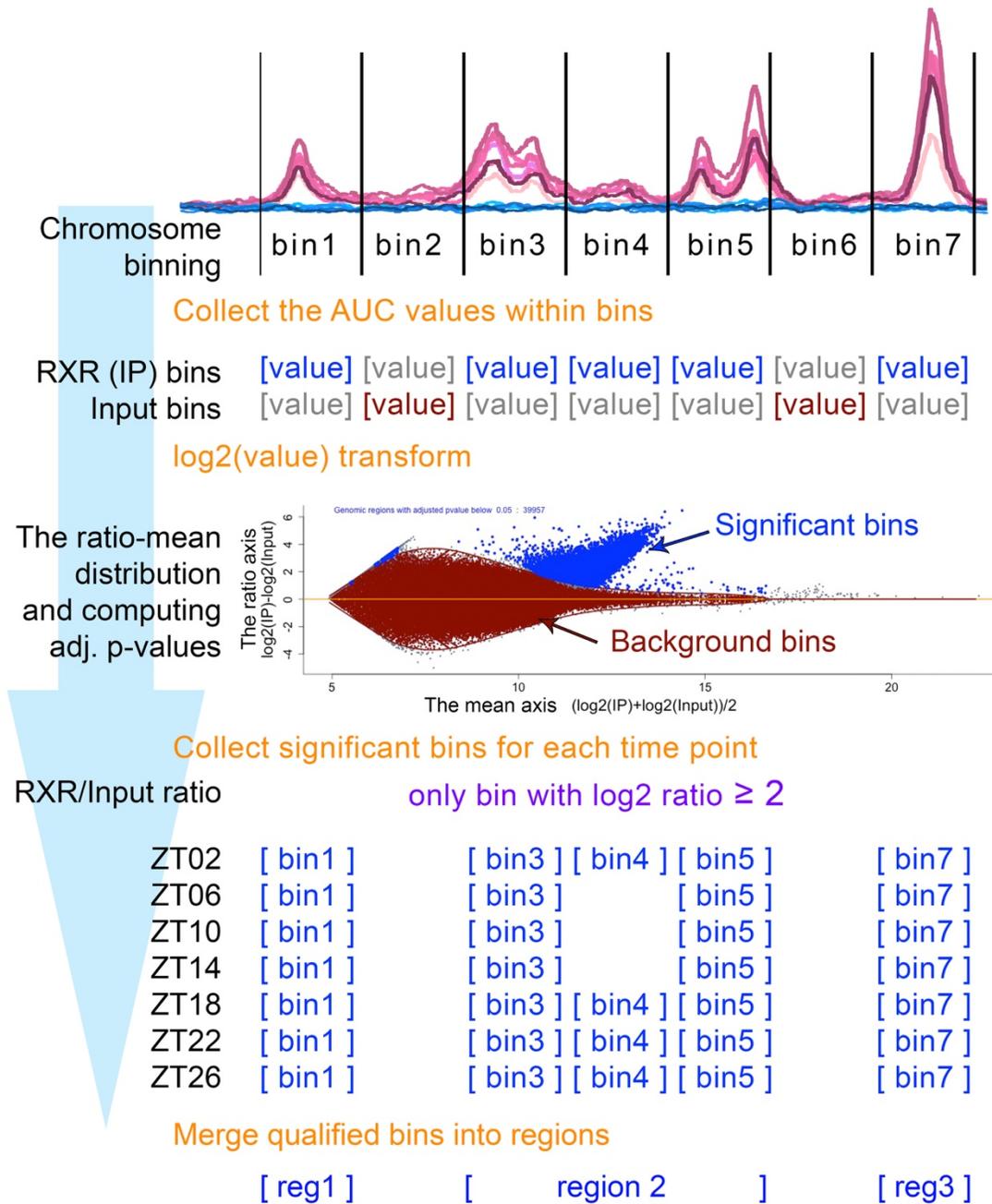


Figure 5: Identification of enriched regions

Work flow of method. We bin the genome into equal 500 bases length bins, calculate p-values from the distribution of the ratio-mean of log₂ tag density of RXR and Input in each of the 7 time points. Significant bins with adjusted p-value ≤ 0.05 and RXR/Input ratio > 4 in at least one time point were collected and merged into regions.

We also repeated the same whole genome analysis using bins shifted of half bin size (250 bases) in order to collect all the regions where the signal would be split over two consecutive 500 nt bins. The significant bins in both bin sets (non-shift and 250-shifted) were pooled into the final bin set was determined for each sample. All the collected bins for all time points were pooled and merged with maximum distant of 250 –two bins separated by 250 bases would be merged– into significant regions.

2.4.2 Peak refinement within the regions

For each enriched region, we used the value of tag density of all 7 time points at each nucleotide to compute the maximum, mean and minimum coverage lines. Loess Curve Fitting (Local Polynomial Regression) function was applied on these lines (*Fig.6A.a*). The subsets of data used for each weighted least square fit in Loess are determined by a nearest neighbor algorithm. The smoothing parameter, α , is a number between $(\lambda+1)/n$ and 1, with λ denoting the degree of the local polynomial.

-
- c) *Prediction by Spectrum function: the best fitted width was used to select the best smoothing degree*
 - B.** *The computation of inflections points (green dots) at 3 data windows from panel A.a*
 - a) *Inflection point is identified when the data curve diverts from the predicted trend of the curve*
 - b) *Example case when the maximum and mean curves failed to produce inflection point, method is applied to the minimum curve.*
 - c) *When all three curves failed to produce inflection point, substitute is identified where the curve crosses input values*
 - C.** *The combination of inflection points to form the peak unit(s) for a) Single, or b) Double peak*
 - D.** *Difficult cases where:*
 - a) *too many small peak units - the low plateau*
 - b) *two peak units separated by the high amplitude “valley”*

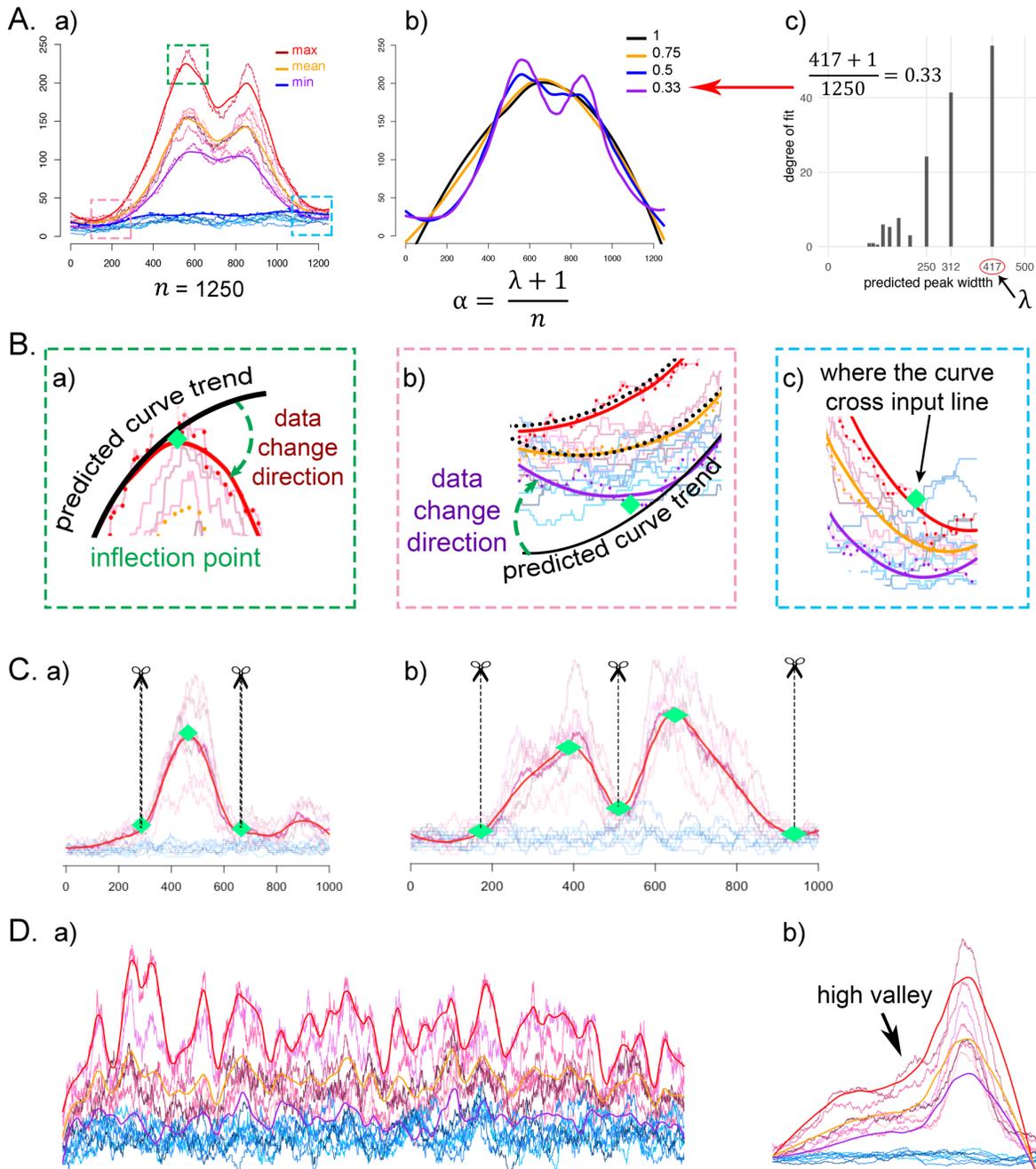


Figure 6: Peak-refinement method

A. Application of Loess function and smoothing parameter prediction

- a) Example of a region with 2 peak units, region width is 1250 nts. The maximum (red), mean (orange), minimum (purple) smooth curves generated by the Loess function.
- b) Effect of different α values on the smoothed lines and consequentially on inflection points identification

The value of α is the proportion of data used in each fit – in our case is the ratio between the expected width containing ChIP-seq peak (λ) over the total width (n) of the region (**Fig.6A.b**). Because the widths of peaks varied greatly, we used the Spectrum function to predict the λ value. (**Fig.6A.c**)

This Loess function produced a set of fitted values, which comprised a smooth line representing the peak. The predict function applied on this curve produces predicted values, obtained by evaluating the regression data of regions. If the predicted values progress with the same linear regression of the smooth curve, we have no inflection. When the predicted values diverge from the smoothed curve, or, in other word, the smoothed curve changes its direction, we could identify at that position an inflection point (**Fig.6B.a**). A combination of 3 inflection points marks a single peak – this is a peak unit. Combinations of 5 or 6 inflection points that make up two peak units indicate the double peak and so on (**Fig.6C**).

To account for the diversity of highest amplitude of each region, the Loess function can be applied to the maximum, mean or minimum value line to retrieve the most optimal inflection points (**Fig.6B.b**). For some regions, the margin cut very close to the peak unit that inflection point could not be detected, we substituted by the location where the data curve crosses the maximal input level (**Fig.6B.c**). When several inflection points were optimal, we fitted the data curve within each peak unit formed by each combination to the Gaussian distribution, then selected the combination of inflection points that yielded the best fit.

Nearly 10% of the enriched regions presented difficulty in peak separation because of the close distance between peak units (**Fig.6D.a**). A similar difficulty was presented for the other 5% highly enriched regions because of the high amplitude of the “valley” between peak units (**Fig.6D.b**). For these regions, we applied the smallest α fraction for the Loess function and used

the rising threshold to collect cross-section points as candidate inflection points. A small subset of regions yielded too many equally small peak units. After visual inspection, they were assigned to a special peak group without refinement called “the low plateau”.

2.5 Motif analysis

All the sequences under RXR peaks were input for motif enrichment analysis using MEME tools with *meme-chip* function, with parameters

- for MEME: *-meme-mod zoops -meme-minw 6 -meme-maxw 30 -meme-nmotifs 50*
- for DREME: *-dreme-e 0.1*
- for CentriMo: *-centrimo-local -centrimo-score 5.0 -centrimo-ethresh 10.0*

The customized hexamer scan, tandem repeats search, and motif position weight matrices (PWM) construction were performed in R using Biostrings package. The statistic scan was done by FIMO (*Grant et al., 2011*) tool, with relax p-value threshold of 0.1, using the PWMs on whole RXR peaks and on ± 15 nts extended summits. Each locus was assigned a Benjamini-Hochberg adjusted q-value (false discovery rate) by FIMO for its significance of occurrence.

2.6 Genes enrichment analysis

All gene sets enrichment, genes family and pathways analysis were performed on R using clusterProfiler package.

Gene networks were generated by ToppCluster as eXtensible Graph Markup and Modeling Language (XGMML) format. We used Cytoscape to perform layout and rearrangement for figure illustration.

2.7 Bioinformatics tools

R studio 3.5.2

BEDTools version 2.26.0:

Samtools version 0.1.19

MACS version 1.4.2.1

HOMER version 4.7

MEME suite version 4.12.0

CycliX toolbox

hgWiggle and phastCons 30way, bedGraphToBigWig from ENCODE

Cytoscape 3.7.0

2.8 List of used data sets

All the experiments were sampled in the circadian 7 time points of mouse liver tissue.

Data set	Data type	Antibody	Set*	Ref.	GEO
Pol-II	ChIP-seq	Santa Cruz Biotechnology, sc-673-18	A	<i>(Le Martelot et al., 2012)</i>	GSE35788
H3K36me3	ChIP-seq	Abcam, ab9050	A	<i>(Le Martelot et al., 2012)</i>	GSE35788
H3K4me3	ChIP-seq	Abcam, ab8580	A	<i>(Le Martelot et al., 2012)</i>	GSE35788
DHSs	DNaseI-seq	Worthington Biochemical Corporation, DPFF	B	<i>(Sobel et al., 2017)</i>	GSE60430
mRNA	mRNA microarray	Invitrogen, TRIzol reagent	A	<i>(Le Martelot et al., 2012)</i>	GSE35789
RXR	ChIP-seq	customized pan-RXR	C		

(* *Experiment sets* A, B and C indicate different sets of mice used for each experiment, each time point pool included 3-5 mice.

3. RESULTS

3.1 The peak-caller method is efficient to identify ChIP-seq enriched regions

The classical peak caller methods — for example MACS (*Zhang et al., 2008*), Homer (*Heinz et al., 2010*), SCICER (*Xu et al., 2014*), ChIPPeak (*Ambrosini et al., 2016*) and many other — can only process one condition/time point at a time, with one immunoprecipitated sample and one input sample, and would produce one set of peaks for each time point. Other methods have been developed to detect differential ChIP-seq peaks in pair-wise conditions comparison — such as DiffBind (*Stark and Brown, 2011*), DBChIP (*Liang and Keles, 2012*), MAnorm (*Shao et al., 2012*), diffReps (*Shen et al., 2013*), ODIN (*Allhoff et al., 2014*), ChIPComp (*Chen et al., 2015*) — and may be better fitted for our study analysis. But due to the study scheme of 7 time-points, using these peak callers would produce a large number of combinations of differential comparisons that would complicate downstream analyses. We thus developed an appropriate method suitable for the combination of multiple timepoints, described in Materials and methods. This method allowed us to identify 25'233 enriched regions. To validate this approach, we compared these enriched regions with the peak sets obtained from using the MACS tool. For a given single RXR binding site, MACS produced 7 different peaks (*Fig. 7A*), while our method created one single region. The computation time to process all the sample with our method is comparatively shorter than using the MACS2 peak caller. We used the binary bam-like files that were developed during the CycliX project, which are lighter, easily manipulated, and efficient for all further analysis steps.

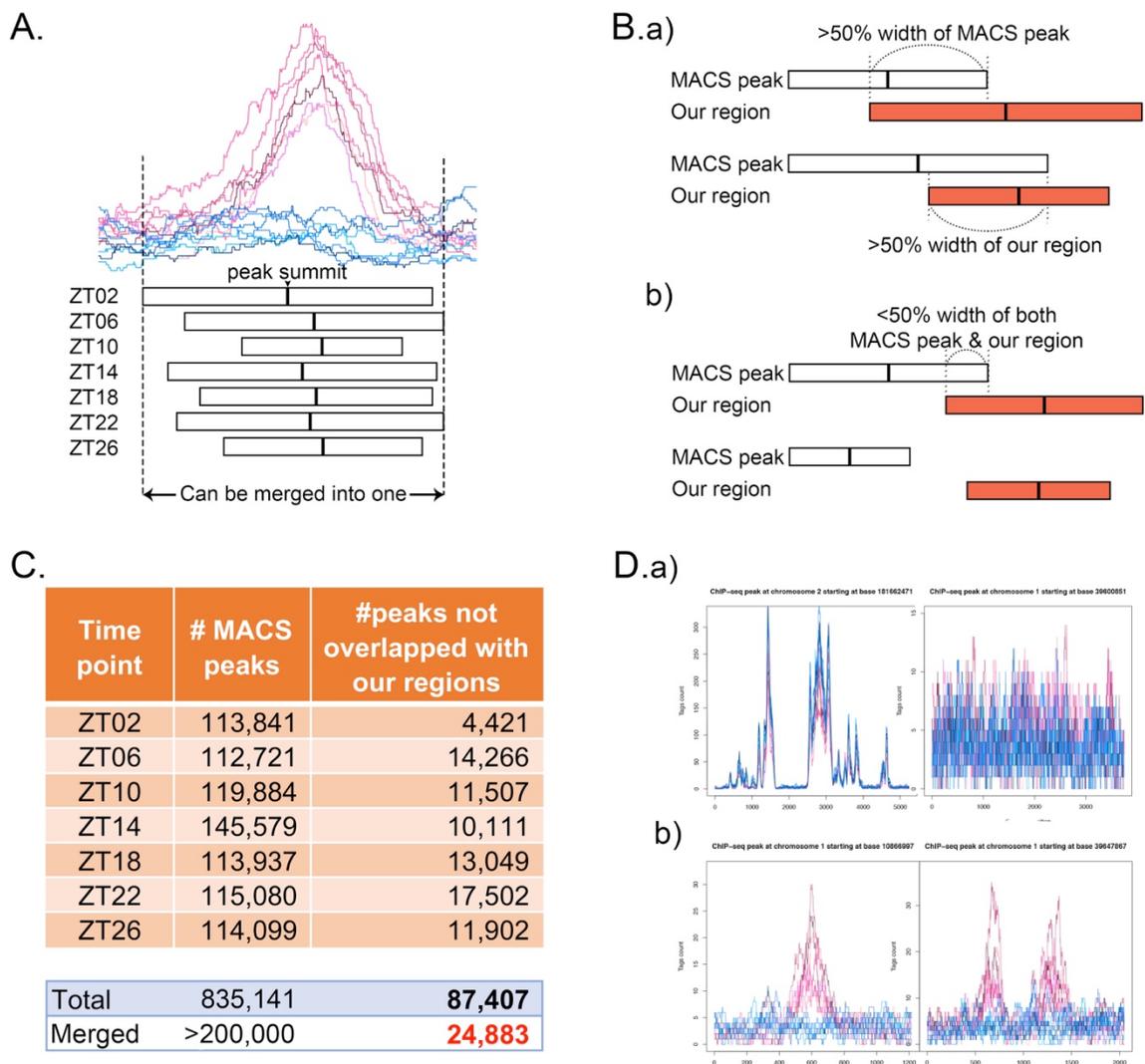


Figure 7: Comparing enriched regions identified by our method and peaks from MACS2

A. Complication created by MACS: each time point produced different peak coordinates and summit positions that actually can be merged into a single region

B. Intersection of MACS peaks with enriched regions identified by our method

a) Threshold to retain overlapping peaks

b) Non-overlapping cases

C. Table showing the number of peaks called by MACS2 for each time point, and the number of peaks that did not overlap with our enriched regions. The total 87'407 non-overlapped peaks represented 24'883 distinct regions

D. 4 examples of those 24'883 regions:

a) False peaks called by MACS: with same RXR signal as the background, or low signal

b) The bona-fide peaks that were overlooked by our method due to RXR/input ratio lower than our threshold

We used BEDTools to intersect each MACS peak set with our enriched regions. Each overlap that encompasses at least 50% of either our region or a MACS peak is considered an overlap (**Fig. 7B.a**), otherwise it is a non-overlap (**Fig. 7B.b**). As anticipated, MACS produced large number of peaks for each time point, which added up to total more than 800'000 peaks. However, as illustrated in **Fig. 7A**, different MACS peaks from different timepoint could have little coordinates variations within the same region, which reduced the total number of MACS identified regions down to over 200'000. 97% of the 25'233 enriched regions identified by our method were also detected by MACS, whereas 24'883 regions (resulting from the merging of 87'407 peaks from MACS) were not reported by our method (**Fig. 7C**). However, it turned out that the vast majority (~99.7%) of these MACS merged peaks have either the same signal as background or a very low signal accumulation spread in wide range regions (**Fig. 7D.a**). These peaks were not detected by our method because of the threshold of 4 fold enrichment over input that we imposed at the beginning, which ensured that we did not collect as many insignificant peaks.

Among these 24'883 regions identified by MACS, our method missed only 82 bona-fide regions, whereas it eliminated 24'801 false positives regions. Interestingly, the 82 bona-fide regions we missed (examples in **Fig. 7D.b**) would have been identified had we used a RXR/input ratio of 3.8 instead of 4.0. However, considering the relatively large number (over 100'000) of false positive peaks that would have been included by decreasing the RXR/input ratio cut-off value, we maintained the stringent threshold which limited the number of false negative peaks (n=82).

3.2 Peak Refinements

3.2.1 Identifying single, double, and multiple peaks

After validating our total enriched regions with the results from MACS, we sought to inspect their general characteristics. The 25'233 regions covered 9.5% of the mouse genome assembly mm9, containing 5.5% total sequenced reads. This fraction of reads retrieved in peaks is quite higher than the estimated 1% usually found in ChIP-seq study (*Landt et al., 2012*). Due to the method, the enriched regions widths were uniformly bins of $500 + n \cdot 250$ nucleotides (nts) (*Fig.8A*). One common problem for all peak callers, especially the ones analyzing only pair conditions, is the fact that original regions usually not only contain the ChIP peak but also some flanking portions. These flanking portions are the source of error in subsequent motif analysis, where longer sequences decrease the precision of result. To refine the peaks in those regions and limit the error, we first employed the established method (*Gilardi et al., 2014*), simply removing small portions of each side of the region until the remaining signal reached 75% of the original signal (*Fig.8B.a*). This method works well with small regions containing only one peak and has the advantage of simplicity. But for larger regions that contain two or more peaks next to each other (20-30%, of our set), this method is not sophisticated enough to retain the peak information. The removed chunks often cut on the flank of peaks; and for regions containing peaks of unequal signal, the smaller peak was often cut at the summit (*Fig.8B.b*) whereas the region should be split to generate two distinct peaks. These features show that more sophisticated method of refinement was required. We thus developed a method to refine the ChIP-seq peaks, which uses the actual curve of the tag density (the peak) to determine the signal boundaries. Using a prediction function on the smoothed curve, we can identify the inflection points, defined as position on the curve where it changes direction (see the method

section for details). For each combination of inflections points indicating that the curve changes direction in an up – down – up fashion, we define a peak.

Based on how many peaks were contained in each region, and their shapes, we identified four separate morphologies (*Fig.8C*):

- the single peaks: only 1 peak in the region
- the double peaks: 2 peaks positioned close to each other in one region. The distance and the peak shape between these 2 peaks are variable.
- the multiple peaks: there are more than 2 peaks located near each other in the same regions
- the plateau peaks: regions where tag accumulate on a wide chromatin region. Most of them present a continuous high tag density along wide genomic ranges (over 1500 nts), a few others contain in addition one or more very tall peak-shape-regions within a high tag density plateau.

For each refined peak, the tag counts falling in the peak at each time point were normalized by total tags of the sample and divided by the width of each peak (*Fig.8D.a*). Thus, the change in peak amplitude through 7-time points can be inspected by a simple plot. Moreover, peak amplitude can be presented and ranked by the maximum, the minimum, the mean, the variation across time points and the relative amplitude (*Fig.8D.b*). These various aspects of peak amplitude were used in different analyses, in particular for the rhythmic analysis. The same 7-values format was used for all the factors studied.

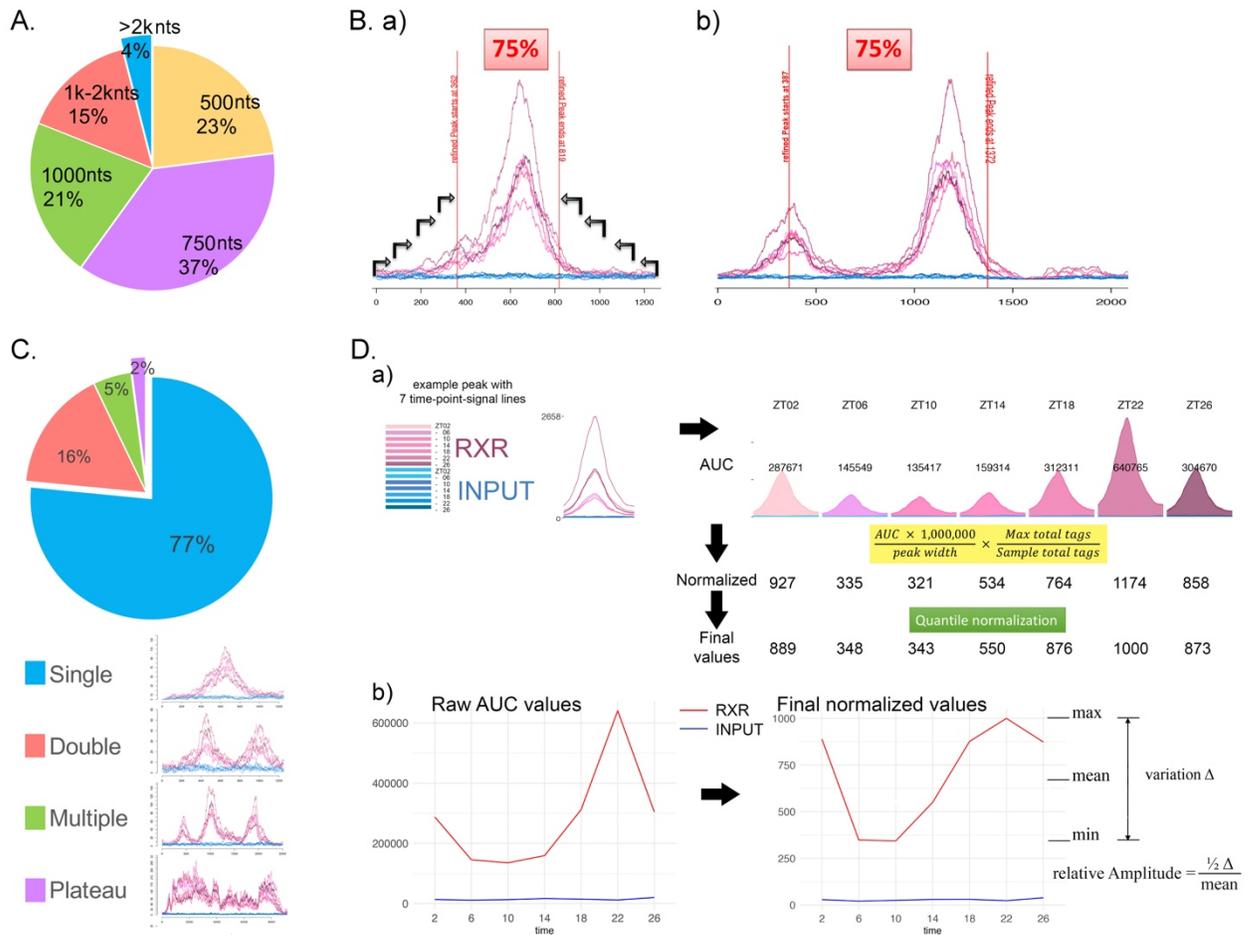


Figure 8: Results from peak refinement

- A.** Pie chart showing the relative proportion of enriched regions' widths
- B.** Identification of peak summit by retaining region encompassing 75% of the signal, vertical red lines indicate where the method refines the peak
 - a) Illustration of a successful case where only a single peak lays within a region
 - b) Examples of failure, where multiple peak units reside within one region
- C.** Pie chart describing the relative proportion of four peak morphology groups based on Loess fit refinement method, and example models of each peak group
- D.** Refined peak value
 - a) Step by step normalization for each refined peak
 - b) Illustrating effect of normalization on the 7 time point values

3.2.2 RXR peaks validation by analyzing the overlaps with DNase hypersensitive sites

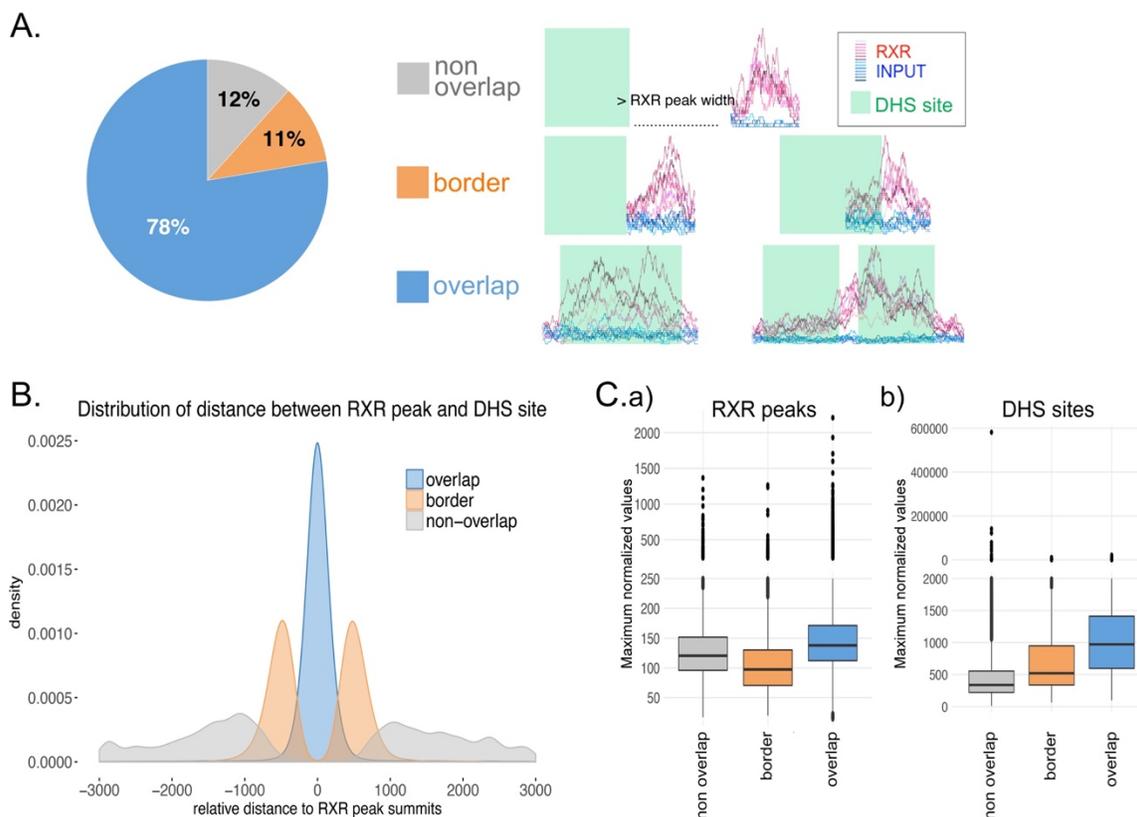


Figure 9: RXR binding sites overlapping DNase I hypersensitive sites

- A.** Pie chart (top) shows proportions of 3 types of overlap between RXR ChIP-seq peaks and DNase I hypersensitive sites (DHS sites). An overlap is considered “real” only when at least half of the RXR peak is covered. Illustrations (below) show examples for each type of overlap:
- The non-overlap (grey): including RXR peaks that do not overlap with any DHS site,
 - The border overlaps where the DHS site is adjacent to RXR peak, and where the overlap regions are narrower than half of the width of RXR peaks.
 - The real overlap (blue): when RXR peaks overlaps with one or more DHS site.
- B.** Box plots of maximum normalized tag densities of a) RXR peaks, and b) of DHS sites within each overlapping type
- C.** Distance between DHS sites with RXR peak summits of 3 groups of overlapping

DNase hypersensitive sites (DHSs) correspond to regions of open chromatin regions, where transcription factors may have access to DNA. In the frame of the CycliX project, a genome-wide analysis of all DHS has been performed on a same time-wise dataset (same time entrained scheme

but different mice; the same time-wise data set (Sobel et al., 2017). We observed that our RXR ChIP-seq peaks are often highly correlated with the DHSs from DNaseI-seq. As a validation for the peak refinement, we intersected the refined RXR peaks with the loci of DHSs. The overlapping fraction required to consider an overlap is the same we used for the MACS comparison (see Fig.8B.a for the principle of comparison).

About 90% of our RXR bindings overlapped with DHS sites. 10% of the overlaps are borderline, where the overlap sequences are less than half of the peak width, but it is possibly due to the difference between our peak-calling method and the one used in (Sobel et al., 2017) or simply due to the fact that different mice were used (Fig.9A). Figure 9B shows the distribution of distances between DHS sites and RXR summits for the 3 types of overlaps defined. There were significant differences in the tag density of RXR peaks between these 3 groups (Fig.9C.a), as well as in the tag density of the DHS sites (Fig.9C.b). While the tag density of DNaseI increased with the degree of overlapping, the borderline RXR peaks appeared to have lowest tag density.

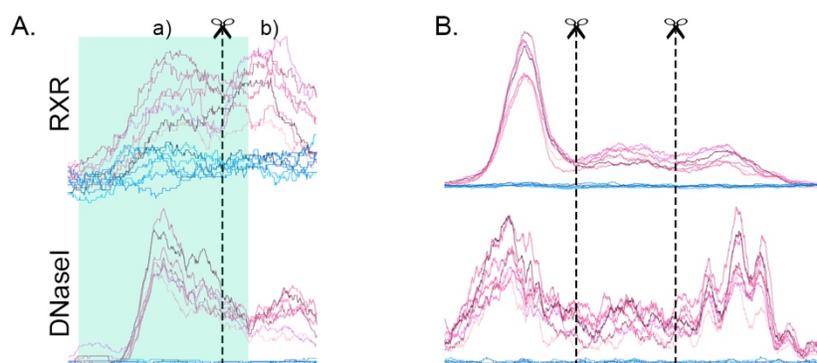


Figure 10: Exploiting the DNaseI-seq tag density track

A. Case of two RXR peaks:

a) one correlates with DNaseI peak (identified by ChIPPeak - green rectangle)

b) one with lower DNaseI tag density peak (not identified by ChIPPeak) and thus assigned into the border overlap (from Fig.8B)

B. Case showing the dynamic correlation of tag density between RXR and DNaseI

The DHS sites from the Cyclix study (Sobel *et al.*, 2017) were called by ChIPPeak (Ambrosini *et al.*, 2016), resulting in exactly 600 nts long sites (± 300 nts extended from the summit). To understand the extent of the difference between our peak-calling method and the one used by Sobel *et al.*, we re-computed the per-base maximum genomic coverage of DNase I and compared it with the RXR coverage under our peaks. This analysis clarified the border intersection group from the previous analysis, which mainly resulted from difference in peak-calling methods (Fig.10A). Following this reanalysis, almost all RXR peaks colocalized with DNase I accumulation, albeit the variable in tag density. As expected, the corresponding tag density between RXR peaks and DHS sites was not well correlated (Fig.10B), some very prominent RXR peaks were in company of very low DNase I tags accumulation. More than two third of identified DHS sites had no RXR signal either, which was expected since RXR does not bind systematically to all open regions, thus verifying the specificity of the RXR antibody. Moreover, although these two data sets were designed on the same time concept, the sample tissues were collected from different batch of mice. This observation, hence served as an additional development step of our peak refinement method, especially for double peaks where the interval between peak units was too high or too short to elicit a separation (Fig.10A). By using the discrepancy between RXR and DNase I tag densities, we could divide different peak units within one enriched region (Fig.10B).

3.2.3 PhastCons analyses and peak shapes

To further characterize the different categories of RXR peaks we identified (single, double, multiple and plateaus), we explore the possible relationship between the various peak shapes and the underneath sequences, as recently proposed (Cremona *et al.*, 2015). In addition, a previous report on RXR:RAR binding properties highlighted the fact that the degree of conservation of the

binding motif sequences across species were higher than neighbor sequences (*Chatagnon et al., 2015*). To further explore these two observations with respect to our data, we retrieved the base-wise phastCons score based on the genomic conservation between the mouse genome and scores for multiple alignments of 30 vertebrate genomes (phastCons30way).

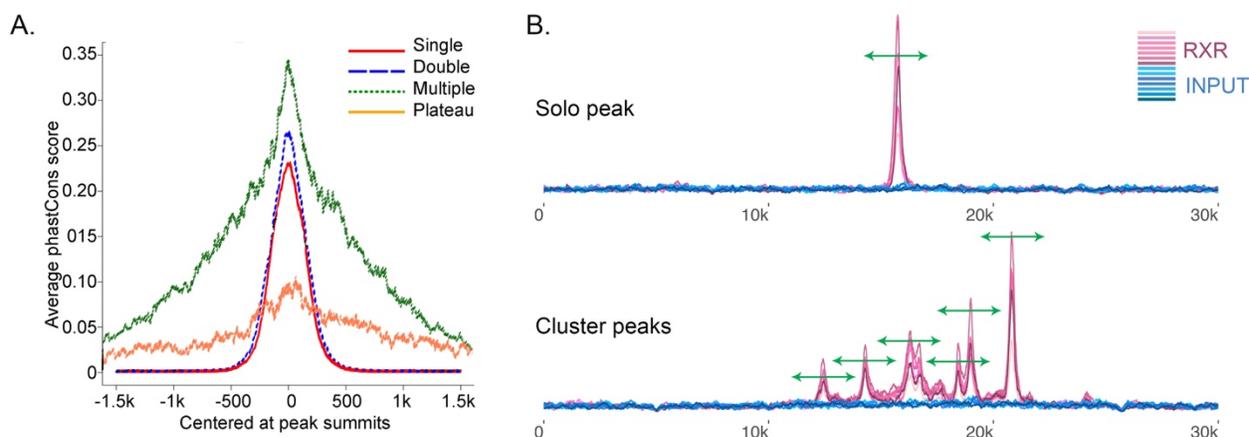


Figure 11: Peak morphologies and peak co-localization

A. The difference in base-wise average phastCons score between four morphology peak groups
B. Topography of cluster and solo peaks: the solo peak (top) has no neighbor peaks on chromatin landscape within a window of 2000 nucleotides from its RXR binding summit (depicted by green arrows), while cluster peaks (bottom) have at least one neighbor peak

The phastCons score represents the probabilities of negative selection and range between 0 and 1. Since the conservation score correlates closely with binding site motifs, we speculated that the phastCons score and/or the motif distribution would have been different between our morphology groups. For each peak, we used hgWiggle command (developed by UCSC bioinformatics group) to retrieve and then calculate the base-wise phastCons average score within each morphology groups +/- 1.5 kb around the peak summit. The average base-wise phastCons score of each peak group shows that the genomic regions containing Multiple peaks have the highest conservation score per nucleotide, followed by the Double, the Single then the plateaus (**Fig.11A**). This shows that RXR tends to accumulate more at DNA regions that are more conserved across

species. The lowest score interestingly belongs to the plateau peaks group, showing that in spite of the continuous high tag density stretch on long regions, these long chromatin regions are not well conserved. However, it is interesting to note that although the maximum conservation level is not as high as for other morphologies, unlike single peaks, the score does not fall to zero +/- 500nt away from the maximum.

Although the morphology classification did reveal several interesting characteristics of the RXR peaks, the subsequent genomic features distribution analysis alerted us of some caveats. Firstly, the morphology classification was based on number of peak(s) which could be found within one enrichment region, without considering the distance between peaks nor the genomic context of the loci. For example, two single peaks located 250 nts apart were closer to each other than a double peak with 500 nts in the valley between. Secondly, we suspected some differences in genomic features within each peak group that would definitely affect the average phastCons scores for each group. It is worth noting that there was a correlation between the “degree of accumulative” RXR peaks and the genomic feature proportions.

Considering these observations, we reclassified the RXR peaks—that are not plateaus—based on the degree of clustering of the peaks on chromatin landscape. The occurrence of a neighboring peak summit within +/-1kb extension from the main summit assigned a peak as “cluster”; and the absence of neighboring peak assigned a peak as “solo” (**Fig.11B**). Altogether, out of 37’143 well-refined RXR peaks, 36’124 were peak-shape, 1019 were plateaus. Among the 36’124 peak-shapes, 20’686 were cluster peaks and 15’438 were solo peaks.

3.3 Peak Annotation

3.3.1 RXR peaks prefer promoter regions, including transcription start sites

With the supposition raised by the phastCons analysis, we sought to investigate how RXR peaks loci distributed on different genomic features. We annotated the refined RXR peak set with Ensembl Mus musculus assembly (NCBI m37, UCSC mm9 database) based on the closest distance from the summit of each RXR peak to the nearest gene (**Fig.12A**). The annotated features were assigned following the hierarchy from top priority “Promoter” (-1.5kb upstream to 0.5kb downstream from TSS), “Intragenic” (on gene body, including “Exon” and “Intron”), “Upstream geneStart” (-1.5 to -25kb upstream TSS), “Downstream geneEnd” (from Polyadenylation site (PAS) site to 5kb downstream), to bottom priority "Intergenic" (remaining peaks). The TSS of each transcript was assigned with corresponding direction of transcript. We also expanded the method of ChIP-seq peak annotation by incorporating the locations of known enhancers of transcription units (ENCODE).

With this approach, 38% of the peaks fell within the Promoter. The same percentage of the peaks fell within gene bodies. About five hundred RXR peaks (2%) were found within 5kb after the end of transcript PAS. The remaining ~20% was equally shared between Upstream and Intergenic annotations (**Fig.12A – pie chart**). The annotation confirmed the differences between peak morphologies groups. The proportion of promoter remarkably increased from single to double to multiple peak group. The plateaus, due to their width, were annotated based on their highest identifiable summits, and bore the same distribution as single peaks (**Fig.12B**). This explained the increasing of conservation phastCons score from single to multiple peaks because the first intron within the promoter usually the most conserved regions (*Park et al., 2014*). The lower score of plateaus can also be explained due to their long ranges on the gene bodies included multiple

elements which are less conserved. Hence, even though the phastCons brought interesting observation for the peak morphologies, it did not serve any further informative and led us to the chicken-and-egg dilemma, and was thus not further used for subsequent analysis.

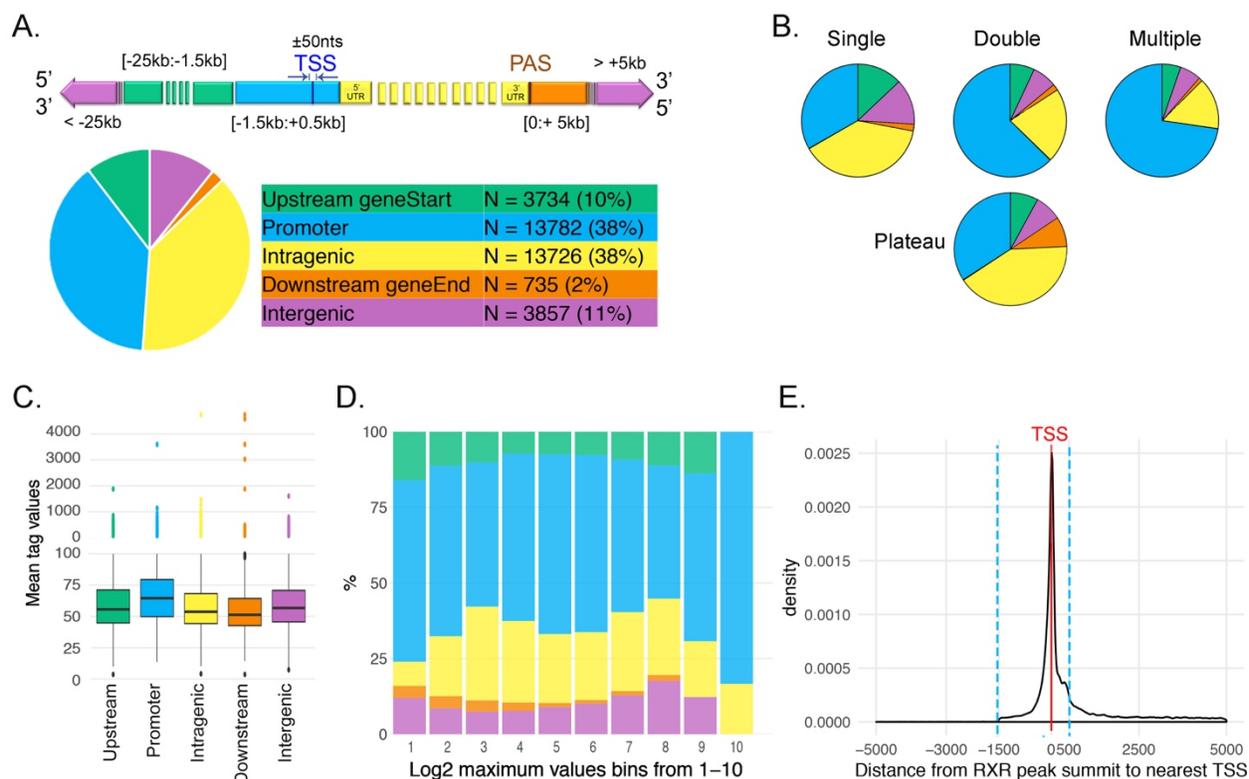


Figure 12: Peak annotation

- A.** RXR peaks were annotated to the genomic features, according to the relative distance from peak summit to the most proximal gene. Transcription unit (top) delimits the distance of each genomic feature to TSS and PAS (polyadenylation site). The hierarchy for peak annotation is: promoter, intragenic, upstream transcript start, downstream from transcript end, and intergenic. Pie chart (bottom left) with frequency table (bottom right) for each annotation
- B.** Pie charts of annotation for four groups of peak morphologies
- C.** Boxplots of average normalized RXR tag density values of each annotation groups
- D.** Fractions of annotation groups across bin of log₂ maximum tag density value
- E.** Distribution of RXR peaks to nearest transcription starting site (TSS)

The differences observed for mean occupancy across time point values between the various genomic annotation groups were not statistically significant (**Fig.12C**). The high binding proportion

of peaks in promoters was found as well in peaks with low RXR tag density as those with high RXR tag density (**Fig.12D**), and at significantly higher percentage (38%) compared to distributions proposed by other studies of RXR binding (*Nielsen et al., 2008, Martens et al., 2010, Zhan et al., 2012, Adhikary et al., 2015*).

Intriguingly, an important part of peaks localized in the promoter were indeed found at the TSS, based on the distribution of distances from RXR summit to nearest TSS – **Fig.12E**). By defining the TSS as the sequence from -50 to +50 bp, 38% of the promoter peaks and 15% of the total number of peaks were located on this particular functional gene region.

3.3.2 Relationship between peak annotation and gene expression levels

Since the binding of a transcription factor to the promoter of a gene usually results in a direct regulation effect, either activating or repressing the transcription of that gene, this preference of RXR to the promoter-rich regions may correlate with certain changes on the downstream genes expression. We can only validate these relationships with genes whose mRNA expression levels were measurable using microarray. Among the total 24'377 measurable genes by mRNA microarray, 15'432 genes were putatively regulated by RXR bindings (**Fig.13A**). There was no statistical difference of maximum time point expression levels in function of the genomic binding location of RXR (**Fig.13B**).

Nevertheless, for those genes bound by RXR at their promoters, the correlation between maximum RXR tag density and maximum expression was statistically higher than the correlations of those genes bound by RXR outside of their promoter. However, all correlation coefficients were too low to statistically validate the direct effect of RXR binding on the expression of genes (**Fig.13C**). Thus, RXR binding sites are found preferentially in promoter-rich regions. It must be

noted that the tag density value of RXR peaks has no statistical correlation with the expression levels of the corresponding gene.

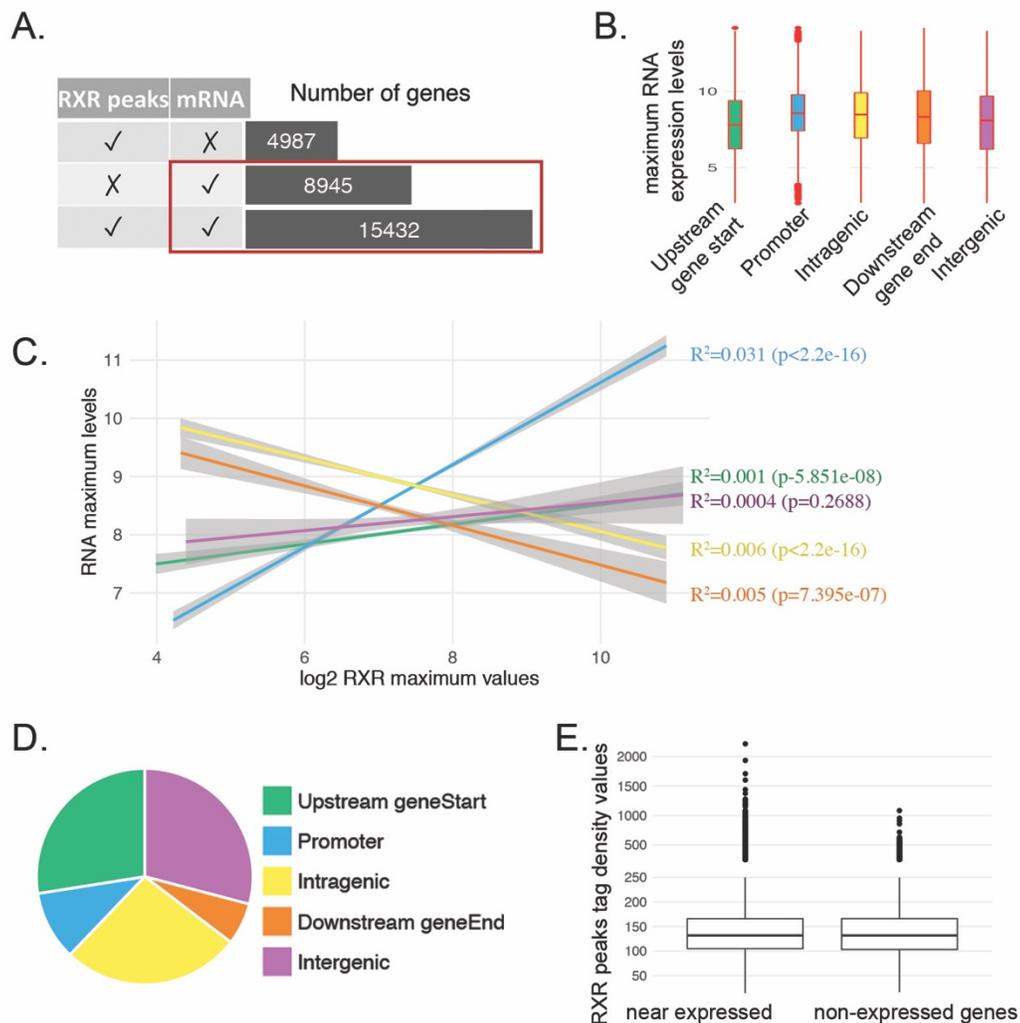


Figure 13: RXR target gene expression

- A.** Number of genes with measurable expression (mRNA) and/or harboring RXR binding loci
- B.** Boxplots of maximum expression levels of genes with RXR peaks from each annotation groups
- C.** Correlation between log₂ maximum tag density of RXR peaks with the maximum expression level of the target genes, separated by annotation groups.
- D.** Pie chart of RXR peaks annotation for the 4987 non-expressed genes from panel A
- E.** Boxplot of RXR tag densities between peaks located near expressed versus non-expressed genes

4987 genes to which we associated with RXR peak were not detected in our mRNA microarray at any time point (**Fig.13A**). However, for these genes, the fractions of our five

annotation groups remarkably change to favor the Upstream and Intergenic compartments (**Fig.13D**) instead of Promoter and intragenic compartments (**Fig.12A**). There was no difference in tag density between RXR peak located near expressed genes and near non-expressed genes (**Fig.13E**).

3.3.3 Clustered RXR peaks and gene expression

Clustered peaks— described in **Fig.11B** as peaks with neighboring peaks within ± 1000 nts window— were found more frequently in promoters than solo peaks, whereas solo peaks were more enriched in intragenic regions (**Fig.14A,B**). We theorized that the clustered bindings found in our study are part of the HOT— “high occupancy target”—regions, or “hotspot” reported by previous transcription factor binding studies performed on various cell types (in *D. Melanogaster* fly (*Moorman et al., 2006*), mouse (*Joshi, 2014, Siersbæk et al., 2014a, Siersbæk et al., 2014b*), and human (*Wang et al., 2012, Xie et al., 2013*) cells).

Such dominant feature of binding called for further investigation, so we analyzed various characteristics comparing cluster and solo peak groups. Taking all genomic regions together there was no statistical difference in RXR tag density distribution between solo and cluster peaks (**Fig.14C**). However, for promoter regions as defined in **Fig.12A**, those containing cluster peaks presented a significantly higher Pol-II tag density than those bound by Solo peaks (**Fig.14D**).

The maximum mRNA microarray levels of transcripts bound by the clustered RXR regions are also significantly higher than transcripts bound by solo RXR bindings, and much higher than those without RXR bindings (**Fig.14E**). Therefore, the density of clustered RXR peaks in a promoter might better reflect the overall transcription activity than the higher tag accumulation of RXR at a given site.

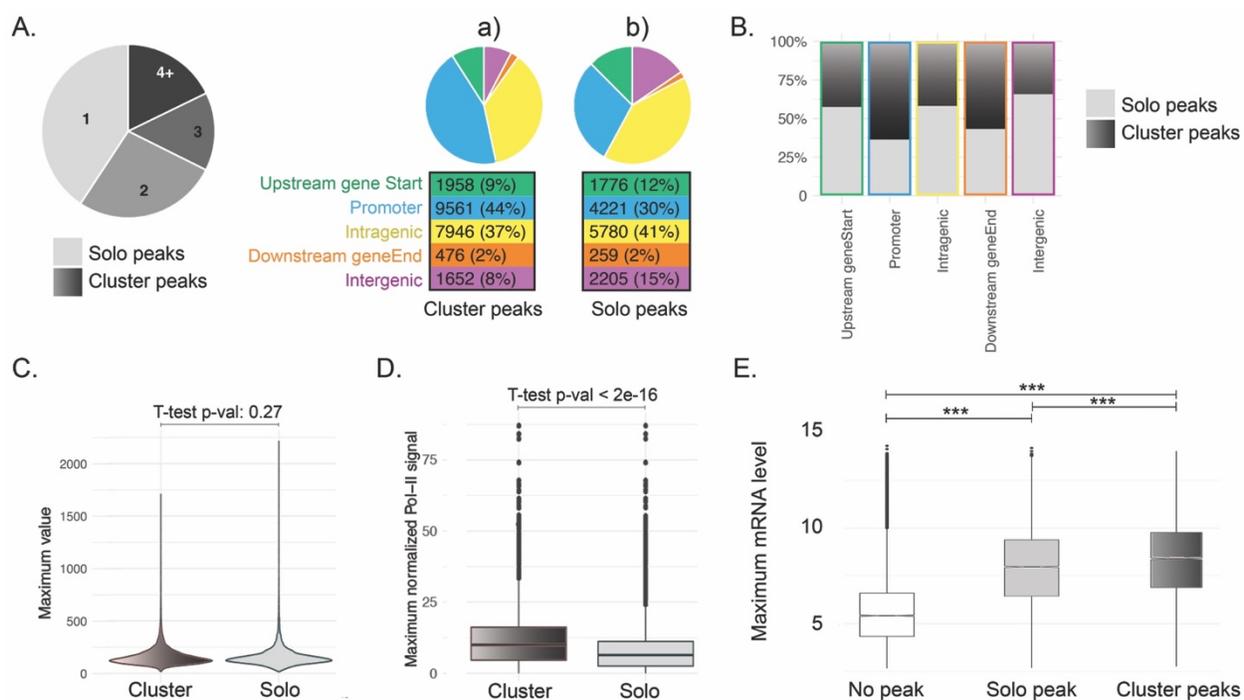


Figure 14: Peak co-localization - The clustering

- A.** ~50% of RXR peaks have no neighbor within a window of 2000 nucleotides from their RXR summit – the “solo”. The other half cluster together in group of 2 or more peaks.
a) Pie chart with numbers on wedges indicate the number of peaks co-localizing together.
b) Pie charts of annotation proportions separated by cluster versus solo peaks
B. Fractions of cluster peaks versus solo peaks across annotation groups
C. Distribution of maximum RXR normalized tag density of cluster and solo peaks
D. Distribution of maximum normalized Pol-II signal at the promoter of genes that harbored cluster and solo peaks
E. Boxplots showing the difference of mRNA level distribution between genes with cluster, solo, and no RXR peak. *** one way ANOVA p value < 2e-16

3.4 The plateaus reveal a promiscuous relationship between RXR and Pol-II

3.4.1 The plateau peaks cover entire genes of small size

Different from the majority of peaks present a typical transcription factor binding peak shape with a single summit and two relatively symmetric slopes (examples pointed by yellow

arrowheads in [Fig.15A.a](#)), the plateaus covered long continuous regions (from few thousands nucleotides up to 20 kb long) with continuously high signal. Within these plateaus, we observed that a small fraction of RXR peaks (309 equals to ~1%) that have multiple exceedingly-high-tag-density summits fused together, we referred to as “High plateau” ([Fig.15A.b](#)). The rest of plateaus shared the lengthy continuous signal pattern, but with lower tag density, were referred to as “Low plateau”. Further inspection of low plateaus revealed that they were often preceded by high peaks. This profile led us to hypothesize that the low plateaus were trailing RXR signal from the original bindings at the high peaks during transcription, and hence not real binding sites. We thus, removed the low plateaus subset from our analysis.

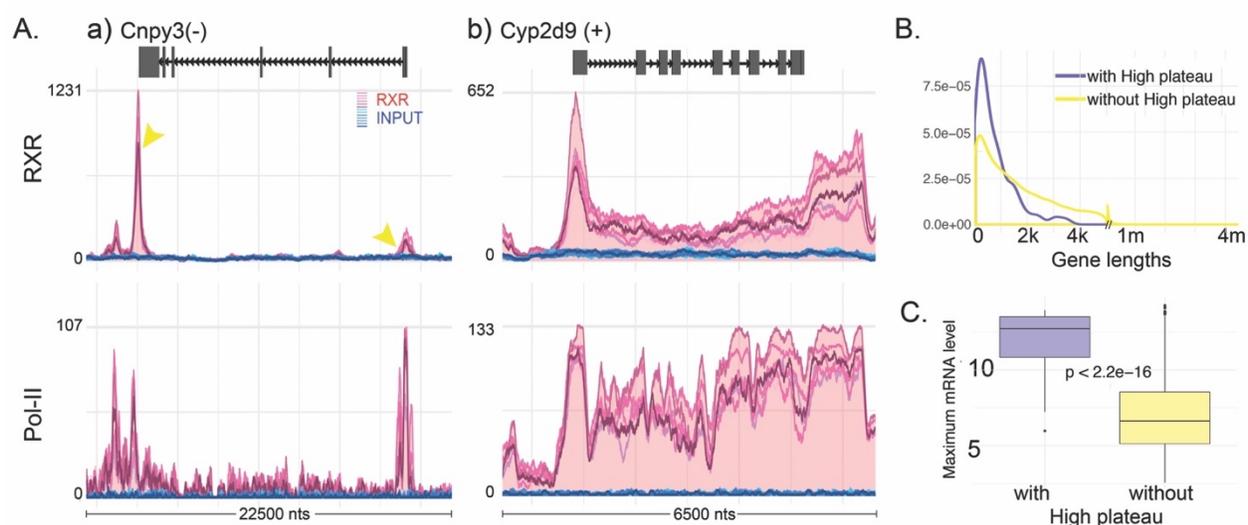


Figure 15: The Plateau

- A.** Genomic view of RXR and RNA polymerase-II ChIP-seq reads mapping to 3 genes:
 a) *Cnpy3* gene with RXR regular peaks (yellow arrow heads) at TSS and PAS
 b) *ApoA1* gene harboring RXR High Plateau signal along the whole gene, from upstream TSS to downstream PAS
- B.** Colored areas are projected maximal signal at each position. Red/blue lines are temporal profiles of ChIP-seq/Input.
- C.** Gene lengths distribution for genes encompassing High plateau (purple), compared to other genes (yellow)
- D.** Boxplot showing the maximum mRNA level of genes encompassing High plateau (purple) compared to others (yellow)

Approximately 50 of the High plateau regions appeared to encompass whole genes (hence termed “High Plateau genes”, example [Fig.15A.b](#)), which were mostly short genes with less than 20kb in size ([Fig.15B](#)). Visual inspection of RXR binding of these genes revealed a special pattern: RXR binding presented a very high density at the promoter, was lower on the gene body and gradually increased to form a second high-density site toward the end of transcript, reminiscent of the Pol-II signal displayed for these genes ([Fig.15A.b](#)).

3.4.2 Colocalization of RNA Polymerase II and RXR on High plateaus and at transcription starting sites

Pol-II immunoprecipitation had been performed along the exact same experimental protocol and time points, in the frame of the CycliX project. We thus could retrieve the corresponding data and analyze the behavior of Pol-II parallel to that of RXR. This allows to uncover a remarkable alignment of RXR and Pol-II signal on the Plateaus. The high percentage of RXR peaks clustered in promoter regions suggested that RXR could be involved in the transcription machinery. Upon inspecting all promoter regions from the mm9 genome, we found a linear correlation between RXR and Pol-II signal, regardless of the occurrence of discernable RXR peaks at these promoters ([Fig.16A](#) – blue and green line). This was true not only for the plateau but across many of RXR peaks. We can only find a few cases where Pol-II signal at the promoters was high without RXR signal ([Fig.16A](#)-red circle), an example of them in [Fig.16B](#). Such occurrences [Fig.16B](#) eliminated the possibility of error cross-link effect from immunoprecipitation step, and suggested that the correlation between the Pol-II signal and the RXR signal was the result of RXR recruiting and then somehow attaching to the transcription machinery. This correlation might definitely affect further analyses, particularly with respect to the motif analyses. Hence, we sought to gate the “Pol-II affected” RXR peaks out. The distribution of log₂ ratio of RXR signal over Pol-II signal within all

RXR peak loci versus only RXR at TSS loci identify a crossing point at the log₂ ratio of 1.4 (Fig.16C).

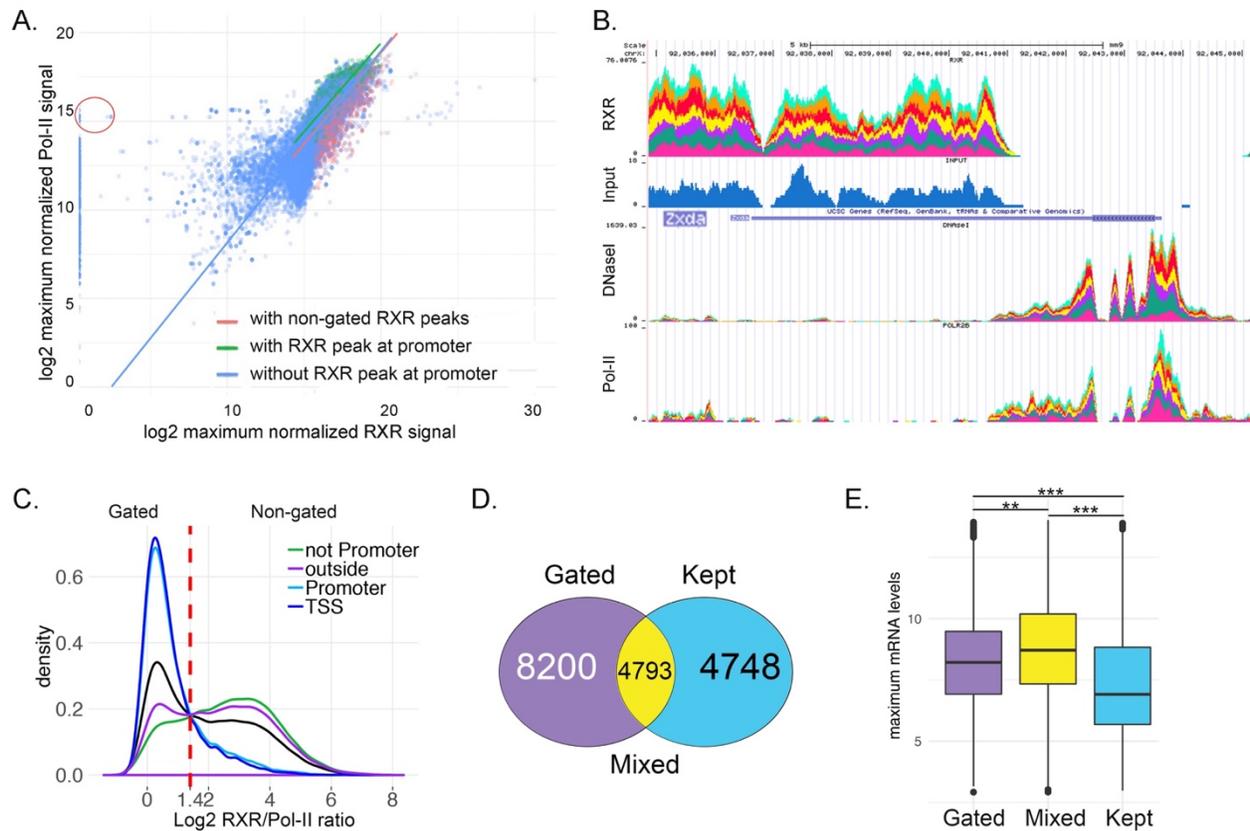


Figure 16: The correlation with Pol-II

- A.** Correlation plot between RXR signal and Pol-II signal within all promoter regions ± 1 kb around TSS, colored by the occurrences of significant RXR peaks at these promoters: without RXR peak - blue, with RXR peak - green, with kept RXR peaks from panel C - red. Red circle indicates a few cases where promoter regions have high Pol-II signal without RXR signal.
- B.** Genome view of one example gene, whose promoter region is enriched with Pol-II and DNaseI signal but devoid of RXR signal
- C.** Distribution of log₂ RXR/Pol-II ratio, considering all the peaks (black line), only peaks in promoters (light blue) or not in promoter (green), or separated by whether the RXR peak is located at the TSS (dark blue) or not (purple). Vertical red dash line shows the coordinate of the crossing point of these curves at a log₂ ratio of 1.42, which was used for gating
- D.** Number of genes within analysis that were gated or included, the mixed group were genes harbor both gated and kept RXR peaks
- E.** Boxplot showing different levels of expression for genes depending on the presence or absence of RXR binding, and whether RXR binding is accompanied by high Pol-II signal.

*** one-way ANOVA p-value < 0.05, ** p-value < 0.1

We took this crossing point value as threshold for gating:

- Gated:
 - RXR peaks with high Pol-II signal (hence, RXR/Pol-II ratio lower than threshold)
 - RXR peaks located exactly ± 50 nts at the TSSs
- Kept: the rest of RXR peaks

This gating did not affect the correlation between RXR and Pol-II (*Fig.16A*-red line). Upon inspecting which RXR target genes were filtered out by our gating, we found that ~8200 genes were eliminated because of the gating, 4748 were not affected by the gating, and 4793 genes—for which one or some gated peaks were removed—were kept because they also contained a validated RXR binding site associated with them somewhere in the gene transcript (*Fig.16D*). Interestingly, the expression of these mixed genes was significantly higher than others (*Fig.16E*). This was in line with a previous study (*Kosters et al., 2013*) that showed significantly higher gene expression of those genes where RXR overlapped with Pol-II signal. *Figure 17* showed an example of the mixed genes on which RXR clustered with different ratios over Pol-II signal: while the “j” peak was gated by the high signal of Pol-II and its overlapping with TSS, there are 7 other validated peaks on the intragenic region of this *PXR* gene. Interestingly, the intragenic peak “a” was located at a previously reported locus containing the IR1 response element of FXR, and able to trigger strong induction after GW4064 – an FXR ligand – treatment in transactivation experiments (*Jung et al., 2006*).

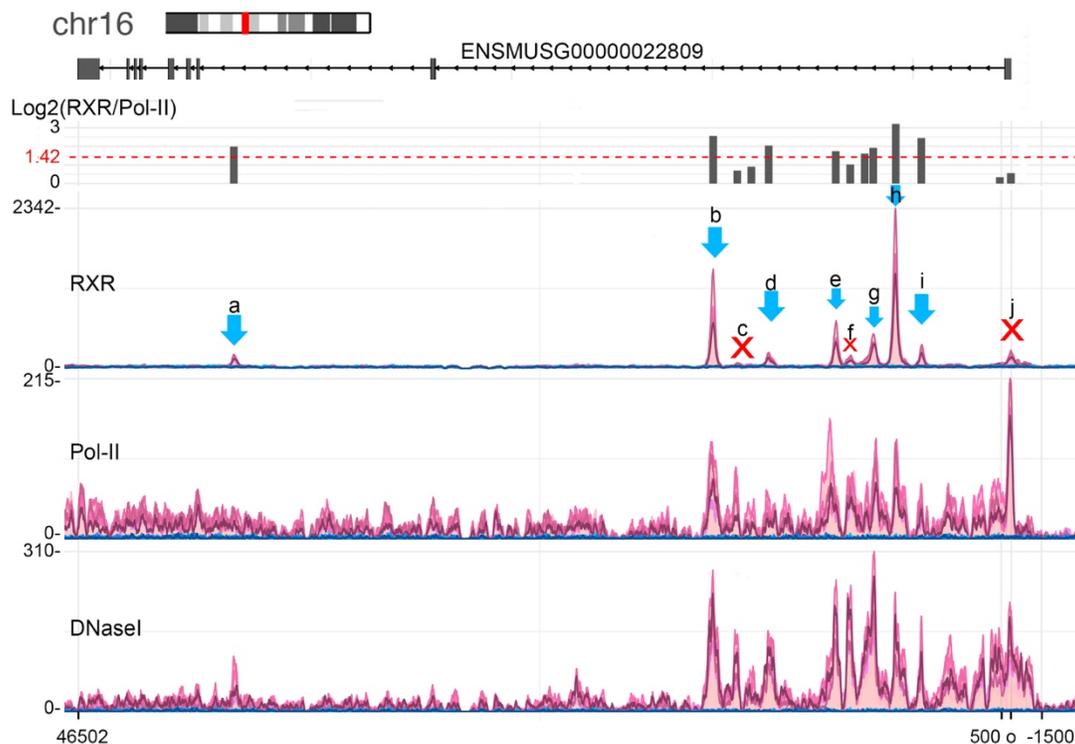


Figure 17: Genome view of *Pxr* gene

An example of mixed genes that harbor both validated peaks (blue arrows) and gated peaks (red crosses) based on the cut-off of ratio. Because of the high tag density of peak h, some other peaks (a, c, d, f, i, j) seemed small but they are significant peaks.

3.5 Circadian analyses

3.5.1 Clustering RXR peaks using peaks tag density values

With the peak-set well refined, our interest focused on dividing RXR peaks in groups of similar behavior. However, the RXR binding poses quite unique challenges. We did not observe any group where the RXR signal enrichment was exclusive of certain time point or certain phase (dark/light-feeding/fasting). Not only could we observe enriched RXR at their chromatin loci at each of the 7 time points probed, but also the change in intensity varied widely between the time points at some sites and very little at others.

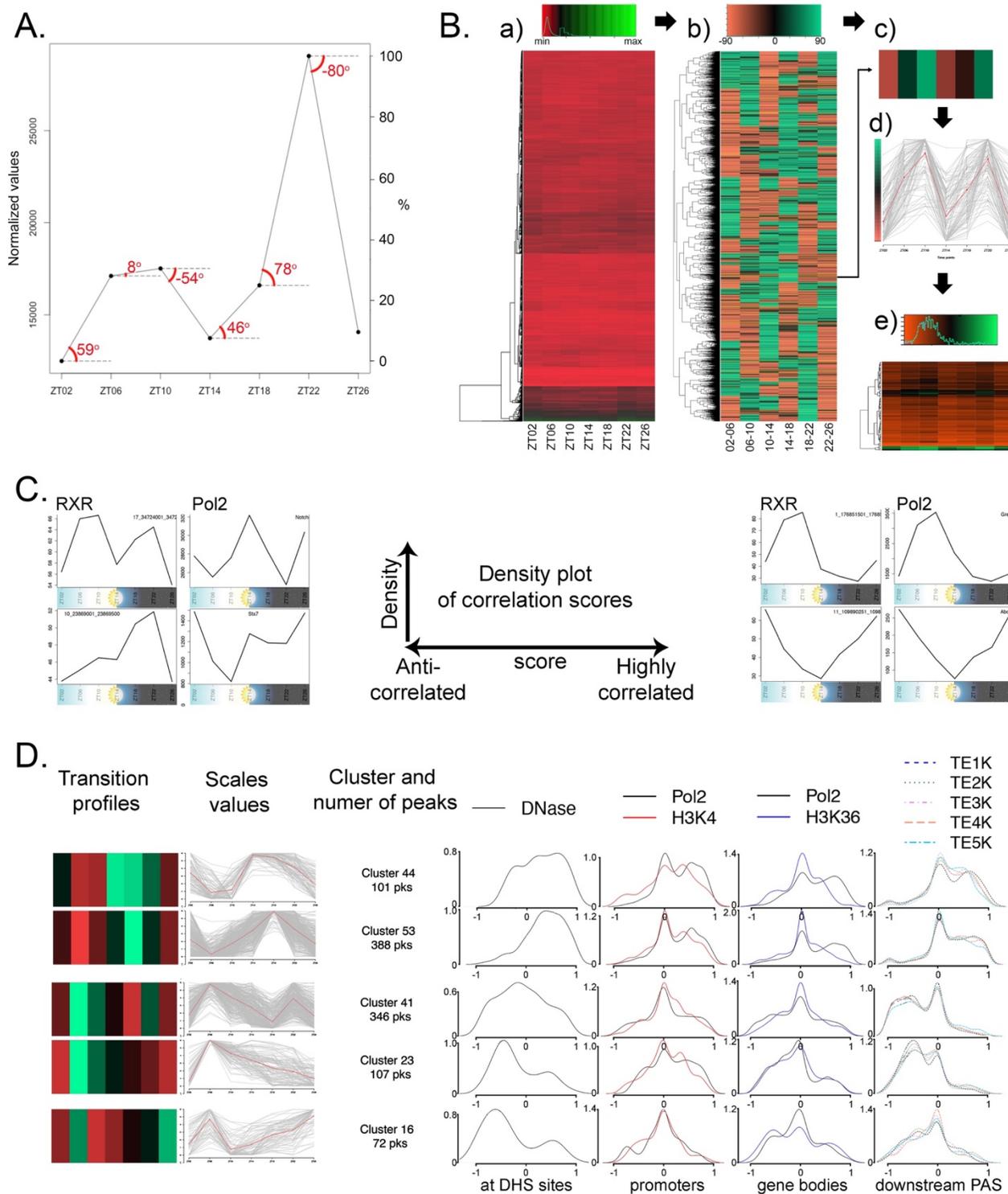


Figure 18: Grouping peaks by tag density values

A. Transition profile: we scaled the 7 time point normalized tag density values to the maximum and minimum of each peak, the transition profile is the 6 angles of two consecutive time points made with the horizontal plane

In order to retain and compare the phase information of our peaks, we normalized the tag density values of each peak at each time point to the total tags of sample and peak width, and rescaled the normalized values to the minimum and maximum tag density across the seven time points (**Fig.8D**). Using the seven time points scaled values, we can calculate the angles created by the regression line between 2 successive time points with the horizontal line (**Fig.18A**). We thus transformed the 7 time point values into 6 transition angle values – which we called the transition profile. Each angle reflects the direction of change between 2 consecutive time points, and also the degree of such change. We then used the transition profiles of RXR peaks as input for complete

-
- B.** *Effect of transition profile on how peaks cluster together in heat-map:*
- a) *The original heat-map by 7 time point normalized tag density values shows no clear clustering*
 - b) *Heat-map on the 6 angles shows better clustering and the dendrogram was used to divide peaks to groups*
 - c) *Example of a group from heat-map b shows time-wise profile displayed as color scheme*
 - d) *Scaled plot for the time-wise profile of example group in c, mean values each time point displayed in red line*
 - e) *Heat-map of non-scaled values from all the peaks in example group in c: values differ greatly between peaks and so subtly between time points that we see 3 sub-groups of low tag density (red), medium density (black) and high density (green)*
- C.** *Correlation plot explained: with anti-correlated examples on the left, highly correlated examples on the right*
- D.** *Examples of five representative groups selected based on transition profiles and their correlation profiles. From left to right for each row:*
- *the transition profile (as B-c)*
 - *the scaled values of 7 time points of all peak within groups*
 - *number of peak within each group*
 - *density plots showing the correlation score between RXR peaks present in each group with four corresponding epigenetic marks that overlapped with RXR: DNase*
 - *correlation score distribution for Pol-II at promoter*
 - *correlation score distribution for H3K4me3 at the promoter regions*
 - *correlation score distribution for Pol-II on gene bodies*
 - *correlation score distribution for H3K36me3 on gene bodies*
 - *correlation score distribution for Pol-II gradually downstream from 1kb to 5kb from PAS and at gene end*

linkage clustering using Euclidian distance. (*Fig18B.a-b*). We used a cutree threshold to separate peaks into 15-20 groups. For each peak group, even though the transition profile of each peak within one group was similar to each other, as imposed by the clustering (example 1 group in *Fig.18B.c-d*), the normalized 7 time point values showed a wide range of binding signal (*Fig18B.e*). We varied the cutree threshold to increase the number of groups, but this did not bring more detail nor reduced the wide spread of values. Thus, by using this clustering approach which centered the groups around the mean values for each variable (the mean tag density of 7 time points in our case), we achieved the peak separation at the cost of real tag density amplitude.

Many studies have suggested that RXR binding has a stable nature, and is not time-dependent (*Simicevic et al., 2013, Boergesen et al., 2012, Minucci et al., 1997b, Brazda et al., 2014, Shen et al., 2011b, Tzameli et al., 2003, Mangelsdorf and Evans, 1995*). The regulatory effect of RXR on transcription may not be reflected clearly by tag density changes, but rather by the changes on epigenetic signal landscape surrounding RXR bindings. Thus, we theorized that certain motifs would appear within groups of RXR peaks having similar matching DHSs, Pol-II and or histone marks profiles, or that those groups may regulate genes involved in specific pathways. To test this, we retrieved data for Pol-II, H3K36me and H3K4me obtained on the same 7 timepoints during a previous Cyclix study (*Le Martelot et al., 2012*). For all Ensembl transcripts of mm9 database, the level of transcription was evaluated by the Pol-II and H3K4me signal within ± 1 kb around TSS, ongoing transcription activities were calculated by Pol-II and H3K36me signal on the transcript bodies, and the amount of Pol-II signal at the end of transcripts (TE1K-TE5K) indicated the recently transcribed genes. (DHS sites described in section [3.2.2](#)).

We calculated the cross correlation between each RXR peak 7 time point values/6angles with the corresponding DHS, Pol-II and histone marks 7 time point values/6 angles, and used the

correlation (*Fig.18C*) to separate groups of RXR peaks presenting the same time-series behavior (*Fig.18D*). We found that the highest positive correlation is observed between RXR and Pol-II, especially at the promoter. This can be partially explained by the high proportion of RXR peaks residing at the promoters (observed in *Fig.12A*) and especially at the TSS (*Fig.12D*). However, in general most results are dominated by a correlation centered around zero indicating randomness. We encountered the same degree of dispersion using alternatively Pol-II, H3K36m3 and H3K4me3 dataset as the master-cluster.

3.5.2 *The sinusoidal fit*

Since one of the focus of our study is on circadian cycle, we hypothesized that each RXR peak was bound with maximum intensity at one certain time of day, and each RXR partner may have a dominating effect at one certain time point of the circadian cycle. We postulated that this could result in an observable segregation of binding motifs. Because of the low resolution of 7 time points per day and the lack of time point replicates of our data sets, it was unadvised to use directly the 7 values profiles for peak set clustering.

To retrieve the phase of each peak, we applied the cosinor function (R package psych), to fit the 7 time point values time series of each RXR peaks - also of the DHSs, Pol II and other histone marks – to one sinusoidal regression function, with alternative period of 12 and 24 hours. This function transformed the 7 time point values of each peak into several curve's parameters: the phase - which is the hour when tag value is highest within the period; the mean and standard deviation of signal amplitude of the curve; and the fit score - which is the correlation between observed and sinusoidal fitted values, ranged from 0 (non-fitted) approaching to 1 (well fitted). Depending on the best fit score, the final period assigned for each peak was determined either 12 or 24 hours (*Fig.19A*).

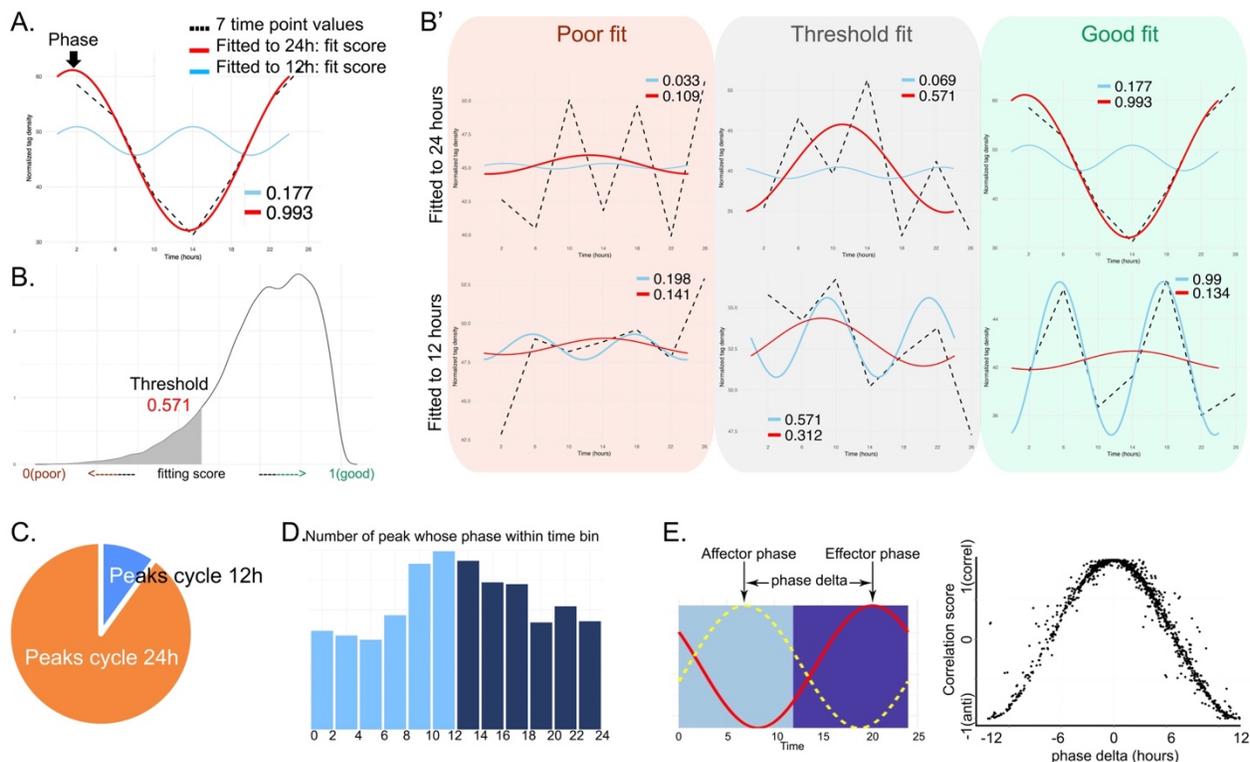


Figure 19: Sinusoidal fitting of 7 time point tag density values

A. Example plot of how 7 time point values fitted to sinusoidal curve of 24h period (red) and of 12h period (blue) and the fit score better for 24h.

Arrow points at the curve phase - where signal is highest

B. Distribution of fit score of all peaks.

B'. Panel of examples showing how data obviously fits well (green) or poorly (orange) to sinusoidal curve of 24h or 12h. The “grey” shows cases where data can or cannot be fitted to a sinusoidal curve depending on how stringent we want the fit to be

C. Pie chart showing the fraction of RXR peaks fitted to 24h or 12h cycling, based on the fit score threshold retained

D. Histogram of number of RXR peaks' phases, binned into 2 hours bins

E. Left: illustration of how to calculate phase delta between 2 curves

Right: distribution of correlation scores across the distribution of phase deltas

The majority of RXR peaks fitted well to either period with a fitting score distributed more toward well-fitted value 1 (**Fig.19B**). The fit score of 0.571 indicated 3-4 time points from our 7 time point values fitted to the curve, which can be consecutive or segmented. Some examples in **Fig.19B'** show examples of poorly fitted peaks (below threshold), just retained peaks, e.g. at the

threshold, where only 3-4 data points fit well, either inconsecutive (top grey panel: ZT02,06,14,22,26) or consecutive (bottom grey panel: ZT10,14,18,22) and well fitted peaks at a 24 hour period (top) or 12 hours period (bottom). Using the 0.571 fit score threshold, 90% of RXR peaks were considered cycling, ~10-15% of which were on 12 hours period, the rest were cycling on a 24 hours period (*Fig.19C*).

The resulting phase-assigned peaks were binned into 2 hours intervals, which showed higher population of RXR peaks that reached their highest intensity at mid-day (*Fig.19D*). Since all of the data sets were transformed into phases grouped by the sinusoidal fitting, we were able to use the phase-delay between RXR phase and their regulatory effect markers phases to split RXR peaks into groups presenting the same regulatory effect (*Fig.19E*). We tried several groupings, from the simplest correlation/anti-correlation to the more concise degrees of delayed effects between RXR and other signal, (RXR – Pol-II at TSS – Pol-II in gene body - microarray)) (illustrated by *Fig.20*).

Though this method gave us the better harmonic grouping of RXR and its regulatory effect, it also hindered us accessing the amplitude of the effect within each group. This grouping effect also gave a dispersion pattern like the previous clustering by heat-map method, although more subtly because of the phase-delay gates we imposed. We speculated that each known regulated cellular function of RXR would manifest on certain time point of the circadian landscape, therefore the different groups of RXR would be enriched for different pathways, which we did not observe. However, our analysis scheme only grouped peaks based on the time of their highest intensity, not taking into account the real oscillating signal amplitudes. We reasoned this might be one of the reasons our pathways enrichment analysis did not yielded any distinct pattern.

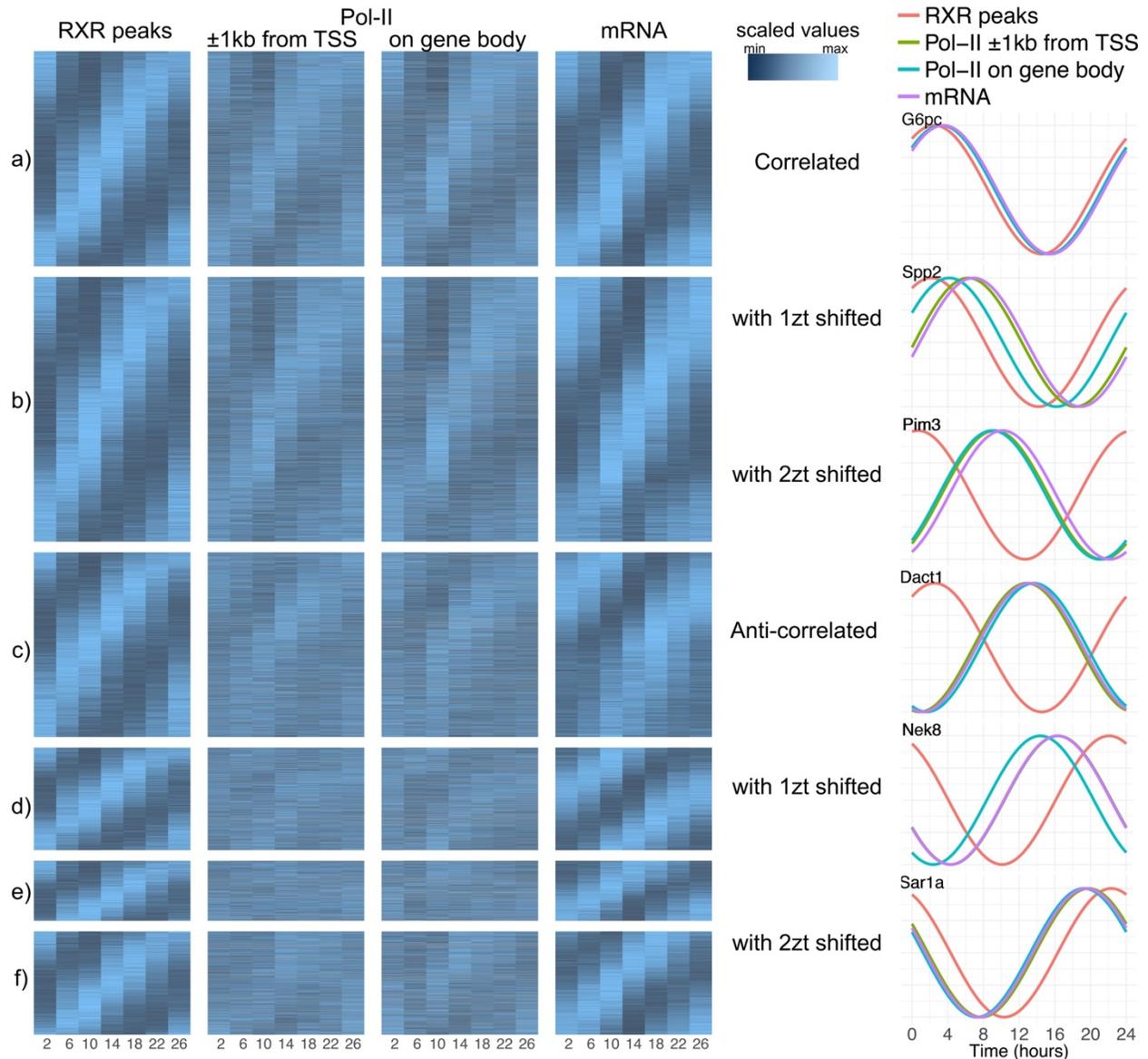


Figure 20: Correlation groups

Heatmaps (left panel) of peaks for which RXR tag density and mRNA expression are both fitted to 24 hours circadian period. Each row shows the oscillation of RXR binding with their target genes modulations for 3 epigenetic and genetic signals: the intensity of DNA polymerase II bindings within $\pm 1\text{kb}$ around the TSS, and on gene body (respectively); and the mRNA expression. We grouped them into 6 regulatory effects:

- The correlated (top) group includes RXR peaks oscillation are well synchronized with their targeted genes;
- the correlated with 1zt shifted (second heatmap) includes target genes that oscillate approximately 4 hours later than their RXR peaks;

We tackled the amplitude obstacle by applying statistical tests on the oscillating amplitude of the whole peak set to retain only those with significantly cycling signal. The fit score from cosinor also alerted us to several cases of fitting where neither 24 hours nor 12 hours period fitted well to the data although the data did display rhythmic pattern. Hence, the restriction of period using only to 24 or 12 hours seemed too stringent. Thus, we used the MetaCycle package (*Wu et al., 2016*), whose combination of time-domain and frequency-domain provided us with flexibility in periodic analysis while introducing a more stringent F-test probability. The MetaCycle incorporates ARSER (ARS) (*Yang and Su, 2010*), JTK_CYCLE (JTK) (*Hughes et al., 2010*), and Lomb-Scargle (LS) (*Glynn et al., 2005*) properly for periodic signal detection.

Based on our study time scheme, we used the suggested function meta2d (*Fig.21A*). We set the periodic prediction with lower limit at 8 hours and upper limit at 26 hours based on prior observation from other studies (*Veen and Gerkema, 2017, Castellana et al., 2018, Houben et al., 2014, Krishnaiah et al., 2017*). JTK reset upper limit as 24 for its own algorithm. ARS produced multiple fits for some of our peaks, which resulted in NULL p-values. For each of these events, we applied a hierarchical selection for the period which produced fitted values with highest match amplitude, highest correlation compared to the true values, and the least distance from the periods

-
- c) the correlated with 2zt shifted (third heatmap) includes target genes that oscillate approximately 8 hours later;*
 - d) the Anti-correlated (fourth heatmap) show RXR peaks are anti-synchronized with their targeted genes;*
 - e) the Anti-correlated with 1zt shifted (fifth heatmap) includes target genes that reverse-oscillate approximately 4 hours later than their RXR peaks;*
 - f) the Anti-correlated with 2zt shifted (bottom heatmap) includes target genes that reverse-oscillate approximately 8 hours later.*

The curve plots (right panel) show examples of the above 6 correlation groups, using the fitted sinusoid curves for RXR peaks tag density, Pol-II intensity and mRNA expression, showing the shifting/delaying of oscillation phases of the corresponding genes compared to the RXR peaks.

produced by JTK and LS. In truth, integrated p-values for these events had lower degree of freedom, hence were less creditable. However, this permitted the discovery of a small group of RXR peaks fitting to rhythms shorter than 12 hours (around 8 hours), while retaining the classical period range of 12h and 24h ([Fig.21B](#)). The integrated p-values and Benjamini-Hochberg (BH) adjusted q-values generated (FDR) by MetaCycle only partially indicated the rhythmic of RXR signal and unequally reflected the degree of fit across the period range. Upon detailed inspection, q-values up to ~ 0.6 still produced nicely fitted sinusoid curves for 12 and 24 hours period rhythmic RXR signal ([Fig.21C,D](#)). For shorter periods, however, q-values stringently correlated with the degree of fit ([Fig.21E,F](#)), but even with the most significant q-value the amplitude of signal oscillation were not great. This was one of the short-coming of having low time resolution and no (biological) replicate data, because the prediction of rhythmic period become unreliable. Thus, although having wide range of period prediction brought new perspective for circadian analysis, the significances by FDR produced by this method seemed to be not suitable for our data set.

We decided next to test the significance of signal oscillation without imposing a periodic fit. For time course data, we used the recommended piecewise-cubic splines function of the *limma* package, although the advantages of applying splines were not up to their potential due to the lack of replicate and low time point resolution of our RXR data set. With the FDR of 0.05, about 10% of our dataset presented a significant oscillation, matching the percentage reported for previous circadian studies ([Akhtar et al., 2002](#), [Bozek et al., 2010](#), [Vollmers et al., 2009](#)). For this subset of data, we retrieved back the sinusoidal fits and the period fits from previous analyses for the downstream functional analysis.

A. **P-value:** Fisher's method integrating multiple p-values into one test statistic (X2)

$$X_{2k}^2 \sim -2 \sum_{i=1}^k \ln(p_i)$$

Period: arithmetic mean value of multiple periods

Phase integration based on mean of circular quantities:

- convert phase values to polar coordinates α_j
- convert polar coordinates to cartesian coordinates $(\cos\alpha_j, \sin\alpha_j)$
- compute the arithmetic mean of these points and its corresponding polar coordinate $\bar{\alpha}$

$$\bar{\alpha} = \text{atan2}\left(\frac{\sum_{j=1}^n \sin\alpha_j}{n}, \frac{\sum_{j=1}^n \cos\alpha_j}{n}\right)$$

- convert the resulting polar coordinate to a integrated phase value

Amplitude:

$$Y_i = B + TRE * (t_i - \frac{\sum_{i=1}^n t_i}{n}) + A * \cos(2 * \pi * \frac{t_i - PHA}{PER})$$

- Y_i : observed value at time t_i
- B : baseline level of the time-series profile
- TRE : trend level of the time-series profile
- A : amplitude of the waveform
- PER : integrated period
- PHA : integrated phase

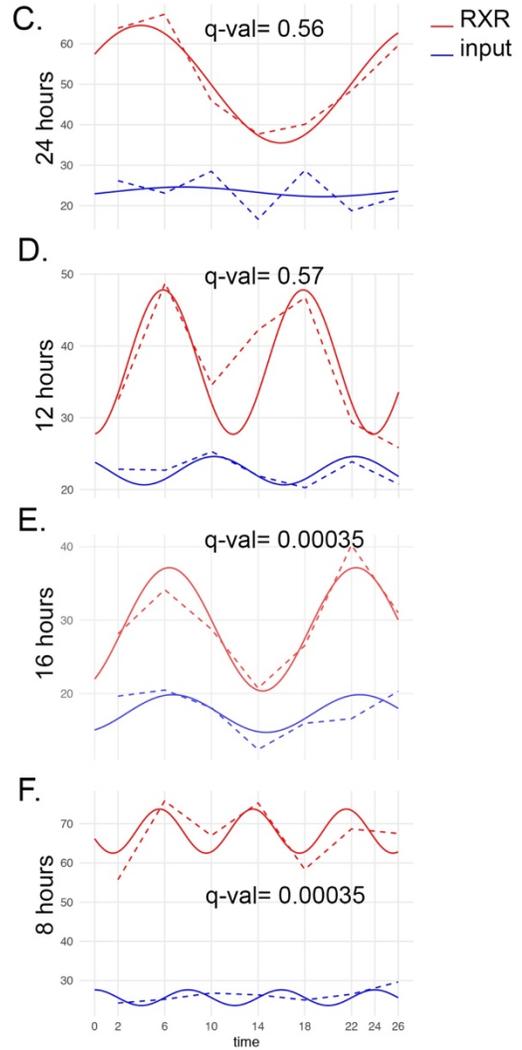
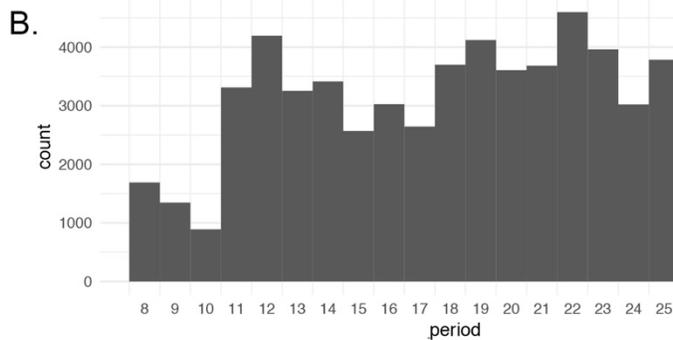


Figure 21: MetaCycle analysis

- A. Integrating methods of MetaCycle
- B. Histogram of periods produced by meta2d function
- C. Example case of RXR tag density well fitted to 24 hours period with q-value ~0.6
- D. Example case of RXR tag density well fitted to 12 hours period with q-value ~0.6
- E. Example case of RXR tag density well fitted to 16 hours period with very stringent q-value
- F. Example case of RXR tag density well fitted to 8 hours period with very stringent q-value

3.5.3 Circadian functional analyses

About 10% of our RXR peaks were considered as cycling. Among them, 3517 presented a circadian periodicity and 532 cycled with a periodicity of about 12 or less hours (the ultradian) (**Fig.22A**). The distributions of their phases are presented in circular histograms in **Fig.22B**. Interestingly, we observe different phases distributions for the gated (e.g. RXR signal with concomitant Pol-II signal, see section 3.4.2) and kept RXR peaks (RXR peaks devoid of Pol-II signal). The kept circadian RXR peaks (**Fig.22B.a**) are mainly observed around the transition of light-dark ZT0, which may suggest the influence of the core clock factors, namely the ZT22-ZT01 by the CLOCK-BMAL1 pair. The gated circadian RXR peaks (**Fig.22B.b**), on the other hand, are observed mainly before or at the beginning of feeding-fasting period, when the liver transcriptions were reportedly more abundant (*Takahashi, 2017*). The ultradian peaks mainly phase around ZT01-ZT02 with an additional population at ZT06 and ZT11 (**Fig.22B.c-d**) that may reflect the second (12h rhythms) and even third (8h rhythm) signal peaking during the 24h course.

The RXR peak annotations for these rhythmic peaks were enriched for promoter-dominant for the gated and intragenic for the kept (**Fig.22C**). Since the distinction of phase distribution between gated and kept RXR peaks was so clear, and that a large proportion of gated peaks was located at promoter, and that gated peaks are, by definition, jointly occupied by RXR and Pol-II, we investigated the rhythmic signal of Pol-II at the promoters of all transcripts. With the same circadian analysis, more than 4000 promoters had significant rhythmic Pol-II signal, whose phases distribution for circadian and ultradian rhythms are presented in **Fig.22D**. There were two populations of phases present for both circadian and ultradian groups: around ZT08 and ZT18 for circadian, ZT02 and ZT6 for ultradian. By splitting these promoters into groups with or without

RXR peaks, we found that promoters devoid of RXR peaks had a Pol-II signal peaking at ZT08 for circadian, ([Fig.22D.c](#)) and ZT1 and ZT5 for ultradian ([Fig.22D.g](#)).

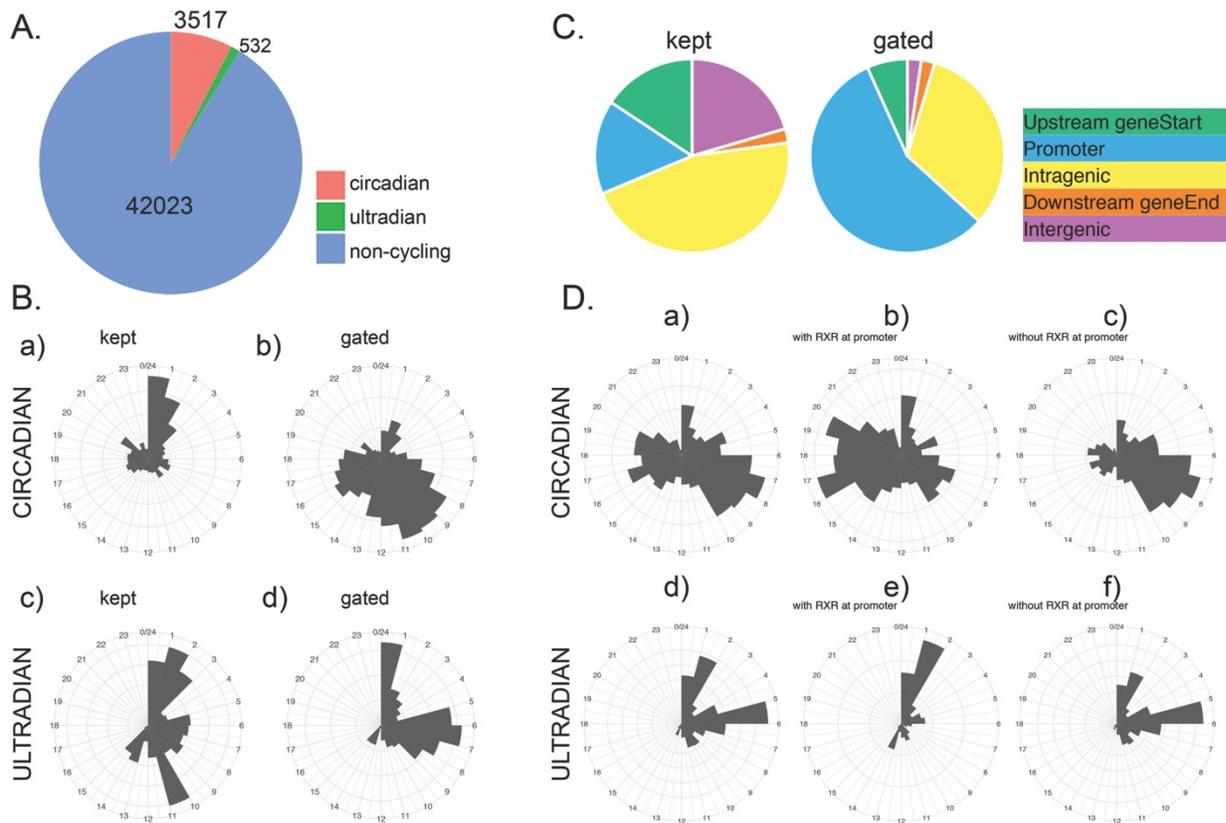


Figure 22: RXR circadian analysis

- A.** Pie chart showing the proportions of RXR peaks cycling in 24-hours (circadian) period, in less than 24h period (ultradian) and of those with constant density (non-cycling).
- B.** Circle histograms of phases distributions for circadian (a,b) and ultradian (c,d) kept peaks (a,c) and gated peaks (b,d)
- C.** Annotation distribution of cycling RXR peaks
- D.** Circle histograms of phases distribution for circadian (top) and ultradian (bottom) Pol-II signal at the 2262 promoters:
- a) the total genes with circadian Pol-II at promoter
 - b) genes with circadian Pol-II with RXR peaks located at promoter
 - c) genes with circadian Pol-II without RXR peaks located at promoter
 - d) the total genes with ultradian Pol-II at promoter
 - e) genes with ultradian Pol-II with RXR peaks located at promoter
 - f) genes with ultradian Pol-II without RXR peaks located at promoter

These distributions interestingly coincided (with 1-2 hours anticipation) with the distributions of gated RXR peaks (**Fig.22B.b,d**). The promoters with RXR peaks, on the other hand, had Pol-II signal peaking toward the end of day, around ZT19-20 for circadian and at beginning of day ZT01 for ultradian. These two patterns of Pol-II, segregated by the presence or absence of RXR peaks at the promoters, are astonishingly very similar to two patterns reported for poised and active epigenetic state by a study on full circadian landscape of mammals core clock (*Takahashi, 2017, Koike et al., 2012*). This suggests that the role of RXR binding at gene promoters might not be simply to initiate transcription but also to keep the ready state of genes for flexible and fast regulation by its partners.

A similar proportion of the mRNA microarray was considered circadian or ultradian cycling by the same analysis (**Fig.23A**). This proportion is consistent with many previous circadian studies (*Akhtar et al., 2002, Panda et al., 2002*), another study reported a higher proportion (22% of transcriptome) but the experiment was performed under other conditions, at different level of mRNA maturity (*Benegiamo et al., 2018*). In contrast to RXR, the peaking of expression of these genes distributed quite evenly along the day (**Fig.23B**), at least for circadian. In our study, not all rhythmic RXR peaks resulted in a rhythmic of transcription activity. Indeed, the majority of RXR target transcripts (11'905 from ~15'000 in **Fig.13A**) are not cycling. 1848 transcripts harbored rhythmic RXR bindings but expressed stable transcription. On the other hand, 738 transcripts are rhythmically transcribed whereas their associated RXR binding is stable. Only 219 genes are cycling both at their RXR binding level and at their mRNA transcription level (**Fig.23C**). The RXR peak annotations of these rhythmic peaks is again enriched for promoter-dominant (**Fig.23D**).

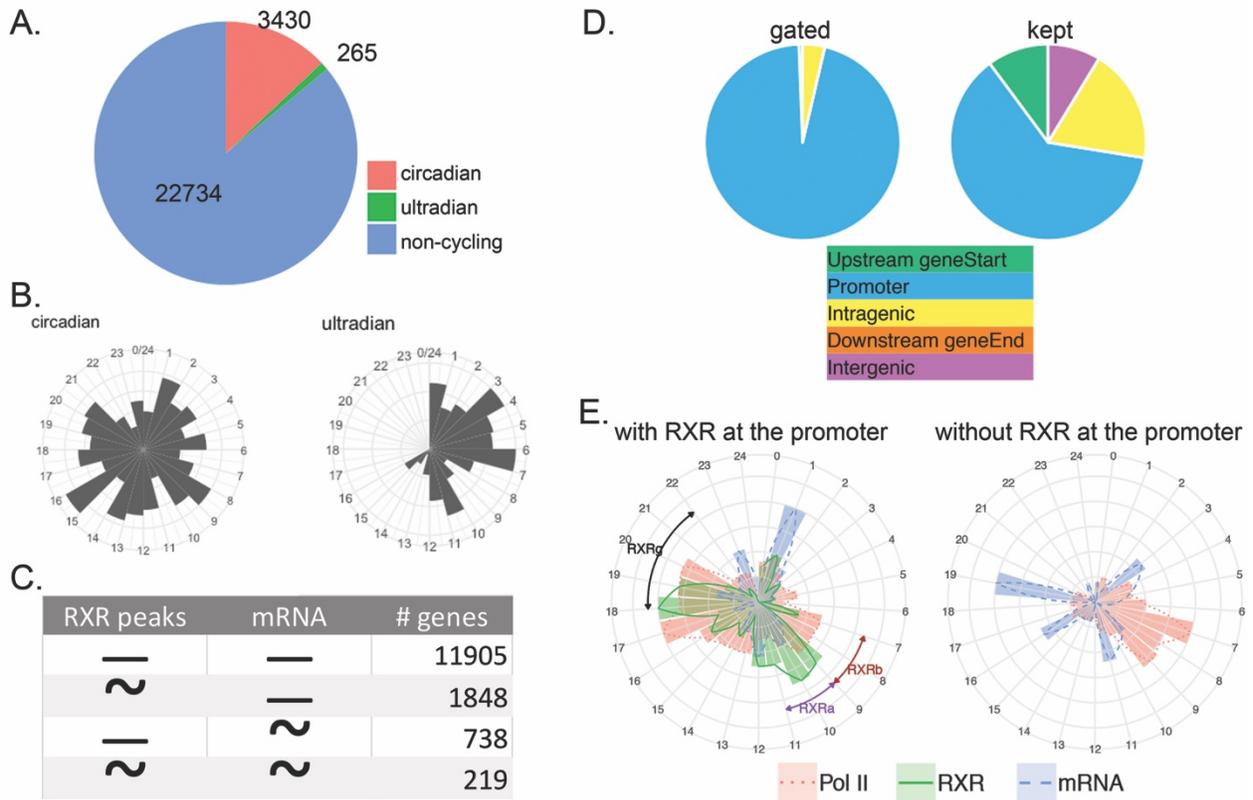


Figure 23: RXR target genes expression circadian analysis

- A.** Pie chart showing proportions of expressed genes that cycling in 24-hours (circadian) period, in less than 24h period (ultradian) and of those with constant density (non-cycling).
- B.** Circle histograms of phases distributions for circadian (left) and ultradian (right) expressed genes.
- C.** Table showing number of genes with cycling or not cycling RXR peaks and expression levels.
- D.** Annotation distribution of those 219 genes for which both RXR peaks and mRNA levels are cycling.
- E.** Circle histograms presenting the phases of 84 genes for which promoter Pol-II signal, promoter RXR signal, and expression all present a circadian rhythmicity. Left circle is for genes without RXR at the promoters. Two-heads arrows indicate the expression phases of 3 RXR isoforms.

The circadian core gene BMAL1 is considered cycling by our analysis, but the CLOCK gene is not, because the change in transcript expression value is low and statistically insignificant. The same mild rhythmic changes were also found for RXR and its partners. RXR α,β (Fig.23E), PPAR γ expressions peaked at ZT06-10, while RAR α,γ have the opposite rhythm with level peaks

at ZT14-ZT18. FXR expression peaked at ZT12 and RXR γ peaked around ZT20. Most of these patterns are consistent with other studies (Yang et al., 2006, Ma et al., 2009); except for RAR α , which peaked at ZT14 in our analysis but at ZT8 in (Yang et al., 2006), and LXR, which has the same 1.3 fold change as RAR α in our analysis but was considered non-rhythmic in (Yang et al., 2006). However, with expression changes less than 1.5-fold change between the highest and lowest level, these genes fall in the grey area just below the 1.5x fold change threshold usually set to be considered rhythmic in most circadian studies.

By overlaying the phases distributions of the promoter Pol-II signal, RXR signal and the mRNA expression levels of the 84 genes with cyclic RXR, Pol-II at promoters and mRNA, with the indications of the time when the three RXR isoforms gene expression peaked (**Fig.23E**), we could see that overall the phase oscillation of RXR, when present in promoters, peaks at ZT10 and ZT18, roughly +/- 8 hours away from the peak of mRNA production. However, the number of genes fitted to this ideal downstream regulation were very small. These analyses indicated that the rhythmic recruitment of one partner to RXR binding at the promoters alone is not sufficient to drive rhythmic transcription for the majority of its target genes, but by further regulated by diversion interplay with other partners.

3.6 Motif search identifies nuclear receptor consensus binding sites

3.6.1 Unbiased motif searches

Transcription factors bind to DNA through the recognition of specific nucleotide patterns. In the case of RXR, studies (Leid et al., 1992a, Mangelsdorf and Evans, 1995, Bastien and Rochette-Egly, 2004) have revealed the presence of the consensus sequence AGGTCA. RXR is capable of heterodimerization with several other nuclear hormone receptors, which results in the

recognition a various specific nucleotide patterns (**Fig.2**). We first performed motif analysis of all sequences below RXR peaks using HOMER v4.7 *findMotifsGenome* with default parameters and mm9 for genome. The result displayed variations from the known consensus of RXR which indicated that our ChIP-seq antibody worked well. The two most significant motifs matched the known consensus of RXR heterodimer with PPAR γ and FXR from Homer motif databases, which comprises two hexamers separated by one nucleotide spacer, either as direct repeat, DR1 or inverted repeat, IR1 (**Fig.24** – motif 1,2). Beside DR1 and IR1, the other most abundant motifs retrieved, were the GC-rich motif of SP1 marking the promoters (**Fig.24** – motif 4) reflecting the high proportion of RXR peaks at promoter (**Fig.12A**), the motif of CCAAT-enhancer binding protein (**Fig.24** – motif 5), and several Homeobox motifs (**Fig.24** – motif 3,6,8). Expecting to find different RXR partners heterodimers at different time, we ran Homer function on different RXR timepoints peak groups (described in section 3.5.1, **Fig.18**). However, all groups showed variations from the motifs from the total peak set analysis, just with different order and roughly similar target percentages. We did not observe any particular peak group with significantly different motifs from the others.

Motif	P-value	log P-value	% of Targets	Best Match/Details
1. GGTCAAAGGTCA	1e-580	-1.34E+03	31.06%	TR4(NR),DR1/Hela-TR4-ChIP-Seq(GSE24685)/Homer(0.935)
2. AGGTCATGACC	1e-315	-7.27E+02	16.57%	FXR(NR),IR1/Liver-FXR-ChIP-Seq(Chong et al.)/Homer(0.974)
3. GTTAATCATTAA	1.00E-150	-3.48E+02	3.03%	Hnf1(Homeobox)/Liver-Foxa2-Chip-Seq(GSE25694)/Homer(0.957)
4. AAGCCCCGCCCF	1.00E-141	-3.25E+02	27.41%	Sp1(Zf)/Promoter/Homer(0.951)
5. TTGCGTAA	1.00E-133	-3.09E+02	24.67%	MF0006.1_bZIP_cEBP-like_subclass/Jaspar(0.896)
6. AATCGATA	1.00E-121	-2.79E+02	16.86%	HNF6(Homeobox)/Liver-Hnf6-ChIP-Seq(ERP000394)/Homer(0.912)
7. CCGGAAGT	1.00E-118	-2.72E+02	11.14%	ELF1(ETS)/Jurkat-ELF1-ChIP-Seq(SRA014231)/Homer(0.982)
8. ETGTTIAC	1.00E-97	-2.24E+02	23.53%	Foxo1(Forkhead)/RAW-Foxo1-ChIP-Seq(Fan et al.)/Homer(0.947)
9. AGGTCAGTTCAG	1.00E-83	-1.93E+02	15.03%	MF0004.1_Nuclear_Receptor_class/Jaspar(0.696)
GCATTCIGGGAAATGTAGTC	1.00E-78	-1.80E+02	2.63%	GFY-Staff(?_Zf)/Promoter/Homer(0.988)

Figure 24: HOMER motif analysis result for total RXR peaks

However, the annotation system of Homer depends on a set of customized motif position weight matrices that the developers collected and curated from only ChIP-seq studies, and the correlation hierarchy of “best-scores”. But not all RXR partners have been equally studied using ChIP-seq, while PPAR and RAR were extensively studied in various conditions from genome-wide to target search, the other partners were less so. Moreover, RXR binding motifs are well known and extensively explored, and the goal of our motif analysis ultimately was to assign each (and hopefully every) RXR peaks with a specific RXR heterodimer partner. Hence, matching our data set motifs with other ChIP-seq studies motifs with completely different sequence content and number, conditions, cell lines, tissues, species, was inequitable.

Thus, we changed to MEME suite, which uses a combination of different motif databases not restricted to only ChIP-seq studies and capable of de-novo motif detection. We used the MEME-ChIP tool, which by default called two different motif discovery algorithms: multiple EM for motif elicitation - MEME, and discriminative regular expression motif elicitation – DREME, to discover novel sequence motifs enriched in particular set of sequences compared to control set. We used default parameters of the DREME motif discovery, which results in short motif 5-8 nts long, while by default MEME searches for motifs between 6 and 30 nucleotides, and obtains longer motifs (*Fig.25A*). The short motifs from DREME identified as top hit the canonical RXR binding hexamer RGGTCA. The MEME suite also directly invokes TOMTOM (*Gupta et al., 2007*) to annotate the de novo motifs identified and assign them to known TF when possible. Nevertheless, we chose to re-annotate using more databases. We collected the known consensus of all the candidate TFs predicted by TOMTOM from JASPAR, SwissRegulon, Jolma, HOCOMOCO databases (*Fig.25A*) to build a complete reference of known NR binding motifs and incorporate it with the newly discovered motifs from DREME and MEME. We used pairwise Pearson correlation coefficient to

calculate the distances between motifs position weight matrices, then used UPGMA from DECIPHER package to create a motif tree. We manually curated them into 25 motif signatures that represented the most enriched consensus of RXR peaks (**Fig.25B**).

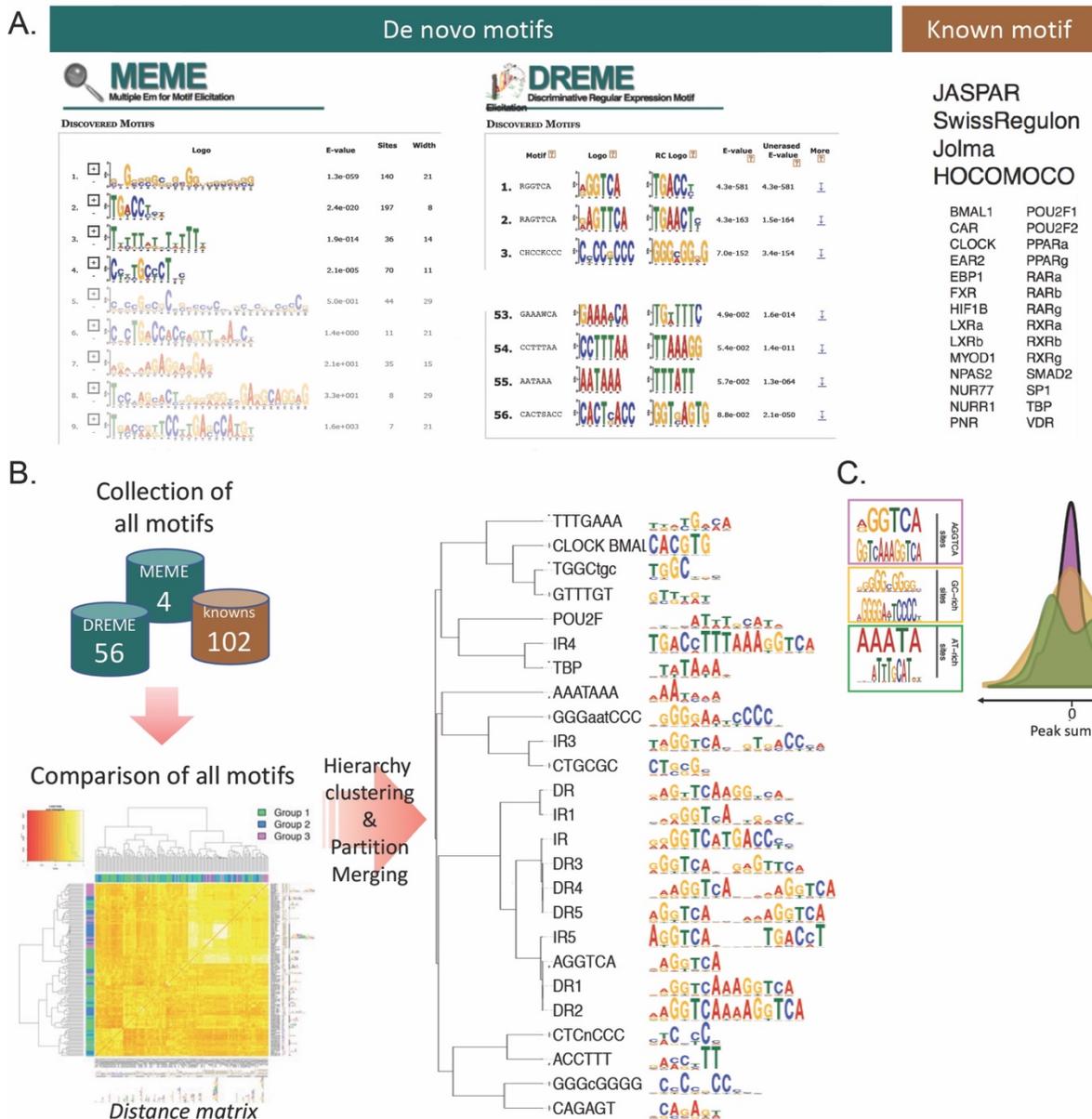


Figure 25: MEME discovered motifs

A. MEME discovery framework and results

B. Method of grouping MEME results and known motifs into 25 significant motif patterns within RXR peaks

C. Distribution of three representative motif patterns around peak summit

We then re-scanned the sequence under each of our peaks specifically for the identification of the motifs found in newly built database. The distributions of 25 motif signatures around peak summit congregated into 3 patterns: both the canonical AGGTCA motifs and the GC-rich motifs accumulated more toward the peak summit, while the AT-rich motifs clustered more toward the flanking of other motifs (**Fig.25C**).

Despite our efforts of tuning the parameters of different tools, using preset pre-annotated motif databases brought the promiscuousness of RXR heterodimer partners into the context of motif construction. Thus, the process of assigning RXR peak with specific distinct motif then grouping distinct motif-peaks together were obscured and hindered by the cross-section between similar motifs. We therefore decided to customize our analysis.

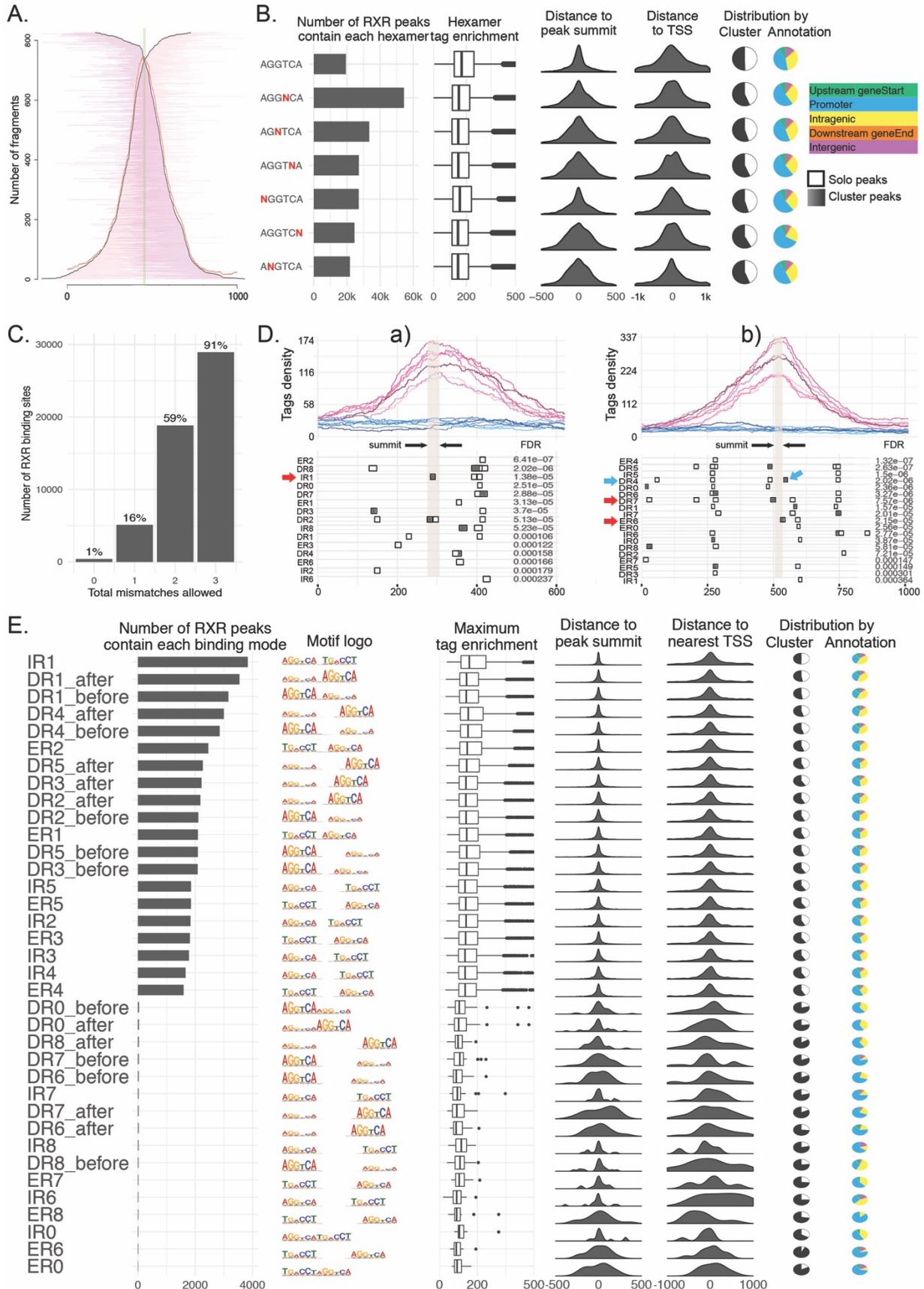
3.6.2 Searching for RXR heterodimer consensus-derived binding sites

Since our motif analysis method using pre-designed algorithms to search for enriched consensus have many times skewed our analysis towards undesirable motifs, we decided to restrict our search with customized scanning. In previous studies, the rule of core half site RGKTCA (*Leid et al., 1992a, Mangelsdorf and Evans, 1995, Bastien and Rochette-Egly, 2004*), and the affinity of RXR and its partners to certain variations of this hexamer have been tested (*Munoz-Canoves et al., 1990, Duester et al., 1991, Suva et al., 1994, Vansant and Reynolds, 1995, Balmer and Blomhoff, 2005*). RXR-partner heterodimers have the ability to bind response elements formed of a tandem repeat composed of two hexamer half-sites. These repeats, that we henceforth deemed heterodimerization binding modes, can be either direct repeat–DR, inverted repeat–IR or everted repeat–ER with 1 to 8 nucleotides spacing between each half site. To identify the potential RXR heterodimer binding mode present under each peak, we chose to scan each peak for the presence of these customized patterns. Since a degenerate hexamer sequence will be found many times under a

given peak, we utilized the advantage of pair-ended CHIP-seq to know precisely the fragment that was immunoprecipitated and only search the binding motifs under the region of the peak covered by most of the fragments (**Fig.26A**). We allowed one mismatch for the AGGTCA hexamer scanning, and attempted to determine for each peak which of the various binding modes (**Fig.2**) was the most likely. The least conserved position was the fourth nucleotide of the hexamer (**Fig.26B** – the most frequent hexamer was AGG**N**CA). With the allowance of at most three mismatches in the 12mer resulting from both hexamers, we could identify at least one binding mode in about 91% of RXR peaks (**Fig.26C**). Using a window of 15 nucleotides flanking peak summits reduced the number of peaks containing at least one binding mode to only ~45%.

Using the coordinates of the hexamers, we constructed a library of sequences for the different binding modes with a number of space-between each hexamer ranging from 0 to 8. All the sequences constituted to each binding modes were pooled to create a position weight matrix and create a custom binding mode loci library. These matrices were used to scan the loci library with FIMO (*Grant et al., 2011*) from the MEME suite to score the significance of each locus.

The overlapping between the binding modes were extensive, which resulted in multiple potential binding modes within one RXR peak/one locus. In order to simplify the overall statistic without leaving out too many peaks, the selection of one single binding mode for each peak was made based on the proximity to peak summit and the lowest FDR produced by Fimo (examples in **Fig.26D**). The most common binding mode is DR1 (heterodimerization with RAR or PPAR), which is present in ~16% of the peaks, closely followed by DR4 (heterodimerization with TR, LXR, CAR).



No significant difference in annotation and distance-to-peak-summit distribution could be found between different binding modes ([Fig.26E](#)), except for all the binding modes with a spacing of 0, 6, 7, or 8 nucleotides between each hexamer, which were very infrequent and are most likely false positives. The detailed classification of binding modes by DR, IR, ER with 1-8 nucleotides spacing encompassed every motif consensus that had been reported in previous studies for RXR and its partners. However, these specific studies using DNA gel electrophoresis in combination with

Figure 26: Customized binding modes scanning

- A.** Illustration of the diverse fragment sizes and positions of all precipitated fragments within one representative RXR peak. The fragments were arranged by coordinates of their 5'-end (left to right) and 3'-end (right to left), which created the borders (black lines) almost fitted with the peak shape (red line). The region with most coverage around peak summit was used for motif assignment.
- B.** Characteristics of each hexamer with 1-mismatch variations found with RXR peaks: number of peaks, tag enrichment, distance to peak summit, distance to nearest TSS, cluster versus solo peaks proportions, and annotation fractions
- C.** Bar-plot showing the number of RXR peaks (percentage on each bar) containing at least one tandem repeat with increasing mismatch threshold for the hexamers
- D.** Examples of occurrences of multiple binding modes within each RXR peak and the assignment of one mode to one peak was based on presence of the mode with lowest FDR reported by FIMO within peak summit (grey band). Top plots show RXR tag density of the sites. Bottom plots show the exact identified loci of different binding modes within the peaks. Each binding mode loci were aligned on each row. When there are multiple loci of a single binding mode, the FDR on the right is the lowest FDR of all the loci within each row, and the corresponding locus was marked black. The row order corresponding to the lowest FDR of all the rows.
 - a)** Intragenic RXR peak from Fig.15a shows simple assignment for IR1 (red arrow) within summit and has lowest FDR
 - b)** RXR peak at TSS of Hgfac gene with complicated assignment because DR7 has lower FDR but ER6 is closer to the summit (red arrow). DR4 is little further away from the summit but has lower FDR than the other two.
- E.** Detail of each repeat consensus found within our peaks, with filters favored for spacing of 1-5 nucleotides. The “before” or “after” indicates the relative position of the hexamer with the lowest mismatch.

sequencing, at very specific target genes and/or genomic loci, were done for only a few specific binding modes. The everted repeats ER0, ER7, ER8 especially, which we did not identify in this study, were only reported from computational scanning studies. In our case, the reported observed consensus only encompass the DR1-5, ER1-5 and IR1-5, because one of these binding mode could always be attributed to an RXR peak with a higher confidence than half-sites with spacers 0, 6,7 or 8. Pairwise overlapping permutation test showed the extensiveness of the overlapping between each binding mode sites identified under the same RXR peak ([Fig.27A](#)). As expected, the DR loci were more clustered within themselves and less with the IR and ER. The ER sites interestingly overlapped more significantly with DR and IR sites (redder cells) than within themselves (more yellow cells). With this pattern of overlapping across binding modes, we performed pairwise comparative motif enrichment analysis for each pair of all binding modes loci libraries, using all the signature motifs from the earlier analysis (in [Fig25.B](#)). Only the a few DR and IR loci libraries had resulted enriched motifs consistent with their binding modes ([Fig.27B.a-c](#)). The loci denoted for DR1-5 and IR1-2 were affirmatively enriched with the correct motif signatures. The loci of other DR and IR showed less significance in the matching signature motifs. The results for ER were especially negative, because there were always pattern of DR and IR signature motifs within loci denoted for ER binding modes (examples in [Fig.27B.d](#)). This confirmed that for these ER binding modes, even when we can find an ER within a locus, there would always be another DR or IR mode overlapped. Therefore, we reformulated the customized scan for only 7 binding modes, which matched the 7 known consensus of RXR partners and minimized the list of binding modes ([Fig.27C](#)). Delightfully, this reformulation scan did not reduce the annotated RXR peaks population. For example, the RXR peak in [Fig.26D.b](#) with this reduced binding modes list was re-assigned with DR4 other than DR7/ER6, and ultimately was gated out from analysis due to overlapping with TSS.

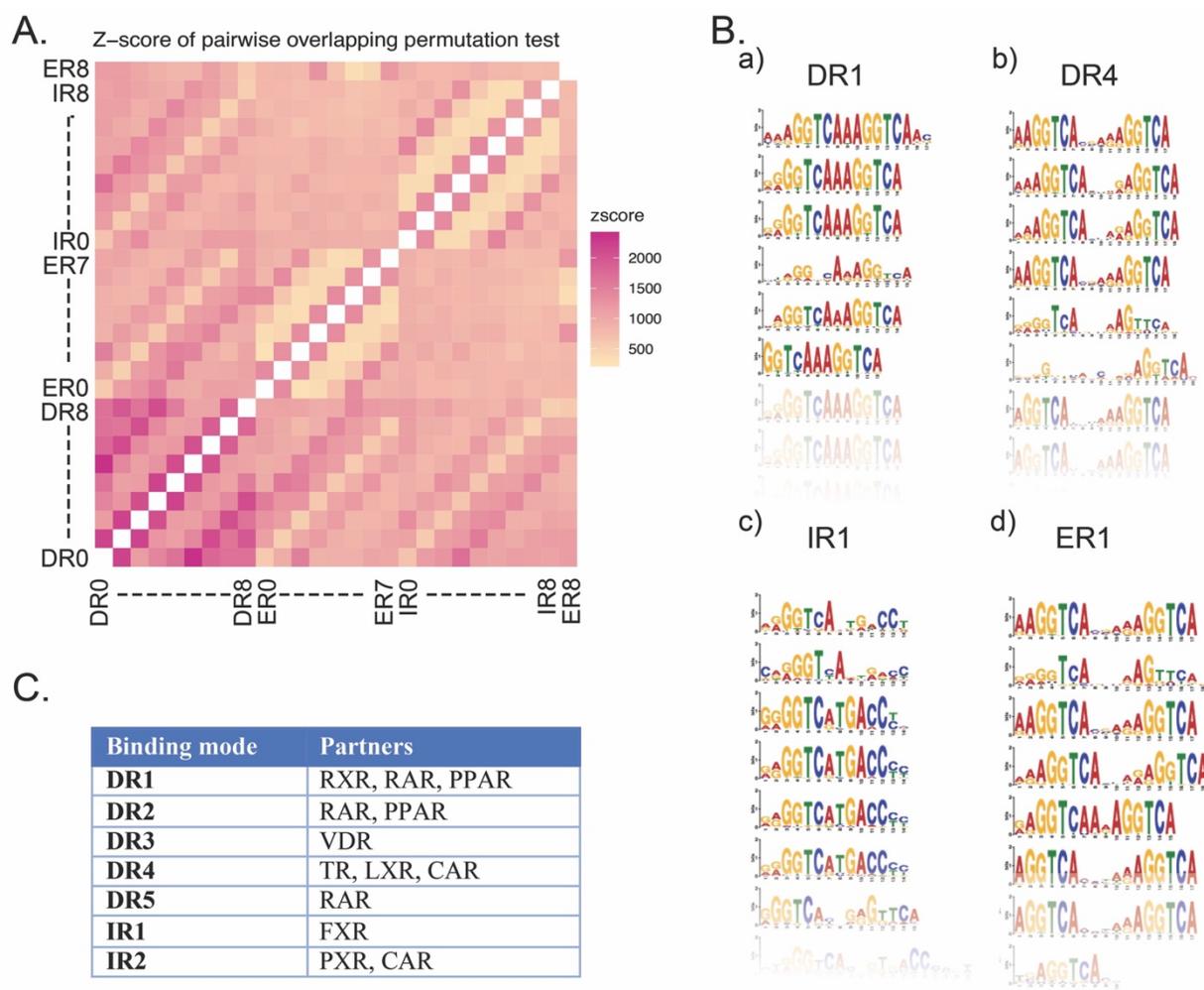


Figure 27: Binding modes comparative analysis

- A.** Heat-map showing all pairwise overlapping sites between each pair of all binding modes. Intensity of the cell colors toward red indicated higher z-score produced by permutation test, meaning that the overlap between the sites were not occurring by chance.
- B.** Examples of the enriched motifs at DR1, DR4, IR1 and ER1 loci when compared with other binding modes loci. The motifs were consistent with the binding mode for DR1, DR4 and IR1. ER1 loci were enriched not with ER1-pattern motif, but with DR4- and DR2-pattern motifs
- C.** Table of set binding modes we reformulated scanning on RXR peaks sequences and their corresponding reported partners identities.

3.6.3 *Distribution of motif patterns across RXR peak groups, time points and annotations*

At each evolution of our analysis, we looked at the general distribution of the resulting motifs across various aspects of RXR peaks and peak groups in the expectation to observe specific motif pattern dominantly displayed within certain profile peak group or certain time point. However, the disproportionally high promoters within RXR binding sites casted a strong effect not only on the target gene expression and rhythmic profiles (described in previous sections 3.3.2, 3.4, 3.5.3) but also on the discovered motif patterns. Half of RXR peaks clustered more toward promoters ([Fig.12A](#)), promoter sequences are generally GC-rich, hence promoter RXR peaks were enriched with high GC-content motif ([Fig.28A](#)). When we analyzed the distribution of variations of the canonical AGGTCA hexamer around RXR peak summits across different profile peak groups (described in section 3.5.1, [Fig.18B](#)), various AGGTCA variations were found clearly within groups containing less promoter annotation, and the occurrence stood abundant even with zoomed in ± 15 nts window at peak summits. Within groups containing a high proportion of promoter annotation, we found less hexamer occurrences ([Fig.28B](#)).

-
- C. Boxplot of %GC content of RXR peaks across the cosinor-fitted phases (section 3.6.2), separated by promoter annotation*
 - D. Distribution of 4 representative signature motifs (section 3.7.1) along the cosinor fitted RXR peaks' phases:*
 - *Top: the DR1 and IR1 signature*
 - *Bottom: The GC-rich signature*
 - E. Pie charts showing differences of annotation and clustering proportions based on the peaks gating.*
 - F. Comparative enriched motif analysis between gated versus non-gated RXR peaks showed enrichment of GC-rich motifs, which confirmed our gating analysis*

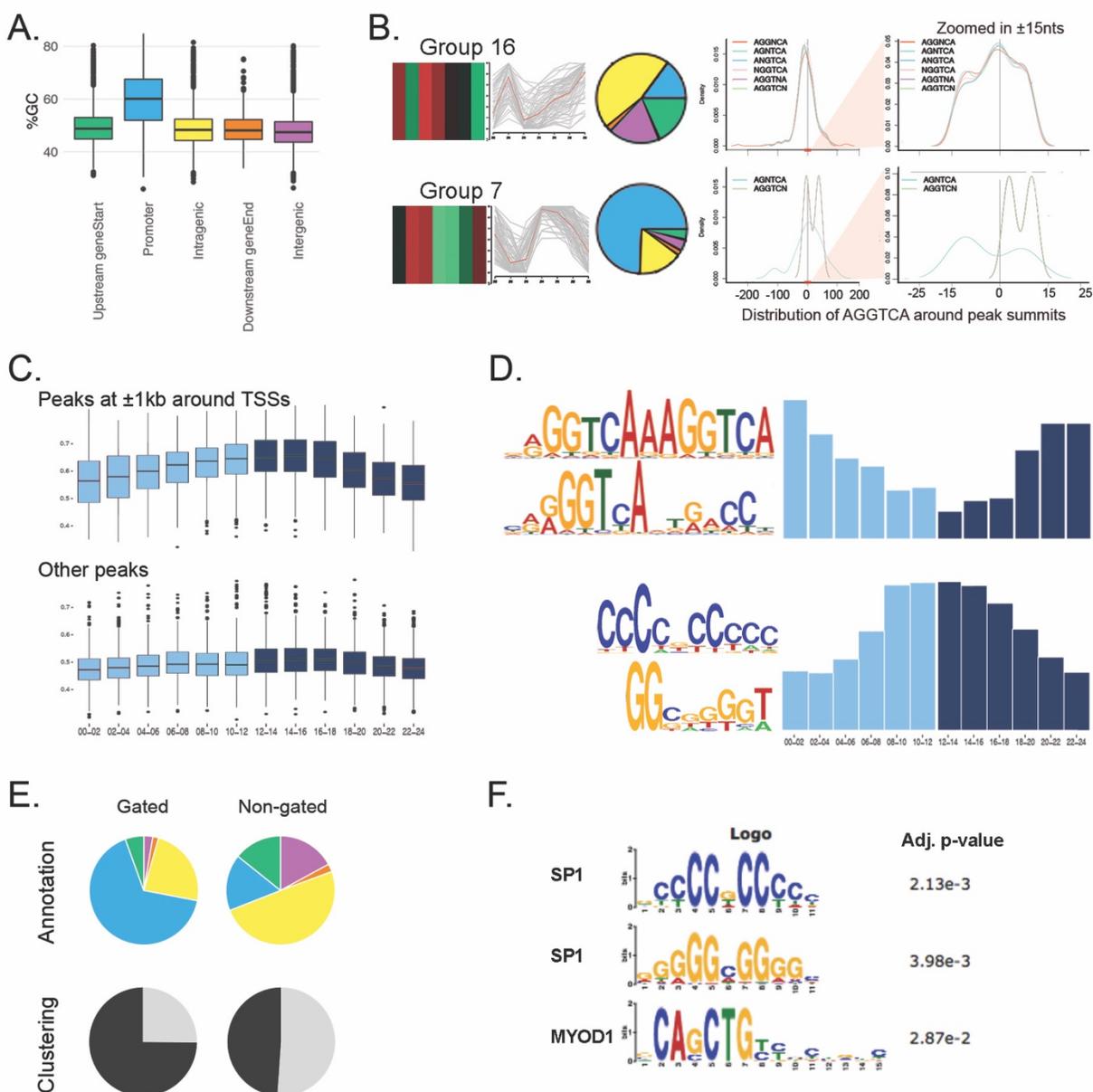


Figure 28: Motif distribution patterns

A. Differences in GC-content of the sequences under RXR peaks, Promoter peaks have significantly higher %GC under their sequences. Because RXR accumulates more around promoters, the most frequent motif discovered resulted in a high GC-content motif

B. Example of two different peak groups (from section 3.6.1) with distinct annotation contents:

- Top: group with low Promoter annotation (pie chart), has strong accumulation of variations of the AGGTCA hexamer at peak summits (middle density plot), and even more visible when we zoom in ± 15 nts around peak summit
- Bottom: group with high Promoter proportion, has only few occurrences of the two hexamers at peak summit

For the cosinor-fitted phases peak groups (described in section 3.5.2), population of RXR peaks that reached highest intensity at mid-day had significantly higher GC content (Fig.28C). We observed slight trend of the GC-rich signature motifs (from 3.6.1) more prominent among the mid-day RXR peak from the sinusoidal fit; while an opposite decrease mid-day trend displayed for those motifs with AGGTCA content (Fig.28D). This corresponded well with the higher promoter annotated proportion of RXR peaks within this population. And the phases distribution of promoter Pol-II signal described in section 3.5.3 and Fig.22D showed the same pattern with the GC-rich motif. These results further confirmed the effect of the co-localization with Pol-II on the signal of RXR: whether these RXR peaks were considered significantly rhythmic or not, the highest RXR signal at these promoters coincided exactly with the signal of Pol-II. And with this involvement of RXR motif in the transcript machinery, any analysis taking these regions into account would be heavily affected by the high GC content of these sequences, clearly displayed through the three signature patterns in Fig.25C. In fact, we looked at the effect of the gating strategy described in 3.4.2, which essentially gated out all the RXR peaks at the TSS, and thus highly “contaminated” by GC content. Since RXR peaks clustered more at the promoters, this gating created two populations with very different proportions of solo/cluster peaks (Fig.28E). Comparative motif enrichment analysis using all the motifs collected from section 3.6.1, showed that the gated RXR peaks’ sequences were enriched with GC-rich motifs such as SP1 and MYOD1 motifs and mostly depleted in AGGTCA RXR binding motifs, compared to the kept RXR peaks’ sequences (Fig.28F), which confirmed the necessity of the gating.

Taking into account the gating strategy and the simplified binding modes list, we reanalyzed binding modes of the kept RXR peaks. The most abundant binding modes found was DR1, same as earlier. DR2-5 presented similar relative abundances but the number of IR1 peaks dropped from

the third most abundant category to the sixth (**Fig.29A**). There were still no significant differences between these various binding modes concerning the genomic, annotation, distance-to-peak-summit, distance-to-TSS (**Fig.29D-G**). However, the distance-to-peak-summit clearly showed more concentrated motifs toward the peak summits than before (**Fig.29F**).

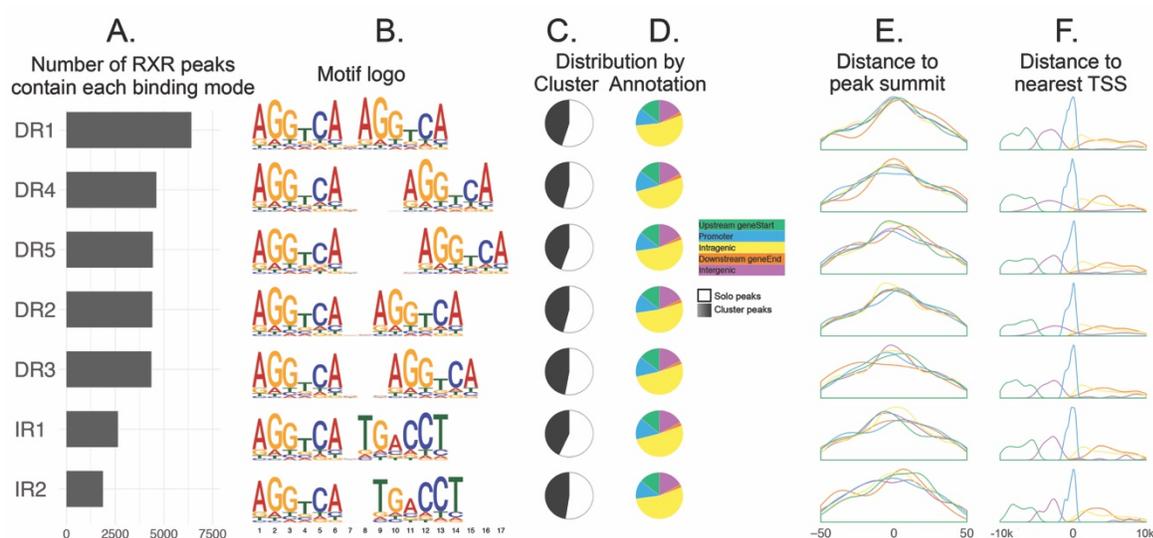


Figure 29: Characteristics of simplified binding modes within kept RXR peaks

- A.** Distribution of motif among RXR peaks: number of RXR peak with each specific binding mode
- B.** The motif logos derived from all the sequenced assigned with each specific binding mode
- C.** Cluster/Solo pie charts for each binding mode (cluster peaks: black, solo peaks: white)
- D.** Annotation pie charts for each binding mode (reference colors on Fig.1)
- E.** Distribution of the distance between found binding mode to its peak's summit
- F.** Distribution of the distance between found binding mode to nearest TSS

Numerous properties of protein-DNA binding “motif” have been under scrutiny, from base compositions, structures, thermodynamic properties (*Ahmad et al., 2008, Ahmad et al., 2004*), expressions (*Pham et al., 2005*), to bonding and force types, TF conservation and mutation (*Luscombe and Thornton, 2002*) and bending of the DNA (*Jones et al., 1999*). In addition, machine learning methods, such as neural networks, support vector machines (*Ofran et al., 2007*) and regressions (*Zhou and Liu, 2008*) have been proposed. Often, each method was set under one

specific protein and a particular data set inspection, as was our method. It has been proposed that multiple transcription factors can compete for binding site in regulating gene transcription (*Carlberg and Seuter, 2010*); on the other hand, the present/or not present of certain motifs can contribute to searching-binding dynamics of transcription factors (*Dror et al., 2016*). Thus, it is highly possible that multiple partners can equally and/or alternately contribute to the transcriptional regulatory at one binding site (examples in *Fig.26D*). Assigning one heterodimer binding mode to one RXR peak, and extrapolation from binding modes to partners obscured such interplay. Here we demonstrated the equal distributions among heterodimer binding modes, and conclude that the various RXR heterodimers do not show any preference for a particular genomic binding location such as promoter or intergenic.

3.6.4 Distribution of the different Nuclear receptor binding modes across target genes

In further attempts to infer which partner coupled with RXR to regulate which gene, we annotated the binding modes using the same hierarchy as the peaks (*Fig.12A*). Before gating and simplified binding modes, the analysis showed that only 38.7% (about 5000 – *Fig.30A*) of the putative RXR target genes were associated with only one unique binding mode, whose abundance and annotation patterns are similar to the analysis within peaks. About 8000 remaining genes were associated with at least 2 binding modes or more (*Fig.30A*). The increased number of binding modes associated with a given gene shifted the genomic annotation proportions from promoter dominant toward Intragenic-dominant, and interestingly, Tukey Honest Significant Differences test in conjunction with One-Way ANOVA showed that the average expression levels significantly increased with the number of associated binding modes (*Fig.30B*).

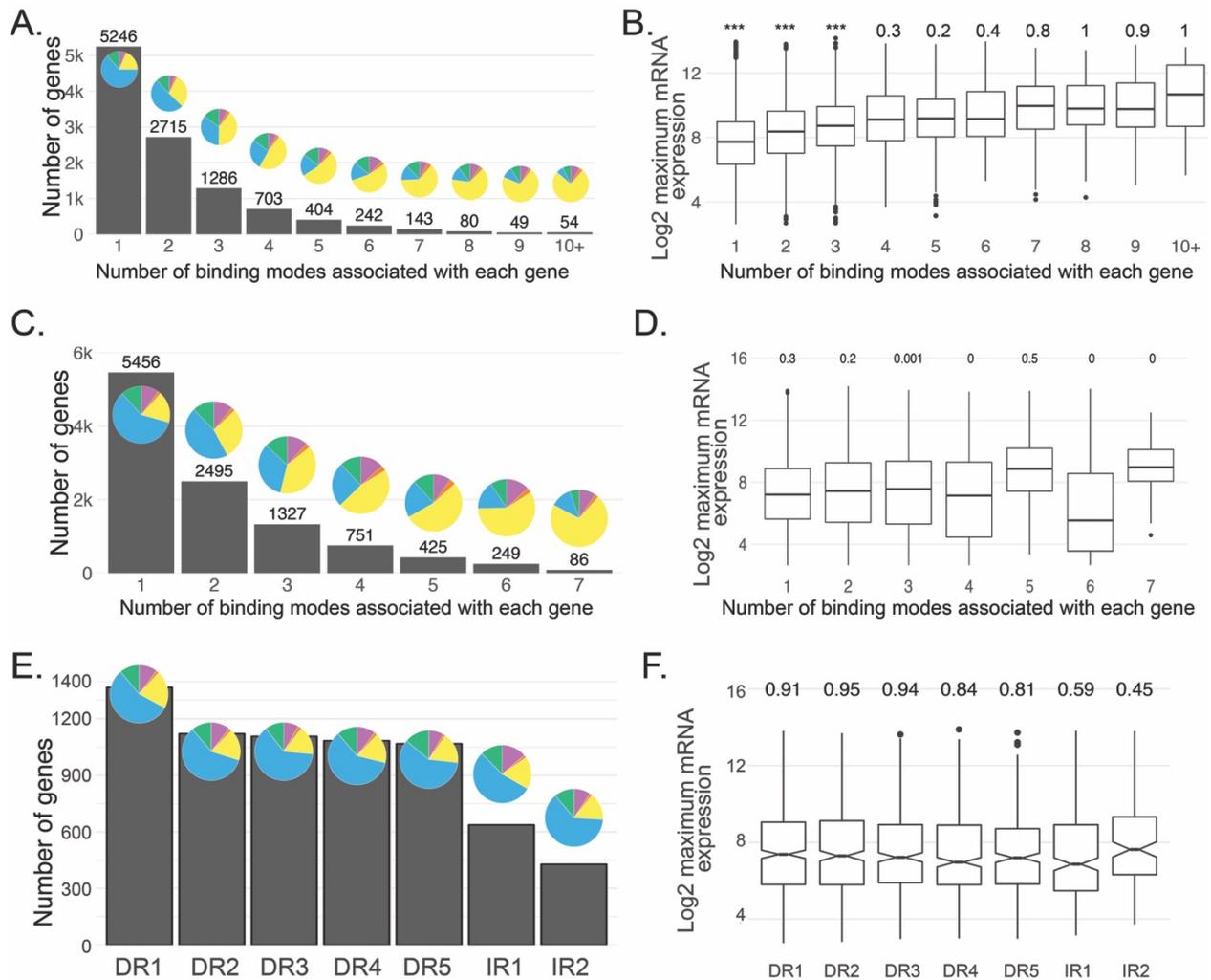


Figure 30: Distribution of binding modes among RXR target genes

- A.** Number of genes regulated by different number of binding mode combinations. Pie chart above each column details the relative proportion of genomic annotations as described in Fig.10A.
- B.** Boxplot of maximum gene expression level for each group of genes based on the number of regulating binding modes.
- C.** Same as panel A but for only kept RXR peaks with simplified binding modes
- D.** Same as panel B but for only kept RXR peaks with simplified binding modes
- E.** Number of genes regulated uniquely by each specific binding mode for only kept RXR peaks with simplified binding modes. Annotation pie chart for each column
- F.** Boxplot of maximum genes expression for each group based on specific regulating binding modes, for only kept RXR peaks with simplified binding modes

After gating and simplified binding modes, the results showed the same distribution trend as before where half of the genes presented a unique binding mode (**Fig.30C**). The distribution of genomic annotation in function of the specific binding mode of this group did not show any difference than before. The increased number of binding modes associated with a given gene also shifted the genomic annotation proportions from promoter dominant toward Intragenic-dominant. However, we no longer observed the correlation between number of associated binding modes with the average expression levels (**Fig.30D**), nor between different binding modes of the monotypic-binding mode genes with their expression levels (**Fig.30E,F**).

The only slight difference was observed for IR1-binding mode, characteristic of FXR-regulated genes, which were overall less expressed than any other binding modes, which is consistent with general repression effect proposed for FXR target genes (*Lee et al., 2012*). Further phase distribution analysis for individual binding modes within the rhythmic components, however, did not yield any distinct pattern. The relative enrichment of motifs among cyclic genes are comparable to the total genes, thus none of the heterodimer partner dominantly contribute to the transcriptional rhythmicity.

3.7 Enriched pathways analyses

3.7.1 *The highly expressed – high plateau genes*

To explore in which liver functions the RXR target genes were involved, we inspected the small group of short and highly expressed genes containing the RXR high plateaus described in section 3.4.1. The plateau genes can be categorized into 4 main functional families (**Fig.31**): the cytochrome P450 genes playing a role in oxidoreductase and oxygenase activity; the Serpin genes with serine-type endopeptidase inhibitor activity; the Apolipoproteins genes; and crucial genes

related to glucose and lipid metabolism (Acox1, Ass1, Alb, Cps1, Fabp1, Hnf4a, Glud1, Mat1a, Pck1). Because of their involvement with such core functions of liver metabolism, it is interesting but not entirely unexpected to find they have significantly higher expressions compared to other genes (**Fig.15C**). In particular, the cytochrome genes family has been shown to be highly RXR α -dependent in male mice (*Cai et al., 2003*). The core glucose and lipid metabolism genes are notorious reported targets of RXR partners. The other three gene families are putatively regulated by other transcription factors who do not hetero-dimerize with RXR, for examples: the Apolipoproteins genes are regulated by the orphan receptor ROR α (*Vu-Dac et al., 1997, Park et al., 2006, Jetten, 2009*), the Cytochrome genes by HNF4A, STAT5B (*Park et al., 2006, Jover et al., 2009*), Serpin genes by NF-IL6 (*Morgan and Kalsheker, 1997*).

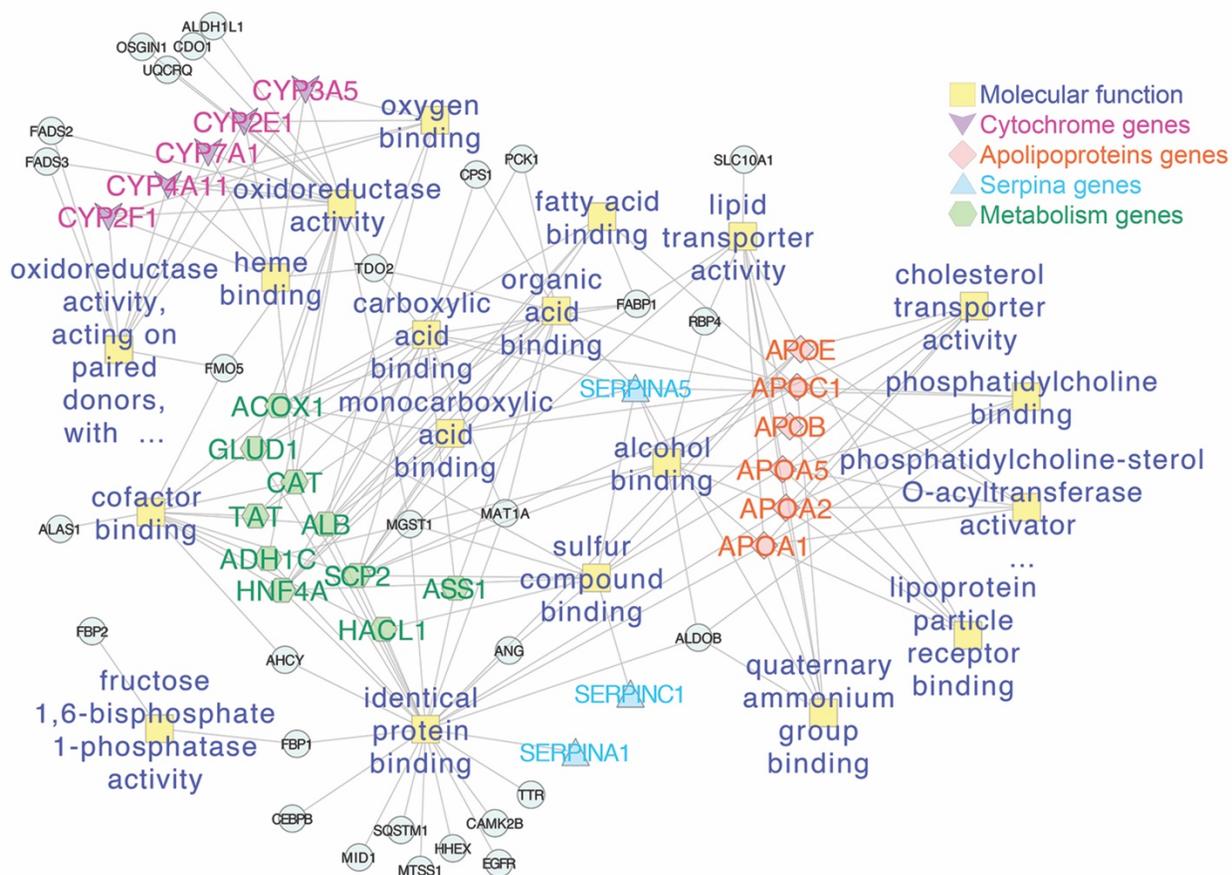


Figure 31: Gene families of the Plateau genes and their functions

3.7.2 Comparative pathways enrichment among RXR binding modes

To compare the targets of RXR different binding modes, we performed pathway enrichment analyses using the expression levels of the ~15 thousands expressed genes putatively regulated by RXR. With the various binding modes assigned to these RXR peaks and their putatively associated target genes, we sought to determine patterns of regulation present within the RXR regulatory network. Pathway enrichment analyses were performed using *enricher* and *clusterCompare* from clusterProfiler R package, using one group of genes containing the total genes in the database as control, or using multiple groups of genes in comparison. From the motif analysis, we observed that one target gene can be associated with one RXR binding mode – the monotypic-binding mode, or with many – the multiple-binding modes (**Fig.30A**).

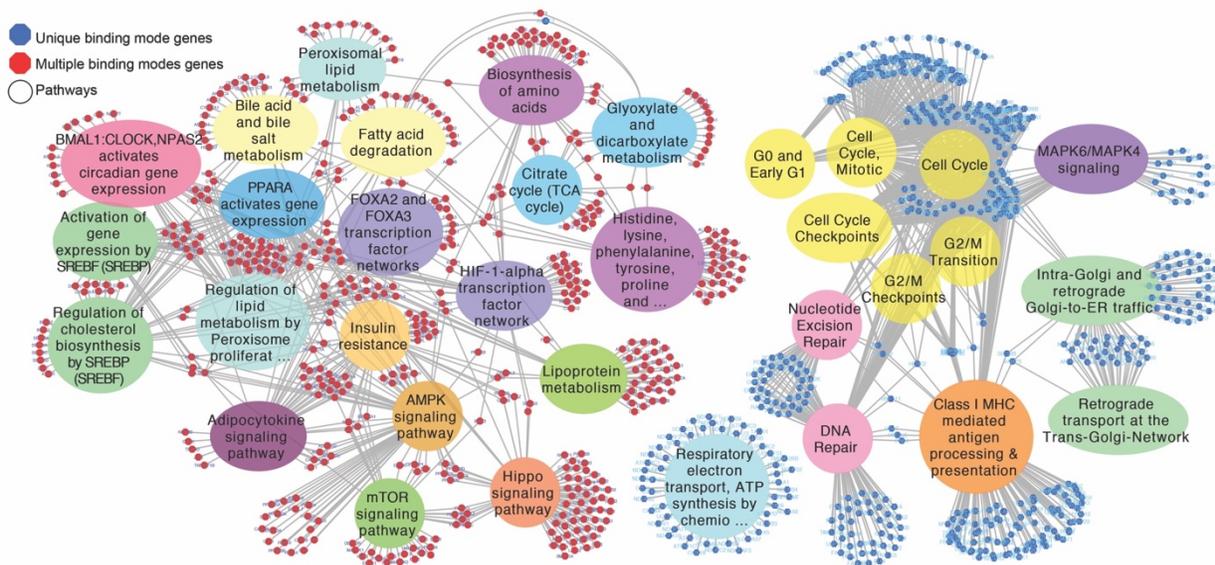


Figure 32: Comparative pathways enrichment shows distinct patterns of functions for genes with unique binding mode (blue) compared to genes with multiple binding modes (red).

Pathway comparative enrichment analysis between the monotypic mode genes and the multiple modes genes showed a clear distinction in the regulatory pathways in which these genes were involved. Genes participating in general metabolic pathways contained multiple binding

modes, consistent with the well-known lipid, glucose regulation functions of RXR and the circadian rhythm maintenance. In contrast, genes participating in very specific pathways, such as cell cycle checkpoints tend to be controlled by monotypic-binding mode ([Fig.32](#)). We ascertained that the majority of regulatory genes of the liver that are modulators for gene expression across different pathways are in fact influenced by many TFs (hence, multiple motifs), and that RXR binding occurs mainly on the promoter and gene body. On the other hand, genes with unique binding mode are in charge of specific pathways. However, we also observed the tight correlation between RXR and Pol-II within the transcription machinery (section [3.4](#)). This led to the suggestion that some group of genes presenting RXR binding precisely at their TSS or covering the whole transcript unit were perhaps not specific targets of RXR, but rather the product of the bulky transcription machinery in which RXR participated. The gating strategy described in [3.4.2](#) helped us separate these genes from the total RXR target genes.

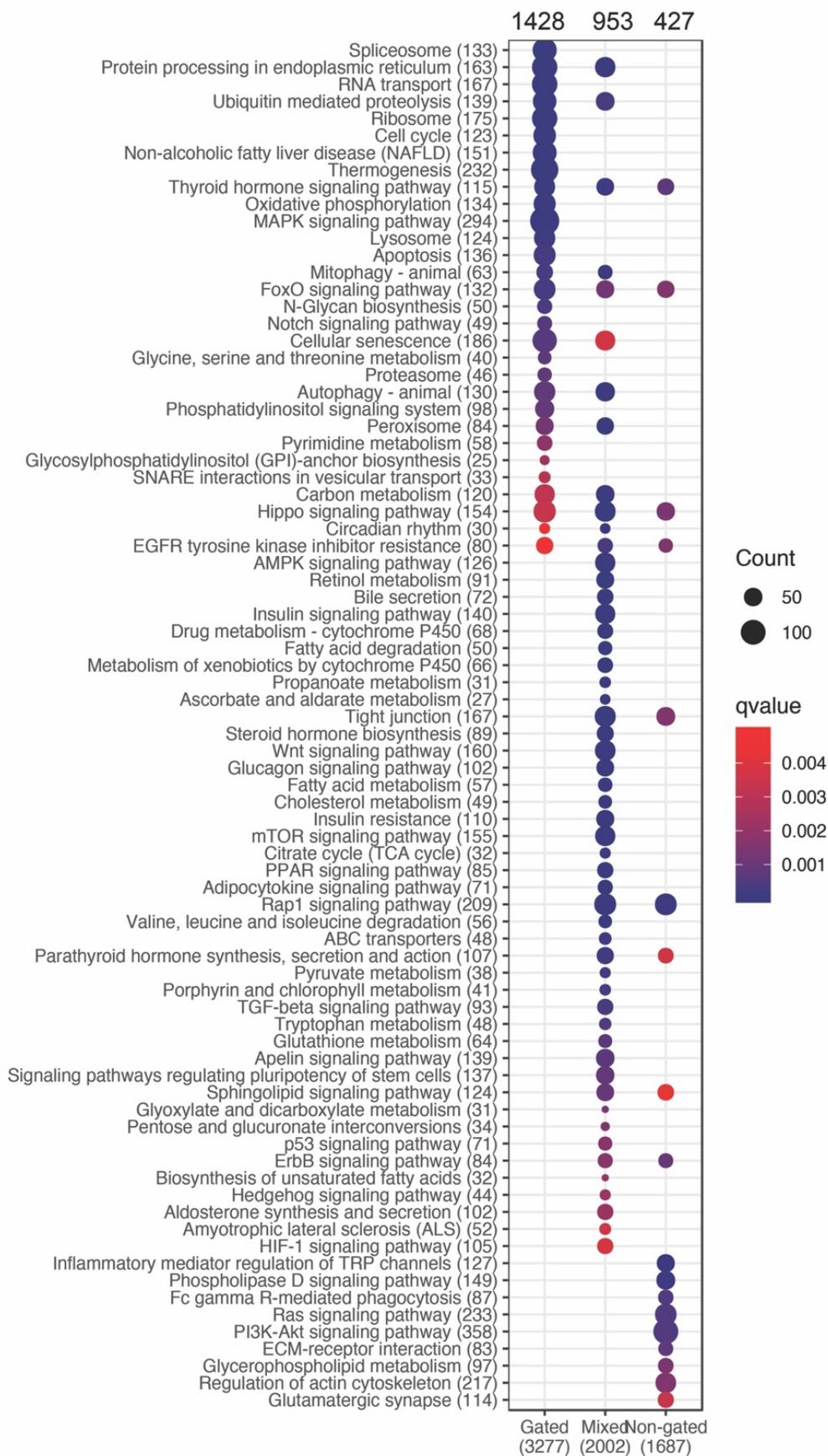
3.7.3 The gating effect on pathway analysis

Pathway comparative enrichment analysis on the three genes groups, the gated, the kept and the mixed ([Fig.16D](#)) showed that we could associate distinct pathways patterns for each genes group ([Fig.33](#)). Interestingly, the pathways enriched for the gated genes were very similar to the pathways specifically enriched for the monotypic binding mode genes before gating ([Fig.32](#)). These pathways, including cell cycle regulations, cell senescence, apoptosis, etc. are mainly general cellular sustaining functions. Diversely, the pathways enriched for the mixed and the kept genes are signaling pathways essential for hepatocyte metabolic functions. Indeed, after gating, pathway comparative enrichment analysis between the monotypic mode genes and the multiple modes genes (from [Fig.30C](#) among the mixed and kept genes) no longer showed the clear pathways distinction. While the same general metabolic pathways with the well-known lipid, glucose regulation functions

of RXR and the circadian rhythm maintenance were still enriched for the multiple binding modes groups, consistent with the previous analysis ([Fig.32](#)), the monotypic binding mode genes group was no longer enriched for any clear pathway pattern.

The gated out monotypic mode genes analyzed earlier in [Fig.32](#) had higher Pol-II signal and bore many GC-rich RXR peaks due to the TSS contamination (demonstrated in [Fig.28F](#)). However, we can still retrieve well-defined binding modes under the sequences of one third of the gated peaks. An example for these monotypic genes is the *Hgfac* gene that has only one RXR peak at the TSS ([Fig.26D.b](#)) with valid DR4 binding mode. This gene was gated out from the analysis only due to the overlapping of RXR peak at TSS. Thus, these analyses showed that RXR binds to target sites in a restricted manner with unique binding mode and highly overlapped Pol-II signal upon the genes involving in general cell sustaining functions. This suggested the selective partnering of RXR with other nuclear receptors for these cell sustaining/survival functions that required no tinkering. For other genes that involved diverse metabolic regulation functions of the liver, RXR heterodimers with various partners for more crosstalk, co-regulated signaling pathways. Such flexible regulation effects are required for the cellular capability to adapt to different modulations such as lightning and feeding. These effects may not always be transcription activities but repression or inhibition, thus the signal of Pol-II were less enriched upon these RXR binding sites.

Figure 33: Comparative pathways differences between 3 gene groups: gated, kept and mixed. Colors of points indicate q-values, sizes the number of genes participated within specific pathway.
Number after each pathway name indicate the number of genes present in the database for a given pathway. Number under group names indicate the total number of genes included for the analysis. Numbers on top give the number of genes present within the pathways identified for each group.



4. DISCUSSION

4.1 Peak-callers are not one-size-fit-all tools

In the realm of ChIP-seq data analysis, there are many currently available peak caller tools. The choice can be made considering the DNA-binding nature of the protein, the presence of replicates, of appropriate controls or external set of regions of interest (*Steinhauser et al., 2016*). But in general, this decision is a matter of preferences, since most of the tools are built with flexible parameters to make themselves compatible with various experimental conditions. However, these methods are mostly built to compare 2 conditions at most, and the input protocol requires feeding data one-by-one or pair-by-pair. When handling multiple conditions, most methods usually perform an individual analysis of each sample followed by different means of overlapping the features of each sample. In our opinion, this workflow is quite heavy computationally and also complicates the data handling, since numerous genomic features are excluded and then re-introduced during the overlapping process. Furthermore, the possible pairwise combination between 7 conditions (time points) in our study, each accompanied by a control sample, leads to over 5000 ways to combine the data and may cumulate errors. The in-house method we employed, which process all the data at the same time has the great advantage of being simple, eliminates the issue raised by the necessity to overlap peaks detected in various samples, saves time and computational power while filtering noisy peaks.

Our comparison with MACS, described in chapter 3.1, established that the majority of peaks falsely called by RXR originate from a low RXR/input ratio. Inevitably, one may argue that with the right parameters, MACS would perform better. For example, one could follow the example instructions on MACS github from the author, which gives quite detailed steps to advance the analysis using subcommands. However, depending on the informatics knowledge depth of the investigator, these customized commands can either be too restricted or too detailed. The newer

version of MACS and other peak caller usually have options for paired-end ChIP-seq, which bypass the strand-shift prediction step and takes into account the advantage of whole fragments. However, based on the distribution of peak widths in our study and the presence of high plateau peaks found in our study by our method, we observed that the profile of RXR binding can be narrow peaks (as general transcription factors), cluster of multiple peaks, as well as broad peaks like those found for histone marks; thus, using a MACS parameter to target only one of these instances would be at the detriment of the detection of the others. A study conducted a benchmark analysis to compare between different peak calling algorithms (*Thomas et al., 2017*) and concluded that methods incorporating input tend to perform worse than the others. As we observed during the comparison, MACS bases its prediction from the deviation from the mean signal of a given sample to model the signal distribution, which leads to the false discovery of peaks also present in the input. Our method accesses the true signal ratio of RXR over input within each chromatin bin to avoid this problem. Thus, the advantage of using the input is to give more weight on the filtering steps without impairing the detection of the real peaks. The author from this study (*Thomas et al., 2017*) also showed that peak callers that use windows of various sizes are more powerful than the ones that do not. Due to the relatively wide peak shapes of RXR bindings, the fixed scanning window size used by MACS might also contribute to its shortcomings. Although we also used a fixed window size for the chromatin bins, this is compensated by the fact that we scanned the genome twice with bins shifted by half the bin size, and merged contiguous significant peaks to create longer regions, which is somewhat equivalent to using various window sizes. This process not only overcame the difficulty raised by the presence of various peak sizes and shapes, but also allows to work using the same referential for all samples, thus eliminating the need to reconcile peaks that would be identified in various samples with slightly different coordinates. This again accentuates the role of detecting

different shapes of transcription factors binding, where the concept of narrow TF peaks versus broad histone marks peaks is too overgeneralized and should be used with caution.

In retrospect, in our case, it might have been sufficient to pool the signal of all time points into one sample and use a conventional peak-caller with extra ratio filtering step. However, this method is only valid if the RXR binding signal is mostly constant across timepoints. Since we expected detecting many RXR binding sites where the RXR signal would change between time points, the pooling of all time points to detect peaks would inevitably have favoured the detection of peaks with high signal present at all the time points and missed the peaks presenting binding at only one or two specific time points. We thus devised a method that could detect at a good resolution peaks present at only one time point as well as peaks present in multiple time points without the need to merge or align peaks between timepoints. Whether this method is versatile remains to be benchmarked, but it served well our specific needs as well those from other projects of our lab with different ChIP-seq experiment scheme. So, the development of our method, which took up an important proportion of the work, was necessary, sufficient and successful for our study.

4.2 Association of an RXR peak with its target gene is not perfect

With the well-developed and fitted methods, we have explored in detail the characteristics of RXR binding in mouse liver within the framework of the diurnal rhythm. We used multiple ChIP-seq signals including Pol-II, H3K4me3, H3K36me3 and DNase-I to probe the epigenetic stage, as well as mRNA microarrays to inspect the transcriptional outcome upon RXR bindings. Our main interest was to associate the motifs found under RXR peaks to a specific RXR heterodimer partner, and identify the associated patterns of regulatory signal and or coordinated transcription oscillations.

The function of transcription factors at promoter sites is largely described as recruitment signal for the transcription initiation complex to form (or to be re-cycled) (Cosma, 2002). It is not surprising to find abundance of RXR binding at the promoter of many highly expressed genes in the liver. However, recent epigenomics studies have shown that transcription factors may assert their action on target genes not only by direct contact or by close distance; their influence may reach far beyond their point of DNA binding, for instance via a looping chromatin (MacQuarrie et al., 2011). When trying to link RXR binding and gene activation, 4987 genes—to which at least one RXR peak was associated—were not detected in our mRNA microarray (Fig.13A). One possible explanation could be that RXR represses these genes. However, approximately 1000 of these 4987 genes were pseudo-genes, which could hardly be considered evidence for RXR “silencing” function. Albeit we cannot rule out the possibility that some of these associated genes were not probed by the microarray, the simplest explanation is that we used the conventional genome-wide annotation system, that associates peaks to the nearest gene promoter. Because of the chromatin looping mentioned above, the RXR peaks detected on the gene body (intragenic) or very further away (intergenic, upstream gene start) might not directly control the expression of the nearest promoter, but rather regulate different gene(s). On the other side, around eight thousand genes that expressed without RXR binding sites (Fig.13A) might really be under influence of RXR trans-regulation. The mechanism of gene regulation via transcription factor binding to promoter is well established. However, the complexity of regulatory mechanism is also highlighted by evidences of enhancers, silencers, not to mention the emerging trans-regulation mechanism. Hence this very simplified annotation method has obvious shortcomings.

4.3 Chromatin – regulators organization for co-expressed/co-regulated genes

In quest of exploring the mechanism underlying transcription factors binding sites architecture regulation on transcription, many studies have focused on the structure of cis-regulatory elements specially around promoters. One of the most commonly described structure is the homotypic cluster, a group of adjacent binding sites for the same transcription factor. These “homotypic clustering” of binding sites has been demonstrated in many organisms (bacteria (*Ezer et al., 2014b, Hermsen et al., 2006*), fruit fly (*Lifanov et al., 2003*), human (*Sinha et al., 2008, Gotea et al., 2010*) and mouse, rat (*Zhang et al., 2006*) genomes). Furthermore, as mentioned earlier, several studies have reported the “hotspots” where several ChIP-seq binding sites of multiple transcription factors overlapped. This contribution to transcription regulation of the order and spacing of transcription factors binding sites has been demonstrated by several synthetic studies (*Cox et al., 2007, Gertz et al., 2009, Sharon et al., 2012, Smith et al., 2013*). This is consistent with the trend of modeling gene regulation considering not only the regulatory sequence but also the architecture and abundance of TFs (*Ezer et al., 2014a*). Similarly, a study (*Cremona et al., 2015*) demonstrated a small percentage of binding loci that were extensively shared by many transcription factors with very high tag density occupancy. The target genes of these loci are expressed in a highly transcription factors binding-dependent manner. With the promiscuous nature of RXR and our observation of RXR peaks clustered densely around promoter-rich regions, it is plausible to consider the intersection between the “homotypic clustering” and the “HOT-spot” concepts. This leads to two hypotheses: first, this kind of binding is a property of every transcription factor, or at least of those which act as ancillary factors requiring association with other transcription factors. The second hypothesis is that these bindings represent the “signature sites” of each cell type, where

multiple transcription factors share their roles in regulation. The two scenarios may co-exist or may exclusively exist. It may be a general principle, but what we could observe in our case, is that the presence of cluster of RXR peaks increases the gene transcription.

The concept of nucleus compartmentalization showed that similar functional genes are arranged together in the three-dimensional space and share the transcription regulation machinery. Corresponding to cellular energy conservation, transcription factors converge via weak protein-protein interaction (potentially by low-complexity domains), their high concentrations act in reciprocity and in competition to regulate transcription activity of the local chromatin. This modern transcription factors hub model is consistent with the presence of regions which bear highly overlapped ChIP-seq signals, from which stemmed the definition of HOT-spot. RXR with its promiscuous heterodimer capability would thus have its share upon these magnetizing spots. The role of RXR in recruiting and regulating transcription activity goes beyond the concentration-dependent kinetics. The ability of RXR to selectively recruit genes to the transcription factory hub was observed through the difference of RXR signal amplitudes in promoter-clusters of highly expressed genes.

The studies of co-regulated genes have uncovered the structure of transcription factories, where transcription factors play pivotal role. Whether factories pre-assemble before active genes migrate in (*Mitchell and Fraser, 2008, Ferrai et al., 2010*), or the genes create the docking sites for dynamic factories to assemble (*Darzacq et al., 2007, Yao et al., 2007*) is still debated with contrasted evidences. Nevertheless, it has been evidenced that transcription factors are specialized for some factories (*Xu and Cook, 2008, Schoenfelder et al., 2010*). With binding of specific transcription factors, genes gain access to the transcription factories. Thus, with multiple bindings of RXR, the target genes are tethered more efficiently at the factories, maintaining the open

chromatin conformation, increasing frequency of transcription activation and initiation, which in turn results in higher produced mRNA. RXR could be bound tightly within a stable factory, the DNA will go through this factory and be extruded. The RXR DNA binding will eventually be disrupted as the initial binding site goes away, and the RXR DNA binding domain may randomly land on other part of the DNA being extruded and be crosslinked at that time. This concept could explain the special profile of the high plateaus.

4.4 Pol-II: the unexpected expected journey

It has been established that some transcription factors exert their action not by binding to chromatin but to RNA polymerase (*Zenkin and Yuzenkova, 2015, Stepanova et al., 2009*). Also the kinetics of formation from the initial transcription complex to the productive transcription complex is shown to be highly dynamic (*Engel et al., 2018*), involving multiple players, co-activator, inhibitor, etc. Moreover, one of the crucial kinetic of nuclear receptors is to attract co-regulators: coactivators and corepressors at the promoter sites of target genes. However, the high percentage (38%) of RXR bound at promoter we found within our data does not match with the proportion reported by other studies. A ChIP-seq study of PPAR:RXR heterodimer study reported only 10% of peaks located from TSS to 5kb upstream in mouse adipocyte (*Nielsen et al., 2008*), the same percentage was reported in mouse liver for LXR:RXR binding sites (*Boergesen et al., 2012*). Without giving percentage another ChIP-seq study of VDR:RXR heterodimer stated that the bindings can localize more frequently at multiple sites many kilobases from target gene promoters than the exact promoters (*Pike et al., 2010*); the same observation was reported for RXR binding in macrophages (*Daniel et al., 2014*). The only ChIP-seq study emphasizing the high proportion of RXR binding in promoters, similar to our study was reported on RAR:RXR in F9 embryonal carcinoma cells (*Chatagnon et al., 2015*). Although different cell lines, tissues, antibodies used for

different studies can contribute to the discordant percentages, this high promoter attribution is still an interesting phenomenon. Another aspect of RXR binding that we did not explore in our study is the effect of possible ligands. As we listed in the Introduction, the presence of endogenous RXR ligand in vivo is still debatable. Lefebvre showed that dephosphorylation decreases the DNA binding of RXR α (Lefebvre et al., 1995). Another study later highlighted that, upon retinoid acid treatment, RAR α recruitment to the target gene promoters was increased and this phenomenon was controlled by phosphorylation cascades initiated by the rapid RA activation of the p38MAPK/MSK1 pathway (Bruck et al., 2009). The relationship with the RA signaling pathway to this promoter targeting and the high promoter percentage in RAR:RXR binding may indicate that these association is specific for the heterodimer of RXR with RAR. However, we did not observe any preference toward RAR motif in promoter peaks within our data set. This later study also demonstrated the co-recruitment of general transcription factor TFIID with RXR α to the promoters. This could contribute to the accumulation of RXR at promoters and particularly at TSS with its high correlation with Pol-II occupancy, that we observed in our data. The co-occurrence of Pol-II and RXR could be explained by the possibility of a crosslink of RXR and Pol-II during immunoprecipitation steps, or of a cross-reaction of the antibody. However, we identified several promoter enriched regions with Pol-II but devoid of RXR signal (Fig.16B), which would in theory exclude a systematic cross-reaction of the antibody, although in fairness, the regions enriched with Pol-II but devoid of RXR signal could also be explained by a batch effect since different sets of mice were used for RXR and Pol-II immunoprecipitations. That being said, inspecting the 27 amino acids sequence used to raise our RXR antibody, did not reveal any similarity with any proteins present in the transcription machinery that could have elicited a cross-reaction.

Many studies have demonstrated great overlapping between RXR and Pol-II, or a high correlation between RXR signal and Pol-II signal (*Nielsen et al., 2008, Daniel et al., 2014, Kusters et al., 2013, Boergesen et al., 2012*). This relationship was never taken as anything else other than as evidence of RXR transcription regulatory ability due to the assumption that determination of the activity of Pol-II is a reliable indicator of transcriptional activation (*Koch and Andrau, 2011, Wang et al., 2011, Bonn et al., 2012*). But is there more than just correlation of the two very dependent transcription drivers? Beside the theory of randomly crosslinked factors at the site where DNA is being extruded, is it possible that RXR actually binds to, or at least attaches somehow to the productive transcription machinery in action? Or the signal of RXR along the gene bodies simply resulted from RXR probing the chromatin and the transcription machinery on its own act? The coincidence theory is becoming unlikely if we consider how synchronized the oscillation of signal between RXR and Pol-II were. Initial inspection of highly transcribed genes showed that less than half of RXR binding have similar binding pattern along the genes, from TSS to PAS with the pattern of Pol-II. During the process of clustering PPAR-RXR target genes, a study identified two classes of genes: first class with decreased Pol-II during adipocyte differentiation including genes involving in cell cycle and cell proliferation; second class with increased Pol-II during adipocyte differentiation including genes involving in lipid and glucose metabolism signaling pathways (*Nielsen et al., 2008*). These two classes are remarkably similar to what we observed in our study: different pathway enrichment analyses have proven that there were clearly distinct functions for the genes where RXR tag along Pol-II, versus those where RXR and Pol-II are uncorrelated (**Fig.32**). Which factor(s) determine(s) when and where RXR attaches to Pol-II and what is the function of such behavior are still to be discovered. One possible theory is that RXR is bound and poised at promoters of crucial, constantly high expressed genes, necessary for cellular survival functions such

as cell-fate/cell-phase check-points, apoptosis, etc. In contrast, other genes involved in adaptation to daily metabolism are governed by different manner that would not need such constant “pre-activated” state.

4.5 What is the function of high plateaus

Several algorithms have been developed for ChIP-seq peak shape identification, mainly for the purpose of merging positive and negative strands peaks produced by single-ended ChIP-seq, or for the comparison between different ChIP-seq marks or cell lines (*Hower et al., 2011, Mendoza-Parra et al., 2013, Schweikert et al., 2013, Rezaeian and Rueda, 2014, Strino and Lappe, 2016*). Nevertheless, few studies have pointed out different biological significance governed by the different shapes of transcription factor bindings (*Cremona et al., 2015*). Hence, it is astounding that a profile similar to our high plateaus has never been reported so far for any transcription factors. The observation of the phenomenon of plateau signal of Pol-II was mentioned in subfigure from an earlier study, which they termed the “waviness” and contributed them to presence of numerous repeat sequences scattered around the genome and is commonly observed in profiles of Pol-II or histone marks (*Nielsen et al., 2008*). This profile may contribute to too small bindings percentage of most ChIP-seq studies and thus usually are discarded as outliers; or the peak-callers algorithms tend to cut around the tops of these plateaus and report them as individual normal peaks. These are possible explanations for why this profile is still “flying under the radar” within ChIP-seq studies. Nevertheless, the observation of the high plateaus definitely brought more supporting evidence for the linkage of RXR with the Pol-II machinery. The usually observed profile of Pol-II machine is described as follow: highly accumulated at the promoters during the poise state when transcription machinery is assembled, then slightly stretched out on the gene body during transcript elongation, and finally a second accumulation site after the PAS when the machinery pauses for 3'-end

processing factor recruitment. The RXR signal mimics precisely this action of Pol-II along the gene. For the short high plateau genes, the tight involvement of RXR within the transcription factory discussed above can explain this imitation profile.

Interestingly, using the genome-wide ChIP-seq track provided in [GSM864674](#) data set where they used sc-553 RXR α antibody from Santa Cruz on female mice (*Boergesen et al., 2012*), we could observe similar high plateau profiles on several high plateau genes detected in our study. The presence of high plateau profiles in another study using a different RXR antibody further supports the conclusion that these plateaus were not the product of a cross-reaction with Pol-II. In more details, the high plateaus can be found on the groups of some fibrinogen and apolipoprotein genes. However, on the serpin and cytochrome genes, this profile was inconsistent with our data. As described by (*Cai et al., 2003*), the transcription of cytochrome genes family has been shown to be RXR α -dependent in male mice. The RXR α genome-wide profile, using LG268 as ligand, was reported to be sexually dimorphic and to determine male-female differences in the expression of hepatic lipid processing genes in mice (*Kosters et al., 2013*). The use of LG268 ligand was shown to increase RXR binding sites in female mice, as well as the expression of many of these “newly” target genes (including *Pnpla3* and *Elovl6*). Given that Boergesen used female mice for their study (*Boergesen et al., 2012*), and that we used male mice, the inconsistency of high plateaus on cytochromes genes between our studies could be due to the sex-dimorphism. Of course, further experiments must be conducted to test this theory, but our list of high plateau genes might act as guideline for future RXR (or other nuclear receptors) target gene studies to focus their search.

4.6 Circadian rhythms

Although many circadian studies have been experimented before, and many have established the rhythmic landscape of liver epigenetic and transcriptional activity (*Wager-Smith and Kay, 2000, Akhtar et al., 2002, Panda et al., 2002, Storch et al., 2002, Ueda et al., 2002, Yang et al., 2007, Yang et al., 2006, Bozek et al., 2010, Koike et al., 2012, Hughes et al., 2009*), there has been no golden path for biologic rhythms analysis. A guideline provided by Hughes group can be considered as initial practices consensus to analyze biologic rhythmic data (*Hughes et al., 2017*). There is large discordance in findings of which genes are considered cycling between different studies. This depends on many experimental factors, such as the stability of mRNA, the time sampling resolution, and also on the analysis method, for example which harmonic regression or threshold is being used. The conclusion of rhythmic cycling of biologic data can shift drastically between studies. For example, the circadian rhythms of three isoforms of RXR genes reported by a study using real-time PCR of the same mouse line with similar age and gender showed only RXR α to be rhythmic and RXR β,γ were deemed non-rhythmic (*Yang et al., 2006*); while in our microarray data set, RXR γ has the most significant expression level change through time compared to the other two. Extending the cycling analysis to the realm of binding signal of transcription factors is a big reach. Not dwelling into the fact that there is not much correspondence between the strength of transactivation effect with the amplitude of TF signal at binding sites, the distributions of signal amplitude oscillation of different DNA-binding proteins never follows the same normal distribution. Thus, applying the same false discovery rate threshold as 0.05 may not be the optimal choice. Despite the best efforts made to standardize the biological data of time-wise analyses, we remain dependent on increasing replicates and sampling time points to minimize and compensate the errors raised by analysis methods. Given the lack of replicates and the relatively low time

sampling resolution of our study scheme, our circadian results could not fully represent the full image of RXR circadian and ultradian rhythms in mouse liver. Nevertheless, we did observe in our data some overlapping pathways with the liver circadian and ultradian transcriptome reported by a study from M.E. Hughes (*Hughes et al., 2009*). Also, the occurrence of different cycling patterns might only be expressed under certain conditions, especially for the ultradian (*van der Veen et al., 2006*). Hence, experiments with alternative perturbations on one or more liver cycles may bring forth more information on the circadian/ultradian regulation of RXR. Although the report from Hughes *et al.* proposed that the 8 and 12 hours ultradian might be the result of combination of opposite-phase circadian cycles in cultured cell, and not be intrinsic for cellular functions (*Hughes et al., 2009*), more and more later studies have proven the autonomous of ultradian rhythm with respect to cellular clock-maintaining functions (*Isomura and Kageyama, 2014, Blum et al., 2014, Guzmán et al., 2017, Veen and Gerkema, 2017*).

A recent study showed the interconnection of the cellular clock and nutrient cycle, where more than 70% of the cycling mouse liver transcriptome loses rhythmicity under arrhythmic feeding (food provided in day time only) (*Greenwell et al., 2019*). Interestingly, *Rxra* expression rhythm became arrhythmic with the arrhythmic food intake, while *Rxrg* normally arrhythmic expression became rhythmic under arrhythmic feeding, and *Rxrb* expression rhythm became arrhythmic even when the mice had food access all day. The three *Rxr* genes were shown to be clock-independent (*Wang et al., 2018*), while in different study *Rxrg* was shown to be among the genes that lose rhythm in *Bmal1*^{-/-} mice (*Sobel et al., 2017*). Regardless of how the RXR genes change their expression rhythms, it is clear that the perturbation of the cellular clock and of the food intake cycle affects the rhythmic expression of RXR isoforms. This emphasizes the importance of experiment

control and normalization for any comparison of RXR regulation effects between different conditions.

Thus far, we have seen evidences of static RXR binding and of the involvement of RXR within the transcription machinery. This concept is on apparent contradiction with the well-known hit-and-run mechanism generally described for transcription factors. However, the question of the ligand effect must be brought up here. The association of RXR with ligand has been discussed as a mean to increase the number of RXR binding sites. This indicates that the kinetic of RXR activity changes with the presence of ligand. The underlying mechanism was explored in a study using single cell imaging on HeLa cells (*Brazda et al., 2014*) which showed that the fraction of RXRs static chromatin-bound population immediately and reversibly increased from 15% to 43% upon agonist LG100268 treatment. This redistribution was observed throughout the nucleus and involved a major contribution from coactivator binding. These results indicated that RXR has a distinct, highly dynamic nuclear behavior and follows hit-and-run kinetics which is reduced by LG100268. So, although our study probed the in vivo status of RXR binding on a tissue rather than a cell line, it is evident that each static RXR binding site has the innate potential for dynamic change. Whether or not we agree with the existence of endogenous ligand for RXR, if RXR binding kinetic changes with induced retinoids, we might anticipate that the same change may occur with some metabolites, who might act as endogenous RXR ligands. Given the fluctuation of metabolites during the cycle of food intake, it is possible that some of RXR static bindings might turn to dynamic bindings in rhythmic fashion. We observed that the high tag density level of RXR observed at the promoters, especially at TSS, did not change through time as much as in other genomic regions. Thus, the different response of reduced RXR mobility to LG100268 in (*Brazda et al., 2014*) may correspond to the static signal of RXR bound at gene promoters. The effect of this ligand on RXR was further

explored in angiogenic activity in macrophages (*Daniel et al., 2014*) where it was found that the RXR cistrome binding sites were not impacted by ligand. Among these binding sites, a subset of liganded RXR-bound sites presented PU.1 binding, an increase of enhancer RNA, and P300 recruitment, thus allowing to assign these sites to active enhancers. These sites were shown to predominantly reside in CTCF/cohesin-limited functional domains, and were validated using chromosome conformation capture (3C) and 3C combined with sequencing (3C-seq). These enhancers were shown to communicate with promoters via stable or RXR-induced loops and some of these enhancers interact with each other, forming an inter-chromosomal network, and provides the macrophage with a novel inducible program. This finding is concordant with several studies where they showed RXR may modulate the accessibility of DNA (*Martens et al., 2010, Menendez-Gutierrez and Ricote, 2017*). Such studies that established pioneer-like function of RXR as well as a recent study showing the same regulation mechanism for the core pacemaker CLOCK/BMAL1 (*Trott and Menet, 2018*), make plausible the hypothesis that RXR function as a circadian keeper with the same permissive but pilot role of priming the target genes for flexible stimulations. Hence, it is tempting to postulate that the mobility of RXR play part in the modulation of circadian management to ensure a flexible daily metabolism adaptation.

4.7 The heterodimer partners: public playground (with rules)

With the single-ended ChIP-seq, only a short section at one end of the immunoprecipitated DNA fragment is sequenced, which makes it impossible to know the exact size of the immunoprecipitated fragment. The peak-caller methods have added computational steps to calculate the mean of the true lengths of such fragments, and use it to shift and combine reads from both DNA strands (*Landt et al., 2012*) in order to assign the signal to the DNA region likely associated with the protein binding site (example tool <https://ccg.epfl.ch/chipseq/>). However, since

the shift is global, this presupposes that the average fragment length at each site is the same, which is not necessarily the case. The paired-end sequencing technology resolved those ambiguities and provided us with precise, high resolution binding sites, an advantage that we exploited for our motif analysis. Compared to ChIP-exo, which allows high resolution binding sites detection but requires a complex procedure, our exploitation of paired-end ChIP-seq is more practical, putting weight on downstream dedicated analysis rather than putting weight on the precision of laboratory procedure.

The ChIP-seq study using RXR antibody (in conjunction with PPAR γ) in 3T3-L1 cells showed that the shared percentage of genome-wide profiles jointly occupied by RXR and PPAR γ increased during adipocyte differentiation (*Nielsen et al., 2008*). However, the shared target genes analysis mainly showed concordance with the PPAR lipid regulation function. The later genome-wide profiling study for RXR α along with LXR and PPAR α in mouse liver, exploited the effect of bexarotene – a retinoid selectively activating RXR (*Boergesen et al., 2012*). Even though they observed extensive overlap of binding sites between these factors, they also showed relatively selective binding of PPAR α -RXR to DR1 and of LXR-RXR to DR4. RXR binding sites in mouse liver showed little overlap with bindings sites in 3T3-L1 from Nielsen study, indicating a different repertoire of RXR regulated targets between the liver and this cell line. They showed the capacity of LXR-agonist T0901317 to induce PPAR-associated genes involved in lipogenesis and fatty acid oxidation, and to repress PPAR-associated genes involved in lipoprotein metabolism. This suggested that a cross talk exists between hepatic LXR- and PPAR-regulated pathways. A different profiling of RXR coupled with that of RAR α was designed, in the presence of all-trans RA as a ligand for RAR, in acute promyelocytic leukemia (APL) cell line and in two APL primary blasts (*Martens et al., 2010*). They found great colocalization of PML-RAR α with RXR, and showed that all-trans RA treatment induces changes in H3 acetylation, but not H3K27me3, H3K9me3, or DNA

methylation at the PML-RAR α /RXR binding sites or at nearby target genes. This suggested that PML-RAR α /RXR carry important functions as a local chromatin modulator for hematopoietic differentiation, RAR signaling, and epigenetic control. Profiling of RXR in heterodimer with RAR were later described again in mouse embryonic cells, also using all-trans RA as ligand (*Chatagnon et al., 2015*). They showed that, similar to the finding reported by Nielson (*Nielsen et al., 2008*), RXR binding sites were found to be more constant during cell differentiation than binding sites of its partner RAR. The shared binding sites of RAR/RXR, however changed significantly. In undifferentiated cells, these regions were characterized by binding of pluripotency-associated factors and a high prevalence of the non-canonical DR0 response element. In differentiated cells, these regions are enriched in functional Sox17 binding sites and are characterized with a higher frequency of the canonical DR5 motif. These suggested that the RAR/RXR action is mediated via two different sets of regulatory regions tightly associated with cell differentiation status. A quite different heterodimer partner of RXR, the constitutive androstane receptor (CAR) was studied recently. CAR-linked genes were showed to be involved in not only the main known function of detoxification, but also in regulating lipid, carbohydrate, and energy metabolism (*Tian et al., 2018*). They also found enriched DR4 motif for the CAR:RXR bound sites. Similar to Nielsen study, they also found certain degree of occurrence of C/EBP motif within the shared binding sites.

As described above, the multiple functions of RXR has been thus far overshadowed by the shared pathways with its partners, to the extent that some researchers called RXR a “silent” partner (*Castillo et al., 2004, Botling et al., 1997, Evans and Mangelsdorf, 2014*). Some studies sought to characterize unique RXR profile, but always under effect of ligand inducement (*Kosters et al., 2013, Daniel et al., 2014, Brazda et al., 2014*). However, the motifs and pathways reported in all of these studies that claimed to be specific might be the products of “you see what you seek”. The

most common finding across all RXR ChIP-seq and its partners ChIP-seq studies are the high overlap of not only the known consensus RXR heterodimer partners, but also consensus of other nuclear receptors and transcription factors. This is similar to the extensive overlapping of binding modes and motifs we found within our studies using different analyses methods. The difficulty we encountered while attempting to extrapolate specific binding modes to specific pathways hinted at the complexity of the regulation networks. Future studies that use target gene approaches may take their regulator-target results with a grain of salt, especially for those core metabolomic genes who displayed multiple binding modes in our analysis. Because they are most likely driven not by a single factor, but by multiple factors and layers of regulation. It is likely that there is no central node of regulation as we hoped, but rather a cooperative function, where RXR acts as a gateway/scaffolding member that prime the targets for all the regulators of the consortium. In response to diverse stimuli, transcription factors alter their interactions with RXR to varying degrees, thereby rewiring the regulation network.

References

- ADACHI, S., OKUNO, M., MATSUSHIMA-NISHIWAKI, R., TAKANO, Y., KOJIMA, S., FRIEDMAN, S. L., MORIWAKI, H. & OKANO, Y. 2002. Phosphorylation of retinoid X receptor suppresses its ubiquitination in human hepatocellular carcinoma. *Hepatology*, 35, 332-40.
- ADAM-STITAH, S., PENNA, L., CHAMBON, P. & ROCHETTE-EGLY, C. 1999. Hyperphosphorylation of the retinoid X receptor alpha by activated c-Jun NH2-terminal kinases. *J Biol Chem*, 274, 18932-41.
- ADHIKARY, T., WORTMANN, A., SCHUMANN, T., FINKERNAGEL, F., LIEBER, S., ROTH, K., TOTH, P. M., DIEDERICH, W. E., NIST, A., STIEWE, T., KLEINESUDEIK, L., REINARTZ, S., MULLER-BRUSSELBACH, S. & MULLER, R. 2015. The transcriptional PPARbeta/delta network in human macrophages defines a unique agonist-induced activation state. *Nucleic Acids Res*, 43, 5033-51.
- AHMAD, S., GROMIHA, M. M. & SARAI, A. 2004. Analysis and prediction of DNA-binding proteins and their binding residues based on composition, sequence and structural information. *Bioinformatics*, 20, 477-86.
- AHMAD, S., KESKIN, O., SARAI, A. & NUSSINOV, R. 2008. Protein-DNA interactions: structural, thermodynamic and clustering patterns of conserved residues in DNA-binding proteins. *Nucleic Acids Res*, 36, 5922-32.
- AKHTAR, R. A., REDDY, A. B., MAYWOOD, E. S., CLAYTON, J. D., KING, V. M., SMITH, A. G., GANT, T. W., HASTINGS, M. H. & KYRIACOU, C. P. 2002. Circadian Cycling of the Mouse Liver Transcriptome, as Revealed by cDNA Microarray, Is Driven by the Suprachiasmatic Nucleus. *Current Biology*, 12, 540-550.
- ALLHOFF, M., SERÉ, K., CHAUVISTRÉ, H., LIN, Q., ZENKE, M. & COSTA, I. G. 2014. Detecting differential peaks in ChIP-seq signals with ODIN. *Bioinformatics*, 30, 3467-3475.
- AMBROSINI, G., DREOS, R., KUMAR, S. & BUCHER, P. 2016. The ChIP-Seq tools and web server: a resource for analyzing ChIP-seq and other types of genomic data. *BMC Genomics*, 17, 938.
- ARAO, Y. & KORACH, K. S. 2018. The F domain of estrogen receptor alpha is involved in species-specific, tamoxifen-mediated transactivation. *Journal of Biological Chemistry*.
- ARNHOLD, T., TZIMAS, G., WITTFOHT, W., PLONAIT, S. & NAU, H. 1996. Identification of 9-cis-retinoic acid, 9, 13-di-cis-retinoic acid, and 14-hydroxy-4, 14-retro-retinol in human plasma after liver consumption. *Life sciences*, 59, PL169-PL177.
- ASHBURNER, M., CHIHARA, C., MELTZER, P. & RICHARDS, G. 1974. Temporal control of puffing activity in polytene chromosomes. *Cold Spring Harb Symp Quant Biol*, 38, 655-62.
- BAIN, D. L., HENEGHAN, A. F., CONNAGHAN-JONES, K. D. & MIURA, M. T. 2007. Nuclear Receptor Structure: Implications for Function. *Annual Review of Physiology*, 69, 201-220.
- 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, 347-54.
- BASTIEN, J., ADAM-STITAH, S., PLASSAT, J.-L., CHAMBON, P. & ROCHETTE-EGLY, C. 2002. The Phosphorylation Site Located in the A Region of Retinoic X Receptor α 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. & ROCHETTE-EGLY, C. 2004. Nuclear retinoid receptors and the transcription of retinoid-target genes. *Gene*, 328, 1-16.
- BEATO, M. 1991. Transcriptional control by nuclear receptors. *The FASEB Journal*, 5, 2044-2051.
- BEIGNEUX, A. P., MOSER, A. H., SHIGENAGA, J. K., GRUNFELD, C. & FEINGOLD, K. R. 2000. The acute phase response is associated with retinoid X receptor repression in rodent liver. *J Biol Chem*, 275, 16390-9.
- BENEGIAMO, G., MURE, L. S., ERIKSON, G., LE, H. D., MORIGGI, E., BROWN, S. A. & PANDA, S. 2018. The RNA-Binding Protein NONO Coordinates Hepatic Adaptation to Feeding. *Cell Metab*, 27, 404-418.e7.

- BILLAS, I. M., MOULINIER, L., ROCHEL, N. & MORAS, D. 2001. Crystal structure of the ligand-binding domain of the ultraspiracle protein USP, the ortholog of retinoid X receptors in insects. *Journal of Biological Chemistry*, 276, 7465-7474.
- BLEDSON, R. K., MONTANA, V. G., STANLEY, T. B., DELVES, C. J., APOLITO, C. J., MCKEE, D. D., CONSLER, T. G., PARKS, D. J., STEWART, E. L. & WILLSON, T. M. 2002. Crystal structure of the glucocorticoid receptor ligand binding domain reveals a novel mode of receptor dimerization and coactivator recognition. *Cell*, 110, 93-105.
- BLOMHOFF, R. & BLOMHOFF, H. K. 2006. Overview of retinoid metabolism and function. *Journal of Neurobiology*, 66, 606-630.
- BLUM, I. D., ZHU, L., MOQUIN, L., KOKOEVA, M. V., GRATTON, A., GIROS, B. & STORCH, K.-F. 2014. A highly tunable dopaminergic oscillator generates ultradian rhythms of behavioral arousal. *eLife*, 3, e05105.
- BOERGESEN, M., PEDERSEN, T. Å., GROSS, B., VAN HEERINGEN, S. J., HAGENBEEK, D., BINDESBØLL, C., CARON, S., LALLOYER, F., STEFFENSEN, K. R., NEBB, H. I., GUSTAFSSON, J.-Å., STUNNENBERG, H. G., STAELS, B. & MANDRUP, S. 2012. Genome-Wide Profiling of Liver X Receptor, Retinoid X Receptor, and Peroxisome Proliferator-Activated Receptor α in Mouse Liver Reveals Extensive Sharing of Binding Sites. *Molecular and Cellular Biology*, 32, 852-867.
- BOMMER, M., BENECKE, A., GRONEMEYER, H. & ROCHETTE-EGLY, C. 2002. TIF2 mediates the synergy between RAR α 1 activation functions AF-1 and AF-2. *J Biol Chem*, 277, 37961-6.
- BONN, S., ZINZEN, R. P., GIRARDOT, C., GUSTAFSON, E. H., PEREZ-GONZALEZ, A., DELHOMME, N., GHAVI-HELM, Y., WILCZYNSKI, B., RIDDELL, A. & FURLONG, E. E. 2012. Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat Genet*, 44, 148-56.
- BOOKOUT, A. L., JEONG, Y., DOWNES, M., YU, R. T., EVANS, R. M. & MANGELSDORF, D. J. 2006. Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network. *Cell*, 126, 789-99.
- BOTLING, J., CASTRO, D. S., ÖBERG, F., NILSSON, K. & PERLMANN, T. 1997. Retinoic Acid Receptor/Retinoid X Receptor Heterodimers Can Be Activated through Both Subunits Providing a Basis for Synergistic Transactivation and Cellular Differentiation. *Journal of Biological Chemistry*, 272, 9443-9449.
- BOURGUET, W., RUFF, M., CHAMBON, P., GRONEMEYER, H. & MORAS, D. 1995. Crystal structure of the ligand-binding domain of the human nuclear receptor RXR- α . 375, 377-382.
- BOZEK, K., ROSAHL, A. L., GAUB, S., LORENZEN, S. & HERZEL, H. 2010. Circadian transcription in liver. *Eighth International Workshop on Information Processing in Cells and Tissues: "From Small Scale Dynamics To Understanding Systems Behavior"*, 102, 61-69.
- BRAZDA, P., KRIEGER, J., DANIEL, B., JONAS, D., SZEKERES, T., LANGOWSKI, J., TÓTH, K., NAGY, L. & VÁMOSI, G. 2014. Ligand Binding Shifts Highly Mobile Retinoid X Receptor to the Chromatin-Bound State in a Coactivator-Dependent Manner, as Revealed by Single-Cell Imaging. *Molecular and Cellular Biology*, 34, 1234-1245.
- BRUCK, N., VITOUX, D., FERRY, C., DUONG, V., BAUER, A., DE THÉ, H. & ROCHETTE-EGLY, C. 2009. A coordinated phosphorylation cascade initiated by p38MAPK/MSK1 directs RAR α to target promoters. *The EMBO journal*, 28, 34-47.
- BRZOZOWSKI, A. M., PIKE, A. C., DAUTER, Z., HUBBARD, R. E., BONN, T., ENGSTROM, O., OHMAN, L., GREENE, G. L., GUSTAFSSON, J. A. & CARLQUIST, M. 1997. Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature*, 389, 753-8.
- BUGGE, A., GRONTVED, L., AAGAARD, M. M., BORUP, R. & MANDRUP, S. 2009. The PPAR γ 2 A/B-domain plays a gene-specific role in transactivation and cofactor recruitment. *Mol Endocrinol*, 23, 794-808.
- BUGGE, T. H., POHL, J., LONNOY, O. & STUNNENBERG, H. G. 1992. RXR α , a promiscuous partner of retinoic acid and thyroid hormone receptors. *The EMBO journal*, 11, 1409-1418.
- BUNGER, M. K., WILSBACHER, L. D., MORAN, S. M., CLENDENIN, C., RADCLIFFE, L. A., HOGENESCH, J. B., SIMON, M. C., TAKAHASHI, J. S. & BRADFIELD, C. A. 2000. Mop3 Is

- an Essential Component of the Master Circadian Pacemaker in Mammals. *Cell*, 103, 1009-1017.
- BURGERMEISTER, E., CHUDERLAND, D., HANOCH, T., MEYER, M., LISCOVITCH, M. & SEGER, R. 2007. Interaction with MEK causes nuclear export and downregulation of peroxisome proliferator-activated receptor gamma. *Mol Cell Biol*, 27, 803-17.
- CAI, Y., DAI, T., AO, Y., KONISHI, T., CHUANG, K. H., LUE, Y., CHANG, C. & WAN, Y. J. 2003. Cytochrome P450 genes are differentially expressed in female and male hepatocyte retinoid X receptor alpha-deficient mice. *Endocrinology*, 144, 2311-8.
- CAI, Y., KONISHI, T., HAN, G., CAMPWALA, K. H., FRENCH, S. W. & WAN, Y. J. 2002. The role of hepatocyte RXR alpha in xenobiotic-sensing nuclear receptor-mediated pathways. *Eur J Pharm Sci*, 15, 89-96.
- CALDERON, F. & KIM, H.-Y. 2007. Role of RXR in neurite outgrowth induced by docosahexaenoic acid. *Prostaglandins, Leukotrienes and Essential Fatty Acids*, 77, 227-232.
- CANELLA, D., BERNASCONI, D., GILARDI, F., LEMARTELOT, G., MIGLIAVACCA, E., PRAZ, V., COUSIN, P., DELORENZI, M., HERNANDEZ, N. & CONSORTIUM, T. C. 2012. A multiplicity of factors contributes to selective RNA polymerase III occupancy of a subset of RNA polymerase III genes in mouse liver. *Genome Research*, 22, 666-680.
- CARLBERG, C. & SEUTER, S. 2010. Dynamics of nuclear receptor target gene regulation. *Chromosoma*, 119, 479-484.
- CASTELLANA, S., MAZZA, T., CAPOCEFALO, D., GENOV, N., BIAGINI, T., FUSILLI, C., SCHOLKMANN, F., RELOGIO, A., HOGENESCH, J. B. & MAZZOCCOLI, G. 2018. Systematic Analysis of Mouse Genome Reveals Distinct Evolutionary and Functional Properties Among Circadian and Ultradian Genes. *Front Physiol*, 9, 1178.
- CASTILLO, A. I., SÁNCHEZ-MARTÍNEZ, R., MORENO, J. L., MARTÍNEZ-IGLESIAS, O. A., PALACIOS, D. & ARANDA, A. 2004. A permissive retinoid X receptor/thyroid hormone receptor heterodimer allows stimulation of prolactin gene transcription by thyroid hormone and 9-cis-retinoic acid. *Molecular and cellular biology*, 24, 502-513.
- CHANDRA, V., HUANG, P., HAMURO, Y., RAGHURAM, S., WANG, Y., BURRIS, T. P. & RASTINEJAD, F. 2008. Structure of the intact PPAR-gamma-RXR- nuclear receptor complex on DNA. *Nature*, 456, 350-6.
- CHANDRA, V., WU, D., LI, S., POTLURI, N., KIM, Y. & RASTINEJAD, F. 2017. The quaternary architecture of RAR β -RXR α heterodimer facilitates domain-domain signal transmission. 8, 868.
- CHARMANDARI, E., RAJI, A., KINO, T., ICHIJO, T., TIULPAKOV, A., ZACHMAN, K. & CHROUSOS, G. P. 2005. A Novel Point Mutation in the Ligand-Binding Domain (LBD) of the Human Glucocorticoid Receptor (hGR) Causing Generalized Glucocorticoid Resistance: The Importance of the C Terminus of hGR LBD in Conferring Transactivational Activity. *The Journal of Clinical Endocrinology & Metabolism*, 90, 3696-3705.
- CHATAGNON, A., TRIQUENEAUX, G., MORIN, V., BENOIT, G., VEBER, P., D'ALCHÉ-BUC, F., BEDO, J., SÉMON, M. & LAUDET, V. 2015. RAR/RXR binding dynamics distinguish pluripotency from differentiation associated cis-regulatory elements. *Nucleic Acids Research*, 43, 4833-4854.
- CHAWLA, A., REPA, J. J., EVANS, R. M. & MANGELSDORF, D. J. 2001. Nuclear Receptors and Lipid Physiology: Opening the X-Files. *Science*, 294, 1866.
- CHEN, J. D. & EVANS, R. M. 1995. A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature*, 377, 454-7.
- CHEN, L., WANG, C., QIN, Z. S. & WU, H. 2015. A novel statistical method for quantitative comparison of multiple ChIP-seq datasets. *Bioinformatics*, 31, 1889-96.
- CHEN, T., CHANG, T. C., KANG, J. O., CHOUDHARY, B., MAKITA, T., TRAN, C. M., BURCH, J. B., EID, H. & SUCOV, H. M. 2002. Epicardial induction of fetal cardiomyocyte proliferation via a retinoic acid-inducible trophic factor. *Dev Biol*, 250, 198-207.
- CHEN, Y., TAKESHITA, A., OZAKI, K., KITANO, S. & HANAZAWA, S. 1996. Transcriptional Regulation by Transforming Growth Factor β of the Expression of Retinoic Acid and Retinoid X Receptor

- Genes in Osteoblastic Cells Is Mediated through AP-1. *Journal of Biological Chemistry*, 271, 31602-31606.
- CHESKIS, B. & FREEDMAN, L. P. 1996. Modulation of Nuclear Receptor Interactions by Ligands: Kinetic Analysis Using Surface Plasmon Resonance. *Biochemistry*, 35, 3309-3318.
- CHO, H., ZHAO, X., HATORI, M., YU, R. T., BARISH, G. D., LAM, M. T., CHONG, L.-W., DITACCHIO, L., ATKINS, A. R., GLASS, C. K., LIDDLE, C., AUWERX, J., DOWNES, M., PANDA, S. & EVANS, R. M. 2012. Regulation of circadian behaviour and metabolism by REV-ERB-[agr] and REV-ERB-[bgr]. *Nature*, 485, 123-127.
- CHO, Y., NOSHIRO, M., CHOI, M., MORITA, K., KAWAMOTO, T., FUJIMOTO, K., KATO, Y. & MAKISHIMA, M. 2009. The basic helix-loop-helix proteins differentiated embryo chondrocyte (DEC) 1 and DEC2 function as corepressors of retinoid X receptors. *Mol Pharmacol*, 76, 1360-9.
- CHOI, J. H., BANKS, A. S., ESTALL, J. L., KAJIMURA, S., BOSTROM, P., LAZNIK, D., RUAS, J. L., CHALMERS, M. J., KAMENECKA, T. M., BLUHER, M., GRIFFIN, P. R. & SPIEGELMAN, B. M. 2010. Anti-diabetic drugs inhibit obesity-linked phosphorylation of PPARgamma by Cdk5. *Nature*, 466, 451-6.
- CLAYTON, G. M., PEAK-CHEW, S. Y., EVANS, R. M. & SCHWABE, J. W. 2001. The structure of the ultraspiracle ligand-binding domain reveals a nuclear receptor locked in an inactive conformation. *Proceedings of the National Academy of Sciences*, 98, 1549-1554.
- CONNORS, R. V., WANG, Z., HARRISON, M., ZHANG, A., WANSKA, M., HISCOCK, S., FOX, B., DORE, M., LABELLE, M., SUDOM, A., JOHNSTONE, S., LIU, J., WALKER, N. P., CHAI, A., SIEGLER, K., LI, Y. & COWARD, P. 2009. Identification of a PPARdelta agonist with partial agonistic activity on PPARgamma. *Bioorg Med Chem Lett*, 19, 3550-4.
- COSMA, M. P. 2002. Ordered recruitment: gene-specific mechanism of transcription activation. *Mol Cell*, 10, 227-36.
- COX, R. S., SURETTE, M. G. & ELOWITZ, M. B. 2007. Programming gene expression with combinatorial promoters. *Molecular Systems Biology*, 3.
- CREMONA, M. A., SANGALLI, L. M., VANTINI, S., DELLINO, G. I., PELICCI, P. G., SECCHI, P. & RIVA, L. 2015. Peak shape clustering reveals biological insights. *BMC Bioinformatics*, 16, 349.
- DANIEL, B., NAGY, G., HAH, N., HORVATH, A., CZIMMERER, Z., POLISKA, S., GYURIS, T., KEIRSSE, J., GYSEMANS, C., VAN GINDERACHTER, J. A., BALINT, B. L., EVANS, R. M., BARTA, E. & NAGY, L. 2014. The active enhancer network operated by liganded RXR supports angiogenic activity in macrophages. *Genes & Development*, 28, 1562-1577.
- DARZACQ, X., SHAV-TAL, Y., DE TURRIS, V., BRODY, Y., SHENOY, S. M., PHAIR, R. D. & SINGER, R. H. 2007. In vivo dynamics of RNA polymerase II transcription. *Nature Structural & Molecular Biology*, 14, 796.
- DAWSON, M. I. & XIA, Z. 2012. The retinoid X receptors and their ligands. *Biochim Biophys Acta*, 1821, 21-56.
- DE COSMO, S. & MAZZOCCOLI, G. 2017. Retinoid X Receptors Intersect the Molecular Clockwork in the Regulation of Liver Metabolism. *Frontiers in Endocrinology*, 8.
- DE URQUIZA, A. M., LIU, S., SJÖBERG, M., ZETTERSTRÖM, R. H., GRIFFITHS, W., SJÖVALL, J. & PERLMANN, T. 2000. Docosahexaenoic acid, a ligand for the retinoid X receptor in mouse brain. *Science*, 290, 2140-2144.
- DESERGNE, B. 2007. RXR: from partnership to leadership in metabolic regulations. *Vitamins & Hormones*, 75, 1-32.
- DESERGNE, B., A, I. J., DEVCHAND, P. R. & WAHLI, W. 1998. The peroxisome proliferator-activated receptors at the cross-road of diet and hormonal signalling. *J Steroid Biochem Mol Biol*, 65, 65-74.
- DEVARAKONDA, S., HARP, J. M., KIM, Y., OZYHAR, A. & RASTINEJAD, F. 2003. Structure of the heterodimeric ecdysone receptor DNA - binding complex. *The EMBO Journal*, 22, 5827.
- DE LERA, Á. R., KREZEL, W. & RÜHL, R. 2016. An Endogenous Mammalian Retinoid X Receptor Ligand, At Last! *ChemMedChem*, 11, 1027-1037.

- DIRENZO, J., SÖDERSTROM, M., KUROKAWA, R., OGLIASTRO, M.-H., RICOTE, M., INGREY, S., HÖRLEIN, A., ROSENFELD, M. G. & GLASS, C. K. 1997. Peroxisome proliferator-activated receptors and retinoic acid receptors differentially control the interactions of retinoid X receptor heterodimers with ligands, coactivators, and corepressors. *Molecular and cellular biology*, 17, 2166-2176.
- DOLLE, P., FRAULOB, V., KASTNER, P. & CHAMBON, P. 1994. Developmental expression of murine retinoid X receptor (RXR) genes. *Mech Dev*, 45, 91-104.
- DREYER, C., KREY, G., KELLER, H., GIVEL, F., HELFTENBEIN, G. & WAHLI, W. 1992. Control of the peroxisomal β -oxidation pathway by a novel family of nuclear hormone receptors. *Cell*, 68, 879-887.
- DROR, I., ROHS, R. & MANDEL-GUTFREUND, Y. 2016. How motif environment influences transcription factor search dynamics: Finding a needle in a haystack. *BioEssays*, 38, 605-612.
- DUESTER, G., SHEAN, M. L., MCBRIDE, M. S. & STEWART, M. J. 1991. Retinoic acid response element in the human alcohol dehydrogenase gene ADH3: implications for regulation of retinoic acid synthesis. *Mol Cell Biol*, 11, 1638-46.
- DUEZ, H. & STAELS, B. 2010. Nuclear Receptors Linking Circadian Rhythms and Cardiometabolic Control. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 30, 1529-1534.
- ECHCHGADDA, I., SONG, C. S., OH, T., AHMED, M., DE LA CRUZ, I. J. & CHATTERJEE, B. 2007. The Xenobiotic-Sensing Nuclear Receptors Pregnane X Receptor, Constitutive Androstane Receptor, and Orphan Nuclear Receptor Hepatocyte Nuclear Factor 4 α in the Regulation of Human Steroid-/Bile Acid-Sulfotransferase. *Molecular Endocrinology*, 21, 2099-2111.
- EGEA, P. F., MITSCHLER, A. & MORAS, D. 2002. Molecular Recognition of Agonist Ligands by RXRs. *Molecular Endocrinology*, 16, 987-997.
- ENGEL, C., NEYER, S. & CRAMER, P. 2018. Distinct Mechanisms of Transcription Initiation by RNA Polymerases I and II. *Annual Review of Biophysics*, 47, 425-446.
- ESCRIVA, H., DELAUNAY, F. & LAUDET, V. 2000. Ligand binding and nuclear receptor evolution. *Bioessays*, 22, 717-727.
- ESCRIVA, H., SAFI, R., HÄNNI, C., LANGLOIS, M.-C., SAUMITOU-LAPRADE, P., STEHELIN, D., CAPRON, A., PIERCE, R. & LAUDET, V. 1997. Ligand binding was acquired during evolution of nuclear receptors. *Proceedings of the National Academy of Sciences*, 94, 6803-6808.
- EVANS, R. M. & MANGELSDORF, D. J. 2014. Nuclear Receptors, RXR, and the Big Bang. *Cell*, 157, 255-266.
- EZER, D., ZABET, N. R. & ADRYAN, B. 2014a. Homotypic clusters of transcription factor binding sites: A model system for understanding the physical mechanics of gene expression. *Computational and Structural Biotechnology Journal*, 10, 63-69.
- EZER, D., ZABET, N. R. & ADRYAN, B. 2014b. Physical constraints determine the logic of bacterial promoter architectures. *Nucleic Acids Research*, 42, 4196-4207.
- FALASCA, L., FAVALE, A., SERAFINO, A., ARA, C. & CONTI DEVIRGILIIS, L. 1998. The effect of retinoic acid on the re-establishment of differentiated hepatocyte phenotype in primary culture. *Cell Tissue Res*, 293, 337-47.
- FEIGE, J. N., GELMAN, L., MICHALIK, L., DESVERGNE, B. & WAHLI, W. 2006. From molecular action to physiological outputs: peroxisome proliferator-activated receptors are nuclear receptors at the crossroads of key cellular functions. *Prog Lipid Res*, 45, 120-59.
- FEIGE, J. N., GELMAN, L., TUDOR, C., ENGELBORGHES, Y., WAHLI, W. & DESVERGNE, B. 2005. Fluorescence Imaging Reveals the Nuclear Behavior of Peroxisome Proliferator-activated Receptor/Retinoid X Receptor Heterodimers in the Absence and Presence of Ligand. *Journal of Biological Chemistry*, 280, 17880-17890.
- FEINGOLD, K., KIM, M. S., SHIGENAGA, J., MOSER, A. & GRUNFELD, C. 2004. Altered expression of nuclear hormone receptors and coactivators in mouse heart during the acute-phase response. *American Journal of Physiology-Endocrinology and Metabolism*, 286, E201-E207.
- FERRAI, C., XIE, S. Q., LURAGHI, P., MUNARI, D., RAMIREZ, F., BRANCO, M. R., POMBO, A. & CRIPPA, M. P. 2010. Poised Transcription Factories Prime Silent uPA Gene Prior to Activation. *PLOS Biology*, 8, e1000270.

- FRANK, C., MAKKONEN, H., DUNLOP, T. W., MATILAINEN, M., VÄISÄNEN, S. & CARLBERG, C. 2005. Identification of Pregnane X Receptor Binding Sites in the Regulatory Regions of Genes Involved in Bile Acid Homeostasis. *Journal of Molecular Biology*, 346, 505-519.
- FREEDMAN, L. P., LUISI, B. F., KORSZUN, Z. R., BASAVAPPA, R., SIGLER, P. B. & YAMAMOTO, K. R. 1988. The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain. *Nature*, 334, 543-6.
- FU, Y., ZHONG, H., WANG, M. H., LUO, D. G., LIAO, H. W., MAEDA, H., HATTAR, S., FRISHMAN, L. J. & YAU, K. W. 2005. Intrinsically photosensitive retinal ganglion cells detect light with a vitamin A-based photopigment, melanopsin. *Proc Natl Acad Sci U S A*, 102, 10339-44.
- FUJIMOTO, Y., SHIRAKI, T., HORIUCHI, Y., WAKU, T., SHIGENAGA, A., OTAKA, A., IKURA, T., IGARASHI, K., AIMOTO, S., TATE, S. & MORIKAWA, K. 2010. Proline cis/trans-isomerase Pin1 regulates peroxisome proliferator-activated receptor gamma activity through the direct binding to the activation function-1 domain. *J Biol Chem*, 285, 3126-32.
- GAMPE JR, R. T., MONTANA, V. G., LAMBERT, M. H., MILLER, A. B., BLEDSOE, R. K., MILBURN, M. V., KLIEWER, S. A., WILLSON, T. M. & XU, H. E. 2000. Asymmetry in the PPAR γ /RXR α crystal structure reveals the molecular basis of heterodimerization among nuclear receptors. *Molecular cell*, 5, 545-555.
- GAMPE, R. T., MONTANA, V. G., LAMBERT, M. H., WISELY, G. B., MILBURN, M. V. & XU, H. E. 2000. Structural basis for autorepression of retinoid X receptor by tetramer formation and the AF-2 helix. *Genes & Development*, 14, 2229-2241.
- GAO, W., LIU, J., HU, M., HUANG, M., CAI, S., ZENG, Z., LIN, B., CAO, X., CHEN, J., ZENG, J. Z., ZHOU, H. & ZHANG, X. K. 2013. Regulation of proteolytic cleavage of retinoid X receptor-alpha by GSK-3beta. *Carcinogenesis*, 34, 1208-15.
- GERMAIN, P., IYER, J., ZECHEL, C. & GRONEMEYER, H. 2002. Co-regulator recruitment and the mechanism of retinoic acid receptor synergy. *Nature*, 415, 187-92.
- GERMAIN, P., STAELS, B., DACQUET, C., SPEDDING, M. & LAUDET, V. 2006. Overview of Nomenclature of Nuclear Receptors. *Pharmacological Reviews*, 58, 685-704.
- GERMAN, O. L., MONACO, S., AGNOLAZZA, D. L., ROTSTEIN, N. P. & POLITI, L. E. 2013. Retinoid X receptor activation is essential for docosahexaenoic acid protection of retina photoreceptors. *Journal of Lipid Research*, 54, 2236-2246.
- GERTZ, J., SIGGIA, E. D. & COHEN, B. A. 2009. Analysis of combinatorial cis-regulation in synthetic and genomic promoters. *Nature*, 457, 215-8.
- GHYSELINCK, N. B., DUPE, V., DIERICH, A., MESSADDEQ, N., GARNIER, J., ROCHETTE-EGLY, C., CHAMBON, P. & MARK, M. 1997. Role of the retinoic acid receptor beta (RARbeta) during mouse development. *The International journal of developmental biology*, 41, 425-447.
- GHYSELINCK, N. B., WENDLING, O., MESSADDEQ, N., DIERICH, A., LAMPRON, C., DECIMO, D., VIVILLE, S., CHAMBON, P. & MARK, M. 1998. Contribution of retinoic acid receptor beta isoforms to the formation of the conotruncal septum of the embryonic heart. *Dev Biol*, 198, 303-18.
- GIANNI, M., TARRADE, A., NIGRO, E. A., GARATTINI, E. & 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, 34458-66.
- GIGUERE, V., HOLLENBERG, S. M., ROSENFELD, M. G. & EVANS, R. M. 1986. Functional domains of the human glucocorticoid receptor. *Cell*, 46, 645-52.
- GIGUERE, V., ONG, E. S., SEGUI, P. & EVANS, R. M. 1987. Identification of a receptor for the morphogen retinoic acid. *Nature*, 330, 624-9.
- GILARDI, F., MIGLIAVACCA, E., NALDI, A., BARUCHET, M., CANELLA, D., LE MARTELOT, G., GUEx, N., DESVERGNE, B. & THE CYCLI, X. C. 2014. Genome-Wide Analysis of SREBP1 Activity around the Clock Reveals Its Combined Dependency on Nutrient and Circadian Signals. *PLoS Genet*, 10, e1004155.
- GLYNN, E. F., CHEN, J. & MUSHEGIAN, A. R. 2005. Detecting periodic patterns in unevenly spaced gene expression time series using Lomb-Scargle periodograms. *Bioinformatics*, 22, 310-316.

- GOLDSTEIN, J. T., DOBRZYN, A., CLAGETT-DAME, M., PIKE, J. W. & DELUCA, H. F. 2003. Isolation and characterization of unsaturated fatty acids as natural ligands for the retinoid-X receptor. *Archives of biochemistry and Biophysics*, 420, 185-193.
- GOTEA, V., VISEL, A., WESTLUND, J. M., NOBREGA, M. A., PENNACCHIO, L. A. & OVCHARENKO, I. 2010. Homotypic clusters of transcription factor binding sites are a key component of human promoters and enhancers. *Genome Research*, 20, 565-577.
- GRANT, C. E., BAILEY, T. L. & NOBLE, W. S. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics*, 27, 1017-1018.
- GREEN, S. & CHAMBON, P. 1987. Oestradiol induction of a glucocorticoid-responsive gene by a chimaeric receptor. *Nature*, 325, 75-8.
- GREEN, S., WALTER, P., KUMAR, V., KRUST, A., BORNERT, J. M., ARGOS, P. & CHAMBON, P. 1986. Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A. *Nature*, 320, 134-9.
- GREENWELL, B. J., TROTT, A. J., BEYTEBIERE, J. R., PAO, S., BOSLEY, A., BEACH, E., FINEGAN, P., HERNANDEZ, C. & MENET, J. S. 2019. Rhythmic Food Intake Drives Rhythmic Gene Expression More Potently than the Hepatic Circadian Clock in Mice. *Cell Reports*, 27, 649-657.e5.
- GRONEMEYER, H., GUSTAFSSON, J. A. & LAUDET, V. 2004. Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov*, 3, 950-64.
- GUNDERSEN, T. E., BASTANI, N. E. & BLOMHOFF, R. 2007. Quantitative high-throughput determination of endogenous retinoids in human plasma using triple-stage liquid chromatography/tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, 21, 1176-1186.
- GUPTA, S., STAMATOYANNOPOULOS, J. A., BAILEY, T. L. & NOBLE, W. S. 2007. Quantifying similarity between motifs. *Genome Biology*, 8, R24.
- GUZMÁN, D. A., FLESIA, A. G., AON, M. A., PELLEGRINI, S., MARIN, R. H. & KEMBRO, J. M. 2017. The fractal organization of ultradian rhythms in avian behavior. *Scientific Reports*, 7, 684.
- HEINZ, S., BENNER, C., SPANN, N., BERTOLINO, E., LIN, Y. C., LASLO, P., CHENG, J. X., MURRE, C., SINGH, H. & GLASS, C. K. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*, 38, 576-89.
- HELSEN, C. & CLAESSENS, F. 2014. Looking at nuclear receptors from a new angle. *Mol Cell Endocrinol*, 382, 97-106.
- HENRICH, V. C., SIITER, T. J., LUBAHN, D. B., MACLINTYRE, A. & GILBERT, L. I. 1990. A steroid/thyroid hormone receptor superfamily member in *Drosophila melanogaster* that shares extensive sequence similarity with a mammalian homologue. *Nucleic acids research*, 18, 4143-4148.
- HERMSEN, R., TANS, S. & TEN WOLDE, P. R. 2006. Transcriptional Regulation by Competing Transcription Factor Modules. *PLOS Computational Biology*, 2, e164.
- HEYMAN, R. A., MANGELSDORF, D. J., DYCK, J. A., STEIN, R. B., EICHELE, G., EVANS, R. M. & THALLER, C. 1992. 9-cis retinoic acid is a high affinity ligand for the retinoid X receptor. *Cell*, 68, 397-406.
- HJORTH, E., ZHU, M., TORO, V. C., VEDIN, I., PALMLAD, J., CEDERHOLM, T., FREUND-LEVI, Y., FAXEN-IRVING, G., WAHLUND, L.-O. & BASUN, H. 2013. Omega-3 Fatty Acids Enhance Phagocytosis of Alzheimer's Disease-Related Amyloid- β 42 by Human Microglia and Decrease Inflammatory Markers. *Journal of Alzheimer's Disease*, 35, 697-713.
- HONG, C. & TONTONOZ, P. 2014. Liver X receptors in lipid metabolism: opportunities for drug discovery. *Nat Rev Drug Discov*, 13, 433-444.
- HOOPES, C. W., TAKETO, M., OZATO, K., LIU, Q., HOWARD, T. A., LINNEY, E. & SELDIN, M. F. 1992. Mapping of the mouse Rxr loci encoding nuclear retinoid X receptors RXR alpha, RXR beta, and RXR gamma. *Genomics*, 14, 611-7.
- HORTON, C. & MADEN, M. 1995. Endogenous distribution of retinoids during normal development and teratogenesis in the mouse embryo. *Developmental Dynamics*, 202, 312-323.

- HOSODA, K., SATO, M. & YANAI, K. 2015. Identification and Characterization of Human Genomic Binding Sites for Retinoic Acid Receptor/Retinoid X Receptor Heterodimers. *Advances in Biological Chemistry*, Vol.05No.02, 15.
- Houben, T., Coomans, C. P. & Meijer, J. H. 2014. Regulation of Circadian and Acute Activity Levels by the Murine Suprachiasmatic Nuclei. *PLOS ONE*, 9, e110172.
- HOWER, V., EVANS, S. N. & PACTER, L. 2011. Shape-based peak identification for ChIP-Seq. *BMC Bioinformatics*, 12, 15.
- HUANG, P., CHANDRA, V. & RASTINEJAD, F. 2010. Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics. *Annu Rev Physiol*, 72, 247-72.
- HUGHES, M. E., ABRUZZI, K. C., ALLADA, R., ANAFI, R., ARPAT, A. B., ASHER, G., BALDI, P., DE BEKKER, C., BELL-PEDERSEN, D., BLAU, J., BROWN, S., CERIANI, M. F., CHEN, Z., CHIU, J. C., COX, J., CROWELL, A. M., DEBRUYNE, J. P., DIJK, D. J., DITACCHIO, L., DOYLE, F. J., DUFFIELD, G. E., DUNLAP, J. C., ECKEL-MAHAN, K., ESSER, K. A., FITZGERALD, G. A., FORGER, D. B., FRANCEY, L. J., FU, Y. H., GACHON, F., GATFIELD, D., DE GOEDE, P., GOLDEN, S. S., GREEN, C., HARER, J., HARMER, S., HASPEL, J., HASTINGS, M. H., HERZEL, H., HERZOG, E. D., HOFFMANN, C., HONG, C., HUGHEY, J. J., HURLEY, J. M., DE LA IGLESIA, H. O., JOHNSON, C., KAY, S. A., KOIKE, N., KORNACKER, K., KRAMER, A., LAMIA, K., LEISE, T., LEWIS, S. A., LI, J., LI, X., LIU, A. C., LOROS, J. J., MARTINO, T. A., MENET, J. S., MERROW, M., MILLAR, A. J., MOCKLER, T., NAEF, F., NAGOSHI, E., NITABACH, M. N., OLMEDO, M., NUSINOW, D. A., PTACEK, L. J., RAND, D., REDDY, A. B., ROBLES, M. S., ROENNEBERG, T., ROSBASH, M., RUBEN, M. D., RUND, S. S. C., SANCAR, A., SASSONE-CORSI, P., SEHGAL, A., SHERRILL-MIX, S., SKENE, D. J., STORCH, K. F., TAKAHASHI, J. S., UEDA, H. R., WANG, H., WEITZ, C., WESTERMARK, P. O., WIJNEN, H., XU, Y., WU, G., YOO, S. H., YOUNG, M., ZHANG, E. E., ZIELINSKI, T. & HOGENESCH, J. B. 2017. Guidelines for Genome-Scale Analysis of Biological Rhythms. *J Biol Rhythms*, 32, 380-393.
- HUGHES, M. E., DITACCHIO, L., HAYES, K. R., VOLLMERS, C., PULIVARTHY, S., BAGGS, J. E., PANDA, S. & HOGENESCH, J. B. 2009. Harmonics of Circadian Gene Transcription in Mammals. *PLOS Genetics*, 5, e1000442.
- HUGHES, M. E., HOGENESCH, J. B. & KORNACKER, K. 2010. JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J Biol Rhythms*, 25, 372-80.
- IJPENBERG, A., JEANNIN, E., WAHLI, W. & DESVERGNE, B. 1997. Polarity and Specific Sequence Requirements of Peroxisome Proliferator-activated Receptor (PPAR)/Retinoid X Receptor Heterodimer Binding to DNA: A FUNCTIONAL ANALYSIS OF THE MALIC ENZYME GENE PPAR RESPONSE ELEMENT. *Journal of Biological Chemistry*, 272, 20108-20117.
- IMAI, T., JIANG, M., CHAMBON, P. & METZGER, D. 2001. Impaired adipogenesis and lipolysis in the mouse upon selective ablation of the retinoid X receptor alpha mediated by a tamoxifen-inducible chimeric Cre recombinase (Cre-ERT2) in adipocytes. *Proc Natl Acad Sci U S A*, 98, 224-8.
- ISOMURA, A. & KAGEYAMA, R. 2014. Ultradian oscillations and pulses: coordinating cellular responses and cell fate decisions. *Development*, 141, 3627-3636.
- ISSEMANN, I. & GREEN, S. 1990. Activation of a member of the steroid hormone receptor superfamily by peroxisome proliferators. *Nature*, 347, 645-50.
- JENSEN, E. V., DESOMBRE, E. R., KAWASHIMA, T., SUZUKI, T., KYSER, K. & JUNGBLUT, P. W. 1967. Estrogen-binding substances of target tissues. *Science*, 158, 529-30.
- JETTEN, A. M. 2009. Retinoid-related orphan receptors (RORs): critical roles in development, immunity, circadian rhythm, and cellular metabolism. *Nuclear Receptor Signaling*, 7, e003.
- JIANG, G., LEE, U. & SLADEK, F. M. 1997. Proposed mechanism for the stabilization of nuclear receptor DNA binding via protein dimerization. *Mol Cell Biol*, 17, 6546-54.
- JONES, S., VAN HEYNINGEN, P., BERMAN, H. M. & THORNTON, J. M. 1999. Protein-DNA interactions: A structural analysis. *J Mol Biol*, 287, 877-96.

- JOSHI, A. 2014. Mammalian transcriptional hotspots are enriched for tissue specific enhancers near cell type specific highly expressed genes and are predicted to act as transcriptional activator hubs. *BMC Bioinformatics*, 15, 412.
- JOVER, R., MOYA, M. & GOMEZ-LECHON, M. J. 2009. Transcriptional regulation of cytochrome p450 genes by the nuclear receptor hepatocyte nuclear factor 4- α . *Curr Drug Metab*, 10, 508-19.
- JUGE-AUBRY, C., PERNIN, A., FAVEZ, T., BURGER, A. G., WAHLI, W., MEIER, C. A. & DESVERGNE, B. 1997. DNA Binding Properties of Peroxisome Proliferator-activated Receptor Subtypes on Various Natural Peroxisome Proliferator Response Elements: IMPORTANCE OF THE 5' -FLANKING REGION. *Journal of Biological Chemistry*, 272, 25252-25259.
- JUNG, D., MANGELSDORF, D. J. & MEYER, U. A. 2006. Pregnane X Receptor Is a Target of Farnesoid X Receptor. *Journal of Biological Chemistry*, 281, 19081-19091.
- KANE, M. A. 2012. Analysis, occurrence, and function of 9-cis-retinoic acid. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, 1821, 10-20.
- KANE, MAUREEN A., CHEN, N., SPARKS, S. & NAPOLI, JOSEPH L. 2005. Quantification of endogenous retinoic acid in limited biological samples by LC/MS/MS. *Biochemical Journal*, 388, 363-369.
- KANE, M. A., FOLIAS, A. E., PINGITORE, A., PERRI, M., OBROCHTA, K. M., KROIS, C. R., CIONE, E., RYU, J. Y. & NAPOLI, J. L. 2010. Identification of 9-cis-retinoic acid as a pancreas-specific autacoid that attenuates glucose-stimulated insulin secretion. *Proceedings of the National Academy of Sciences*, 107, 21884-21889.
- KANE, M. A., FOLIAS, A. E., WANG, C. & NAPOLI, J. L. 2008. Quantitative Profiling of Endogenous Retinoic Acid in Vivo and in Vitro by Tandem Mass Spectrometry. *Analytical Chemistry*, 80, 1702-1708.
- KANG, J. O. & SUCOV, H. M. 2005. Convergent proliferative response and divergent morphogenic pathways induced by epicardial and endocardial signaling in fetal heart development. *Mech Dev*, 122, 57-65.
- KASTNER, P., GRONDONA, J. M., MARK, M., GANSMULLER, A., LEMEURE, M., DECIMO, D., VONESCH, J. L., DOLLE, P. & CHAMBON, P. 1994. Genetic analysis of RXR α developmental function: convergence of RXR and RAR signaling pathways in heart and eye morphogenesis. *Cell*, 78, 987-1003.
- KATAGIRI, Y., TAKEDA, K., YU, Z. X., FERRANS, V. J., OZATO, K. & GUROFF, G. 2000. Modulation of retinoid signalling through NGF-induced nuclear export of NGFI-B. *Nat Cell Biol*, 2, 435-40.
- KAUPPI, B., JAKOB, C., FÄRNEGÄRDH, M., YANG, J., AHOLA, H., ALARCON, M., CALLES, K., ENGSTRÖM, O., HARLAN, J. & MUCHMORE, S. 2003. The three-dimensional structures of antagonistic and agonistic forms of the glucocorticoid receptor ligand-binding domain ru-486 induces a transconformation that leads to active antagonism. *Journal of Biological Chemistry*, 278, 22748-22754.
- KERSTEN, S., SEYDOUX, J., PETERS, J. M., GONZALEZ, F. J., DESVERGNE, B. & WAHLI, W. 1999. Peroxisome proliferator-activated receptor α mediates the adaptive response to fasting. *The Journal of Clinical Investigation*, 103, 1489-1498.
- KIM, D. H., XIAO, Z., KWON, S., SUN, X., RYERSON, D., TKAC, D., MA, P., WU, S. Y., CHIANG, C. M., ZHOU, E., XU, H. E., PALVIMO, J. J., CHEN, L. F., KEMPER, B. & KEMPER, J. K. 2015. A dysregulated acetyl/SUMO switch of FXR promotes hepatic inflammation in obesity. *The EMBO Journal*, 34, 184-199.
- KITAREEWAN, S., BURKA, L. T., TOMER, K. B., PARKER, C. E., DETERDING, L. J., STEVENS, R. D., FORMAN, B. M., MAIS, D. E., HEYMAN, R. A., MCMORRIS, T. & WEINBERGER, C. 1996. Phytol metabolites are circulating dietary factors that activate the nuclear receptor RXR. *Molecular Biology of the Cell*, 7, 1153-1166.
- KLIEWER, S. A., LEHMANN, J. M. & WILLSON, T. M. 1999. Orphan Nuclear Receptors: Shifting Endocrinology into Reverse. *Science*, 284, 757.
- KLIEWER, S. A., UMESONO, K., MANGELSDORF, D. J. & EVANS, R. M. 1992a. Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone and vitamin D3 signalling. *Nature*, 355, 446-9.

- KLIEWER, S. A., UMESONO, K., NOONAN, D. J., HEYMAN, R. A. & EVANS, R. M. 1992b. Convergence of 9-cis retinoic acid and peroxisome proliferator signalling pathways through heterodimer formation of their receptors. *Nature*, 358, 771-4.
- KOCH, F. & ANDRAU, J.-C. 2011. Initiating RNA Polymerase II and TIPs as hallmarks of enhancer activity and tissue-specificity. *Transcription*, 2, 263-268.
- KOIKE, N., YOO, S. H., HUANG, H. C., KUMAR, V., LEE, C., KIM, T. K. & TAKAHASHI, J. S. 2012. Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science*, 338, 349-54.
- KOSTERS, A., SUN, D., WU, H., TIAN, F., FELIX, J. C., LI, W. & KARPEN, S. J. 2013. Sexually dimorphic genome-wide binding of retinoid X receptor alpha (RXRalpha) determines male-female differences in the expression of hepatic lipid processing genes in mice. *PLOS ONE*, 8, e71538.
- KREZEL, W., DUPE, V., MARK, M., DIERICH, A., KASTNER, P. & CHAMBON, P. 1996. RXR gamma null mice are apparently normal and compound RXR alpha +/-RXR beta -/-RXR gamma -/- mutant mice are viable. *Proc Natl Acad Sci U S A*, 93, 9010-4.
- KRISHNAIAH, S. Y., WU, G., ALTMAN, B. J., GROWE, J., RHOADES, S. D., COLDREN, F., VENKATARAMAN, A., OLARERIN-GEORGE, A. O., FRANCEY, L. J., MUKHERJEE, S., GIRISH, S., SELBY, C. P., CAL, S., ER, U., SIANATI, B., SENGUPTA, A., ANAFI, R. C., KAVAKLI, I. H., SANCAR, A., BAUR, J. A., DANG, C. V., HOGENESCH, J. B. & WELJIE, A. M. 2017. Clock Regulation of Metabolites Reveals Coupling between Transcription and Metabolism. *Cell Metab*, 25, 961-974.e4.
- KUBALAK, S. W., HUTSON, D. R., SCOTT, K. K. & SHANNON, R. A. 2002. Elevated transforming growth factor beta2 enhances apoptosis and contributes to abnormal outflow tract and aortic sac development in retinoic X receptor alpha knockout embryos. *Development*, 129, 733-46.
- KUMAR, V., GREEN, S., STACK, G., BERRY, M., JIN, J. R. & CHAMBON, P. 1987. Functional domains of the human estrogen receptor. *Cell*, 51, 941-51.
- LAFFITTE, B. A., KAST, H. R., NGUYEN, C. M., ZAVACKI, A. M., MOORE, D. D. & EDWARDS, P. A. 2000. Identification of the DNA binding specificity and potential target genes for the farnesoid X-activated receptor. *J Biol Chem*, 275, 10638-47.
- LAMMI, J., PERLMANN, T. & AARNISALO, P. 2008. Corepressor interaction differentiates the permissive and non-permissive retinoid X receptor heterodimers. *Arch Biochem Biophys*, 472, 105-14.
- LANDT, S. G., MARINOV, G. K., KUNDAJE, A., KHERADPOUR, P., PAULI, F., BATZOGLOU, S., BERNSTEIN, B. E., BICKEL, P., BROWN, J. B., CAYTING, P., CHEN, Y., DESALVO, G., EPSTEIN, C., FISHER-AYLOR, K. I., EUSKIRCHEN, G., GERSTEIN, M., GERTZ, J., HARTEMINK, A. J., HOFFMAN, M. M., IYER, V. R., JUNG, Y. L., KARMAKAR, S., KELLIS, M., KHARCHENKO, P. V., LI, Q., LIU, T., LIU, X. S., MA, L., MILOSAVLJEVIC, A., MYERS, R. M., PARK, P. J., PAZIN, M. J., PERRY, M. D., RAHA, D., REDDY, T. E., ROZOWSKY, J., SHORESH, N., SIDOW, A., SLATTERY, M., STAMATOYANNOPOULOS, J. A., TOLSTORUKOV, M. Y., WHITE, K. P., XI, S., FARNHAM, P. J., LIEB, J. D., WOLD, B. J. & SNYDER, M. 2012. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Research*, 22, 1813-1831.
- LANZ, R. B. & RUSCONI, S. 1994. A conserved carboxy-terminal subdomain is important for ligand interpretation and transactivation by nuclear receptors. *Endocrinology*, 135, 2183-2195.
- LE MARTELOT, G., CANELLA, D., SYMUL, L., MIGLIAVACCA, E., GILARDI, F., LIECHTI, R., MARTIN, O., HARSHMAN, K., DELORENZI, M., DESVERGNE, B., HERR, W., DEPLANCKE, B., SCHIBLER, U., ROUGEMONT, J., GUEx, N., HERNANDEZ, N. & NAEF, F. 2012. Genome-wide RNA polymerase II profiles and RNA accumulation reveal kinetics of transcription and associated epigenetic changes during diurnal cycles. *PLoS Biol*, 10.
- LE MARTELOT, G., CLAUDEL, T., GATFIELD, D., SCHAAD, O., KORNMANN, B., SASSO, G. L., MOSCHETTA, A. & SCHIBLER, U. 2009. REV-ERB α Participates in Circadian SREBP Signaling and Bile Acid Homeostasis. *PLOS Biology*, 7, e1000181.

- LEE, H. Y., SUH, Y. A., ROBINSON, M. J., CLIFFORD, J. L., HONG, W. K., WOODGETT, J. R., COBB, M. H., MANGELSDORF, D. J. & KURIE, J. M. 2000. Stress pathway activation induces phosphorylation of retinoid X receptor. *J Biol Chem*, 275, 32193-9.
- LEE, J., SEOK, S., YU, P., KIM, K., SMITH, Z., RIVAS-ASTROZA, M., ZHONG, S. & KEMPER, J. K. 2012. Genomic analysis of hepatic farnesoid X receptor binding sites reveals altered binding in obesity and direct gene repression by farnesoid X receptor in mice. *Hepatology (Baltimore, Md.)*, 56, 108-117.
- LEE, J. W., GULICK, T. & MOORE, D. D. 1992. Thyroid hormone receptor dimerization function maps to a conserved subregion of the ligand binding domain. *Molecular Endocrinology*, 6, 1867-1873.
- LEFEBVRE, P. 2001. Molecular basis for designing selective modulators of retinoic acid receptor transcriptional activities. *Curr Drug Targets Immune Endocr Metabol Disord*, 1, 153-64.
- LEFEBVRE, P., GAUB, M. P., TAHAYATO, A., ROCHETTE-EGLY, C. & FORMSTECHEP, P. 1995. Protein phosphatases 1 and 2A regulate the transcriptional and DNA binding activities of retinoic acid receptors. *J Biol Chem*, 270, 10806-16.
- LEID, M., KASTNER, P. & CHAMBON, P. 1992a. Multiplicity generates diversity in the retinoic acid signalling pathways. *Trends Biochem Sci*, 17, 427-33.
- LEID, M., KASTNER, P., LYONS, R., NAKSHATRI, H., SAUNDERS, M., ZACHAREWSKI, T., CHEN, J.-Y., STAUB, A., GARNIER, J.-M. & MADER, S. 1992b. Purification, cloning, and RXR identity of the HeLa cell factor with which RAR or TR heterodimerizes to bind target sequences efficiently. *Cell*, 68, 377-395.
- LEMOTE, P. K., KEIDEL, S. & APFEL, C. M. 1996. Phytanic acid is a retinoid X receptor ligand. *European journal of biochemistry*, 236, 328-333.
- LENGQVIST, J., MATA DE URQUIZA, A., BERGMAN, A.-C., WILLSON, T. M., SJÖVALL, J., PERLMANN, T. & GRIFFITHS, W. J. 2004. Polyunsaturated Fatty Acids Including Docosahexaenoic and Arachidonic Acid Bind to the Retinoid X Receptor α Ligand-binding Domain. *Molecular & Cellular Proteomics*, 3, 692-703.
- LEVIN, A. A., STURZENBECKER, L. J., KAZMER, S., BOSAKOWSKI, T., HUSELTON, C., ALLENBY, G., SPECK, J., KRATZEISEN, C., ROSENBERGER, M., LOVEY, A. & ET AL. 1992. 9-cis retinoic acid stereoisomer binds and activates the nuclear receptor RXR α . *Nature*, 355, 359-61.
- LI, G., WALCH, E., YANG, X., LIPPMAN, S. M. & CLIFFORD, J. L. 2000a. Cloning and characterization of the human retinoid X receptor α gene: conservation of structure with the mouse homolog. *Biochem Biophys Res Commun*, 269, 54-7.
- LI, G., YIN, W., CHAMBERLAIN, R., HEWETT-EMMETT, D., ROBERTS, J. N., YANG, X., LIPPMAN, S. M. & CLIFFORD, J. L. 2006. Identification and characterization of the human retinoid X receptor α gene promoter. *Gene*, 372, 118-27.
- LI, M., INDRA, A. K., WAROT, X., BROCARD, J., MESSADDEQ, N., KATO, S., METZGER, D. & CHAMBON, P. 2000b. Skin abnormalities generated by temporally controlled RXR α mutations in mouse epidermis. *Nature*, 407, 633-6.
- LI, Y., LAMBERT, M. H. & XU, H. E. 2003. Activation of nuclear receptors: a perspective from structural genomics. *Structure*, 11, 741-6.
- LIANG, K. & KELES, S. 2012. Detecting differential binding of transcription factors with ChIP-seq. *Bioinformatics (Oxford, England)*, 28, 121-122.
- LIFANOV, A. P., MAKEEV, V. J., NAZINA, A. G. & PAPATSENKO, D. A. 2003. Homotypic Regulatory Clusters in Drosophila. *Genome Research*, 13, 579-588.
- LOHNES, D., KASTNER, P., DIERICH, A., MARK, M., LEMEURE, M. & CHAMBON, P. 1993. Function of retinoic acid receptor γ in the mouse. *Cell*, 73, 643-58.
- LOU, X., TORESSON, G., BENOD, C., SUH, J. H., PHILIPS, K. J., WEBB, P. & GUSTAFSSON, J. A. 2014. Structure of the retinoid X receptor α -liver X receptor β (RXR α -LXR β) heterodimer on DNA. *Nat Struct Mol Biol*, 21, 277-81.
- LUFKIN, T., LOHNES, D., MARK, M., DIERICH, A., GORRY, P., GAUB, M. P., LEMEURE, M. & CHAMBON, P. 1993. High postnatal lethality and testis degeneration in retinoic acid receptor α mutant mice. *Proc Natl Acad Sci U S A*, 90, 7225-9.

- LURIA, A. & FURLOW, J. D. 2004. Spatiotemporal retinoid-X receptor activation detected in live vertebrate embryos. *Proceedings of the National Academy of Sciences*, 101, 8987-8992.
- LUSCOMBE, N. M. & THORNTON, J. M. 2002. Protein-DNA interactions: amino acid conservation and the effects of mutations on binding specificity. *J Mol Biol*, 320, 991-1009.
- MA, K., XIAO, R., TSENG, H.-T., SHAN, L., FU, L. & MOORE, D. D. 2009. Circadian Dysregulation Disrupts Bile Acid Homeostasis. *PLOS ONE*, 4, e6843.
- MACQUARRIE, K. L., FONG, A. P., MORSE, R. H. & TAPSCOTT, S. J. 2011. Genome-wide transcription factor binding: beyond direct target regulation. *Trends in genetics : TIG*, 27, 141-148.
- MADER, S., CHEN, J. Y., CHEN, Z., WHITE, J., CHAMBON, P. & GRONEMEYER, H. 1993. The patterns of binding of RAR, RXR and TR homo- and heterodimers to direct repeats are dictated by the binding specificities of the DNA binding domains. *The EMBO journal*, 12, 5029-5041.
- MANGE, F., PRAZ, V., MIGLIAVACCA, E., WILLIS, I. M., SCHUTZ, F. & HERNANDEZ, N. 2017. Diurnal regulation of RNA polymerase III transcription is under the control of both the feeding-fasting response and the circadian clock. *Genome Res*, 27, 973-984.
- MANGELSDORF, D. J., BORGMEYER, U., HEYMAN, R. A., ZHOU, J. Y., ONG, E. S., ORO, A. E., KAKIZUKA, A. & EVANS, R. M. 1992. Characterization of three RXR genes that mediate the action of 9-cis retinoic acid. *Genes & Development*, 6, 329-344.
- MANGELSDORF, D. J. & EVANS, R. M. 1995. The RXR heterodimers and orphan receptors. *Cell*, 83.
- MANGELSDORF, D. J., ONG, E. S., DYCK, J. A. & EVANS, R. M. 1990. Nuclear receptor that identifies a novel retinoic acid response pathway. *Nature*, 345, 224-9.
- MANGELSDORF, D. J., UMESONO, K., KLIEWER, S. A., BORGMEYER, U., ONG, E. S. & 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.
- MANO, H., MORI, R., OZAWA, T., TAKEYAMA, K., YOSHIZAWA, Y., KOJIMA, R., ARAO, Y., MASUSHIGE, S. & KATO, S. 1994. Positive and negative regulation of retinoid X receptor gene expression by thyroid hormone in the rat. Transcriptional and post-transcriptional controls by thyroid hormone. *J Biol Chem*, 269, 1591-4.
- MARK, M., GHYSELINCK, N. B. & CHAMBON, P. 2009. Function of retinoic acid receptors during embryonic development. *Nuclear receptor signaling*, 7, e002-e002.
- MARTENS, J. H., BRINKMAN, A. B., SIMMER, F., FRANCOIJS, K. J., NEBBIOSO, A., FERRARA, F., ALTUCCI, L. & STUNNENBERG, H. G. 2010. PML-RARalpha/RXR Alters the Epigenetic Landscape in Acute Promyelocytic Leukemia. *Cancer Cell*, 17, 173-85.
- MASCREZ, B., GHYSELINCK, N. B., CHAMBON, P. & MARK, M. 2009. A transcriptionally silent RXRalpha supports early embryonic morphogenesis and heart development. *Proc Natl Acad Sci U S A*, 106, 4272-7.
- MASCREZ, B., GHYSELINCK, N. B., WATANABE, M., ANNICOTTE, J. S., CHAMBON, P., AUWERX, J. & MARK, M. 2004. Ligand-dependent contribution of RXRbeta to cholesterol homeostasis in Sertoli cells. *EMBO Rep*, 5, 285-90.
- MASCREZ, B., MARK, M., DIERICH, A., GHYSELINCK, N. B., KASTNER, P. & CHAMBON, P. 1998. The RXRalpha ligand-dependent activation function 2 (AF-2) is important for mouse development. *Development*, 125, 4691-707.
- MASCREZ, B., MARK, M., KREZEL, W., DUPE, V., LEMEURE, M., GHYSELINCK, N. B. & CHAMBON, P. 2001. Differential contributions of AF-1 and AF-2 activities to the developmental functions of RXR alpha. *Development*, 128, 2049-62.
- MATSUSHIMA-NISHIWAKI, R., OKUNO, M., ADACHI, S., SANO, T., AKITA, K., MORIWAKI, H., FRIEDMAN, S. L. & KOJIMA, S. 2001. Phosphorylation of retinoid X receptor alpha at serine 260 impairs its metabolism and function in human hepatocellular carcinoma. *Cancer Res*, 61, 7675-82.
- MATT, N., DUPE, V., GARNIER, J. M., DENNEFELD, C., CHAMBON, P., MARK, M. & GHYSELINCK, N. B. 2005. Retinoic acid-dependent eye morphogenesis is orchestrated by neural crest cells. *Development*, 132, 4789-800.

- MCDERMOTT, N. B., GORDON, D. F., KRAMER, C. A., LIU, Q., LINNEY, E., WOOD, W. M. & HAUGEN, B. R. 2002. Isolation and functional analysis of the mouse RXRgamma1 gene promoter in anterior pituitary cells. *J Biol Chem*, 277, 36839-44.
- MCNAMARA, P., SEO, S. B., RUDIC, R. D., SEHGAL, A., CHAKRAVARTI, D. & FITZGERALD, G. A. 2001. Regulation of CLOCK and MOP4 by nuclear hormone receptors in the vasculature: a humoral mechanism to reset a peripheral clock. *Cell*, 105, 877-89.
- MENDOZA-PARRA, M. A., NOWICKA, M., VAN GOOL, W. & GRONEMEYER, H. 2013. Characterising ChIP-seq binding patterns by model-based peak shape deconvolution. *BMC Genomics*, 14, 834.
- MENENDEZ-GUTIERREZ, M. P. & RICOTE, M. 2017. The multi-faceted role of retinoid X receptor in bone remodeling. *Cell Mol Life Sci*, 74, 2135-2149.
- MERKI, E., ZAMORA, M., RAYA, A., KAWAKAMI, Y., WANG, J., ZHANG, X., BURCH, J., KUBALAK, S. W., KALIMAN, P., IZPISUA BELMONTE, J. C., CHIEN, K. R. & RUIZ-LOZANO, P. 2005. Epicardial retinoid X receptor alpha is required for myocardial growth and coronary artery formation. *Proc Natl Acad Sci U S A*, 102, 18455-60.
- MERROW, M., MAZZOTTA, G., CHEN, Z. & ROENNEBERG, T. 2006. The right place at the right time: regulation of daily timing by phosphorylation. *Genes & Development*, 20, 2629-2633.
- MIESFELD, R., RUSCONI, S., GODOWSKI, P. J., MALER, B. A., OKRET, S., WIKSTROM, A. C., GUSTAFSSON, J. A. & YAMAMOTO, K. R. 1986. Genetic complementation of a glucocorticoid receptor deficiency by expression of cloned receptor cDNA. *Cell*, 46, 389-99.
- MILLER, B. H., MCDEARMON, E. L., PANDA, S., HAYES, K. R., ZHANG, J., ANDREWS, J. L., ANTOCH, M. P., WALKER, J. R., ESSER, K. A., HOGENESCH, J. B. & TAKAHASHI, J. S. 2007. Circadian and CLOCK-controlled regulation of the mouse transcriptome and cell proliferation. *Proceedings of the National Academy of Sciences*, 104, 3342-3347.
- MINUCCI, S., LEID, M., TOYAMA, R., SAINT-JEANNET, J. P., PETERSON, V. J., HORN, V., ISHMAEL, J. E., BHATTACHARYYA, N., DEY, A., DAWID, I. B. & OZATO, K. 1997a. Retinoid X receptor (RXR) within the RXR-retinoic acid receptor heterodimer binds its ligand and enhances retinoid-dependent gene expression. *Mol Cell Biol*, 17, 644-55.
- MINUCCI, S., LEID, M., TOYAMA, R., SAINT-JEANNET, J. P., PETERSON, V. J., HORN, V., ISHMAEL, J. E., BHATTACHARYYA, N., DEY, A., DAWID, I. B. & OZATO, K. 1997b. Retinoid X receptor (RXR) within the RXR-retinoic acid receptor heterodimer binds its ligand and enhances retinoid-dependent gene expression. *Molecular and Cellular Biology*, 17, 644-655.
- MINUCCI, S., ZAND, D. J., DEY, A., MARKS, M. S., NAGATA, T., GRIPPO, J. F. & OZATO, K. 1994. Dominant negative retinoid X receptor beta inhibits retinoic acid-responsive gene regulation in embryonal carcinoma cells. *Mol Cell Biol*, 14, 360-72.
- MITCHELL, J. A. & FRASER, P. 2008. Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes Dev*, 22, 20-5.
- MOHAWK, J. A., GREEN, C. B. & TAKAHASHI, J. S. 2012. Central and Peripheral Circadian Clocks in Mammals. *Annual Review of Neuroscience*, 35, 445-462.
- MOHIDEEN-ABDUL, K., TAZIBT, K., BOURGUET, M., HAZEMANN, I., LEBARS, I., TAKACS, M., CIANFÉRANI, S., KLAHOLZ, B. P., MORAS, D. & BILLAS, I. M. L. 2017. Importance of the Sequence-Directed DNA Shape for Specific Binding Site Recognition by the Estrogen-Related Receptor. *Frontiers in Endocrinology*, 8.
- MOORMAN, C., SUN, L. V., WANG, J., DE WIT, E., TALHOUT, W., WARD, L. D., GREIL, F., LU, X.-J., WHITE, K. P., BUSSEMAKER, H. J. & VAN STEENSEL, B. 2006. Hotspots of transcription factor colocalization in the genome of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences*, 103, 12027.
- MORGAN, K. & KALSHEKER, N. A. 1997. Regulation of the serine proteinase inhibitor (SERPIN) gene alpha 1-antitrypsin: a paradigm for other SERPINS. *Int J Biochem Cell Biol*, 29, 1501-11.
- MOUTIER, E., YE, T., CHOUKRALLAH, M. A., URBAN, S., OSZ, J., CHATAGNON, A., DELACROIX, L., LANGER, D., ROCHEL, N., MORAS, D., BENOIT, G. & DAVIDSON, I. 2012. Retinoic acid receptors recognize the mouse genome through binding elements with diverse spacing and topology. *J Biol Chem*, 287, 26328-41.

- MUNOZ-CANOVES, P., VIK, D. P. & TACK, B. F. 1990. Mapping of a retinoic acid-responsive element in the promoter region of the complement factor H gene. *J Biol Chem*, 265, 20065-8.
- MURAI, J., IKEGAMI, D., OKAMOTO, M., YOSHIKAWA, H. & TSUMAKI, N. 2008. Insulation of the Ubiquitous Rxb Promoter from the Cartilage-specific Adjacent Gene, Col11a2. *Journal of Biological Chemistry*, 283, 27677-27687.
- NAGAO, Y., FRENCH, B. A., CAI, Y., FRENCH, S. W. & WAN, Y. J. 1998. Inhibition of PPAR alpha/RXR alpha-mediated direct hyperplasia pathways during griseofulvin-induced hepatocarcinogenesis. *J Cell Biochem*, 69, 189-200.
- NAKAHATA, Y., AKASHI, M., TRCKA, D., YASUDA, A. & TAKUMI, T. 2006. The in vitro real-time oscillation monitoring system identifies potential entrainment factors for circadian clocks. *BMC Mol Biol*, 7, 5.
- NAKAMURA, K., INOUE, I., TAKAHASHI, S., KOMODA, T. & KATAYAMA, S. 2008. Cryptochrome and Period Proteins Are Regulated by the CLOCK/BMAL1 Gene: Crosstalk between the PPARs/RXRalpha-Regulated and CLOCK/BMAL1-Regulated Systems. *PPAR Res*, 2008, 348610.
- NAKATSUKA, A., WADA, J. & MAKINO, H. 2013. Cell cycle abnormality in metabolic syndrome and nuclear receptors as an emerging therapeutic target. *Acta Med Okayama*, 67, 129-34.
- NAVIGATORE-FONZO, L. S., GOLINI, R. L., PONCE, I. T., DELGADO, S. M., PLATEO-PIGNATARI, M. G., GIMENEZ, M. S. & ANZULOVICH, A. C. 2013. Retinoic acid receptors move in time with the clock in the hippocampus. Effect of a vitamin-A-deficient diet. 24, 859-867.
- NIELSEN, R., PEDERSEN, T. A., HAGENBEEK, D., MOULOS, P., SIERSBAEK, R., MEGENS, E., DENISSOV, S., BORGESEN, M., FRANCOIJS, K. J., MANDRUP, S. & STUNNENBERG, H. G. 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, 2953-67.
- NOHARA, A., KOBAYASHI, J. & MABUCHI, H. 2009. Retinoid X receptor heterodimer variants and cardiovascular risk factors. *J Atheroscler Thromb*, 16, 303-18.
- NUNEZ, V., ALAMEDA, D., RICO, D., MOTA, R., GONZALO, P., CEDENILLA, M., FISCHER, T., BOSCA, L., GLASS, C. K., ARROYO, A. G. & RICOTE, M. 2010. Retinoid X receptor alpha controls innate inflammatory responses through the up-regulation of chemokine expression. *Proc Natl Acad Sci U S A*, 107, 10626-31.
- NÄÄR, A. M., BOUTIN, J.-M., LIPKIN, S. M., VICTOR, C. Y., HOLLOWAY, J. M., GLASS, C. K. & ROSENFELD, M. G. 1991. The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. *Cell*, 65, 1267-1279.
- OFRAN, Y., MYSORE, V. & ROST, B. 2007. Prediction of DNA-binding residues from sequence. *Bioinformatics*, 23, i347-53.
- ORLOV, I., ROCHEL, N., MORAS, D. & KLAHOLZ, B. P. 2012. Structure of the full human RXR/VDR nuclear receptor heterodimer complex with its DR3 target DNA. *The EMBO Journal*, 31, 291.
- ORO, A. E., MCKEOWN, M. & EVANS, R. M. 1990. Relationship between the product of the *Drosophila* ultraspiracle locus and the vertebrate retinoid X receptor. *Nature*, 347, 298.
- OSZ, J., MCEWEN, A. G., POUSSIN-COURMONTAGNE, P., MOUTIER, E., BIRCK, C., DAVIDSON, I., MORAS, D. & ROCHEL, N. 2015. Structural Basis of Natural Promoter Recognition by the Retinoid X Nuclear Receptor. *Scientific Reports*, 5, 8216.
- OSZ, J., PETHOUKHOV, M. V., SIRIGU, S., SVERGUN, D. I., MORAS, D. & ROCHEL, N. 2012. Solution Structures of PPARγ2/RXRα Complexes. *PPAR research*, 2012, 701412-701412.
- OWEN, G. I. & ZELENT, A. 2000. Origins and evolutionary diversification of the nuclear receptor superfamily. *Cellular and Molecular Life Sciences CMLS*, 57, 809-827.
- PANDA, S., ANTOCH, M. P., MILLER, B. H., SU, A. I., SCHOOK, A. B., STRAUME, M., SCHULTZ, P. G., KAY, S. A., TAKAHASHI, J. S. & HOGENESCH, J. B. 2002. Coordinated Transcription of Key Pathways in the Mouse by the Circadian Clock. *Cell*, 109, 307-320.
- PARK, S.-H., WIWI, CHRISTOPHER A. & WAXMAN, DAVID J. 2006. Signalling cross-talk between hepatocyte nuclear factor 4α and growth-hormone-activated STAT5b. *Biochemical Journal*, 397, 159-168.

- PARK, S. G., HANNENHALLI, S. & CHOI, S. S. 2014. Conservation in first introns is positively associated with the number of exons within genes and the presence of regulatory epigenetic signals. *BMC genomics*, 15, 526-526.
- PATEL, S. R. & SKAFAR, D. F. 2015. Modulation of nuclear receptor activity by the F domain. *Mol Cell Endocrinol*, 418 Pt 3, 298-305.
- PERLMANN, T., RANGARAJAN, P. N., UMESONO, K. & EVANS, R. M. 1993. Determinants for selective RAR and TR recognition of direct repeat HREs. *Genes & Development*, 7, 1411-1422.
- PETKOVICH, M., BRAND, N. J., KRUST, A. & CHAMBON, P. 1987. A human retinoic acid receptor which belongs to the family of nuclear receptors. *Nature*, 330, 444-50.
- PHAM, T. H., CLEMENTE, J. C., SATOU, K. & HO, T. B. 2005. Computational discovery of transcriptional regulatory rules. *Bioinformatics*, 21 Suppl 2, ii101-7.
- PIKE, J. W., MEYER, M. B., MARTOWICZ, M. L., BISHOP, K. A., LEE, S. M., NERENZ, R. D. & GOETSCH, P. D. 2010. Emerging regulatory paradigms for control of gene expression by 1,25-dihydroxyvitamin D3. *J Steroid Biochem Mol Biol*, 121, 130-5.
- PREDKI, P. F., ZAMBLE, D., SARKAR, B. & GIGUERE, V. 1994. Ordered binding of retinoic acid and retinoid-X receptors to asymmetric response elements involves determinants adjacent to the DNA-binding domain. *Mol Endocrinol*, 8, 31-9.
- RAJARAM, S. 2014. Health benefits of plant-derived α -linolenic acid. *The American journal of clinical nutrition*, 100, 443S-448S.
- RASTINEJAD, F., WAGNER, T., ZHAO, Q. & KHORASANIZADEH, S. 2000. Structure of the RXR-RAR DNA - binding complex on the retinoic acid response element DR1. *The EMBO Journal*, 19, 1045-1054.
- REZAEIAN, I. & RUEDA, L. 2014. CMT: A Constrained Multi-Level Thresholding Approach for ChIP-Seq Data Analysis. *PLOS ONE*, 9, e93873.
- RIB, L., VILLENEUVE, D., MINOCHA, S., PRAZ, V., HERNANDEZ, N., GUEX, N. & HERR, W. 2018. Cycles of gene expression and genome response during mammalian tissue regeneration. *Epigenetics Chromatin*, 11, 52.
- RIPPERGER, J. A. & SCHIBLER, U. 2006. Rhythmic CLOCK-BMAL1 binding to multiple E-box motifs drives circadian Dbp transcription and chromatin transitions. *Nat Genet*, 38, 369-74.
- ROCA-SAAVEDRA, P., MARINO-LORENZO, P., MIRANDA, J., PORTO-ARIAS, J., LAMAS, A., VAZQUEZ, B., FRANCO, C. & CEPEDA, A. 2017. Phytanic acid consumption and human health, risks, benefits and future trends: a review. *Food chemistry*, 221, 237-247.
- ROCHEL, N., CIESIELSKI, F., GODET, J., MOMAN, E., ROESSLE, M., PELUSO-ILTIS, C., MOULIN, M., HAERTLEIN, M., CALLOW, P. & MÉLY, Y. 2011. Common architecture of nuclear receptor heterodimers on DNA direct repeat elements with different spacings. *Nature structural & molecular biology*, 18, 564.
- 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, 909-22.
- ROENNEBERG, T. & MERROW, M. 2005. Circadian clocks [mdash] the fall and rise of physiology. *Nat Rev Mol Cell Biol*, 6, 965-971.
- ROSBASH, M. 2009. The Implications of Multiple Circadian Clock Origins. *PLOS Biology*, 7, e1000062.
- RUHL, R., KRZYZOSIAK, A., NIEWIADOMSKA-CIMICKA, A., ROCHEL, N., SZELES, L., VAZ, B., WIETRZYCH-SCHINDLER, M., ALVAREZ, S., SZKLENAR, M., NAGY, L., DE LERA, A. R. & KREZEL, W. 2015. 9-cis-13,14-Dihydroretinoic Acid Is an Endogenous Retinoid Acting as RXR Ligand in Mice. *PLoS Genet*, 11, e1005213.
- RÜHL, R. 2006. Method to determine 4-oxo-retinoic acids, retinoic acids and retinol in serum and cell extracts by liquid chromatography/diode-array detection atmospheric pressure chemical ionisation tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, 20, 2497-2504.
- SAFE, S. & KIM, K. 2008. Non-classical genomic estrogen receptor (ER)/specificity protein and ER/activating protein-1 signaling pathways. *J Mol Endocrinol*, 41, 263-75.

- SCHMIDT, C. K., BROUWER, A. & NAU, H. 2003. Chromatographic analysis of endogenous retinoids in tissues and serum. *Analytical biochemistry*, 315, 36-48.
- SCHOENFELDER, S., SEXTON, T., CHAKALOVA, L., COPE, N. F., HORTON, A., ANDREWS, S., KURUKUTI, S., MITCHELL, J. A., UMLAUF, D., DIMITROVA, D. S., ESKIW, C. H., LUO, Y., WEI, C. L., RUAN, Y., BIEKER, J. J. & FRASER, P. 2010. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat Genet*, 42, 53-61.
- SCHWEIKERT, G., CSEKE, B., CLOUAIRE, T., BIRD, A. & SANGUINETTI, G. 2013. MMDiff: quantitative testing for shape changes in ChIP-Seq data sets. *BMC Genomics*, 14, 826.
- SHAO, D. & LAZAR, M. A. 1999. Modulating nuclear receptor function: may the phos be with you. *The Journal of Clinical Investigation*, 103, 1617-1618.
- SHAO, Z., ZHANG, Y., YUAN, G.-C., ORKIN, S. H. & WAXMAN, D. J. 2012. MAnorm: a robust model for quantitative comparison of ChIP-Seq data sets. *Genome Biology*, 13, R16.
- SHARON, E., KALMA, Y., SHARP, A., RAVEH-SADKA, T., LEVO, M., ZEEVI, D., KEREN, L., YAKHINI, Z., WEINBERGER, A. & SEGAL, E. 2012. Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat Biotechnol*, 30, 521-30.
- SHEN, L., SHAO, N.-Y., LIU, X., MAZE, I., FENG, J. & NESTLER, E. J. 2013. diffReps: Detecting Differential Chromatin Modification Sites from ChIP-seq Data with Biological Replicates. *PLOS ONE*, 8, e65598.
- SHEN, Q., BAI, Y., CHANG, K. C., WANG, Y., BURRIS, T. P., FREEDMAN, L. P., THOMPSON, C. C. & NAGPAL, S. 2011a. Liver X receptor-retinoid X receptor (LXR-RXR) heterodimer cistrome reveals coordination of LXR and AP1 signaling in keratinocytes. *J Biol Chem*, 286, 14554-63.
- SHEN, Q., BAI, Y., CHANG, K. C. N., WANG, Y., BURRIS, T. P., FREEDMAN, L. P., THOMPSON, C. C. & NAGPAL, S. 2011b. Liver X Receptor-Retinoid X Receptor (LXR-RXR) Heterodimer Cistrome Reveals Coordination of LXR and AP1 Signaling in Keratinocytes. *The Journal of Biological Chemistry*, 286, 14554-14563.
- SHIAU, A. K., BARSTAD, D., LORIA, P. M., CHENG, L., KUSHNER, P. J., AGARD, D. A. & GREENE, G. L. 1998. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell*, 95, 927-37.
- SHULMAN, A. I., LARSON, C., MANGELSDORF, D. J. & RANGANATHAN, R. 2004. Structural determinants of allosteric ligand activation in RXR heterodimers. *Cell*, 116, 417-29.
- SIERSBÆK, R., BAEK, S., RABIEE, A., NIELSEN, R., TRAYNOR, S., CLARK, N., SANDELIN, A., JENSEN, OLE N., SUNG, M.-H., HAGER, GORDON L. & MANDRUP, S. 2014a. Molecular Architecture of Transcription Factor Hotspots in Early Adipogenesis. *Cell Reports*, 7, 1434-1442.
- SIERSBÆK, R., RABIEE, A., NIELSEN, R., SIDOLI, S., TRAYNOR, S., LOFT, A., POULSEN, LARS LA C., ROGOWSKA-WRZESINSKA, A., JENSEN, OLE N. & MANDRUP, S. 2014b. Transcription Factor Cooperativity in Early Adipogenic Hotspots and Super-Enhancers. *Cell Reports*, 7, 1443-1455.
- SIMICEVIC, J., SCHMID, A. W., GILARDONI, P. A., ZOLLER, B., RAGHAV, S. K., KRIER, I., GUBELMANN, C., LISACEK, F., NAEF, F., MONIATTE, M. & DEPLANCKE, B. 2013. Absolute quantification of transcription factors during cellular differentiation using multiplexed targeted proteomics. *Nat Meth*, 10, 570-576.
- SINHA, S., ADLER, A. S., FIELD, Y., CHANG, H. Y. & SEGAL, E. 2008. Systematic functional characterization of cis-regulatory motifs in human core promoters. *Genome Res*, 18, 477-88.
- SMITH, R. P., TAHER, L., PATWARDHAN, R. P., KIM, M. J., INOUE, F., SHENDURE, J., OVCHARENKO, I. & AHITUV, N. 2013. Massively parallel decoding of mammalian regulatory sequences supports a flexible organizational model. *Nat Genet*, 45, 1021-1028.
- SOBEL, J. A., KRIER, I., ANDERSIN, T., RAGHAV, S., CANELLA, D., GILARDI, F., KALANTZI, A. S., REY, G., WEGER, B., GACHON, F., DAL PERARO, M., HERNANDEZ, N., SCHIBLER, U., DEPLANCKE, B., NAEF, F. & CYCLI, X. C. 2017. Transcriptional regulatory logic of the diurnal cycle in the mouse liver. *PLOS Biology*, 15, e2001069.

- SOLOMIN, L., JOHANSSON, C. B., ZETTERSTRÖM, R. H., BISSONNETTE, R. P., HEYMAN, R. A., OLSON, L., LENDAHL, U., FRISEN, J. & PERLMANN, T. 1998. Retinoid-X receptor signalling in the developing spinal cord. *Nature*, 395, 398.
- SOLOMON, C., WHITE, J. H. & KREMER, R. 1999. Mitogen-activated protein kinase inhibits 1, 25-dihydroxyvitamin D 3–dependent signal transduction by phosphorylating human retinoid X receptor α . *The Journal of clinical investigation*, 103, 1729-1735.
- STARK, R. & BROWN, G. 2011. DiffBind differential binding analysis of ChIP-Seq peak data. Bioconductor.
- STEINEGER, H. H., ARNTSEN, B. M., SPYDEVOLD, O. & SORENSEN, H. N. 1997. Retinoid X receptor (RXR α) gene expression is regulated by fatty acids and dexamethasone in hepatic cells. *Biochimie*, 79, 107-110.
- STEINHAUSER, S., KURZAWA, N., EILS, R. & HERRMANN, C. 2016. A comprehensive comparison of tools for differential ChIP-seq analysis. *Briefings in Bioinformatics*.
- STEPANOVA, E. V., SHEVELEV, A. B., BORUKHOV, S. I. & SEVERINOV, K. V. 2009. [Mechanisms of action of RNA polymerase-binding transcription factors that do not bind to DNA]. *Biofizika*, 54, 773-90.
- STORCH, K.-F., LIPAN, O., LEYKIN, I., VISWANATHAN, N., DAVIS, F. C., WONG, W. H. & WEITZ, C. J. 2002. Extensive and divergent circadian gene expression in liver and heart. *Nature*, 417, 78-83.
- STRINO, F. & LAPPE, M. 2016. Identifying peaks in *-seq data using shape information. *BMC Bioinformatics*, 17 Suppl 5, 206.
- SUBBARAYAN, V., MARK, M., MESSADEQ, N., RUSTIN, P., CHAMBON, P. & KASTNER, P. 2000. RXR α overexpression in cardiomyocytes causes dilated cardiomyopathy but fails to rescue myocardial hypoplasia in RXR α -null fetuses. *J Clin Invest*, 105, 387-94.
- SUCOV, H. M., DYSON, E., GUMERINGER, C. L., PRICE, J., CHIEN, K. R. & EVANS, R. M. 1994. RXR α mutant mice establish a genetic basis for vitamin A signaling in heart morphogenesis. *Genes & Development*, 8, 1007-1018.
- SUCOV, H. M., IZPISUA-BELMONTE, J. C., GANAN, Y. & EVANS, R. M. 1995. Mouse embryos lacking RXR α are resistant to retinoic-acid-induced limb defects. *Development*, 121, 3997-4003.
- SUGAWARA, A., URUNO, A., NAGATA, T., TAKETO, M. M., TAKEUCHI, K. & ITO, S. 1998. Characterization of mouse retinoid X receptor (RXR)- β gene promoter: negative regulation by tumor necrosis factor (TNF)- α . *Endocrinology*, 139, 3030-3.
- SUINO, K., PENG, L., REYNOLDS, R., LI, Y., CHA, J. Y., REPA, J. J., KLIEWER, S. A. & XU, H. E. 2004. The nuclear xenobiotic receptor CAR: structural determinants of constitutive activation and heterodimerization. *Mol Cell*, 16, 893-905.
- SUVA, L. J., TOWLER, D. A., HARADA, S., GAUB, M. P. & RODAN, G. A. 1994. Characterization of retinoic acid- and cell-dependent sequences which regulate zif268 gene expression in osteoblastic cells. *Mol Endocrinol*, 8, 1507-20.
- SUZUKI, S., SASAKI, S., MORITA, H., OKI, Y., TURIYA, D., ITO, T., MISAWA, H., ISHIZUKA, K. & NAKAMURA, H. 2010. The role of the amino-terminal domain in the interaction of unliganded peroxisome proliferator-activated receptor γ -2 with nuclear receptor co-repressor. *J Mol Endocrinol*, 45, 133-45.
- SZANTO, A., NARKAR, V., SHEN, Q., URAY, I. P., DAVIES, P. J. & NAGY, L. 2004. Retinoid X receptors: X-ploring their (patho)physiological functions. *Cell Death Differ*, 11 Suppl 2, S126-43.
- TACHIKAWA, M., AKANUMA, S.-I., IMAI, T., OKAYASU, S., TOMOHIRO, T., HATANAKA, Y. & HOSOYA, K.-I. 2018. Multiple Cellular Transport and Binding Processes of Unesterified Docosahexaenoic Acid in Outer Blood–Retinal Barrier Retinal Pigment Epithelial Cells. *Biological and Pharmaceutical Bulletin*, 41, 1384-1392.
- TAKAHASHI, J. S. 2017. Transcriptional architecture of the mammalian circadian clock. *Nature reviews. Genetics*, 18, 164-179.

- THOMAS, R., THOMAS, S., HOLLOWAY, A. K. & POLLARD, K. S. 2017. Features that define the best ChIP-seq peak calling algorithms. *Briefings in bioinformatics*, 18, 441-450.
- THOMPSON, C. L., SELBY, C. P., VAN GELDER, R. N., BLANER, W. S., LEE, J., QUADRO, L., LAI, K., GOTTESMAN, M. E. & SANCAR, A. 2004. Effect of vitamin A depletion on nonvisual phototransduction pathways in cryptochromeless mice. *J Biol Rhythms*, 19, 504-17.
- THOMPSON, P. D., JURUTKA, P. W., HAUSSLER, C. A., WHITFIELD, G. K. & HAUSSLER, M. R. 1998. Heterodimeric DNA Binding by the Vitamin D Receptor and Retinoid X Receptors Is Enhanced by 1,25-Dihydroxyvitamin D₃ and Inhibited by 9-cis-Retinoic Acid: EVIDENCE FOR ALLOSTERIC RECEPTOR INTERACTIONS. *Journal of Biological Chemistry*, 273, 8483-8491.
- TIAN, J., MARINO, R., JOHNSON, C. & LOCKER, J. 2018. Binding of Drug-Activated CAR/Nr1i3 Alters Metabolic Regulation in the Liver. 9, 209-228.
- TROTT, A. J. & MENET, J. S. 2018. Regulation of circadian clock transcriptional output by CLOCK:BMAL1. *PLoS genetics*, 14, e1007156-e1007156.
- TZAMELI, I., CHUA, S. S., CHESKIS, B. & MOORE, D. D. 2003. Complex effects of rexinoids on ligand dependent activation or inhibition of the xenobiotic receptor, CAR. *Nuclear Receptor*, 1, 2.
- UEDA, H. R., CHEN, W., ADACHI, A., WAKAMATSU, H., HAYASHI, S., TAKASUGI, T., NAGANO, M., NAKAHAMA, K.-I., SUZUKI, Y., SUGANO, S., IINO, M., SHIGEYOSHI, Y. & HASHIMOTO, S. 2002. A transcription factor response element for gene expression during circadian night. *Nature*, 418, 534-539.
- ULVEN, S. M., GUNDERSEN, T. E., SAKHI, A. K., GLOVER, J. C. & BLOMHOFF, R. 2001. Quantitative axial profiles of retinoic acid in the embryonic mouse spinal cord: 9-Cis retinoic acid only detected after all-trans-retinoic acid levels are super-elevated experimentally. *Developmental Dynamics*, 222, 341-353.
- UMESONO, K. & EVANS, R. M. 1989. Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell*, 57, 1139-46.
- UMESONO, K., MURAKAMI, K. K., THOMPSON, C. C. & EVANS, R. M. 1991. Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D₃ receptors. *Cell*, 65, 1255-1266.
- VAN DER VEEN, D. R., MINH, N. L., GOS, P., ARNERIC, M., GERKEMA, M. P. & SCHIBLER, U. 2006. Impact of behavior on central and peripheral circadian clocks in the common vole *Microtus arvalis*, a mammal with ultradian rhythms. *Proc Natl Acad Sci U S A*, 103, 3393-8.
- VAN TILBORG, P. J. A., MULDER, F. A. A., DE BACKER, M. M. E., NAIR, M., VAN HEERDE, E. C., FOLKERS, G., VAN DER SAAG, P. T., KARIMI-NEJAD, Y., BOELENS, R. & KAPTEIN, R. 1999. Millisecond to Microsecond Time Scale Dynamics of the Retinoid X and Retinoic Acid Receptor DNA-Binding Domains and Dimeric Complex Formation. *Biochemistry*, 38, 1951-1956.
- VANSANT, G. & REYNOLDS, W. F. 1995. The consensus sequence of a major Alu subfamily contains a functional retinoic acid response element. *Proceedings of the National Academy of Sciences*, 92, 8229.
- VEEN, D. R. V. D. & GERKEMA, M. P. 2017. Unmasking ultradian rhythms in gene expression. *The FASEB Journal*, 31, 743-750.
- VOLLMERS, C., GILL, S., DITACCHIO, L., PULIVARTHY, S. R., LE, H. D. & PANDA, S. 2009. Time of feeding and the intrinsic circadian clock drive rhythms in hepatic gene expression. *Proceedings of the National Academy of Sciences*, 106, 21453-21458.
- VU-DAC, N., GERVOIS, P., GROTZINGER, T., DE VOS, P., SCHOONJANS, K., FRUCHART, J. C., AUWERX, J., MARIANI, J., TEDGUI, A. & STAELS, B. 1997. Transcriptional regulation of apolipoprotein A-I gene expression by the nuclear receptor ROR α . *J Biol Chem*, 272, 22401-4.
- WAGER-SMITH, K. & KAY, S. A. 2000. Circadian rhythm genetics: from flies to mice to humans. *Nat Genet*, 26, 23-27.
- WAHLI, W., BRAISSANT, O. & DESVERGNE, B. 1995. Peroxisome proliferator activated receptors: transcriptional regulators of adipogenesis, lipid metabolism and more. *Chem Biol*, 2, 261-6.

- WALLÉN-MACKENZIE, Å., MATA DE URQUIZA, A., PETERSSON, S., RODRIGUEZ, F. J., FRILING, S., WAGNER, J., ORDENTLICH, P., LENGQVIST, J., HEYMAN, R. A., ARENAS, E. & PERLMANN, T. 2003. Nurr1-RXR heterodimers mediate RXR ligand-induced signaling in neuronal cells. *Genes & Development*, 17, 3036-3047.
- WALTHER, R. F., ATLAS, E., CARRIGAN, A., ROULEAU, Y., EDGEcombe, A., VISENTIN, L., LAMPRECHT, C., ADDICKS, G. C., HACHE, R. J. & LEFEBVRE, Y. A. 2005. A serine/threonine-rich motif is one of three nuclear localization signals that determine unidirectional transport of the mineralocorticoid receptor to the nucleus. *J Biol Chem*, 280, 17549-61.
- WAN, Y. J., CAI, Y., LUNGO, W., FU, P., LOCKER, J., FRENCH, S. & SUCOV, H. M. 2000. Peroxisome proliferator-activated receptor alpha-mediated pathways are altered in hepatocyte-specific retinoid X receptor alpha-deficient mice. *J Biol Chem*, 275, 28285-90.
- WAN, Y. J., HAN, G., CAI, Y., DAI, T., KONISHI, T. & LENG, A. S. 2003. Hepatocyte retinoid X receptor-alpha-deficient mice have reduced food intake, increased body weight, and improved glucose tolerance. *Endocrinology*, 144, 605-11.
- WANG, D., GARCIA-BASSETS, I., BENNER, C., LI, W., SU, X., ZHOU, Y., QIU, J., LIU, W., KAIKKONEN, M. U., OHGI, K. A., GLASS, C. K., ROSENFELD, M. G. & FU, X. D. 2011. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature*, 474, 390-4.
- WANG, D., RENDON, A., OUWEHAND, W. & WERNISCH, L. 2012. Transcription factor co-localization patterns affect human cell type-specific gene expression. *BMC Genomics*, 13, 263.
- WANG, J., SYMUL, L., YEUNG, J., GOBET, C., SOBEL, J., LÜCK, S., WESTERMARK, P. O., MOLINA, N. & NAEF, F. 2018. Circadian clock-dependent and -independent posttranscriptional regulation underlies temporal mRNA accumulation in mouse liver. *Proceedings of the National Academy of Sciences*, 115, E1916.
- WEIGEL, N. L. 1996. Steroid hormone receptors and their regulation by phosphorylation. *The Biochemical journal*, 319 (Pt 3), 657-667.
- WESTIN, S., KUROKAWA, R., NOLTE, R. T., WISELY, G. B., MCINERNEY, E. M., ROSE, D. W., MILBURN, M. V., ROSENFELD, M. G. & GLASS, C. K. 1998. Interactions controlling the assembly of nuclear-receptor heterodimers and co-activators. *Nature*, 395.
- WIENS, M., BATEL, R., KORZHEV, M. & MÜLLER, W. E. 2003. Retinoid X receptor and retinoic acid response in the marine sponge *Suberites domuncula*. *Journal of Experimental Biology*, 206, 3261-3271.
- WIETRZYCH-SCHINDLER, M., SZYSZKA-NIAGOLOV, M., OHTA, K., ENDO, Y., PÉREZ, E., DE LERA, A. R., CHAMBON, P. & KREZEL, W. 2011. Retinoid x receptor gamma is implicated in docosahexaenoic acid modulation of despair behaviors and working memory in mice. *Biological psychiatry*, 69, 788-794.
- WILLY, P. J. & MANGELSDORF, D. J. 1997. Unique requirements for retinoid-dependent transcriptional activation by the orphan receptor LXR. *Genes & development*, 11, 289-298.
- WONGSIRIROJ, N., JIANG, H., PIANTEDOSI, R., YANG, K. J. Z., KLUWE, J., SCHWABE, R. F., GINSBERG, H., GOLDBERG, I. J. & BLANER, W. S. 2014. Genetic dissection of retinoid esterification and accumulation in the liver and adipose tissue. *Journal of Lipid Research*, 55, 104-114.
- WU, G., ANAFI, R. C., HUGHES, M. E., KORNACKER, K. & HOGENESCH, J. B. 2016. MetaCycle: an integrated R package to evaluate periodicity in large scale data. *Bioinformatics*, 32, 3351-3353.
- WU, S., ZHANG, Z. P., ZHANG, D., SOPRANO, D. R. & SOPRANO, K. J. 1997. Reduction of both RAR and RXR levels is required to maximally alter sensitivity of CA-OV3 ovarian tumor cells to growth suppression by all-trans-retinoic acid. *Exp Cell Res*, 237, 118-26.
- XIE, D., BOYLE, A. P., WU, L., ZHAI, J., KAWLI, T. & SNYDER, M. 2013. Dynamic trans-acting factor colocalization in human cells. *Cell*, 155, 713-24.
- XU, H. E., STANLEY, T. B., MONTANA, V. G., LAMBERT, M. H., SHEARER, B. G., COBB, J. E., MCKEE, D. D., GALARDI, C. M., PLUNKET, K. D., NOLTE, R. T., PARKS, D. J., MOORE, J.

- T., KLIEWER, S. A., WILLSON, T. M. & STIMMEL, J. B. 2002. Structural basis for antagonist-mediated recruitment of nuclear co-repressors by PPARalpha. *Nature*, 415, 813-7.
- XU, M. & COOK, P. R. 2008. Similar active genes cluster in specialized transcription factories. *The Journal of Cell Biology*, 181, 615.
- XU, S., GRULLON, S., GE, K. & PENG, W. 2014. Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells. *Methods in molecular biology (Clifton, N.J.)*, 1150, 97-111.
- YAMANAKA, M., ISHIKAWA, T., GRIEP, A., AXT, D., KUMMER, M. P. & HENEKA, M. T. 2012. PPAR γ /RXR α -Induced and CD36-Mediated Microglial Amyloid- β Phagocytosis Results in Cognitive Improvement in Amyloid Precursor Protein/Presenilin 1 Mice. *The Journal of Neuroscience*, 32, 17321-17331.
- YANG, R. & SU, Z. 2010. Analyzing circadian expression data by harmonic regression based on autoregressive spectral estimation. *Bioinformatics*, 26, i168-74.
- YANG, X., DOWNES, M., YU, R. T., BOOKOUT, A. L., HE, W., STRAUME, M., MANGELSDORF, D. J. & EVANS, R. M. 2006. Nuclear Receptor Expression Links the Circadian Clock to Metabolism. *Cell*, 126, 801-810.
- YANG, X., GUO, M. & WAN, Y. J. 2010. Deregulation of growth factor, circadian clock, and cell cycle signaling in regenerating hepatocyte RXRalpha-deficient mouse livers. *Am J Pathol*, 176, 733-43.
- YANG, X., LAMIA, K. A. & EVANS, R. M. 2007. Nuclear Receptors, Metabolism, and the Circadian Clock. *Cold Spring Harbor Symposia on Quantitative Biology*, 72, 387-394.
- YAO, J., ARDEHALI, M. B., FECKO, C. J., WEBB, W. W. & LIS, J. T. 2007. Intranuclear distribution and local dynamics of RNA polymerase II during transcription activation. *Mol Cell*, 28, 978-90.
- YASMIN, R., KANNAN-THULASIRAMAN, P., KAGECHIKA, H., DAWSON, M. I. & NOY, N. 2010. Inhibition of mammary carcinoma cell growth by RXR is mediated by the receptor's oligomeric switch. *J Mol Biol*, 397, 1121-31.
- YASMIN, R., YEUNG, K. T., CHUNG, R. H., GACZYNSKA, M. E., OSMULSKI, P. A. & NOY, N. 2004. DNA-looping by RXR tetramers permits transcriptional regulation "at a distance". *J Mol Biol*, 343, 327-38.
- YEGANEH, M., PRAZ, V., CARMELI, C., VILLENEUVE, D., RIB, L., GUEX, N., HERR, W., DELORENZI, M. & HERNANDEZ, N. 2019. Differential regulation of RNA polymerase III genes during liver regeneration. *Nucleic Acids Res*, 47, 1786-1796.
- YOUNG, M. E., RAZEGHI, P. & TAEGTMEYER, H. 2001. Clock genes in the heart: characterization and attenuation with hypertrophy. *Circ Res*, 88, 1142-50.
- YU, V. C., DELSERT, C., ANDERSEN, B., HOLLOWAY, J. M., DEVARY, O. V., NAAR, A. M., KIM, S. Y., BOUTIN, J. M., GLASS, C. K. & ROSENFELD, M. G. 1991. RXR beta: a coregulator that enhances binding of retinoic acid, thyroid hormone, and vitamin D receptors to their cognate response elements. *Cell*, 67, 1251-66.
- ZENKIN, N. & YUZENKOVA, Y. 2015. New Insights into the Functions of Transcription Factors that Bind the RNA Polymerase Secondary Channel. *Biomolecules*, 5, 1195-1209.
- ZHAN, Q., FANG, Y., HE, Y., LIU, H. X., FANG, J. & WAN, Y. J. 2012. Function annotation of hepatic retinoid x receptor alpha based on genome-wide DNA binding and transcriptome profiling. *PLoS One*, 7, e50013.
- ZHANG, C., XUAN, Z., OTTO, S., HOVER, J. R., MCCORKLE, S. R., MANDEL, G. & ZHANG, M. Q. 2006. A clustering property of highly-degenerate transcription factor binding sites in the mammalian genome. *Nucleic Acids Res*, 34, 2238-46.
- ZHANG, X.-K., HOFFMANN, B., TRAN, P. B. V., GRAUPNER, G. & PFAHL, M. 1992. Retinoid X receptor is an auxiliary protein for thyroid hormone and retinoic acid receptors. *Nature*, 355, 441-446.
- ZHANG, Y., LIU, T., MEYER, C. A., ECKHOUTE, J., JOHNSON, D. S., BERNSTEIN, B. E., NUSBAUM, C., MYERS, R. M., BROWN, M., LI, W. & LIU, X. S. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology*, 9, R137.

- ZHAO, Q., CHASSE, S. A., DEVARAKONDA, S., SIERK, M. L., AHVAZI, B. & RASTINEJAD, F. 2000. Structural basis of RXR-DNA interactions. *Journal of molecular biology*, 296, 509-520.
- ZHAO, Y., QIN, S., ATANGAN, L. I., MOLINA, Y., OKAWA, Y., ARPAWONG, H. T., GHOSN, C., XIAO, J. H., VULIGONDA, V., BROWN, G. & CHANDRARATNA, R. A. 2004. Casein kinase 1alpha interacts with retinoid X receptor and interferes with agonist-induced apoptosis. *J Biol Chem*, 279, 30844-9.
- ZHENG, W., LU, Y., TIAN, S., MA, F., WEI, Y., XU, S. & LI, Y. 2018. Structural insights into the heterodimeric complex of the nuclear receptors FXR and RXR. *Journal of Biological Chemistry*.
- ZHOU, Q. & LIU, J. S. 2008. Extracting sequence features to predict protein-DNA interactions: a comparative study. *Nucleic acids research*, 36, 4137-4148.
- ZOMER, A. W., VAN DER BURG, B., JANSEN, G. A., WANDERS, R. J. & VAN DER SAAG, P. T. 2000. Pristanic acid and phytanic acid: naturally occurring ligands for the nuclear receptor peroxisome proliferator-activated receptor α . *Journal of lipid research*, 41, 1801-1807.