



Cellular factors controlling human L1 retrotransposition

Ramona Nicoleta Galantonu

► To cite this version:

Ramona Nicoleta Galantonu. Cellular factors controlling human L1 retrotransposition. Cellular Biology. COMUE Université Côte d'Azur (2015 - 2019), 2017. English. NNT : 2017AZUR4128 . tel-03058907

HAL Id: tel-03058907

<https://theses.hal.science/tel-03058907>

Submitted on 12 Dec 2020

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

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

École doctorale des Sciences de la Vie et de la Santé

Unité de recherche : IRCAN - CNRS UMR 7284 - INSERM U1081

THÈSE DE DOCTORAT

Présentée en vue de l'obtention du grade de Docteur en Sciences
de l'Université Côte d'Azur - UFR Sciences

Mention : Interactions moléculaires et cellulaires

Défendue publiquement par

Ramona Nicoleta GALANTONU

Cellular factors controlling human L1 retrotransposition

Facteurs cellulaires contrôlant la rétrotransposition du L1 humain

Dirigée par Gaël Cristofari, Directeur de Recherche, INSERM, IRCAN

Soutenue le 11 Decembre 2017

Devant le Jury composé de :

Jean-Ehrland Ricci	Directeur de Recherche, INSERM	Président du Jury
Pascale Lesage	Directeur de Recherche, CNRS	Rapporteur
Séverine Chambeyron	Directeur de Recherche, CNRS	Rapporteur
Gaël Cristofari	Directeur de Recherche, INSERM	Directeur de thèse

This work was supported by:



(ANR-11-LABX-0028-01)



CNRS UMR 7284 - INSERM U1081

“Seek simplicity and then distrust it.”

Alfred North Whitehead

Acknowledgements

I would like to thank Gael, my supervisor, for the patient guidance and advice he has provided throughout my PhD. For his time, his prompt answers, attention to details, for the professional integrity which he instilled in the lab. For asking my opinion and being willing to change his mind if there was a good argument to do so. However, we've never agreed on the same color of the graphs.

I would like to thank all the current and previous members of the team that I had the honor to work with. Thank you to Aurelien for helping me in the lab, for initiating me in the culture of the comic books, for the good music and for inviting me to spend time with his beautiful family. Thanks to Claude for the scientific insights, for his calm, empathy, for being a model of strength and positive energy.

I thank Julian, "my French teacher", for lifting up my moral with his contagious kindness towards people. Thanks to Paula for being such an authentic person and a good friend, to Arpita for her encouragements and to Vivien for her daily attention.

Thanks to the alumni of the lab: to Javi for guiding me in my first year of PhD, Seb for not giving up to our friendship, to Jorge for his laughs, Pilvi for her kindness, Nadira for her singing talents and help in the lab, to Tania for sharing the ups and downs of the PhD process. Also, thanks to Natacha, Dulce, Ashfaq, Baptiste, Christopher for creating a good atmosphere at work.

I would like also to express my gratitude to other people from the institute, for all the great discussions, smiles and jokes, specially Telomeres, Nematostella, Yeast and Transcription factors teams for their enjoyable company.

Special thanks go to Delphine for her kindness and useful suggestions to nuclear receptors chapter of this thesis. I would like to thank "bella Sabrina" for her warm presence, her chocolate donations, her laugh, her optimistic words and help, especially in difficult times.

To Nadir, my desk colleague, for always being willing to help. To Marie-Jo, for her support and for reminding me to take breaks from my work. To Aaron, for his scientific assistance and for the meaningful discussions. To Alex, for always having a story to tell and for reminding me that I need to buy some groceries. To Julien from whom I taught that whatever happens in life, you have to keep moving forward. To Jerome for his humor, candies and understanding. To Alice for being so kind and for making me feel like home when I was her guest. To Rita, Charlene, Serge, Ludo for their help and contribution to a nice work environment.

Other special thanks go to Liudmyla, a dear friend and my western blot guru, for being a great listener, for all her help and advices and for the uncountable ad hoc discussions about everything. Also, thanks to Ben for his amity, joviality and for giving me many rides with his Ben-mobile.

I am thankful to my entire Signalife family for being an example of breaking down the barriers between people from different cultures. Special thanks to Derya for allowing me to isolate in her apartment where I wrote a big part of this thesis, to Rania who knows that our silence means sometimes more than words, to Gaia, Tomas, Sanya, Racha, Hereroa, Margo, Johan and George for feeling good and energetic around them.

A deep gratitude goes to "Maria del mi corazon", my forever great flatmate, my thesis writing buddy, my example of stability and balance.

Elena, my regular dose of motivation and inspiration, receives a big thank you for her valuable professional and personal insights, as well as practical advice when my work/life balance was challenged.

I am so grateful to "Scorpiuta mea" for her continuous help and assistance when the big changes from my life occurred. Her unconditional kindness is hard to describe in words, but her deeds speak for themselves. No lapse of time or distance can lessen our friendship.

I would like also to gratefully acknowledge my teachers and the former lab in Romania for initiating me into scientific research and giving me the courage to pursue a PhD.

I would like to express a deep sense of gratitude to my parents for their trust and for the freedom to choose what I want to do in life. Rarely seeing my family in the last four years was one of the biggest sacrifices I did for this thesis. Thanks for always welcoming me home with open arms.

It is really hard to put into words (mostly because we have already used all of them in our endless phone calls or e-mails:)) the help of my sister, for teaching me how to walk the rough path of personal development. She has been a model to follow in life. If my parents gave me wings, Tia taught me how to fly.

Finally, I would like to thank Jonathan for his unconditional support and love. For understanding the late hours of work, for all the rides from the lab back home, for all his help and for being an open-minded person.

Table of contents

List of figures.....	2
List of abbreviations	3
Abstract.....	5
Résumé	6
Overview of the study	7
1. The world of mobile genetic elements	9
1.1 Varieties of transposable elements	10
1.1.1 DNA transposons move by a cut-and-paste mechanism and are molecular fossils in the human genome.....	11
1.1.2 Retrotransposons mobilize by a copy-and-paste mechanism and are the predominant class of TEs in most mammalian genomes.....	12
1.2 In the spotlight: L1, the only active transposable element in humans.....	19
1.2.1 L1 has a bidirectional promoter in its 5' UTR.....	20
1.2.2 L1 5'UTR contains many binding sites for transcription factors	21
1.2.3 Activities and structure of L1-encoded proteins.....	23
1.2.4 Mechanism of L1 retrotransposition: from L1 transcription to L1 integration	24
1.2.5 The endonuclease-independent pathway EN(i) provides an alternative mechanism of integration.....	27
1.2.6 L1 is the main driver of retrotransposition in humans and impacts human health	27
1.3 Host-L1 interaction and cohabitation. Many LINEs of defense restrict retrotransposition	31
1.3.1 L1 self-control.....	32
1.3.2 Antiviral pathways used against L1 retrotransposon	32
1.3.3 Cellular factors acting on L1 in the nucleus	40
2. Introduction to the world of nuclear receptors.....	47
2.1 The nuclear receptor nomenclature: six subfamilies based on sequence homology.....	48
2.2 The common domain structure of nuclear receptors	54
2.2.1 The N-terminal domain, the least conserved domain of NRs, is transcriptionally active without a ligand.....	55
2.2.2 The DNA Binding Domain recognizes specific genomic template and has a highly organized structure.....	56
2.2.3 The hinge region links the DBD and the LBD and assures the synergy between them	61
2.2.4 The structure and function of the Ligand Binding Domain.....	61
2.2.5 Nuclear receptors expression	52
2.3 Steroid nuclear receptors	64
2.3.1 The signaling pathway of steroid receptors	65
2.3.2 Estrogen related receptor (ERR) family	68
2.4 Functional redundancy and cross-talk between nuclear receptors.....	75

2.4.1	Functional redundancy of nuclear receptors	75
2.4.2	Cross-talk between nuclear receptors	79
3.	Results	83
3.1	Thesis objectives	83
3.2	Research article - in preparation for submission: A subset of steroid receptors interacts with LINE-1 ORF2p and regulate retrotransposition.....	84
4.	General discussion and perspectives	125
4.1	In the tethering model, the integration of transposable elements is guided to specific chromosomal regions by cellular factors	126
4.1.1	A broad diversity of tethering factors contributes to the diversity of integration site preferences among TEs	126
4.1.2	Properties to be considered when evaluating a putative tethering factor.....	127
4.2	Association between steroid receptor and retrotransposon activities.....	131
5.	References	136

List of figures

Chapters 1-2

Figure 1. The composition of the human genome and the distribution of transposable elements (adapted from {Rollins:2005hz}).	10
Figure 2. The structure of a prototype DNA transposon.	12
Figure 3. Schematic representation of LTR-retrotransposons.	13
Figure 4. Non-LTR-retrotransposons in the human genome.	17
Figure 5. Scheme of L1 structure, promoters and 5'UTR binding sites for transcription factors.	22
Figure 6. L1 life cycle.	24
Figure 7. Overview of the cellular factors that impact L1 retrotransposition cycle at different steps. (legend on the next page)	31
Figure 8. The phylogenetic tree of the 48 human nuclear receptors superfamily.	51
Figure 9. A summary of tissue-frequency profiles of the nuclear receptor superfamily in mice.	52
Figure 10. Nuclear receptors and the connection between their expression profile, function and physiological pathways	53
Figure 11. Domain arrangement of nuclear receptors.	54
Figure 12. Nuclear receptor DNA recognition.	58
Figure 13. Variety of nuclear receptors dimerization and binding to DNA response elements.	60
Figure 14. Structure of nuclear receptor ligand binding domain.	62
Figure 15. Schematic of structural changes induced on the estrogen receptor by agonists and antagonists.	64
Figure 16. Phylogenetic tree of steroid nuclear receptors (subfamily 3, NR3).	65
Figure 17. The distinct signaling pathways used in the regulatory actions of estrogen receptors.	66
Figure 18. Structural organization and domain homology of Estrogen Related Receptors (ERRs).	70
Figure 19. A schematic representation of the physiological responses of ERRs.	72

List of abbreviations

APE: Apurinic/Apyrimidinic endonuclease
bp: Base pair
cDNA: Complementary DNA
DNA: Deoxyribonucleic acid
EN: Endonuclease
L1HS: Human-specific L1
L1/LINE1: Long Interspersed Element-1
LTR: Long terminal repeat
MYA: Million year ago
nt: Nucleotide
ORF: Open reading frame
PCR: Polymerase chain reaction
RNA Pol II: RNA polymerase II
RC-L1: Retrotransposition-competent L1
RLE: Restriction-like endonuclease
RNA: Ribonucleic acid
RNase H: Ribonuclease H
RNP: Ribonucleoprotein particle
RT: Reverse transcriptase
SINE: Short Interspersed Element
SVA: SINE-VNTR-Alu
SUMO: Small Ubiquitin-like MOdifier
TPRT: Target-Primed Reverse Transcription
TSD: Target site duplication
UTR: Untranslated region
VLP: Virus-like particle
TE: Transposable element

Abbreviations related to nuclear receptors

AF-1: activation function-1
AF-2: activation function-2
AR: androgen receptor
Agrp: agouti-related peptide
CAR: constitutive androstane receptor
COUP-TF: chicken ovoalbumin upstream promoter receptors
CTE: C-terminal extension
DNA: DNA binding domain
DSS-AHC critical region on the X chromosome, protein 1 (dosage- sensitive sex reversal, adrenal hypoplasia critical region, on chromosome X, protein 1)
DRs: direct repeats
DES: diethylstilbesterol

ER α and ER β : estrogen receptor α and β
ERR: estrogen-related receptors
ERRE: ERR response element
FXR: farnesoid receptor
GR: glucocorticoid receptor
GCNF1: Germ Cell Nuclear Factor variant 1
HNF4: Hepatocyte Nuclear Factor 4
H12: helix 12
HSPs: heat shock proteins
LBD: ligand binding domain
LXR: the liver X receptor
LRH1: liver receptor homolog
LBP: ligand binding pocket
MR: mineralocorticoid receptor
NPY: neuropeptide Y
NTD: N-terminal domain
NBRE: NGFI-B DNA responsive element
ONRs: orphan nuclear receptors
PPAR: peroxisome proliferator-activated receptor
PXR: pregnane X receptor
PR: progesterone receptor
PNR: photoreceptor specific
PPAR γ : peroxisome proliferator-activated receptor γ
RAL: raloxifene
ROR: related orphan receptor
RAR: retinoic acid receptor
REs: response elements
RXR: retinoid X receptor
SRE: steroid response elements
SF1: steroidogenic factor-1
SHP: small heterodimer partner
SRC-1: steroid receptor coactivator-1
SRC-2: steroid receptor coactivator-2
SUMO-1ylation: sumoylation by SUMO1-small ubiquitin-like modifier 1 protein
SHP: small heterodimer partner
TR: thyroid receptor
TLX: tailless homolog
TR2 and TR4: testicular receptor 2 and 4
TCR: T-cell receptor
tRA: trans retinoic acid
VDR: vitamin D receptor
Nur77: Nerve growth factor-induced clone B (IB)
9C-RA: 9-cis retinoic acid

Abstract

The abundance of genetic mobile elements in our DNA has a critical impact on the evolution and function of the human genome. Even if most transposable elements are inactive due to the accumulation of mutational events, the Long INterspersed Element-1 (LINE-1 or L1) retrotransposon continues to diversify and impact our genome, being involved in the evolution of modern humans and in the appearance of genetic diseases or in tumorigenesis. L1 forms 17% of human DNA. It is autonomously active being replicated through an RNA-mediated ‘copy-and-paste’ mechanism. The L1 element encodes two proteins, ORF1p and ORF2p, which associate with the L1 mRNA to form L1 ribonucleoprotein particles, the core of the retrotransposition machinery. However, little is known about the cellular pathways involved in L1 replication. Our laboratory has discovered by yeast 2-hybrid screens an interaction between L1 ORF2p and the estrogen-related receptor α (ERR α), a member of the nuclear receptor family. Here, we confirmed and extended this observation to several other members of the steroid receptor superfamily using a fluorescent two-hybrid assay (F2H) in human cultured cells. To get further insight into the potential role of ERR α in L1 replication cycle, we performed ERR α siRNA-mediated knock-down and overexpression experiments, which suggest that ERR α is a positive regulator of retrotransposition. Moreover, the artificial tethering and concentration of ERR α to a large and repetitive genomic array inhibits retrotransposition. Collectively, these data link steroid signaling pathways with the post-translational regulation of L1 retrotransposition, suggesting a model by which ERR α , and probably several other nuclear receptors, can recruit the L1 RNP to specific chromosomal locations, acting as tethering factors.

Keywords

transposable element, nuclear receptor, host factor, tethering, integration

Résumé

L'abondance d'éléments génétiques mobiles dans le génome humain a un impact critique sur son évolution et son fonctionnement. Même si la plupart des éléments transposables sont inactifs en raison de l'accumulation de mutations, le rétrotransposon LINE-1 (pour *Long Interspersed Element-1* ; ou élément L1) continue de se mobiliser et d'influer sur notre génome. Il a ainsi contribué à l'évolution de l'homme moderne, mais aussi à l'apparition de maladies génétiques et est parfois impliqué au cours de la tumorigénèse. Les séquences du rétrotransposon L1 correspondent à 17% de la masse totale de l'ADN humain. Une copie active de L1 est capable de se mobiliser de manière autonome par un mécanisme de type «copier-coller» qui met en jeu un intermédiaire ARN et une étape de transcription inverse. L'élément L1 code deux protéines, ORF1p et ORF2p, qui s'associent à l'ARNm de L1 pour former des particules ribonucléoprotéiques, le cœur de la machinerie de rétrotransposition. Cependant, peu de choses sont connues sur les voies cellulaires impliquées dans la mobilité de L1. Notre laboratoire a découvert, par des cribles double-hybride, une interaction entre la protéine ORF2p de L1 et le récepteur α associé aux œstrogènes (ERR α), un membre de la famille des récepteurs nucléaires. Ici, nous avons confirmé et étendu cette observation à plusieurs autres membres de la superfamille des récepteurs de stéroïdes en utilisant un test de double-hybride fluorescent (F2H) en culture cellulaire. Pour mieux comprendre le rôle potentiel d'ERR α dans le cycle de rétrotransposition de L1, nous avons effectué des expériences de suppression et de surexpression qui suggèrent qu'ERR α est un régulateur positif de la rétrotransposition. En outre, la liaison et la concentration d'ERR α à un locus répété artificiel (LacO array) inhibe la rétrotransposition. Collectivement, ces données relient les voies de signalisation des stéroïdes avec la régulation post-traductionnelle de la rétrotransposition de L1, ce qui suggère un modèle par lequel ERR α et probablement plusieurs autres récepteurs nucléaires peuvent recruter le RNP L1 vers des emplacements chromosomiques spécifiques, agissant comme facteurs de liaison et d'adressage.

Mots clés : génome, éléments transposables, rétrotransposon, cribles, récepteurs nucléaires, récepteurs de stéroïdes, ERR α , régulation, post-traductionnelle.

Overview of the study

With almost half of the human genome derived from (retro)transposition, transposable elements are accepted now as an important evolutionary force which shapes our genome. Even if most transposable elements are inactive due to a variety of mutation events, the Long INterspersed Element-1 (LINE-1 or L1) retrotransposons form 17% of our genome and their ongoing activity diversifies and impacts our genome. Through insertional mutagenesis, L1s can be agents of both evolution and disease. They can disrupt genes and provoke mutations in the germline, but they can also occasionally fulfill positive functions for the host. Moreover, the most recent advances in deep-sequencing technologies have revealed that L1 is also able to mobilize in somatic cells. Most of known somatic retrotransposition occurs in the brain or in epithelial cancers. Disclosing the mechanisms that regulate L1 activity is very important to better understand the way retrotransposons and their host co-evolved maintaining a balance between retrotransposon proliferation and host genome stability.

However, little is known about the mechanisms used by its cellular host to regulate L1 replication and expression. Thus, one of the goals of the team where I pursue my doctoral studies, is to discover cellular factors and pathways involved in the regulation of L1 retrotransposon mobility and consequently in genome dynamics and instability. In search of regulators influencing L1 retrotransposon, our laboratory has performed yeast 2-hybrid screens to identify cellular factors interacting with the human L1 retrotransposition machinery and identified several potential hits. One of the most robust hit is the estrogen-related receptor α (ERR α), a member of the nuclear receptor superfamily, which are transcription factors controlled by environmental and hormonal signals, and more specifically from the steroid receptor family. Our results, show that other steroid nuclear receptors can physically interact with L1. These observations and the characterization of the interaction, suggest a model by which these cellular factors could regulate retrotransposition at a post-transcriptional level, possibly by tethering the L1 machinery to specific genomic locations. With this work, we aim to study the involvement of steroid nuclear receptors in L1 retrotransposon activity.

Given the importance of steroid receptors in the physiology of mammals and their adaptation to environmental changes, our results raise the intriguing possibility that

steroid nuclear receptors could modulate the landscape of L1 integration in the genome and link environmental and physiological signals with genomic plasticity.

This dissertation starts with two chapters dedicated to literature synthesis, for understanding in details the area of the research project conducted during my PhD training. **Chapter one** gives an overview on the transposable elements field, with emphasis on L1 retrotransposons. **Chapter two** describes the nuclear receptors superfamily, underlining the characteristics of the steroid receptors subfamily and estrogen-related receptors group in particular. **Chapter three** contains the goal of this study and the results, presented in the form of a research article, in preparation for submission. **Chapter four** comprises a general discussion which integrates the implications and conclusions of this study, as well as limitations and perspectives.

1. The world of mobile genetic elements

Contents

1.1	Varieties of transposable elements.....	10
1.1.1	DNA transposons move by a cut-and-paste mechanism and are molecular fossils in the human genome	11
1.1.2	Retrotransposons mobilize by a copy-and-paste mechanism and are the predominant class of TEs in most mammalian genomes	12
1.2	In the spotlight: L1, the only active transposable element in humans	19
1.2.1	L1 has a bidirectional promoter in its 5' UTR	20
1.2.2	L1 5'UTR contains many binding sites for transcription factors.....	21
1.2.3	Activities and structure of L1-encoded proteins	23
1.2.4	Mechanism of L1 retrotransposition: from L1 transcription to L1 integration.....	24
1.2.5	The endonuclease-independent pathway EN(i) provides an alternative mechanism of integration	27
1.2.6	L1 is the main driver of retrotransposition in humans and impacts human health ...	27
1.3	Host-L1 interaction and cohabitation. Many LINEs of defense restrict retrotransposition	31
1.3.1	L1 self-control	32
1.3.2	Antiviral pathways used against L1 retrotransposon	32
1.3.3	Cellular factors acting on L1 in the nucleus	40

“It might seem unfair, however, to reward a person for having so much pleasure over the years, asking the maize plant to solve specific problems and then watching its responses.”

Barbara McClintock

From the classical PNAS article of Barbara McClintock in 1950 on mutable loci in maize, to the complete sequencing of the human genome, it became clearly proved that genomes are not static. The first draft of the human genome reference revealed not only that as few as 2% of our genome is constituted by protein-coding genes, but also that almost half consists of transposable elements (TEs) (Lander et al. 2001). TEs are DNA sequences, which have the ability to move from one part of the genome to another. Barbara McClintock was the first scientist who proposed the existence of genic unities capable of mobility within genomes, (McCLINTOCK 1950). Her landmark work, awarded by the

Nobel prize for medicine in 1983, suggested that TEs, “controlling elements” in maize, respond to environmental influences and induce genomic changes essential for survival. The studies made in maize, showed that transposition induced upon radiation caused chromosomal rearrangements in maize which resulted in survival of the cells facing DNA damage (McCLINTOCK 1950; Fedoroff 2012).

Far from being “junk DNA”, TEs are important genome shapers and widespread across species, being found in all the domains of cellular life division: archaea, bacteria and eukaryote.

1.1 Varieties of transposable elements

Transposable elements can be classified into DNA transposons and retrotransposons based on their requirement for a reverse transcription step in their replication cycle. Their replicative mechanisms are often compared to the computer command: ‘cut-and-paste’ for DNA transposons and ‘copy-and-paste’ for retrotransposons.

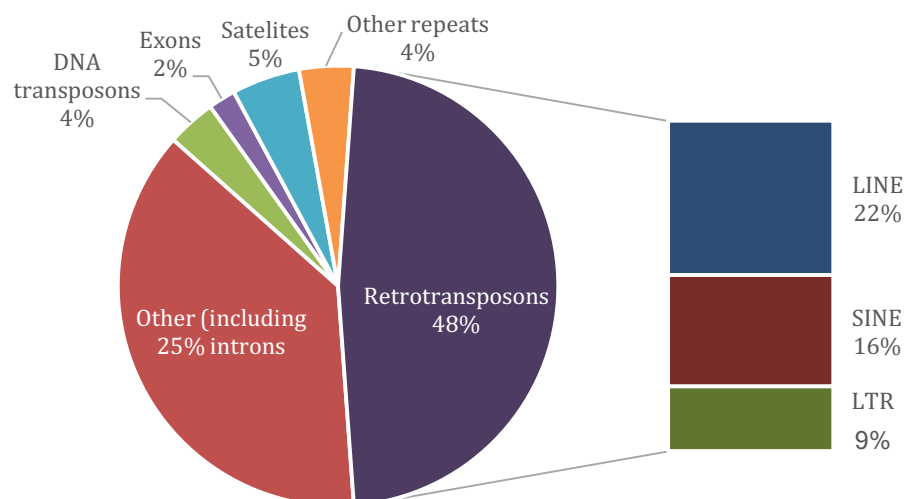


Figure 1. The composition of the human genome and the distribution of transposable elements (adapted from (Rollins 2005)).

1.1.1 DNA transposons move by a cut-and-paste mechanism and are molecular fossils in the human genome

DNA transposons, also known as class II elements, are found in almost all eukaryotes and are mobilized through a DNA intermediate. Classical examples of such mobile elements include the Associator/Dissociator transposons, discovered originally by Barbara McClintock (McCLINTOCK 1950). DNA transposons represent 3% of the human genome (Lander et al. 2001) and transpose by excision from their original location and re-insertion into a new one, using their encoded transposase (Craig 2002). In humans, a recent activity of DNA transposons cannot be detected and they are considered molecular fossils (Pace and Feschotte 2007; Lander et al. 2001). However, this is not true for all Mammals, since highly active DNA transposons have colonized the genome of some bat species (Mitra et al. 2013).

These type of TEs do not require a reverse transcription step to replicate. The simplest form of DNA transposons contains an open reading frame that encodes the transposase protein surrounded by inverted terminal repeat sequences (IR), as seen in Figure 2 (Craig 2002; Richardson et al. 2015). In order to transpose, the mRNA resulting from the transcription of the element is exported to the cytoplasm where its translation leads to the synthesis of the transposase, which has nuclear localization signal, DNA binding and integrase activities (Ivics et al. 1997). Once in the nucleus, the transposase enzyme recognizes and binds near or within the transposon inverted terminal repeats and cuts both strands at each end, to promote transposition by a “cut-and-paste” mechanism (Richardson et al. 2015). The new inserted DNA transposon is generally flanked by short target-site duplications of variable length typical of the type of element (Craig 2002; Richardson et al. 2015).

These TEs can increase their occupancy in the genome by transposing during chromosome replication from a position that has already been replicated to another where the replication fork has not yet passed or they can exploit gap repair following excision to create an extra copy at the donor site (Wicker et al. 2007).

DNA transposons are used as genetic tools in molecular biology applications and in gene therapy, as for *Sleeping Beauty*, a resurrected fish DNA transposon, or *Piggyback*, a cabbage looper moth transposon (Ivics and Izsvák 2010).



Figure 2. The structure of a prototype DNA transposon.

This type of elements occupies 3% of the human genome. In their simplest form, they contain inverted repeats (IR) that surround an open reading frame coding for the transposase, which is essential for their mobility (adapted from (Garcia-Perez, Widmann, and Adams 2016)).

1.1.2 Retrotransposons mobilize by a copy-and-paste mechanism and are the predominant class of TEs in most mammalian genomes

Retrotransposons require an RNA intermediate and the activity of a reverse transcriptase (RT) to expand in the genome, by a replicative copy-and-paste mechanism named retrotransposition (Boeke et al. 1985; Richardson et al. 2015). The original retrotransposon is maintained *in situ*, while a copy is integrated to a new genetic location. This group of TEs, also known as class I mobile elements, represent the predominant class of TEs in most mammalian genomes and they vary by their mechanism of transposition. Two main classes can be distinguished based on the presence of long terminal repeats (LTR) at their extremities: LTR and non-LTR retrotransposons.

1.1.2.1 LTR retrotransposons and endogenous retroviruses are inactive in humans

LTR retrotransposons, and the related endogenous retroviruses (ERVs), are abundant in eukaryotes, mostly in plants where they are the dominant type of transposons. In humans, they comprise approximately 8% of the genome (Lander et al. 2001). LTR-retrotransposons and ERVs share a structure similar to exogenous retroviruses, but LTR-retrotransposons do not possess an envelope gene which obliges them to an intracellular life (see Figure 3) (Bannert and Kurth 2006).

The ERVs, which are permanently integrated retroviruses, accumulated in the genomes through retroviral infections of the germline from the ancient times, have expended in genomes by retrotransposition or reinfection, and are transmitted vertically (Coffin et al. 1997).

It is considered that all human ERVs (HERVs) are inactive due to major deletions and nonsense mutations. However, some recent studies showed that the proviruses of youngest HERV subfamily, HERV-K (where K is the lysine tRNA needed to prime the negative strand cDNA synthesis from an ERV RNA template) can be reanimated using recombinant DNA technology and show infectious activity in human cultured cells (Dewannieux et al. 2006; Y. N. Lee and Bieniasz 2007; Richardson et al. 2015). A possible explanation for this could stand from the trans-complementation of the retrovirus-encoded proteins, which would lead to the formation of functional virus-like particles from partially defective HERVs, resulting in new retrotransposition events (Richardson et al. 2015).

Several studies identified polymorphic HERV-K presence in human population (Belshaw et al. 2005; Macfarlane and Simmonds 2004; Hughes and Coffin 2004; Richardson et al. 2015) and polymorphic ERVs in chimpanzee and gorilla genomes (Yohn et al. 2005), suggesting that ERVs have been active since the divergence of humans and chimpanzees, including in ancestral human species (Richardson et al. 2015).

In contrast with the human genome, the mouse genome contains many active ERV subfamilies, which are responsible for ~10% of all spontaneously mutations arising in mouse (Maksakova et al. 2006).



Figure 3. Schematic representation of LTR-retrotransposons.

These elements represent 8% of the human genome. Full length ERVs are flanked by LTRs necessary for the transcription and maturation of ERV RNAs. They also contain *gag* and *pol* protein coding genes required for retrotransposition. However, they lack a functional *env* gene, used by retroviruses to exit from and re-enter into cells. The mobility of an active ERV includes an RNA intermediate and a copy-and-paste mechanism that is similar to retroviral replication, although they can also multiply intracellularly (adapted from (Garcia-Perez, Widmann, and Adams 2016)).

1.1.2.2 Non-LTR retrotransposons can be autonomous and non-autonomous

Non-LTR retrotransposons do not have LTR sequences and can be divided into two major groups, based on their ability to code the proteins necessary for retrotransposition: autonomous LINEs (Long interspersed elements) and non-autonomous SINEs (Short interspersed elements).

I) **LINE-1 is the only active transposable element in humans**

LINEs comprise approximately 22% of our genome. Most transposable elements, including retrotransposons, in the human genome are inactive, due to a variety of mutations accumulated in the genome during its evolution. With this respect, LINE-1, or L1, represents an exception being the only active and autonomous retrotransposon in modern humans.

L1 codes for proteins (ORF1p and ORF2p) required for its replication. ORF1p is a homotrimeric RNA-binding protein and ORF2p possesses endonuclease (EN) and reverse transcriptase (RT) activities (Figure 4). Altogether, they assemble with the L1 RNA into a ribonucleoprotein particle, which forms the core of the retrotransposition complex.

L1 ORF1p and ORF2p proteins show a strong *cis*-preference and bind their own encoding mRNA (Esnault, Maestre, and Heidmann 2000; Wei et al. 2001). However, they can also mobilize in *trans* non-autonomous non-LTR retrotransposons, as Alu and SVA (Raiz et al. 2012; Hancks et al. 2011; Hancks and Kazazian 2012) and other mature cellular mRNAs, resulting in the formation of processed pseudogenes (Esnault, Maestre, and Heidmann 2000). The hallmarks of L1-mediated insertions include the presence of target site duplications (TSD) that can vary in size and sequence, a 3' poly(dA) tail of variable length, an L1 endonuclease recognition sequence related to the TT/AAAA consensus at the preintegration locus, possible 5' truncations or inversions, and the lack of introns (Beck et al. 2011; Jurka 1997).

More details about the L1 element structure and life cycle will be presented in the next section, 1.2.

II) **SINEs are non-coding retrotransposons mobilized by L1 in trans**

SINEs represent 25% of our genome, they are represented by Alu and SVA elements and require the L1 replicative machinery for their mobilization.

i) Alu elements are the most abundant class of non-autonomous retrotransposons in human

Alu elements contribute to 11% of our genome. With more than 1 million copies, they form the most abundant class of non-autonomous retrotransposons in humans (Lander et al. 2001). Alu elements originate from the cellular 7SL RNA, the RNA moiety of the signal recognition particle (Ullu and Tschudi 1984). They are approximately 300 bp long (Rubin et al. 1980), with a bipartite structure, being composed of two highly similar left and right monomers, separated by a central adenosine-rich region (Figure 4).

The left monomer (most 5' region of the element) contains conserved A and B box sequences needed for RNA polymerase III dependent transcription (Batzer and Deininger 2002; W. M. Chu, Liu, and Schmid 1995). The right monomer ends in a poly(A) tract of variable length and the genomic DNA flanking the 3' end of an Alu element contains an RNA polymerase III terminator sequence which is a stretch of four to six consecutive thymidines (Batzer and Deininger 2002; Richardson et al. 2015). The RNA polymerase III bypasses the poly(A) tract and continues in the downstream flanking genome sequence until it meets a stretch of thymidines (W. M. Chu, Liu, and Schmid 1995). The size of the poly (A) tract and that of the downstream genomic sequence will influence its expression and retrotransposition potential (Comeaux et al. 2009; Dewannieux and Heidmann 2005; Batzer and Deininger 2002; Richardson et al. 2015).

Alu elements contain three major subfamilies and each of them can be further subdivided: Alu J - the oldest lineage with peak activity ~65 million years ago, S - the second oldest lineage mostly active 30 million years ago, and Y- the youngest Alu lineage (Batzer and Deininger 2002). The active Alu “core elements” in the modern human genome belong to the AluY and AluS subfamilies (E. A. Bennett et al. 2008). The most recent AluY subfamilies (AluYa5, AluYb5, AluYd8) are sufficiently polymorphic (presence or absence in a given locus) within the human population, to be used, together with polymorphic L1Hs, as genetic markers in phylogenetic and population genetics studies (Witherspoon et al. 2006; Richardson et al. 2015).

Alu elements do not code for any protein. Instead, they hijack the L1 machinery for their mobilization. The poly(A) tail of the Alu RNA competes with L1 RNA for binding ORF2p in order to replicate successfully (Doucet, Wilusz, et al. 2015; Dewannieux, Esnault, and Heidmann 2003; Mills et al. 2007). In contrast with L1 RNA where the poly (A) tail is added through the classical polyadenylation pathway, the Alu poly(A) is

genetically encoded in the sequence itself. The poly(A) is the site for reverse transcription priming and this is a feature common for all L1-mobilized template RNAs — L1, Alu, SVA and cellular mRNAs (Doucet, Wilusz, et al. 2015; Esnault, Maestre, and Heidmann 2000; Kajikawa and Okada 2002; Monot et al. 2013).

ii) Human SVA elements have a composite structure and are mobilized in trans by L1

SVA elements represent only 0,2% of the human genome, with approximatively 2700 copies (Lander et al. 2001). They originated 25 million years ago and their 2 kb sequence consists of several elements: a hexameric CCCTCT repeat, an inverted Alu-like element repeat, a variable number of GC-rich tandem repeats (VNTRs), a SINE-R sequence similar to a LTR retrotransposon and a canonical cleavage polyadenylation specificity factor (CPSF) binding site, followed by a poly(dA) tract, as seen in Figure 4(Richardson et al. 2015).

The SVA mRNA is an RNA polymerase II transcript, even if no internal RNA pol II promoter has been detected so far (Hui Wang et al. 2005). Importantly, SVA integration in the genome shows hallmarks of L1-mediated mobility (Hancks and Kazazian 2012; Raiz et al. 2012).

SVA elements can also be divided into subfamilies, based on sequence similarities. As for Alu, the youngest SVA subfamilies - SVA-D, SVA-E, SVA-F - show insertional polymorphism in the human population, suggesting that they can be mobilized in modern humans (Hancks and Kazazian 2012; Richardson et al. 2015). Consistently, SVA insertions can cause human genetic diseases (Ostertag et al. 2003), as exemplified in cases of X-linked Dystonia-Parkinsonism or Fukuyama-type congenital muscular dystrophy (Kobayashi et al. 1998).

iii) The L1 machinery can mobilize cellular mRNAs in trans and form processed pseudogenes

The L1 retrotransposition machinery can mobilize its own RNA, the RNAs of non-autonomous retrotransposons, but also cellular protein coding RNAs (Esnault, Maestre, and Heidmann 2000) and small nuclear RNAs, such as U6 (Doucet, Droc, et al. 2015). The integrated copies of the mobilized mRNA show the hallmarks of L1-mediated

insertions. Since they lack introns and promoter, they are referred as processed (retro)pseudogenes (Figure 4) (Vanin 1984; Wei et al. 2001).

Due to the loss of the promoter in the retrotransposition process, to 5' truncations and other rearrangements, processed pseudogenes are in general non-functional. However, some of them can evolve to become functional and support new cellular functions. For instance, the integration of the cyclophilinA mRNA inside the TRIM5 results in a functional fusion protein which provides a new defense mechanism to restrict HIV mobility in owl monkeys (Malfavon-Borja et al. 2013; Nisole et al. 2004; Sayah et al. 2004). Such genes are referred as retrogenes.

Analysis of the human reference genome revealed that there are approximately 8000 to 17000 pseudogenes, and 70% of them have a retrotranspositional origin (processed), while the rest arose by segmental duplication (nonprocessed) (Torrents et al. 2003; Zhang 2003).

Most of the pseudogenes are derived from ribosomal protein and their presence and formation continues to contribute to the genomic diversity between individuals (Kazazian 2014; Ewing et al. 2013). The successful recapitulation of U6/L1 chimeric pseudogene formation in cultured human cells, suggests that U6/L1 pseudogene formation is continuing in the human genome (Garcia-Perez et al. 2007; Gilbert et al. 2005).

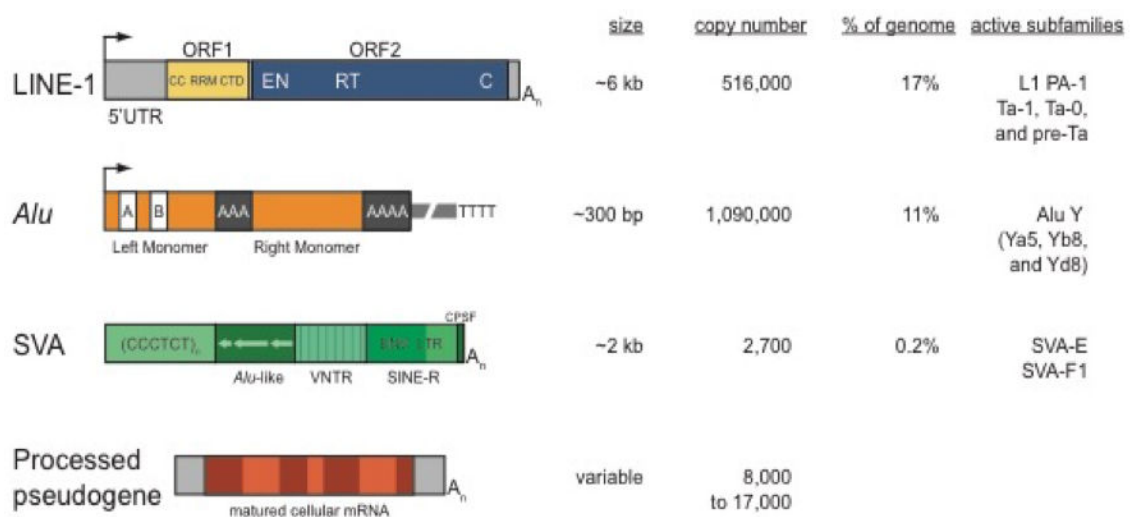


Figure 4. Non-LTR-retrotransposons in the human genome. (legend on the next page)

For each non-LTR retrotransposon class is mentioned its name, structure, average size, copy number, percentage in the reference genome sequence, and the name of the active subfamilies. Human LINE-1 structure is presented again to give an overview on different non-LTR-retrotransposons and consists of: a 5' untranslated region (UTR) containing sense and antisense internal promoters (black arrows); ORF1, which encodes an RNA-binding protein with a coiled-coil domain (CC), an RNA recognition motif (RRM), and a carboxyl-terminal domain (CTD); an inter-ORF spacer (grey box between ORF1 and ORF2); ORF2, which encodes a large protein with endonuclease (EN), reverse transcriptase (RT), and cysteine-rich domains (C); a 3' UTR with a polyadenylation signal, and a poly(dA) tract (A_n downstream of 3' UTR). Alu elements have the following structure: 7SL-derived monomers (orange boxes); RNA polymerase III transcription start site (black arrow) and conserved cis-acting sequences required for transcription (A and B white boxes in left 7SL-derived monomer); adenosine-rich fragment (AAA grey box between left and right 7SL-derived monomers); a terminal poly(dA) tract (AAAA grey box); a variable sized flanking genomic DNA (interrupted small grey rectangle) followed by the RNA pol III termination signal (TTTT). SVA elements are composite elements made of: hexameric CCCTCT repeats (CCCTCT)_n; inverted Alu-like repeat (green box with backward arrows); a GC-rich VNTR; a SINE-R sequence sharing homology with HERVK-10, (envelope (ENV) and long terminal repeat (LTR)); a cleavage polyadenylation specific factor (CPSF) binding site; and a terminal poly (A) tail (A_n). The structure of a processed retropseudogene consists of a spliced cellular mRNA with its 3' UTR and sometimes a 5' UTR (grey boxes) and coding ORF (red boxes for human and purple boxes for mouse, boxes are interrupted by exon-exon junctions (vertical black lines) (Richardson et al. 2015). Boundaries of all these elements are defined by the target-site duplications, which are generated during the integration process (not shown).

1.2 In the spotlight: L1, the only active transposable element in humans

L1 retrotransposon represents ~17% of our genome and continues to shape it, being the only active and autonomous retrotransposon in humans. Still, the majority of L1s are not competent for retrotransposition, because they contain mutations or are 5' truncated (Dombroski et al. 1991; Goodier and Kazazian 2008; Beck et al. 2010). L1 retrotransposons have been amplifying in mammalian genomes for more than 160 million years and in humans since the divergence of the ancestral mouse and human lineages approximately 65 to 75 million years ago (Smit et al. 1995; Richardson et al. 2015). Analysis of sequence comparison revealed 16 primate-specific L1 subfamilies (from L1PA1 to L1PA16, where a high number reflects an older family) (Smit et al. 1995; H. Khan, Smit, and Boissinot 2006). Older L1 subfamilies are replaced over evolutionary time with new ones, in a process known as subfamily succession which might be driven by host proteins that restrict L1 expression (Jacobs et al. 2015; Castro-Diaz et al. 2014; Richardson et al. 2015).

It is estimated that, currently, only 80-100 copies of the human-specific L1 (L1Hs, another name for the youngest L1PA1 subfamily), are still retrotransposition-competent in our genomes (Brouha et al. 2003). These active elements belong to a small group named transcribed active subset (Ta-subset) (Boissinot, Chevret, and Furano 2000).

The L1Hs-Ta-subset can be discriminated from older members of the L1Hs family by a diagnostic nucleotides in the 3' untranslated region (UTR), more specifically ACA and G at positions 5930-5932 and 6015, respectively, instead of GAG and A in older elements (Dombroski et al. 1991; Richardson et al. 2015). Based on additional diagnostic nucleotides, the L1 Ta-subset can be further subdivided into Ta-1, Ta-0 and pre-Ta (Boissinot, Chevret, and Furano 2000). These lineages reflect the mode of L1 amplification and the existence of a limited number of highly active progenitor sequences.

General overview on L1 structure and life cycle

An active full length human-specific L1 element is 6kb long and contains a 5' untranslated region (5' UTR), two open reading frames – ORF1 and ORF2 – separated by a short inter-ORF spacer, a 3'UTR and ending with a poly(dA) tail. L1 sequence itself is

generally flanked by short target site duplications (rarely longer than 20 bp). Recently, it was discovered that L1 5' UTR contains an additional ORF, named ORF0, in opposite orientation relative to L1, which codes for a 70 amino-acid long peptide named ORF0p and can form 5' fusion transcripts with nearby genes by splicing. It is not very well characterized so far, but its overexpression *in trans* slightly enhances L1 retrotransposition in cultured cells (Denli et al. 2015).

The process of L1 genomic amplification is named retrotransposition (Figure 6). As a general overview, it starts by the RNA Pol II-dependent transcription of L1 from its internal promoter which is located in the 5' UTR (Swergold 1990). The resulting mRNA is exported in the cytoplasm, translated into two proteins, ORF1p and ORF2p, which once synthesized bind their own mRNA, forming a ribonucleoprotein particle (RNP). Next, the L1 RNP is imported in the nucleus where it cleaves the target locus and initiates reverse transcription in a coordinated process termed target-primed reverse transcription (TPRT) (Beck et al. 2011).

1.2.1 L1 has a bidirectional promoter in its 5' UTR

The 5' UTR it is ~900 bp long and contains a bidirectional promoter (Figure 5). The first 155 nt are important for the activity of the sense promoter from which full length L1 transcription starts, as shown by deletion analysis (Swergold 1990). Initially promoter characterization suggested that sense transcription starts exclusively at position +1 (Swergold 1990), but a subsequent study showed that transcription initiation can occur at variable sites close to the 5' end (Athanikar, Badge, and Moran 2004; Lavie et al. 2004). Moreover, the upstream chromosomal region can also influence negatively or positively the activity of the promoter (Lavie et al. 2004).

Besides this promoter, the 5'UTR houses an antisense promoter (ASP) which can initiate transcripts that extend into the sequences upstream of L1. The ASP between nucleotides 400-600 of the 5'UTR and is not essential for retrotransposition. However, its activity can produce chimeric transcripts with neighboring genes or sequences (Speek 2001; Nigumann et al. 2002). The combined sense and antisense promoter activities can produce double-stranded RNA, which triggers RNA interference (RNAi) mechanisms, either by generating small interfering RNAs (siRNAs) which will repress

retrotransposition (Nuo Yang and Kazazian 2006), or by the direct cleavage of L1 sense RNA by the Microprocessor complex (Heras et al. 2013).

1.2.2 L1 5'UTR contains many binding sites for transcription factors

The 5'UTR of L1 contains cis-acting elements, binding motifs for transcription and other cellular factors, able to regulate L1 transcription (Figure 5).

The ubiquitous nuclear transcription factor YY1 (ying yang 1) binds close to L1 transcription start site (TSS), at position 13-21 (Minakami et al. 1992; Becker et al. 1993; Athanikar, Badge, and Moran 2004). YY1 is known to regulate gene transcription, both positively and negatively (Austen, Lüscher, and Lüscher-Firzlaff 1997). Located near the 5'terminus of L1, YY1 binding is important for the precision of transcription initiation at position +1 (Athanikar, Badge, and Moran 2004).

L1 promoter accommodates also two binding sites for SOX transcription factors, members of the SRY family (Tchénio, Casella, and Heidmann 2000). The identified consensus sequences in the L1 promoter, are functional, being required for SOX binding and SOX-mediated transactivation of a luciferase reporter gene initiated from the L1 promoter (Tchénio, Casella, and Heidmann 2000).

RUNX3 (Runt-related transcription factor 3) binds in the 5' UTR of L1 at position 83-101, where a consensus Runt-domain core binding sequence was identified (Nuo Yang et al. 2003), and can modulate L1 transcription and retrotransposition. The overexpression of RUNX3 increases L1 activity. In contrast, its depletion by siRNA, the expression of a RUNX3 dominant-negative form, or mutations of the RUNX3 binding site, strongly suppress L1 transcription and retrotransposition. An additional RUNX3 motif identified at position 526-508 could regulate L1 ASP activity (Nuo Yang et al. 2003).

The T-cell factor/lymphoid enhancer factor (TCF/LEF) can bind SOX and SOX/LEF DNA regulatory elements within the body of L1 sequence. This could regulate cryptic bidirectional promoters and the activity of nearby neuronal genes during adult neurogenesis (Kuwabara et al. 2009).

Finally, the transcription factor p53 is another DNA-binding protein capable of activating or repressing transcription. It is a critical tumor-suppressor, with many target genes and

multiple binding sites in the entire genome (Hollstein et al. 1991; Zhao et al. 2000). In particular, p53 binding sites have been identified in the promoter of many L1 sequences, where this transcription factor can bind a short 15-nt site, as shown by gel shift assay, mutational analyses and chromatin immunoprecipitation experiments, leading to p53-dependent transcriptional activation (C. R. Harris et al. 2009). This might seem contradictory, given the role of p53 as guardian of genome stability in somatic cells. p53 activity and its apoptotic activity are repressed in the germ line when critical recombination intermediates appear. This would correspond to the time when L1 is also active in the germline and would not trigger any DNA damage response, being able to integrate and pass to the next generation. In normal somatic cells, p53 is fully functional and protects the genome from double-stranded breaks as those possibly generated by L1 endonuclease (C. R. Harris et al. 2009).

But their working model proposes a positive feedback loop for the genome. In this scenario, p53 activates L1 transcription which leads to more L1 endonuclease mediated dsDNA breaks which in turn will create enough p53 damage response activity to induce apoptosis of the cell. This mechanism would decrease L1 retrotransposition in somatic cells (C. R. Harris et al. 2009). Consistently, p53 was found to inhibit somatic retrotransposition in many different model organisms (Wylie et al. 2016).

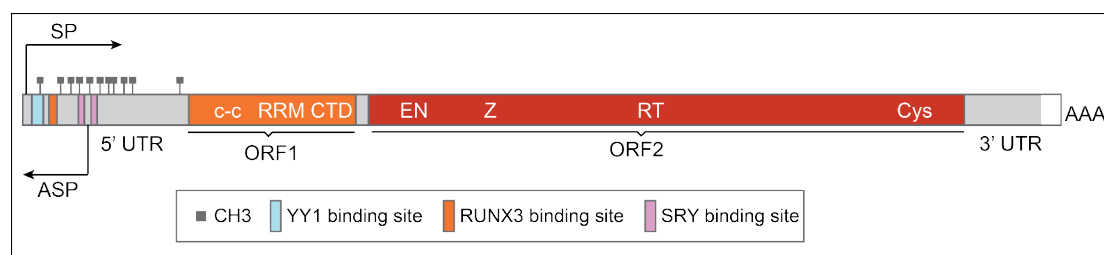


Figure 5. Scheme of L1 structure, promoters and 5'UTR binding sites for transcription factors.

The full-length L1 element consists of five main components: the 5'UTR, the two open reading frames (ORF1 and ORF2), inter-ORF region, the 3' UTR and poly-A tract (AAA). The 5' UTR contains the self-transcribing promoter (SP) function, antisense promoter (ASP) activity, a CpG island that is usually methylated (represented as CH3) and binding sites for transcription factors: one RUNX3 (orange), one YY1 (light blue), and two SRY (pink). ORF1 (orange box) has a coiled-coil (c-c) domain, an RNA recognition motif (RRM), and a C-terminal domain (CTD). ORF2 (red box) contains the endonuclease (EN) and the reverse transcriptase (RT) domains and a

cysteine-rich (cys) motif, whose exact role is unknown, but is important for the L1 retrotransposition. The 3' UTR contains the poly-A signal (adapted from (Ade et al. 2017)).

1.2.3 Activities and structure of L1-encoded proteins

The studies focused on the open reading frames showed that L1 mRNA is bicistronic and produces ORF1p and ORF2p, both necessary for retrotransposition.

ORF1p is a 40 kDa protein (Leibold et al. 1990) which binds RNA (Hohjoh and Singer 1997) and has nucleic acid chaperone function (Martin and Bushman 2001). ORF1p contains three major domains: a coiled coil domain, an RNA recognition motif (RRM domain) and a carboxyl terminal domain (CTD). The coiled-coil domain drives the trimerization of ORF1p proteins (Martin et al. 2003; Khazina and Weichenrieder 2009). This process is mediated by the regular organization of leucines and other hydrophobe amino-acids, which leads to the formation of a leucine zipper structure (Holmes, Singer, and Swergold 1992; Hohjoh and Singer 1996). The N-terminal region of ORF1p is poorly conserved, in contrast with the well-conserved CTD.

The second reading frame encodes a 150 kDa protein, named ORF2p with both endonuclease (EN) (Q. Feng et al. 1996) and reverse transcriptase (RT) activities (Mathias et al. 1991), critical for L1 retrotransposition. ORF2p comprises three main domains: an N-terminal apurinic/apyrimidic endonuclease like domain (APE), a reverse transcriptase domain and a C-terminal cysteine-rich domain of unclear function (Fanning and Singer 1987).

The bicistronic L1 mRNA can be translated to produce large amounts of ORF1p, but very limiting quantities of ORF2p (Wei et al. 2000). The translation of ORF1p initiates at the first AUG codon (Leibold et al. 1990). ORF2p is translated also from the bicistronic mRNA, from the first in-frame AUG codon. ORF2p translation does not need ORF1 sequence, nor the spacer sequence (Alisch et al. 2006). It was suggested that ORF2p is produced by an unconventional termination/reinitiation mechanism, in which the 40S ribosomal subunit would stay bound to the L1 RNA after reaching ORF1 stop codon and would scan the RNA until finding in the 5' UTR of ORF2 a cis-acting sequence where the ribosomes can reassembly and continue translation (Alisch et al. 2006).

1.2.4 Mechanism of L1 retrotransposition: from L1 transcription to L1 integration

The expansion of L1 in the genome happens through the continuous replication of functional progenies. The process of L1 genomic amplification is called retrotransposition and consists in three principal steps: L1 transcription, ORF1p and ORF2p translation, reverse transcription and integration (Figure 6).

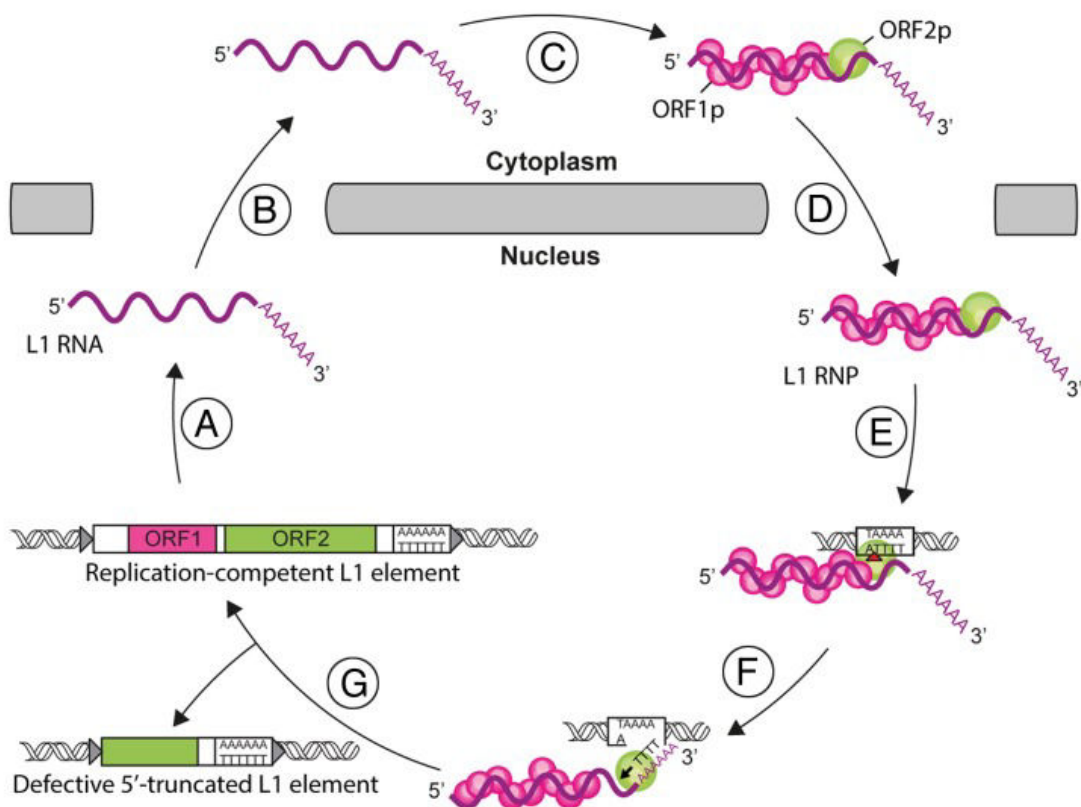


Figure 6. L1 life cycle.

The retrotransposition of a L1 replication-competent element starts by the transcription of a bicistronic mRNA (A). The L1 RNA is exported to the cytoplasm (B). ORF1p and ORF2p proteins are synthesized, bind to their parental L1 RNA and form L1 ribonucleoprotein particles (RNP) (C). The L1 RNP is imported into the nucleus (D). Reverse transcription and integration occur at the genomic target site. (E). First, the L1 endonuclease (EN) activity nicks the target DNA (red arrowhead). (F). Then, the L1 reverse transcriptase (RT) initiates the reverse transcription of L1 RNA (black arrowhead) (G). The mechanisms involved in the final steps of

this process and the resolution of the integration are unresolved yet. Partial reverse transcription can lead to 5'-truncated L1 copies (Viollet, Monot, and Cristofari 2014).

1. L1 transcription.

As previously mentioned, the retrotransposition is initiated from the generation of a full-length L1 mRNA from its internal promoter located in the 5'UTR.

2. ORF1p and ORF2p translation.

The L1 mRNA is transported in the cytoplasm and used by the host cellular translational machinery to synthesize L1 proteins. Once synthesized, the two ORFp proteins bind their own mRNA and form a ribonucleoprotein particle (RNP) which will be the machinery used for the generation and integration of a new L1 copy (Kulpa and Moran 2006; Martin 1991). The association of the L1 proteins with their mRNA is known as cis-preference (Wei et al. 2001; Esnault, Maestre, and Heidmann 2000). How the two L1 proteins associate with their own mRNA was easier to study for ORF1p, which is most abundant. Some studies showed that ORF1p polymerizes at the translation site, facilitating the binding to their own mRNA (Callahan et al. 2012; Furano 2000). ORF2p which is present at much lower quantities was detectable bound to L1 RNA using an epitope and RNA tagging strategy of an engineered human L1 (Doucet et al. 2010). L1 RNPs accumulate in cytoplasmic foci closely related to stress granules (Doucet et al. 2010; Goodier et al. 2007).

More recently, it was shown that ORF2p preferentially associates with the 3' poly A tract of L1 RNA in cis, but also with the polyA tract of Alu RNA and rarely other cellular mRNAs, for trans-mobilization. Moreover, the replacement of the L1 polyA by a triple helix from the long-non-coding RNA MALAT1, which substitute the polyA stabilizing function at the end of the RNA, abolishes L1 mobilization in cis, but not the trans-mobilization of Alu sequences, reinforcing the notion that the polyA tract is necessary for ORF2p recruitment into the L1 RNP (Doucet, Wilusz, et al. 2015).

3. Reverse transcription and integration

The last step is taking place in the nucleus. The transport of the L1 RNP in the nucleus is not yet understood. In theory, it might require passing the nuclear pore complex (NPC) or

might happen during mitosis when the nuclear membrane is disrupted (Görlich and Kutay 1999). However, two studies showed that L1 retrotransposition can take place independently of cell divisions (Kubo et al. 2006; Macia et al. 2017) even if it could be more efficient in dividing cells (Y. Xie et al. 2013).

L1 integration occurs by target-primed reverse transcription (TPRT), a classical endonuclease dependent mechanism (Luan et al. 1993; Cost et al. 2002). TPRT is a process where the reverse transcription and integration are coupled. The endonuclease domain of ORF2p recognizes and cleaves DNA at degenerated consensus genomic sequences 5'-TTTT/A-3', exposing a 3' hydroxyl end (Luan et al. 1993; Q. Feng et al. 1996; Jurka 1997; Eickbush and Jamburuthugoda 2008). Then, the 3'OH end serves as a primer for ORF2p reverse transcriptase activity, and L1 RNA as a template, to generate the first strand cDNA. The second strand, synthesized after a second EN-mediated cleavage on the other strand, could be synthesized by L1 RT activity, which also show DNA-dependent DNA-polymerase activity, or by a cellular DNA polymerase, completing the formation of this additional L1 copy at this new position. Because the two EN cleavage sites are staggered, the new L1 copy is flanked by target site duplications, of variable size from one insertion to another, but rarely longer than 20 bp. Importantly, most L1 insertions are 5' truncated, being unable to retrotranspose again and only 5-6% are full-length (Gilbert, Lutz-Prigge, and Moran 2002; Gilbert et al. 2005). It is not known exactly why the new L1 insertions are often truncated. A connection with DNA-repair mechanisms might be involved (Coufal et al. 2011; J. Suzuki et al. 2009).

Also, some L1 insertions can exhibit a 5' inversion as a result of a mechanism called twin priming (Ostertag and Kazazian 2001). In this situation, one of the two single-strand overhangs anneals to the internal part of L1 RNA, while the other hybridizes to its poly(A) tail, both being used as primers. The two L1 cDNA are completed through microhomology-driven single-strand annealing.

ORF2p is a critical component of the retrotransposition machinery. Mutations in the endonuclease domain invalidates the cleavage in the genomic DNA and consequently prevent L1 retrotransposition in most situations (see next paragraph) (Q. Feng et al. 1996; Moran et al. 1996). On the same note, mutations in the RT domain abolishes the reverse transcriptase activity and retrotransposition.

1.2.5 The endonuclease-independent pathway EN(i) provides an alternative mechanism of integration

An alternative mechanism of retrotransposition is the endonuclease-independent pathway -EN(i). In this pathway, a 3'OH extremity at a dysfunctional telomere or at genomic DNA lesions is directly used as a primer for reverse transcription, without the need for an EN cut (Morrish et al. 2002; Morrish et al. 2007; Sen et al. 2007). EN(i) insertions are usually 3'truncated and lack TSDs and can be experimentally observed when using EN-defective L1 elements. Interestingly, EN(i) is only observed in cells defective for NHEJ (non-homologous end-joining) and with impaired p53 function (Morrish et al. 2002; Coufal et al. 2011). An important study (Morrish et al. 2007) pointed on similarities between the EN(i) and telomerase activity because both use a 3'OH for priming reverse transcription at DNA lesion or chromosome ends. It was proposed that EN(i) retrotransposition is an old RNA-mediated DNA repair mechanism, which was associated with non-LTR retrotransposons before they incorporated their endonuclease domain (Morrish et al. 2007).

1.2.6 L1 is the main driver of retrotransposition in humans and impacts human health

With almost half of the human genome derived from (retro)transposition, TEs are accepted now as an important evolutionary force which shapes our genome. Moreover, because they can insert into genes and provoke mutations, some of which can result in disease, TEs have the potential to be harmful to their hosts. Consistently, 124 cases of human genetic diseases caused by L1-mediated insertions have been reported (Hancks and Kazazian 2016), and include L1, Alu, SVA, and even processed pseudogenes insertions. The insertions cause a variety of diseases, such as cystic fibrosis, hemophilia, muscular dystrophy, autoimmune disorders, neurofibromatosis and occur mainly in coding exons, but also in introns and in promoters, disrupting gene function (Hancks and Kazazian 2016). The number of disease-causing insertions is likely underestimated due to the inherent limits of the most common approaches currently used to identify causative mutations in genetic disease (exome sequencing). Historically, a screen of hemophilia A patients for pathogenic mutations in the Factor VIII gene (F8) on chromosome X revealed

two *de novo* germline L1 insertions causing the disease. This discovery, made in 1988, was the first evidence for ongoing L1 activity in humans. Most probably, the L1 insertions in the F8 gene occurred in the gametes or in early embryogenesis, as the parents of the patients did not have the insertions (Kazazian et al. 1988).

L1 retrotransposition takes place mainly in the germline and early embryo, being inherited by the next generation (Kano et al. 2009). It was approximated that a new L1 integration takes place between 1/100 and 1/200 births (Ewing and Kazazian 2010; Cordaux and Batzer 2009) and most of these events results from the activity of the highly active and polymorphic L1s, named ‘hot L1s’ (Brouha et al. 2003). Each individual has, besides some of the reference L1 copies, hundreds of non-reference L1HS-Ta, which contribute to the diversity of the human genome (Beck et al. 2010; Mir, Philippe, and Cristofari 2015).

The most recent advances in deep-sequencing technologies have revealed that L1HS-Ta is also able to mobilize also in somatic cells. The discovery of L1 activity in adult somatic tissues, as in brain (Muotri et al. 2010; Coufal et al. 2009; Evrony et al. 2012; Upton et al. 2015) and epithelial somatic tumors (Miki et al. 1992; Iskow et al. 2010; Tubio et al. 2014; Ewing et al. 2015; Solyom et al. 2012; Scott and Devine 2017), changed the misconception that L1 retrotransposes is silenced in adult somatic cells. Still, the detection of somatic insertions remains challenging due to the difficulty to isolate rare events within a large population of cells. However, the clonal expansion of tumor cells has rendered this search easier.

Through their insertional activity, L1s can be agents of both evolution and disease. They can disrupt genes and provoke mutations, but they can also fulfill positive functions for the host. For example, when telomere function is disrupted, L1 preferentially inserts into the region of telomeres, protecting their function (Morrish et al. 2007). Moreover, telomerases likely derive from ancient reverse transcriptases encoded by retrotransposons (Belfort, Curcio, and Lue 2011).

1.2.6.1 Somatic L1 insertions are a characteristic of human epithelial cancers

The first somatic L1 insertion was identified in colorectal cancer (and by definition was absent from the adjacent normal colon tissue), suggesting that it occurred during the tumorigenesis (Miki et al. 1992). This insertion disrupted the last exon of the APC tumor suppressor gene, and constitutes a driver mutation in this context. However, it remained difficult to determine the developmental timing of a retrotransposition event (Richardson et al. 2015). Since then, next-generation sequencing brought a tremendous contribution to L1 discovery, enabling the identification of thousands of somatic L1 insertions in all types of human epithelial cancer (Iskow 2010; Helman 2012) (E. Lee et al. 2012; Scott and Devine 2017; Solyom et al. 2012; Rodić et al. 2014; Strevia et al. 2015). Yet, the overall contribution of L1 mutagenic activity, both as a source of driver mutations, and as contributor of cancer genome instability, remains to be defined.

A clear example has been recently described in colon cancer by Scott Devine's laboratory. Indeed, they identified an L1 insertion in the same exon of the APC tumor suppressor gene in which in 1992 Miki and colleagues found the first somatic L1 insertion (Scott et al. 2016). This insertion occurred at a different position in the exon, spanned 1,4 kb and is clinically significant, serving as an early driver mutation in the development of the tumor, as suggested by promoter methylation and RNA-sequencing analysis which revealed that the source element escaped silencing in the adjacent, normal tissue (Scott et al. 2016).

L1, Alu and SVA insertions are occasionally drivers of genetic diseases, but in cancers the mechanisms which keep under control the mobile DNA are dysregulated. This allows L1 to be expressed and be a dynamic component of cancer genomes (Burns 2017).

The silencing of TEs results also from DNA methylation and L1 promoter is typically methylated (Yoder, Walsh, and Bestor 1997). Consistently, the promoter of some L1Hs copies can become hypomethylated (Alves, Tatiro, and Fanning 1996) and this correlates with increased expression and retrotransposition (Shukla et al. 2013).

L1 promoter hypomethylation has been associated with genomic instability (Daskalos et al. 2009), poor outcomes in colon cancer (Ogino et al. 2008) and oesophageal squamous cell carcinoma, poor prognosis in non-small-cell lung cancer (K. Saito et al. 2010), decreased survival and drug resistance in young patients with breast cancer (van Hoesel et al. 2012), recurrence of hepatocellular carcinoma after resection and poor survival in ovarian cancer (reviewed in (Burns 2017)).

1.2.6.2 L1 ORF1p overexpression is a hallmark of human cancers

L1 ORF1p protein is overexpressed in human cancers and constitutes a hallmark, being detected by immunohistochemistry in half of all epithelial tumors (Rodić et al. 2014). Detection of ORF2p is rendered more difficult due to its very low level of expression and the lack of properly validated antibodies. However, immunostaining with a recently developed antibody against ORF2p (De Luca et al. 2016) suggests that it could be expressed in most intestinal adenomas, colon, lung and prostate adenocarcinomas and breast carcinoma (De Luca et al. 2016). Thus, ORF1p and ORF2p could be used as biomarkers for screening or predicting clinical outcomes, even if more investigation is needed to validate their clinical relevance (Burns 2017).

1.3 Host-L1 interaction and cohabitation. Many LINEs of defense restrict retrotransposition

The mutagenic potential of retrotransposons is a threat to genomic stability both in the germ line and in somatic cells, and therefore should be limited and strictly controlled. Hence the way retrotransposons and their hosts co-evolved reflects their intimate relationships. The maintenance of retrotransposon sequences in genomes primarily relies on their selfish ability to proliferate. However, retrotransposons derived-sequences are sometimes exapted by the host, providing new cellular functions, such as the rewiring of host gene regulatory networks.

The classification of L1 regulators can follow diverse styles, from cellular localization, to factor function or how it acts on L1 retrotransposition (Figure 7).

Although, I will give a broad overview of the most studied L1 regulators, I will particularly detail the known regulators of retrotransposition acting at the latest stages of replication, once the L1 RNP is formed and has entered into the nucleus.

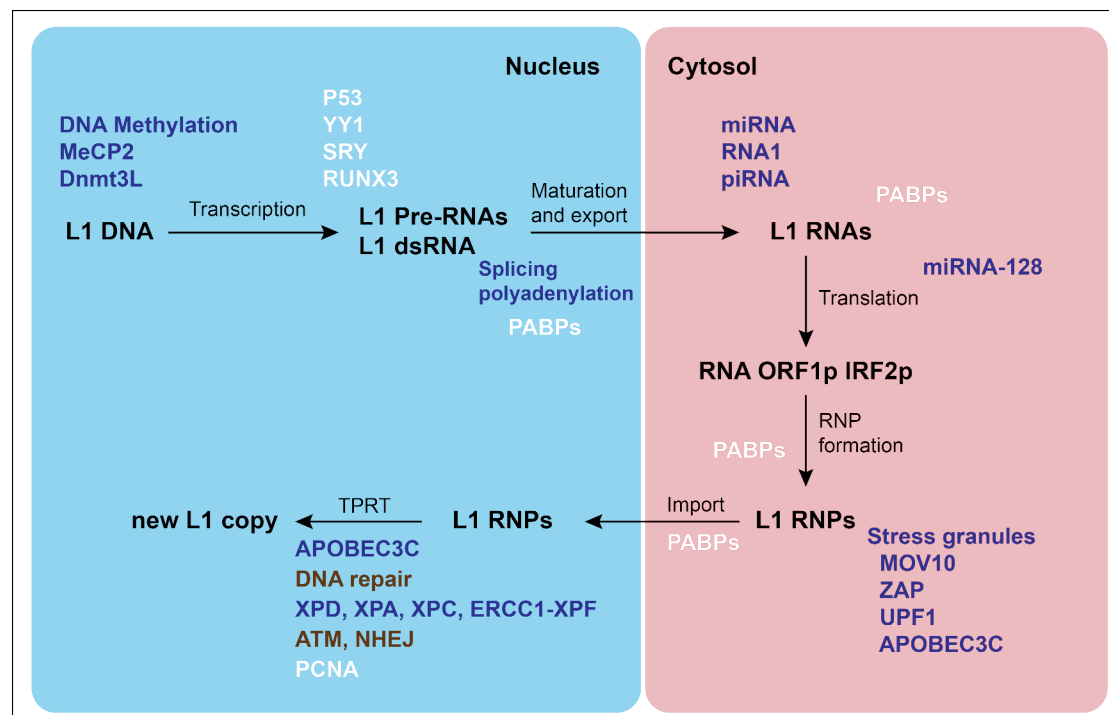


Figure 7. Overview of the cellular factors that impact L1 retrotransposition cycle at different steps. (legend on the next page)

The cellular factors with positive effects on L1 mobility are shown in white, the factors that reduce L1 activity are in blue and the factors with dual effect (both positive and negative) are shown in white (adapted from (Ade et al. 2017)).

1.3.1 L1 self-control

As mentioned before, most L1s become dead at the moment of insertion, through 5' truncations, inversions and internal deletions. The ones which successfully insert in their full-length form, might subsequently suffer mutations and DNA recombinations (Ostertag and Kazazian 2001). Parts of ORFs or 5'UTR are prone to elimination by cryptic splice sites (Belancio, Hedges, and Deininger 2008). L1 contains a weak polyadenylation tail and most L1 transcripts are prematurely truncated due to cryptic polyadenylation signals along the A-rich L1 sequence (Perepelitsa-Belancio and Deininger 2003). Also, L1 is poorly expressed because of its own inadequate transcription elongation (Han and Boeke 2004). Post-translational modifications, as the required phosphorylation of ORF1p, add to the regulation of retrotransposition (Cook, Jones, and Furano 2015). L1 body contains many binding sites for transcription factors as p53, YY1, RUNX3, SOX, as previously mentioned and their mutagenesis affects retrotransposition (Minakami et al. 1992; Becker et al. 1993; C. R. Harris et al. 2009).

1.3.2 Antiviral pathways used against L1 retrotransposon

To alleviate the effect of L1 insertions on genome, the cell is using diverse lines of defense. Many are also antiviral mechanisms, used against both viruses and retrotransposons. These act as a first protective barrier for genome safety and the way they function is mainly based on metabolizing nucleic acids. Also, most of the times they are sensed by the immune system and induced by type I interferons (Goodier 2016).

1.3.2.1 MOV10 helicase binds L1 RNA and reduces its accumulation

Moloney leukemia virus 10 (MOV10) is an ATP-dependent RNA helicase with active role in the RNA interference pathway. It was identified as an antiviral protein, preventing mice from being infected by the Moloney leukemia virus (Mooslehner et al. 1991). It was also shown to reduce infectivity of lentiviruses, such as HIV-1 (Furtak et al. 2010;

Xiaojun Wang et al. 2010; Huang et al. 2015). Several studies have also identified MOV10 as a negative regulator of non-LTR retrotransposons.

First, Goodier and colleagues identified MOV10 as being associated with L1 RNPs by mass spectrometry analyses, using samples from 293T cells transiently expressing an engineered retrotransposition-competent L1 element in which ORF1p was epitope-tagged (Goodier, Cheung, and Kazazian 2012). This association is RNA dependent, being lost after RNAase treatment of the immunoprecipitate, suggesting that it affects L1 RNA metabolism. Moreover, MOV10 and the L1 RNP colocalize in cytoplasmic foci, related to stress granules, where L1 RNPs were previously shown to accumulate (Doucet et al. 2010; Goodier et al. 2010; Goodier, Cheung, and Kazazian 2012). In the proposed hypothesis, MOV10 recruits L1 RNPs to the stress granules where they might be degraded by small RNA pathways (Goodier, Cheung, and Kazazian 2012).

Other mechanisms could be involved, as RNAi silencing, given the known association between MOV10 and RISC (Meister et al. 2005). Consistent with both possibilities, knockdown of endogenous MOV10 increases L1 RNA levels and L1 retrotransposition, while its overexpression reduces L1 RNA levels and inhibits retrotransposition, which makes MOV10 a potent inhibitor of retrotransposition (Goodier, Cheung, and Kazazian 2012; X. Li et al. 2013). Similarly, MOV10 depletion strongly increases SVA and Alu retrotransposition, which relies on the L1 machinery.

1.3.2.2 Zinc finger Antiviral Protein restricts L1 in cytoplasmic stress granules

Another antiviral factor restricting L1 is the Zinc finger Antiviral Protein (ZAP) (Goodier et al. 2015; J. B. Moldovan and Moran 2015).

Goodier et al. identified ZAP as a potential LINE-1 inhibitor by screening interferon stimulated antiviral proteins for their potential effect in a cellular L1 retrotransposition assay (Goodier et al. 2015). In an independent study, published at the same time, Moldovan and Moran performed affinity chromatography and mass spectrometry experiments to identify cellular partners of L1 ORF1p, also leading to the identification of ZAP (J. B. Moldovan and Moran 2015). Both studies showed that ZAP inhibits human L1, but acts on other non-LTR elements as well, such as mouse L1, zebrafish LINE-2 and Alu. Making use of biochemical, genetic and fluorescence microscopy approaches, both they showed that ZAP colocalizes with ORF1p in stress granules, associates with L1 RNP in an RNA-dependent manner, and restricts L1 mobility.

1.3.2.3 TREX1-mediated L1 repression

To provide protection against virus, host cells can detect foreign nucleic acids and activate type I interferon (IFN) antiviral mechanisms which will stop the infection (Stark et al. 1998). Among the factors induced by this antiviral response, three-prime repair exonuclease 1 (TREX1) functions in a cytosolic antiviral pathway that detects DNA, named IFN-stimulatory DNA (ISD) response (Stetson et al. 2008; Okabe et al. 2005; Ishii et al. 2006). TREX1 is one of the most abundant 3' to 5' DNA exonucleases in cells and it depletes both single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA) (Lindahl, Gally, and Edelman 1969; Höss et al. 1999; Mazur 2001). Releasing digested ssDNA/dsDNA from the nucleus might prevent TREX1-mediated immune activation. If depletion of the DNA fails, an autoimmune response will be activated (Grieves et al. 2015). Mutations which affect TREX1 enzymatic activity, as well as mutations in *SAMHD1* and *RNASEH2* genes, are associated with Aicardi-Goutieres syndrome, a neurodegenerative disease present in infancy. This disease is characterized by autoimmunity activated by the presence of nucleic acids which typically are eliminated from the cells (Crow et al. 2006). The syndrome mimics a congenital viral infection, similar with cytomegalovirus or rubella virus (Crow et al. 2003) with no pathogens detected in patients.

In the search of finding TREX1 substrates that could trigger autoimmunity, it was shown that in TREX1 deficient heart cells, single-stranded DNA derived from L1 accumulates. The detected ssDNA mainly maps to the 3' end of the consensus L1 sequence, which can reflect the abundance of defective, 5'-truncated L1 elements in the genome (Ostertag and Kazazian 2001; Stetson et al. 2008). The premise was that TREX1 is using its exonuclease activity to inhibit L1 by digesting the reverse-transcribed cDNA, given that in TREX1 knockout cells activity of L1 is upregulated (Stetson et al. 2008).

However, the use of D200N TREX1 catalytic mutant, but exonucleolytically active, did not have an effect on L1 (Stetson et al. 2008). This can suggest that catalytic mutants can act as dominant negative and inhibit endogenous TREX1, or can reflect a different working model.

Recently, another study reported an exonuclease-independent implication of TREX1 in preventing DNA damage responses generated by L1 retrotransposons (P. Li et al. 2017). TREX1 reduces ORF1p levels through a post-translational mechanism: it binds ORF1p

and target it to proteasome. Since ORF1p is an essential component of L1 RNPs, TREX1 could suppress L1 by disrupting RNP formation.

TREX1 is localized in the endoplasmic reticulum (ER) (Stetson et al. 2008). Thus understanding how ORF1p reach TREX1 remains unknown. It was proposed that TREX1 localized to ER might act as guardian and digest L1 DNA at the nucleus exit before it activates cytosolic immune response (P. Li et al. 2017).

Interestingly, reverse transcriptase inhibitor AZT treatment had no effect on Trex1 deficient mice. This could be due to the presence in mice of different RT enzymes in the genome, not all sensitive to AZT. A combination of different reverse transcriptase inhibitors could be more effective. In a different study it was shown how two clinically approved RT inhibitors used against HIV, are efficient against MLV and showed ameliorated myocarditis in Trex1-knockout mice (Beck-Engeser, Eilat, and Wabl 2011).

1.3.2.4 SAMHD1 reduces cellular pool of dNTPs

As mentioned before, mutations in *TREX1*, *SAMHD1* and *RNASEH2* genes are associated with the Aicardi-Goutieres syndrome disease. Remarkably, SAMHD1 and TREX1 were both found to repress L1 retrotransposition, but this was not explored for RNaseH2 yet. Of note, SAMHD1 is a nuclear protein, but because of its antiviral functions and association with TREX1 and the Aicardi-Goutieres syndrome disease, will be presented in this section and not in the next one, which focuses on cellular factors acting on L1 in the nucleus.

SAMHD1, SAM domain and HD domain containing protein 1 (SAMHD1) is a nuclear factor, initially identified as human homologue of mouse IFN- γ induced protein in dendritic cells (N. Li, Zhang, and Cao 2000). Its role in immunity was then reconfirmed through studies which show that Samhd1 is upregulated in viral infections and can restrict retroviruses (Lahouassa et al. 2012). For example, it inhibits HIV and herpes simplex virus 1 replication, by interfering with the synthesis of viral DNA by hydrolysis of triphosphate deoxynucleoside at the α -phosphate position (Goldstone et al. 2011; Kim et al. 2012; Powell et al. 2011). More precisely, SAMHD1 rapidly and specifically hydrolyses dGTP and has no nuclease activity against ssDNA, dsDNA or RNA, being a deoxyguanosine triphosphate triphosphohydrolase (Goldstone et al. 2011). As a result of hydrolysis, dGTP is converted to guanosine and inorganic triphosphate. Importantly,

SAMHD1 can also hydrolyse other deoxynucleotides, but only upon activation by hydrolysed dGTP, leading to the global depletion of dNTP cellular pool (Goldstone et al. 2011). SAMHD1 is localized in the nucleus (Rice et al. 2009), mainly expressed in myeloid lineages - dendritic cells, monocytes and macrophages -, but is found at basal levels expressed in most tissues (Laguette et al. 2011).

In their study, Goldstone and colleagues, show that SAMHD1 degrades dNTP pools. Together SAMHD1 and TREX1 might cooperate to control synthesis and degrade cytosolic nucleic acids derived from endogenous retroelements. The loss of function of TREX1 leads to accumulation of ssDNA which might be the cause of the autoimmunity observed in AGS (Stetson et al. 2008; Thomas et al. 2017). In a similar way, a functional SAMHD1 represses reverse transcription by limiting dNTPs pools, but a loss of function will result in accumulation of cytosolic DNA and trigger the inappropriate interferon response reported in ASG (Goldstone et al. 2011) .

Zhao 2013 shows for the first time that SAMHD1 represses endogenous retroelements, notably L1. SAMHD1 overexpression leads to reduced ORF2p expression and L1 reverse-transcription in purified L1 RNP is inhibited in presence of SAMHD1, suggesting that ORF2p might be the primary target of this host factor. SAMHD1 also affects ORF2p-dependent trans-mobilization of Alu and SVA elements. It should be underlined that the repression of retroviruses and retrotransposons might use different pathways, because SAMHD1 represses HIV only in nondividing cells, but can inhibit L1 in dividing cells.

SAMHD1 is a nuclear protein, but triggers RNP sequestration in cytoplasmic stress granules

Another study reported an alternative mechanism by which SAMHD1 can inhibit retrotransposition (S. Hu et al. 2015). SAMHD1 stimulates stress granules formation where L1 RNP will be sequestered. This might indicate that SAMHD1 acts on multiple and different cellular compartments and pathways to inhibit as much as possible, many pathogens. It is not clearly understood how the development of cytoplasmic stress granules could be intensified by SAMHD1, a nuclear protein. One hypothesis proposed that fine-tuning the phosphorylation of eIF4a and/or interaction of eIF4G and eIF4A could be involved. As mentioned before, L1 was found in other studies associated with

cytoplasmic foci colocalizing with stress granules (Goodier et al. 2007; Goodier et al. 2010; Doucet et al. 2010). Depletion of stress granules components such as TIA1 or G3BP1 abrogates the ability of SAMHD1 to inhibit L1 retrotransposition, consistent with a role of stress granules in regulating L1 mobilization.

1.3.2.5 RNase L cleaves L1 RNA and blocks RNP formation

RNase L is known for inhibiting viral replication through cleavage of viral RNA, being part of an antiviral system, regulated by interferon. OAS (oligoadenylate (2-5A) synthetase) genes encode IFN inducible enzymes which act as response to viral dsRNAs (Silverman 2007). When activated, OAS will modulate the switch from dormant RNaseL to its enzymatically active dimer (Dong and Silverman 1995). RNase L cuts single-stranded viral RNA, repressing viral amplification (Silverman 2007).

The retrotransposon restriction activity of RNaseL was revealed in human cultured cells where RNase L can suppress retrotransposition of L1, but also of IAP (A. Zhang et al. 2014). The catalytically inactive mutant does not have any effect on L1 or IAP, and the reduction of endogenous RNase L through siRNA results in almost a 2-fold augmentation of L1 retrotransposition. The data suggests that RNase L cleaves L1 RNA post-transcriptionally, leading to a decrease in expression of the two L1 proteins. No colocalization was observed between RNase L and L1 cytoplasmic foci, which is not surprising given that L1 RNA degradation could eliminate L1 RNP (A. Zhang et al. 2014).

1.3.2.6 The APOBEC proteins cytidine-deaminate L1 DNA

The APOBEC proteins (Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like) are DNA cytosine deaminases which accomplish diverse roles in immunity. The APOBEC3 subfamily encodes seven proteins in humans (A, B, C, DE, F, G and H), which inhibit the replication of diverse set of retroviruses (Chiu and Greene 2008; Jarmuz et al. 2002; Sheehy et al. 2002; Alce and Popik 2004; Bishop et al. 2008). One way these proteins act is by making mutations in their target single-stranded DNA through deamination of cytosines and their conversion into uracils by the cytidine deaminase (CDA) domains.

The discovery of APOBEC3G (A3G) as a specific antagonist of HIV-1 Vif, opened the way to a new class of antiviral drugs (Bishop et al. 2008) and several studies focused on revealing a possible effect of APOBEC proteins on retrotransposons.

The ability of A3 proteins to repress L1 was shown by studies of A3 proteins overexpression and knockdown in cultured cells. Some studies suggested that L1 inhibition can be a feature of either some A3 members (H. Chen et al. 2006; Muckenfuss et al. 2006; Bogerd et al. 2006) or of the entire subfamily (Niewiadomska et al. 2007; Kinomoto et al. 2007).

So far it was shown that APOBEC3A (A3A), A3B and A3F seem to be the most potent to restrict L1 retrotransposition.

- **A3A is a nuclear protein and deaminates L1 RNA during TPRT**

It was suggested that deaminase activity is not involved in L1 inhibition by A3A, because G to A mutation were not found in L1 sequence to indicate deamination (Muckenfuss et al. 2006). However, more recently it appears that A3A inhibits L1 retrotransposition through deamination of transiently exposed single-strand L1 DNA intermediates during TPRT (Richardson et al. 2014). Still, these mutations can be repaired by the cellular apurinic/apyrimidinic endonuclease (APE) and uracil DNA glycosylase (UNG) (Krokan, Drabløs, and Slupphaug 2002). Inhibition of UNG restored L1 activity in presence of A3A. De novo L1 insertions in presence of A3A and inhibitors of UNG, showed that the insertions have mutations generated by A3A deamination. Thus, after A3A deamination, the mutations are detected and repaired by UNG and L1 is integrated without containing the scars of A3A deamination (Richardson et al. 2014).

- **A3B restricts L1 in a broad range of cell types**

A3B is also a nuclear protein (Bogerd et al. 2006) and can restrict engineered L1 retrotransposition. shRNA-mediated knockdown of endogenous A3B increases L1 retrotransposition efficiency in HeLa, hESCs and iPSCs cells (Wissing et al. 2011; Marchetto et al. 2013). Heritable retrotransposition events might be the target of A3B retrotransposition restriction, since it is highly expressed in the early embryo where heritable integrations can also occur (van den Hurk et al. 2007; Wissing et al. 2011; Richardson et al. 2014).

- **A3C interaction with ORF1p mediates L1 restriction in a deaminase-independent manner**

A3C is the most abundantly expressed of all A3 proteins in many cell types. Their antiviral role against viruses as HIV and HBV involves cytidine deamination function, as indicated by the mutations found in these viral genomes (Henry et al. 2009).

However, as for L1 restriction, A3C does not require a catalytically active CDA domain. The CDA domain of A3C contains an RNA-binding pocket, essential for L1 RNA binding and L1 repression. Also, A3C dimerization is necessary for L1 inhibition. It was suggested that the complex formation between L1 ORF1p and A3C would mediate the transport of L1 RNPs into stress granules for degradation.

- **A3G oligomerizes to restrict L1 by a different mechanism than cytidine deaminase**

Several studies of A3 proteins on L1 retrotransposition, reported that A3G does not inhibit L1 (Muckenfuss et al. 2006; Stenglein and Harris 2006; H. Chen et al. 2006). However, other studies showed that A3G is also inhibiting L1, but probably by a different mechanism than cytidine deamination given that sequence analysis did not show any DNA deamination effect of hA3G on *de novo* L1 insertions (Kinomoto et al. 2007). Or maybe the mechanism is the same but the rapid complementary degradation of the mutated L1 transcripts by other cellular enzymes makes them undetectable (Kinomoto et al. 2007; Bogerd et al. 2006). Moreover, A3G oligomerization and amino acids in positions 24–28 are very important for its inhibitory activity against L1 and *Alu* retrotransposons (Koyama et al. 2013).

A3F has a cytoplasmic expression as A3G and it was also reported to act on L1 repression (Muckenfuss et al. 2006; Stenglein and Harris 2006). A3F is highly expressed in testes, a profile which might implicate A3F as a male germ cell-specific barrier for L1 retrotransposition (Stenglein and Harris 2006).

A3DE showed very little effect on L1 mobility (Stenglein and Harris 2006).

The expression pattern of A3 proteins seems to differ from one cell line to another and to differentially repress L1 retrotransposition. It is not very clear if the subcellular distribution of A3 proteins is playing a role in L1 inhibition, since both cytosolic and nuclear A3 can suppress L1 retrotransposition. However, A3 cytoplasmic localization might be connected to their antiviral activity (Kinomoto et al. 2007; Lackey et al. 2012; Stenglein, Matsuo, and Harris 2008; Pak et al. 2011). A3A, A3C, A3H are small proteins at 23kDa, but A3DE A3F A3G are 46 kDa and larger than the limit size to passively diffuse into the nucleus (Görlich and Kutay 1999). For the cytoplasmic A3s inhibiting L1 replication at nuclear steps of its life cycle, the proposed explanation was that A3 would associate with the L1 RNA in the cytosol and would be carried with the L1 RNP to the nucleus, blocking the action of ORF2p and/or the TPRT process (Bogerd et al. 2006; Kinomoto et al. 2007).

1.3.3 Cellular factors acting on L1 in the nucleus

Beside the cellular antiviral factors that also repress L1 mobilization, a number of cellular factors function as negative or positive regulators of L1 retrotransposon, in particular in the nucleus after L1 RNP entry (Figure 7).

1.3.3.1 PCNA binds L1 ORF2p through a canonical PIP box and can modulate L1 life cycle

Proliferating cell nuclear antigen (PCNA) is a cofactor of cellular DNA polymerases, and is important for genomic stability (G.-L. L. Moldovan, Pfander, and Jentsch 2007).

PCNA increases the processivity of DNA polymerases and offers in the same time a moving platform where diverse factors can act during replication. Proteins interacting with PCNA contain a PCNA-interacting peptide (PIP) box (Xu et al. 2001; G.-L. L. Moldovan, Pfander, and Jentsch 2007). PCNA is a member of the sliding clamp protein family. It forms a ring which encircles DNA and can slide on both directions freely. PCNA monomer is made from α helices and β strands forming similar globular domains at N terminal and C terminal extremities, linked by an connecting loop. Three molecules of PCNA connect their heads and tails to form the ring complex, with positively charged α helices facing DNA and external surface made by β sheets (Krishna et al. 1994; G.-L.

L. Moldovan, Pfander, and Jentsch 2007). Once PCNA is loaded around DNA, it tethers polymerases to DNA, preventing their dissociation from the DNA template.

Interestingly, PCNA was identified as an interactor of L1 ORF2p in a proteomic study of L1 RNPs (Taylor et al. 2013). Moreover a conserved PIP box was found between the EN and RT domains of ORF2p. Four amino-acids within the PIP domain are conserved in the ORF2p sequence in many species. Co-purification of ORF2p and PCNA is abolished when the conserved residues are mutated. The same mutations decrease L1 retrotransposition efficiency. Consistently, shRNA mediated knockdown of PCNA reduces L1 activity. Interestingly, intact EN and RT ORF2p activities are required for the association of ORF2p with PCNA, suggesting that this process takes place only after EN cleavage and initiation of reverse transcription (Taylor et al. 2013). Collectively, the results highlight the importance of the interaction between ORF2p and PCNA through PIP box for L1 retrotransposition.

Beside a possible ORF2p processivity factor, PCNA could help recruiting repair enzymes, such as DNA ligases (Ulrich 2011). PCNA also recruits RNase H2 to degrade RNA/DNA hybrids formed during DNA replication (RNA primers, misincorporated rNTPs) (Bubeck et al. 2011). Noteworthy, mammalian L1 does not code for such RNase H activity as many other retroelements (Xiong and Eickbush 1990), although it could be important to release the first strand cDNA to allow second strand synthesis. Thus, it is tempting to speculate that PCNA could recruit RNase H2 at the site of TPRT to achieve this goal.

1.3.3.2 UPF1 is a dual regulator of L1 retrotransposon

UPF1 co-purifies with both ORF1p and ORF2p, but the L1-UPF1 interaction is sensitive to RNase treatment. UPF1 is key player of the nonsense-mediated decay (NMD) pathway, known to recognize and scan mRNAs with premature termination codons and target them for degradation by other factors (Shigeoka et al. 2012). Since the L1 mRNA is bicistronic and intronless, it represents a very good potential UPF1 target which could recognize the three stop codons in the inter-ORF region (Hogg and Goff 2010). Using different parts of L1 ORFs, interORF and 3'UTR, it was showed that the ORF2 sequence or protein is required for L1-UPF1 interaction. Interestingly, the silencing of UPF1 increases L1 mRNA levels, but decreases retrotransposition. This has led to the proposal that UPF1 has a dual role in L1 regulation. On the one hand, Upf1 is a positive regulator

of L1 activity in RNP function and on the other hand, it acts as a negative regulator repressing retrotransposition.

1.3.3.3 DNA repair pathways and L1 endonuclease-mediated DNA breaks

Target-primed reverse transcription (TPRT), the current model of non-LTR retrotransposon reverse transcription, involves nicks made by the endonuclease domain of ORF2p, in the genomic DNA at the target site of insertion. These cleavages could be sensed as DNA lesions or double-strand breaks (DSB) which threaten the stability of the genome and activate DNA repair pathways, such as non-homologous end-joining (NHEJ) (Gasior et al. 2006). It was reported also that L1 integration process shows to be quite inefficient because the DSB induced during TPRT exceeds the final insertions rate (Gasior et al. 2006).

H2AX's phosphorylated form (γ -H2AX) localizes to chromatin and forms foci as a mark for double-strand breaks resulted upon L1 expression and activation of its endonuclease which generates the DSB-like intermediate of TPRT. Human L1 overexpression in HeLa cells leads to genomic DSBs and ORF2p mutation in its endonuclease domain results in loss of H2AX foci, in line with the critical role played by the EN for retrotransposition (Moran et al. 1996; Gasior et al. 2006). The γ -H2AX immunolocalization of DSB and their disappearance during a time-course analysis points toward a DNA repair response involvement.

The role of NHEJ is to preserve the integrity of the chromosomes and consists in the ligation of two broken ends. This process results in DNA junctions with little or no homology, often creating small modifications as deletions or insertions (Mehta and Haber 2014; Chiruvella, Liang, and Wilson 2013).

The canonical NHEJ mechanism can religate DNA fragments ending with a 3'-overhang using mainly Ku proteins (heterodimer of two proteins- Ku70 and Ku80), DNA ligase IV, the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs) and Xrcc4 (Walker, Corpina, and Goldberg 2001; Chiruvella, Liang, and Wilson 2013; Spagnolo et al. 2006). Ku heterodimer binds to the DNA break where DNA-PKcs is recruited through an interaction with Ku80 (Hammel et al. 2010; S. M. Bennett et al. 2012). DNA-PKcs binds to the DNA ends and keeps them in proximity (Hammarsten, DeFazio, and Chu 2000). The activity of DNA-PK implies self-phosphorylation and phosphorylation and

recruitment of other NHEJ proteins (Spagnolo et al. 2006; Douglas et al. 2005; Yu et al. 2008). DNA ligase IV in a complex with Xrrc4 is responsible for the ligation of the broken ends (Grawunder et al. 1998). Several other proteins are implicated in NHEJ, as Artemis which will process the incompatible broken ends to make them suitable for ligation. Importantly, the DNA breaks are still repaired by alternative NHEJ methods when the activity of canonical key components as Ku or DNA Ligase IV is deficient (Yan et al. 2007).

The role of DNA repair mechanisms in L1 retrotransposition is not clearly solved, different studies reporting contradictory results. DNA repair mediated by retrotransposition independent of its endonuclease activity has been reported in Chinese Hamster Ovary (CHO) cells deficient in NHEJ pathway (Morrish et al. 2002). Consistent with this, two independent studies brought arguments for ATM (Ataxia-Telangiectasia Mutated), a key DNA repair protein in NHEJ, helping the successful integration of L1 (Gasior et al. 2006; Wallace et al. 2013). ATM is a serine/threonine protein kinase activated by dsDNA breaks (Lempiäinen and Halazonetis 2009; Chiruvella, Liang, and Wilson 2013).

ATM KO cells inhibited L1 retrotransposition leading the two possible roles for L1 dependent of ATM: ATM is part of NHEJ repair pathway and might be involved in the microhomology process during L1 DNA synthesis (Gasior et al. 2006; Zingler et al. 2005) or ATM might play a role in phosphorylation of ORF1p and ORF2p, modification needed for DNA cleavage or pre/post L1 insertion (Gasior et al. 2006).

In another study, using systems of knockout for different components of NHEJ (Ku70, Artemis, DNA ligase IV)) in chicken cells, Suzuki and colleagues studied the effect on retrotransposition of two LINEs: human L1 and Zf12-2 zebrafish element. The results show an opposition between DNA repair and retrotransposition. Zf12-2 retrotransposition is decreasing in a deficient NHEJ system, suggesting an involvement of the canonical NHEJ proteins. However, the characterization of Zf12-2 insertions showed that the absence of NHEJ factors allowed more Zf12-2 full-length insertions to occur compared with the restricted full-length insertions in a functional NHEJ system. An explanation could be offered by the prior processing of the 5' overhang (formed by second-strand cleavage) by Artemis protein, followed by ligation to the 5' end of Zf12-2. Truncated

insertions have less potential to undergo subsequent amplification, whereas longer insertions are generated by more active retrotransposon (Farley, Luning Prak, and Kazazian 2004).

In NHEJ deficient cells, Zf12-2 insertions were found to truncate the target site by generation of long chromosomal DNA deletions. Knowing that Ku proteins defend from exonucleolytic degradation, they might assure that no genomic information is lost during Zf12-2 retrotransposition (J. Suzuki et al. 2009).

Also, when NHEJ disrupted, even if Zf12-2 retrotransposition is decreased, is not completely abolished, which might imply the role of an alternative NHEJ pathway. This alternative could be used to connect the end of the target genomic DNA and the end of the retrotransposon at the 5' junction, via microhomology, independent of NHEJ canonical factors (J. Suzuki et al. 2009; Zingler et al. 2005).

Regarding Zf12-2 element, the hypothesis of the study published by Suzuki and colleagues, is that NHEJ proteins might be recruited to the retrotransposon integration site for the repair of the endonuclease induced DSB, but they are used for the integration of Zf12-2 and to protect the genomic target site when retrotransposons integrate.

The deficiencies of NHEJ pathway had an effect on human L1 also, by decreasing retrotransposition. However, the results show different degrees of retrotransposition, which was much more decreased for Zf12-2 than for human. This suggests that each LINE has distinct variations in their mechanism of retrotransposition, depending more or less of NHEJ or other repair pathways, host specific L1 (J. Suzuki et al. 2009).

a) Ataxia telangiectasia modulates L1 retrotransposition in the brain

Mutations of ATM cause ataxia telangiectasia, a progressive neurodegenerative disease associated with a compromised immune system, infertility and predisposition to cancer (Shiloh 2001; Barlow et al. 1996).

ATM showed a supportive function for retrotransposition when knockout of ATM or reduced ATM protein levels by use of human papillomavirus E6 oncoprotein, reduced retrotransposition levels (Gasior et al. 2006; Wallace et al. 2013). Also, NHEJ deficient chicken cells show decreased retrotransposition of both human L1 and Zf12-2 element (J. Suzuki et al. 2009) .

Contradictory with these studies, Coufal and colleagues argued that ATM-deficient cells and the brains of Atm KO mice showed consistent two to four fold increased activity

of an engineered human L1, containing enhanced green fluorescent protein EGFP as reporter.

ATM deficiency effect on L1 retrotransposition was tested in hESC, neuronal precursor cells, cancer cells (HCT116) and showed that L1 insertions might be longer in ATM-deficient than in ATM-proficient cells, consistent with the study of Suzuki in chicken cells.

Moreover post-mortem brain analysis of patients with ataxia telangiectasia showed increased L1 copies in AT neurons compared with controls detected by PCR (Coufal et al. 2011). However, a recognized caveat of the study is the lack of estimation of changes in L1 copy number on other somatic tissues from AT patients.

It was highlighted that the different experimental approaches in the study of Gasior and Coufal might explain the contrast between their results. Gasior used fibroblast cells and G418 antibiotic for L1 cellular retrotransposition assay and observed decreased resistant-foci formation. The low levels of retrotransposition in fibroblast might be also a result of the sensitivity of these cells to experimental manipulation, normal decreased retrotransposition in this cell type, as well as toxicity caused by G418 antibiotic (Kubo et al. 2006; Goodier 2016; Coufal et al. 2011; Gasior et al. 2006; Thomas, Paquola, and Muotri 2012). Noteworthy, GFP can also cause cell toxicity and affect the results of the experiments (H. S. Liu et al. 1999; Goodier 2016).

b) ERCC1/XPF endonuclease interferes with the integration of the newly formed cDNA

Another DNA repair pathway found to limit L1 retrotransposition is the ERCC1/XPF heterodimer. XPF is known also as ERCC4, and the ERCC1/XPF complex is involved in DNA repair, processing 5' DNA of genomic lesions (Gillet and Schärer 2006; Sijbers et al. 1996).

As already mentioned, L1 TPRT involves one nick in the target DNA to create a primer for reverse transcription initiation and DNA synthesis occurs 5' to 3' leading to a 3' flap extension. This could be a substrate recognized by the ERCC1/XPF activity and cleaved before retrotransposition is completed. Hence, ERCC1/XPF could contribute to limit L1 retrotransposition frequency by excising newly formed cDNA from the genome before the final resolution of the insertion.

Collectively, the NER pathway play an important role as guardian of mammalian cells and might sense retrotransposition intermediates as dangers potentially affecting the fitness of the genome (Gasior, Roy-Engel, and Deininger 2008).

2. Introduction to the world of nuclear receptors

Contents

2.1	Nuclear receptor nomenclature: six subfamilies based on sequence homology	48
2.2	Nuclear receptors expression.....	52
2.3	The common domain structure of nuclear receptors.....	54
2.3.1	The N-terminal domain, the least conserved domain of NRs, is transcriptionally active without a ligand	55
2.3.2	The DNA Binding Domain recognizes specific genomic template and has a highly organized structure	56
2.3.3	The hinge region links the DBD and the LBD and assures the synergy between them 61	
2.3.4	The structure and function of the Ligand Binding Domain	61
2.4	Steroid nuclear receptors.....	64
2.4.1	The signaling pathway of steroid receptors	65
2.4.2	Estrogen related receptor (ERR) family.....	68
2.5	Functional redundancy and cross-talk between nuclear receptors	75
2.5.1	Functional redundancy of nuclear receptors	75
2.5.2	Cross-talk between nuclear receptors	79

The research project presented in this thesis is based on an original discovery made in the laboratory, which identified by yeast two hybrid, an interaction between the ORF2 protein of L1 retrotransposon and ERR α (estrogen related receptor alpha), a member of the nuclear receptor family. In this chapter I will present a detailed characterization of these particular cellular factors, to better understand L1 interaction partners.

Highlights:

- Nuclear receptors (NRs) are transcription factors that can regulate gene transcription in response to a wide range of internal or external stimuli.
- Structurally, NRs comprise an N-terminal (NTD) domain, a DNA-binding domain (DBD), a hinge, and a ligand-binding domain (LBD).
- NRs activity depends on the binding of a specific ligand (eg. hormone) to the LBD for ligand-dependent NRs, or on a constitutively active conformation.
- NRs require coregulators to modulate gene transcription.
- NRs are involved in a myriad of diseases (metabolic diseases, cancer, etc.), thus making them therapeutic targets.

Nuclear receptors are transcription factors which fulfill important functions in development, growth, metabolism and homeostasis (Evans and Mangelsdorf 2014). They accomplish this through chromatin modification or by direct interaction with the transcription machinery, in response to a wide range of internal and external signals such as nutrition and stress (McKenna, Lanz, and O'Malley 1999). They work together with coregulators (coactivators and corepressors) to stimulate or inhibit the transcription complexes at target genes. The human genome codes for 48 nuclear receptors (NRs) (see Figure 8).

Nuclear receptors are mainly known as ligand-regulated transcription factors which are activated by steroid hormones, such as progesterone, estrogen, testosterone, cortisol or other small molecules, as retinoic acid, thyroid hormone, oxysterols, vitamin D. The binding of the hormone to its receptor generally results in a complex which can bind to specific DNA sites and regulate transcription (Mangelsdorf, Evans 1995). Many NRs are called orphan nuclear receptors (ONRs). This term regroups two distinct situations:

- (i) receptors that need a ligand, but the latter has not been identified so far;
- (ii) receptors that do not need a ligand, and constitutively adopt an active conformation due to the presence of a hydrophobic amino acid from the receptor within the ligand binding pocket. The activity of these receptors is generally regulated by classical post-translational modifications (phosphorylation, sumoylation, etc.).

2.1 Nuclear receptor nomenclature: six subfamilies based on sequence homology

The nuclear receptor superfamily is organized in six subfamilies. Nuclear receptors are characterized by a conserved DNA-binding domain (DBD) which recognizes and binds to specific DNA sequences and by a ligand binding domain (LBD) essential for hormone binding and specificity of the physiologic response (Mangelsdorf, Evans 1995).

The nomenclature presented here is based on the sequence alignment of the conserved LBD and DBD domains (Gallastegui et al. 2015). Of note, the groups within a subfamily consist of related genes with paralogous relationship in vertebrates (e.g. RARA, RARB, RARG).

Subfamily 1 (NR1) Thyroid hormone receptor-like

This subfamily consists of the following transcription factors: the thyroid receptor (TR), the retinoic acid receptor (RAR), the peroxisome proliferator-activated receptor (PPAR), the liver X receptor (LXR), the farnesoid receptor (FXR), the vitamin D receptor (VDR), the RAR-related orphan receptor (ROR), REV-ERB, the pregnane X receptor (PXR), and the constitutive androstane receptor (CAR). ROR and REV-ERB are still orphan.

Subfamily 2 (NR2) Retinoid X receptor-like

Members of this family can function both as homodimers and heterodimers. The family comprises retinoid X receptor (RXR), *Hepatocyte Nuclear Factor 4* (HNF4), *tailless* homolog (TLX), photoreceptor specific (PNR), testicular receptor 2 and 4 (TR2 and TR4), and chicken ovoalbumin upstream promoter (COUP-TF) receptors. With the exception of RXR, all the receptors of this family are orphan.

Subfamily 3 (NR3) Estrogen receptor-like

This subfamily comprises the steroid hormone receptors. These NRs are presented in more details in the next section, as they represent the group of interest in the present study.

The ER-like subfamily comprises the estrogen receptors (ER α /NR3A1 and ER β /NR3A2), the glucocorticoid receptor (GR/NR3C1), the mineralocorticoid receptor (MR/NR3C2), the progesterone receptor (PR/NR3C3) and the androgen receptor (AR/NR3C4). The corresponding steroid ligand and the physiological relevance of all these receptors are well known and characterized.

This subfamily includes also the estrogen-related receptors (ERR) with their three paralogs ERR α /NR3B1, ERR β /NR3B2, and ERR γ /NR3B3. The ERRs are often described as orphan receptors since they do not bind estrogen or any known natural ligand, but they are actually constitutively active and not regulated by direct ligand binding as other steroid receptors.

Subfamily 4 (NR4) Nerve growth factor-induced clone B (IB) like

The NRs of this subfamily includes the nerve growth factor-induced clone B group of receptors: Nurr77 (NGFI-B/NR4A1), Nurr1 (NR4A2) and Nor1 (NR4A3). Structural studies indicates that Nurr77 and Nurr1 have a constitutive active conformation, due to the presence of hydrophobic amino acids in their ligand binding pockets (Zhan et al. 2012; Zhulun Wang et al. 2003).

Subfamily 5 (NR5) Steroidogenic factor-like

This is a small subfamily which comprises two constitutively active NRs: steroidogenic factor-1 (SF1/NR5A1) and the liver receptor homolog-1 (LRH1/NR5A2).

Subfamily 6 (NR6)

This subfamily contains only the orphan Germ Cell Nuclear Factor variant 1 (GCNF1) receptor. Neither its structure, nor its regulation have been elucidated.

Subfamily 0. NRs without DBD

Dax1, NR0B1, known also as DSS-AHC critical region on the X chromosome, protein 1 (dosage- sensitive sex reversal, adrenal hypoplasia critical region, on chromosome X, protein 1) (NR0B1) and SHP (small heterodimer partner), NR0B2, were first catalogued as co-regulators of other NRs because they contain in their N terminal domain, a conserved motif used for interaction with other receptors (LxxLL) (Ehrlund and Treuter 2012; Gallastegui et al. 2015). But they are actually, two atypical NRs lacking DBD.

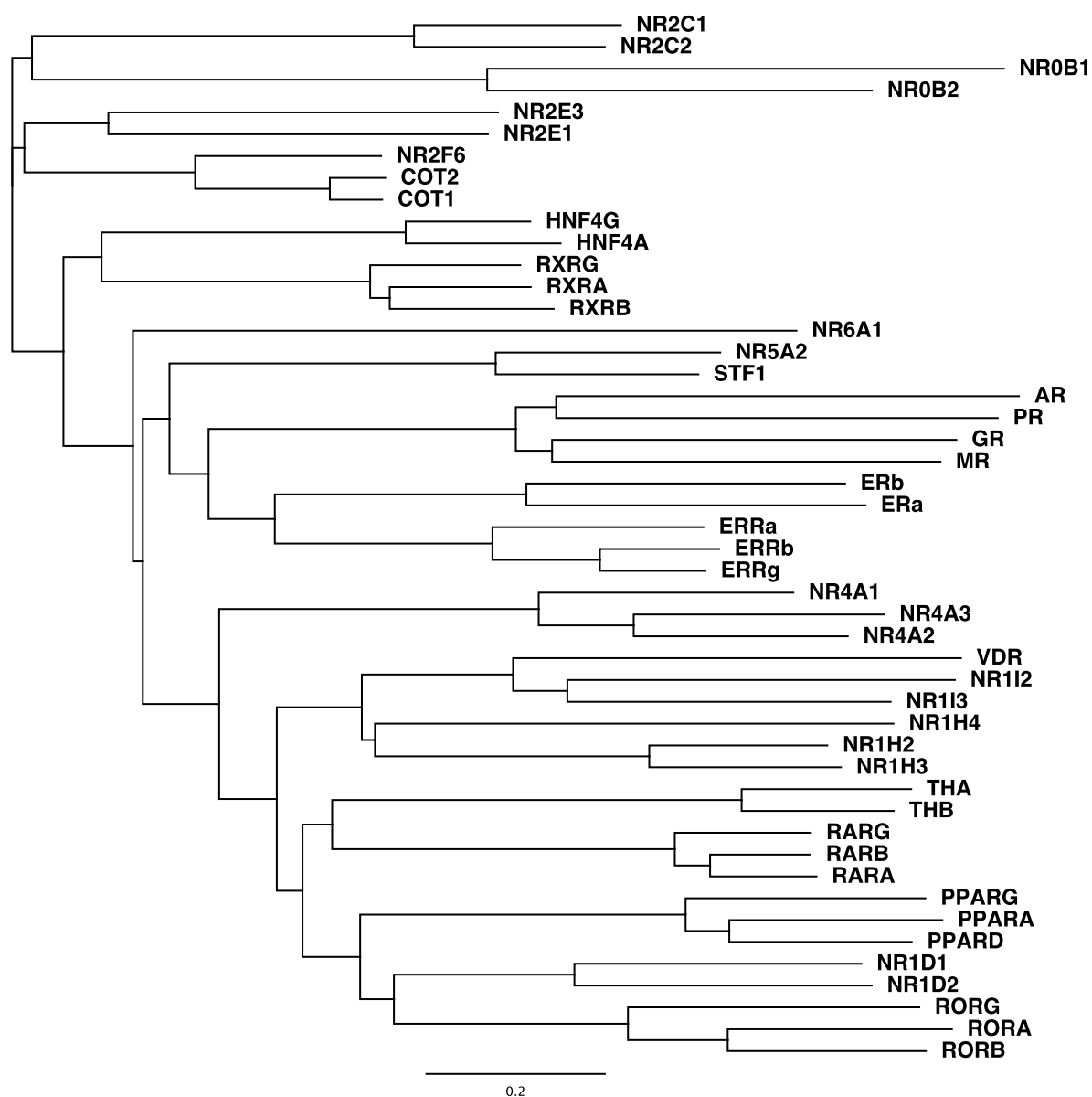


Figure 8. The phylogenetic tree of the 48 human nuclear receptors superfamily.

Distance tree was created by the neighbor-joining method using the Geneious software. The sequence of each NR was downloaded from www.uniprot.org.

2.2 Nuclear receptors expression

NRs modulate a wide variety of physiological pathways, as development, proliferation, homeostasis, reproduction, and metabolism (Bookout et al. 2006). Their expression in a normal or disease state, governs complex transcriptional programs. The expression and activity of NRs was and still is of high importance, given their extensive roles.

The NRs are broadly expressed in most cell types (see Figure 9) (Bookout et al. 2006). However, the most completed spatial and temporal expression profiles of nuclear receptors was done for the murine NR superfamily, which consists of 49 receptors (Bookout et al. 2006; Xiaoyong Yang et al. 2006).

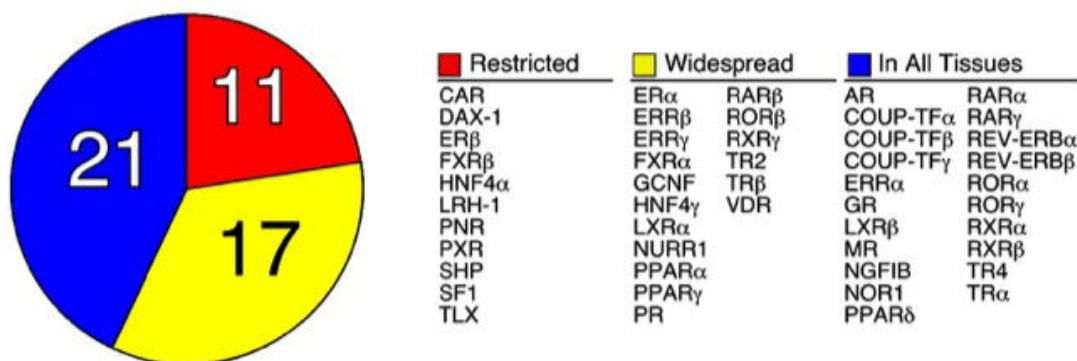


Figure 9. A summary of tissue-frequency profiles of the nuclear receptor superfamily in mice.

TaqMan-based real-time quantitative RT-PCR (qPCR) was used to measure transcript levels of 49 known mouse NRs, in 39 different tissues isolated from two strains of mice widely used in genetic manipulation. The results showed that 21 NRs were expressed in all tissues, 17 were found in more than half of the tissues analyzed and 11 were restricted to only some tissues (Bookout et al. 2006).

From the same study, it was compiled an atlas of the NR expression in every tissue system (Figure 10). The results which show the diversity of NR expression profile within an organism, suggest that while each individual receptor plays an important role in physiologic processes, the NRs operate also together, in specific groups, as part of a high-order regulatory network (Bookout et al. 2006).

However, the possible discovery of wider organizational rules could indicate us how the nuclear receptors superfamily evolved to form the physiology of the body and this would serve to better understand diseases and to better adapt the pharmacology (Evans 2005).

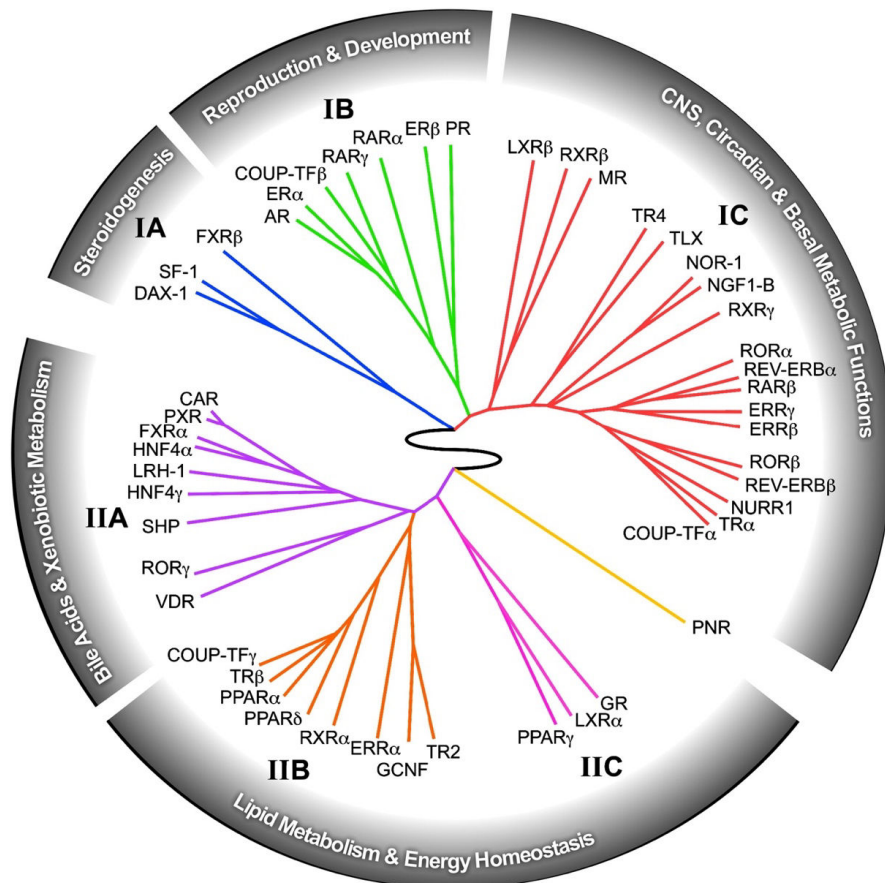


Figure 10. Nuclear receptors and the connection between their expression profile, function and physiological pathways
(Bookout et al. 2006).

2.3 The common domain structure of nuclear receptors

Nuclear receptors have a common basic structure, consisting of an N-terminal (NTD) domain, a DNA-binding domain (DBD), a hinge region, a ligand-binding domain (LBD) and a variable F-domain (Brélivet, Rochel, and Moras 2012).

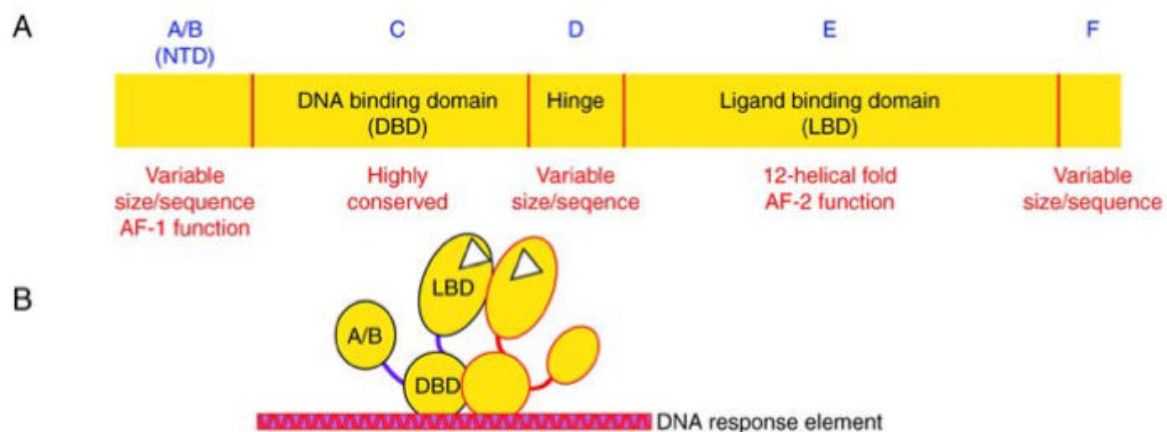


Figure 11. Domain arrangement of nuclear receptors.

A. Letters from A to F represent the nuclear receptor domains from N- to C-terminus. Details relative to the structure and function of each domain are explained in the main text. Nuclear receptors consists of a N-terminal domain (NTD), poorly conserved and of variable size which contains the ligand-independent transcription activation function-1 (AF-1); a highly conserved DNA Binding Domain (DBD) which recognizes and binds the NRs to specific DNA response elements, which can be direct repeats, inverted repeats or single half sites of the hexamers 5'-AGAACA-3' or 5'-AGGTCA-3; a variable and flexible hinge region; a Ligand Binding Domain (LBD) made by 12 helices responsible for ligand binding, recruitment of coregulators and transcriptional regulation through the activation function-2 (AF-2). In addition, some NRs, such as ERs, contain a variable F-domain positioned at their C-terminus and with a poorly understood role. B. Schematic representation of a dimeric NR, bound to DNA response element, after activation by a ligand (white triangle) and conformational changes (Rastinejad et al. 2013).

2.3.1 The N-terminal domain, the least conserved domain of NRs, is transcriptionally active without a ligand

The N-terminal domain (NTD) is the least conserved domain of NRs, showing variable length and sequence (Kumar and McEwan 2012). However, sequence analysis revealed that some residues in the NTD are conserved across species, especially the ones necessary for receptor-dependent gene regulation (Kumar and McEwan 2012; Lavery and McEwan 2005). Due to its high flexibility and intrinsic disorder, this region is difficult to characterize biochemically and its tertiary structure has not yet been obtained (Pawlak, Lefebvre, and Staels 2012).

Within the NTD, the region named activator function 1 (AF-1) is important for both transcriptional activation and repression. Moreover, AF-1 is ligand independent, as shown by studies on steroid receptors in which the deletion of the LBD results in a constitutively active receptor in reporter gene assays (Lavery and McEwan 2005). Also, the binding of the ligand to PPAR γ LBD, does not affect the positioning of the NTD in the 3D structure (Chandra et al. 2008).

In the case of AR, the deletion of entire NTD results in a transcriptionally weak receptor, highlighting the importance of the NTD in the transactivation function (Jenster et al. 1995; Simental et al. 1991). Similarly, deletions of different regions within the NTD of PR and MR blunts receptor-dependent transactivation (Takimoto et al. 2003; Govindan and Warriar 1998; Fischer et al. 2010).

2.3.1.1 The NTD binds coregulators and is involved in inter-domain communication

The NTD region is also important for binding to coregulatory proteins such as chromatin modifiers, transcription factors, coactivators and corepressors. For instance, steroid receptor coactivator-1 (SRC-1/NCoA1), steroid receptor coactivator-2 (SRC-2/TIF2/NCoA2) and p300, allow the collaboration between the AF-1 and AF-2 regions of PPAR γ and RARs, which results in enhanced transactivation of gene expression (Bommer et al. 2002).

Moreover, the NTD is involved in inter-domain communication, regulating the interaction of co-repressors with the LBD in the absence of a ligand (S. Suzuki et al.

2010). Similarly, the AF-1 domain of PR necessitates the DBD to be functional, highlighting the importance of inter-domain communication (Takimoto et al. 2003).

2.3.1.2 NTD is modified by post-translational modifications

The NTD domain is prone to post-translational modifications such as phosphorylation or sumoylation, which modify the transactivation potential of NRs. For instance, phosphorylation of the AF-1 domain of GR by p38 MAPK, promotes a stable and active conformation of the intrinsically disordered NTD, favoring the interaction between the receptor and its coregulatory proteins which are necessary for activation or repression of gene expression (Garza, Khan, and Kumar 2010). SUMO-1ylation (sumoylation by SUMO1-small ubiquitin-like modifier 1 protein) of the AF-1 domain in the androgen receptor inhibits androgen signaling (Poukka et al. 2000). In contrast, a SUMO-1ylated NTD in PPAR γ , stimulates its transactivation capacity (Ohshima, Koga, and Shimotohno 2004).

2.3.2 The DNA Binding Domain recognizes specific genomic template and has a highly organized structure

The DNA binding domain (DBD) is the most conserved domain which recognizes and binds to specific DNA response elements (REs) (Rastinejad et al. 1995).

The DBD has a compact, globular structure and consists of a core region of approximately 66 amino acids, a pair of perpendicular α -helices stabilized by two C4 zinc-binding domains, a short β -sheet, and a few stretches of amino acids (see Figure 12.A) (Hard et al. 1990; Claessens and Gewirth 2004).

The first zinc-finger module contains a proximal box (P-box) which contains three amino acids necessary for DNA sequence specificity. The second zinc domain contains a five residues sequence, named D-box, involved in dimerization of NRs on DNA (Claessens and Gewirth 2004; Smit-McBride 1994).

The precise nucleotide sequence of a given RE defines the specificity of binding of the receptor to its cognate genomic sites, but single nucleotide differences can also determine cofactor specificity and thus the mode of transcriptional regulation (Meijsing et al. 2009).

Homodimers, heterodimers or monomers of nuclear receptors bind the DNA response elements which are arranged as direct repeats (DRs), palindromic (inverted) repeats or single half sites of the hexamers 5'-AGAACA-3' or 5'-AGGTCA-3 (see Figure 12.B). Importantly, the orientation and distance between the repeats is essential for receptor specificity. After binding to REs, NRs recruit other complexes that lead to repression or activation of their target genes (Rastinejad et al. 2013).

2.3.2.1 The basis for DNA recognition

The study of ER and GR homodimers bound to DNA by X-ray crystallography revealed that one DBD α -helix inserts directly into the major groove of the conserved hexamers present in the DNA response elements (Luisi et al. 1991; Schwabe et al. 1993). There are two receptor groups which use one of the two consensus DNA half sites: 5'-AGAACA-3' half-element for AR, GR, MR, PR and 5'-AGGTCA-3' to which ER and most receptors bind. Importantly, the two groups of NRs also use different sets of amino acids on the exposed face of their DNA recognition α -helices. Exactly three residues of the P-box differentiate NR that will bind to one of the two DNA half sites (Rastinejad et al. 1995; Gronemeyer and Moras 1995; Claessens and Gewirth 2004; Rastinejad et al. 2013).

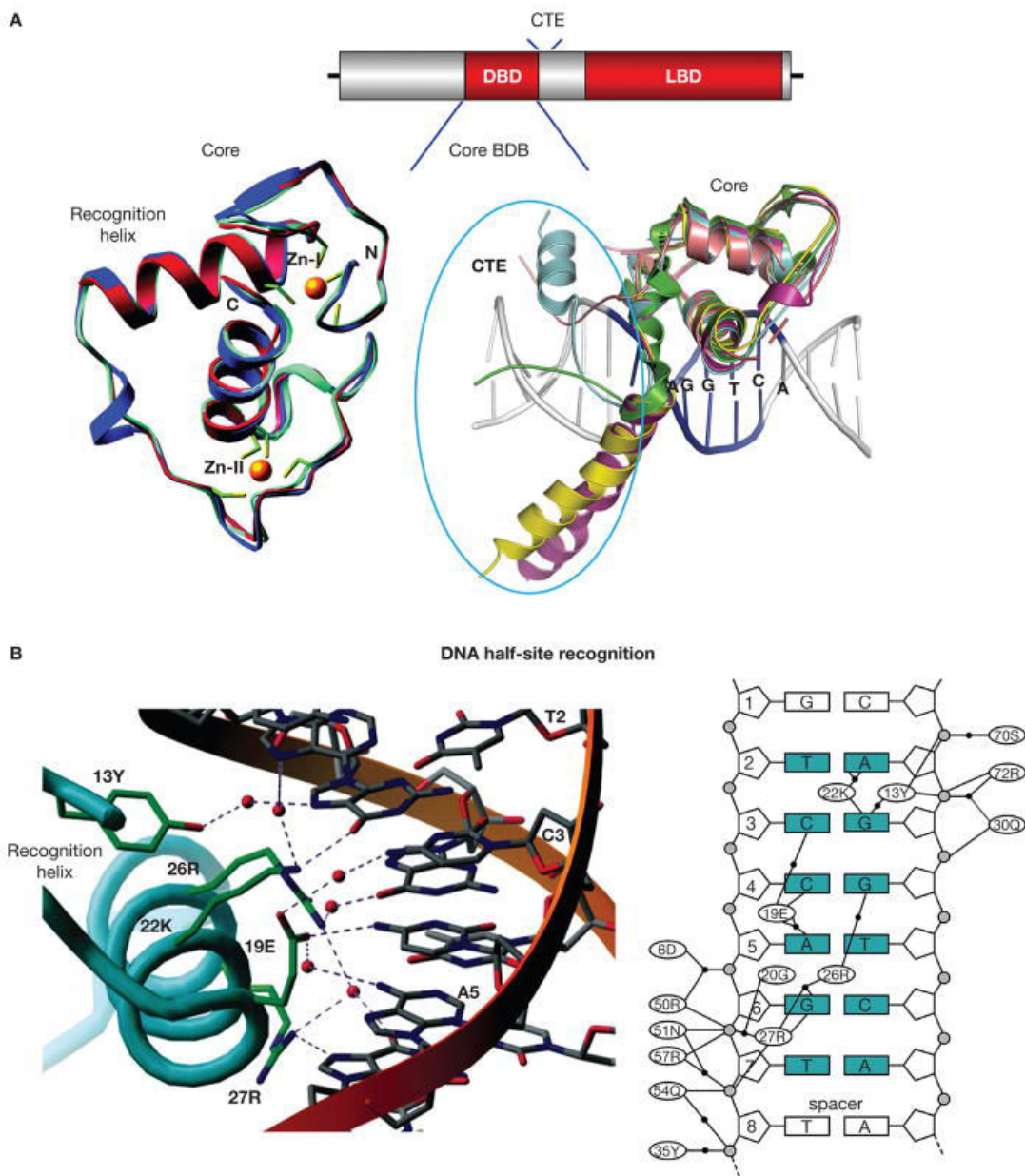


Figure 12. Nuclear receptor DNA recognition.

A) The domain responsible for hexameric DNA half-sites is the core DBD which consists from a 66 residue region and two zinc binding modules. On the left are shown superimposed, the core DBD structures of RXR (in blue) and RAR (in red). The C-terminal extension (CTE) of the DBD, can also have a role in DNA binding and spacer recognition. To highlight the divergence of CTE sequences compared with the conserved DBDs, on the right is the superposition of the DBD-CTE segments of several nuclear receptors: NGFI-B (green), LRH-1 (cyan), VDR (yellow), TR (magenta) and Rev-Erb (salmon). Notably, the CTEs can act as discriminators of DNA spacing.

B) DNA half-site recognition by the DBD of RAR. The interaction between the DBD and the DNA hexamer implies several hydrophilic residues positioned on the same face of the DBD recognition helix. These residues read the DNA base-pair sequence at the major groove. The DNA binding is stabilized by additional basic amino-acids which interact with the phosphate backbone of the DNA. Also, the DNA contacts are mediated often by water molecules (shown as red or black circles), as shown by crystallographic studies (Rastinejad et al. 2013).

However, the recognition of half-site binding is not enough for response element selectivity of NR, given the high degree of sequence conservation in the DBD and that there are only two main types of consensus DNA binding sites. The geometry of the RE, when the two hexamers are arranged in various ways is essential for DNA target specificity. Each spacer between the repeats shifts with some degrees the half sites, which allows only specific NR pair to efficiently interact (Umesono et al. 1991; Mader et al. 1993; Rastinejad et al. 2013).

However, it still remains elusive how the selective binding-site occurs for steroid receptor homodimers, as most of them use the same symmetry, spacing, and consensus half-site sequences (Luisi et al. 1991; Roemer et al. 2006).

The crystal structure of PR DBD bound to DNA suggests that other unique sites flanking the 3' or 5' half sites of these symmetric repeats might influence the selectivity (Roemer et al. 2006; Rastinejad et al. 2013; Nelson et al. 1999). In addition, other geometries of the hexamers or collaboration with other transcription factors might grant the response element selectivity (Rastinejad et al. 2013).

In the figure below (Figure 13), it is represented how the retinoid X receptor (RXR) forms heterodimers with several non-steroid receptors, especially with vitamin D₃ receptor (VDR), thyroid hormone receptor (TR) or PPAR (Rastinejad et al. 2013). The RXR heterodimers bind to response elements made by direct repeats (DRs) separated by distinct half-site spacers. DR elements are named DR1 to DR5. The selectivity is based on the spacing of DRs, a property known as the 1-5 rule (Rastinejad et al. 2013; Umesono et al. 1991). For example, RXR can form DBD–DBD interaction with TR on DR4, but a change in the spacing size would suppress the interaction, which would oblige RXR to heterodimerize with a different partner (Rastinejad et al. 2013). Also, RXR can alternate its half-site position relative to the heterodimerization partner and this can also influence the different responses to ligands and corepressors (Kurokawa et al. 1994; Rastinejad et al. 2013).

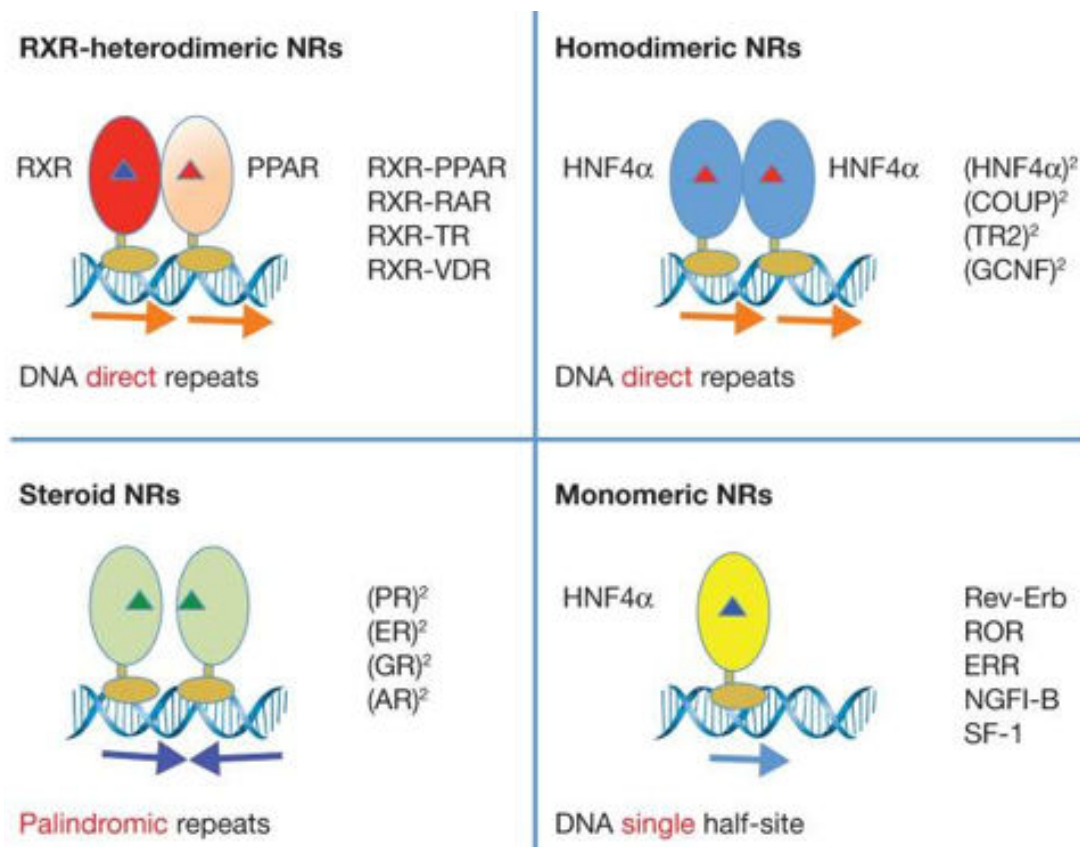


Figure 13. Variety of nuclear receptors dimerization and binding to DNA response elements.

NRs can adopt diverse dimerization states and can bind to a variety of DNA response elements consisting of direct repeat (DR) elements, palindromic (inverted) repeats, or extended monomeric sites. Some non-steroid receptors which form heterodimers with the common partner retinoid X receptor (RXR), as well as many homodimers, bind to direct repeat response elements with different hexamer (half sites) spacers. Steroid receptors bind as homodimers to palindromic binding sites, where the half sites are in an inverted repeat arrangement. Some NRs do not have any partner and bind to DNA as monomers (Rastinejad et al. 2013).

2.3.2.2 Other functions of DBD

NRs do not reside long on their DNA response elements. Indeed, studies showed that the glucocorticoid receptor (GR) binds to its glucocorticoid response elements no longer than 10 seconds (McNally et al. 2000; Voss and Hager 2013).

The DBD of NRs carry out other functions, as nuclear export and interactions with other TFs and chaperones. For instance, it was proposed that the export of NRs from the nucleus is mediated by a necessary and sufficient 15 residue sequence, localized between the two zinc finger domains of the DBD (Black et al. 2001).

2.3.3 The hinge region links the DBD and the LBD and assures the synergy between them

The hinge region is a poorly conserved polypeptide linker between DBD and LBD (see Figure 11.A). However, the post-translational modifications of this region are critical for translocation to the nucleus, DNA binding and ultimately gene transactivation (Rastinejad et al. 2013; Gallastegui et al. 2015).

Mutagenesis studies made on the hinge domain, affected the synergy between the AF-1 and AF-2 domains of ER, highlighting its role as link between the two functional domains of NRs, DBD and LBD (Zwart et al. 2010). In addition, a mutated hinge sequence of farnesoid X receptor (FXR) determines distinct DNA binding affinities of the receptor (Y. Zhang, Kast-Woelbern, and Edwards 2003).

Post-translational modifications of the hinge region can impact gene transactivation. For example, acetylation of the hinge of ER α by p300 can control ligand sensitivity and gene transactivation (Chenguang Wang et al. 2001). SUMOylation of the hinge of PPAR α supports the recruitment of co-repressors and subsequent decreased transcription of the target genes (Pourcet et al. 2010).

2.3.4 The structure and function of the Ligand Binding Domain

As the name states, the ligand binding domain (LBD) of a nuclear receptor is responsible for the accommodation of the hormones, vitamins or other small molecules acting as ligand, but also for the recruitment of transcriptional coregulators and for the transcriptional activity of NRs (Brélivet, Rochel, and Moras 2012).

Ligand dependent NR in particular, are in the center of attraction for pharmacology, because their association with several diseases as cancer or metabolic dysfunctions, makes them biomedical targets (Gallastegui et al. 2015).

2.3.4.1 The global fold of the LBD determines its structure and function

The ligand binding domain of a nuclear receptor is made of 12 α -helices and a variable number of β -turns grouped in a three layer helical sandwich with a cryptic ligand binding pocket (LBP) within the middle layer (see Figure 14) (Rastinejad et al. 2013; Gallastegui et al. 2015; Bourguet et al. 1995; Renaud et al. 1995). In the absence of the ligand, the

LBD has an open structure (Bourguet 1995), while it adopts a compact structure upon ligand binding, which will be buried in the hydrophobic interior of the protein (Tanenbaum et al. 1998).

The helix 12 (H12 or AF2 helix) lies across the LBP and is in direct contact with the ligand. The orientation of H12 relative to the rest of the LBD will determine changes in the volume of the ligand binding pocket, from almost nonexistent to large cavities (see Figure 14). This will mark an active conformation or a repressed one (Ingraham and Redinbo 2005; Gallastegui et al. 2015). But, the LBP conformation is influenced by the global fold of the LBD, due to the presence or absence of specific helices and β -turns, to helix repositioning, and to the influence of a ligand, if present (Gallastegui et al. 2015).

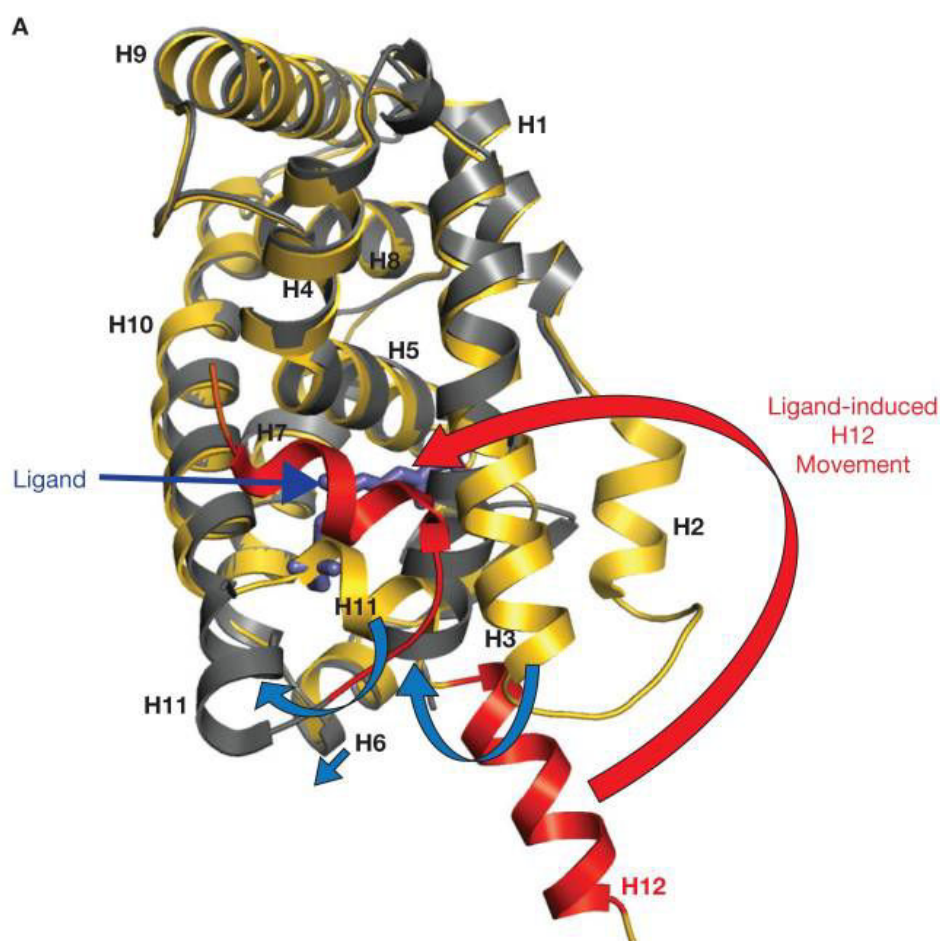


Figure 14. Structure of nuclear receptor ligand binding domain.

The ligands can induce changes in the LBD conformation, mostly in the positioning of helix-12. The movement of H-12 (shown in red) allows the ligand binding pocket to encircle the ligand (sometimes described as a trapping mechanism). Additionally, other smaller rearrangements can be induced by ligand binding (blue arrows) (Rastinejad et al. 2013).

2.3.4.2 NRs interact with coregulators through consensus motifs

For an active LBD, H12, is almost perpendicular to H3 and H5, creating a triangular settlement with access to the activation function 2 (AF-2), which is part of H12 (Darimont et al. 1998). The formed hydrophobic groove, stabilized by helices, permits the interaction between LBD and the signature motif present in NR coactivators, called the NR box (LxxLL, where L is leucine and X is any amino acid) (Darimont et al. 1998; Heery et al. 1997). Studies of isolated LBDs bound to coactivators showed that the LxxLL motif adopts a helical structure when bound to the LBD and makes contact along a hydrophobic binding groove created by helices 3,4,5 and 12. Thus, H12 shifts to an active conformation on ligand binding and is crucial for coactivator binding (Nagy and Schwabe 2004).

If no ligand binds, helix H12 can pose in different repressive conformations, such as folding against the LBD to cover H3 and H5, blocking the coactivator assembly, or overhanging in solution, letting AF2 incomplete (Ingraham and Redinbo 2005; Gallastegui et al. 2015). Notably, only the active LBD allows coactivator binding, while the repressed one favors corepressor binding (Gallastegui et al. 2015).

The recruitment of corepressors to nuclear receptors occurs through the consensus motif LxxH/IIxxxI/L (CoRNR box) (X. Hu and Lazar 1999; Perissi et al. 1999; Nagy et al. 1999). When the LBD binds a corepressor peptide, the CoRNR box adopts helical fold and its binding obstructs H12 from seizing an active conformation. Also, crystal structures revealed that coactivator and corepressor binding is mutually exclusive because both bind to the same surface of the LBD (Xu et al. 2002).

2.3.4.3 The flexibility of the LBD and co-regulator binding impact the structure of NRs

The binding of coregulators brings structural adaptations which influence the LBP cavity (Togashi et al. 2005). X-ray studies showed that NRs rearrange their LBP structure to bind different agonists, compared with unliganded LBDs. Flexibility is a key feature of the LBD and important for the selection of agonist and antagonists. For instance, ligands of the thyroid hormone receptor have been modified and extensions which should disturb H12, are recognized as antagonists, even if they are agonists (Togashi et al. 2005; Webb

et al. 2002). For example, replacing the 5' hydrogen with a bulky isopropyl group, transforms an TR agonist ligand to an antagonist (Togashi et al. 2005).

The above mentioned structural and mechanistic details represent the canonical image of how a ligand-dependent NR is functioning (Mangelsdorf et al. 1995). However, the LBP of orphan nuclear receptors (ONRs) can be different, and as stated before, nonexistent or with a large empty chamber (Mullican, DiSpirito, and Lazar 2013).

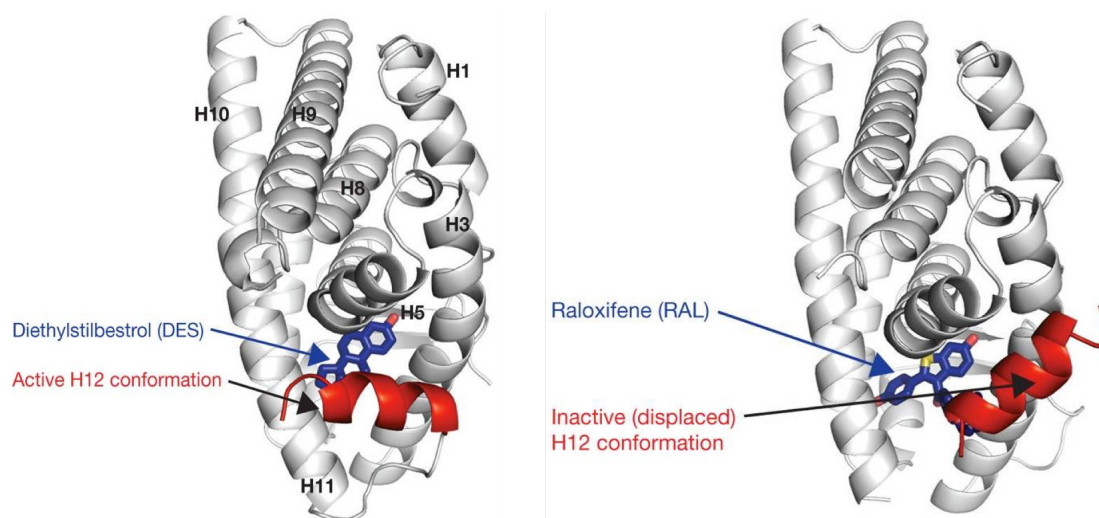


Figure 15. Schematic of structural changes induced on the estrogen receptor by agonists and antagonists.

At the left, the binding of the agonist synthetic molecule diethylstilbestrol (DES, shown in blue) places helix-12 (H12) in the agonist conformation. At the right, raloxifene (RAL), an antagonist, pushes helix-12 out of its active conformation (Rastinejad et al. 2013).

2.4 Steroid nuclear receptors

Steroid receptors are fundamental for the processes regulation of an organism and they are highly evolutionarily conserved (Keay and Thornton 2009). Structural domains responsible for ligand, DNA and co-regulator binding, receptor dimerization, and transactivation are conserved across the family. Hence, this subfamily of NRs responds to hormones based on a relatively conserved chemical structure, the steroid scaffold (Margolis and Christakos 2010).

Based on sequence homology and DNA binding specificity, the NR3 subfamily contains the classical steroid receptors, including the estrogen receptors α and β (ER α , ER β), the

progesterone receptor (PR), the mineralocorticoid receptor (MR), the glucocorticoid receptor (GR), the androgen receptor (AR), as well as a structurally related group of three orphan receptors: $ERR\alpha$, $ERR\beta$ and $ERR\gamma$ (see the phylogenetic tree in Figure 16) (Nuclear Receptors Nomenclature Committee 1999).

From all NRs, the steroid receptors have been the most intensively studied, due to their critical role in development and adult physiology, their dysfunction in many diseases.

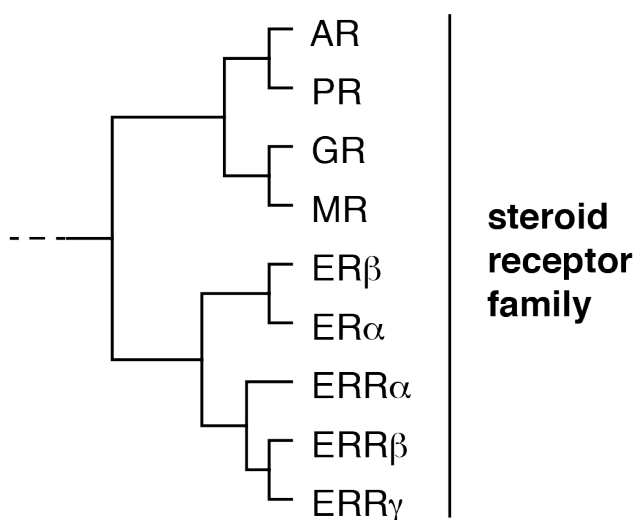


Figure 16. Phylogenetic tree of steroid nuclear receptors (subfamily 3, NR3).

Distance tree was created by the neighbor-joining method using the Geneious software. The sequence of each NR was downloaded from www.uniprot.org.

2.4.1 The signaling pathway of steroid receptors

Steroid receptors have the same general structure and mode of action, as other NRs. All the domains contribute, together with coregulators, to the binding of NRs to specific DNA sites in the nucleus and to the subsequent regulation of gene expression.

The canonical signaling pathway of steroid receptors is based on a simple model in which the ligands (endocrine steroids or exogenous molecules) enter into cells through the cell membrane through mechanisms which are not yet understood and bind to their cognate receptors, either in the cytoplasm or in the nucleus. Thus, the steroid receptors can be located in the cytoplasm (as the glucocorticoid and androgen receptor), bound to heat shock proteins (HSPs), or already present in the nucleus, even if a part of them can also

be found in the cytoplasm, bound to HSPs (estrogen receptor α and β). When the steroids bind their receptors, the cytoplasmic receptors are released by the HSPs, they dimerize, adopt diverse conformations and translocate into the nucleus. All the dimers located in the nucleus can bind to their specific DNA response elements and recruit coregulators to enhance or suppress gene transcription (Levin and Hammes 2016).

Besides this canonical mode of action, transcription can be regulated by steroid receptors through a tethered model, in which receptors do not bind directly to the DNA response elements, but interact with other TFs bound to DNA or coregulators (Barkhem, Nilsson, and Gustafsson 2004; Levin and Hammes 2016).

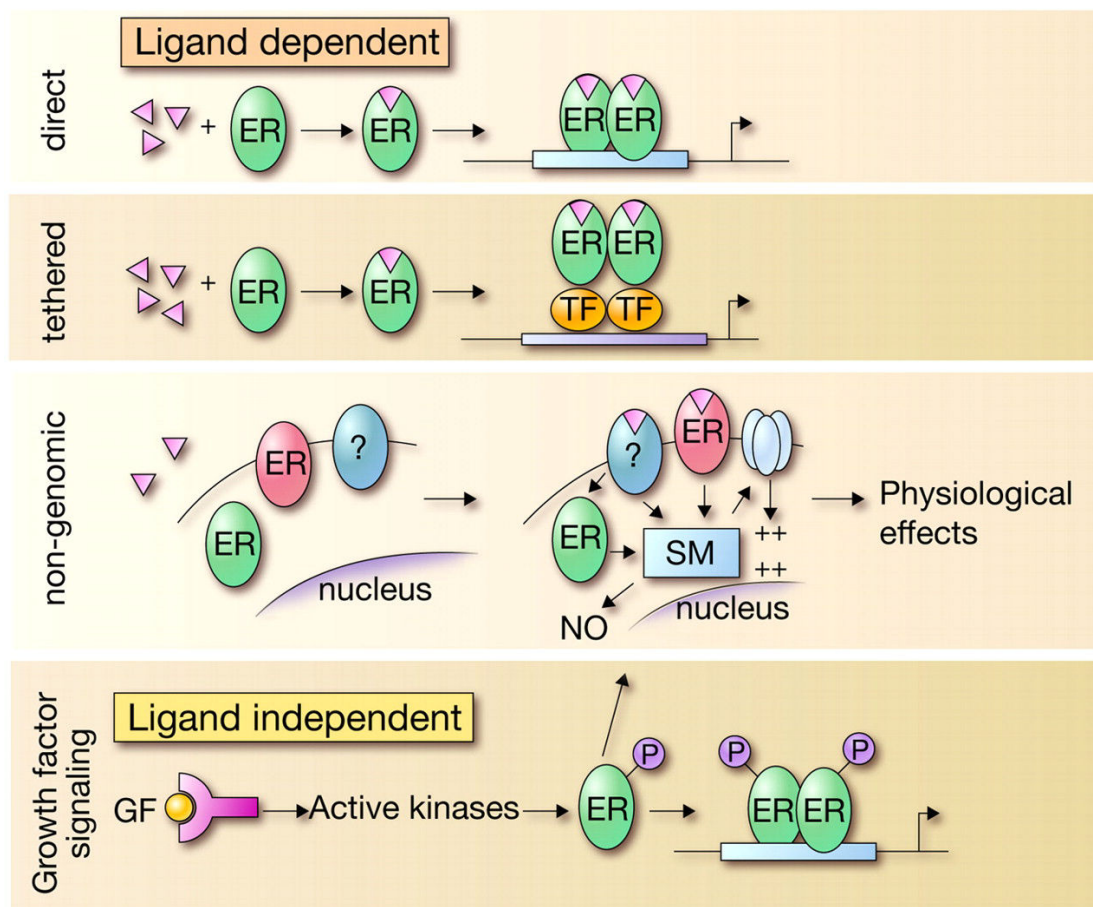


Figure 17. The distinct signaling pathways used in the regulatory actions of estrogen receptors.

For simplicity, it is shown only the estrogen steroid receptor (ER). In the classical (direct) pathway, the ligand binding to the receptor will trigger dimerization, binding to their specific DNA response elements and possible recruitment of coregulators (not shown) to modulate (enhance or suppress) gene transcription. In the tethered signaling model, steroid NRs can affect

transcription without direct DNA contact. In contrast, they interact with other transcription factors and thereby the gene regulation is regulated by indirect DNA binding. The third model is known as nongenomic with rapid effects and is not completely understood as the genomic mechanism, but has been noticed in many tissues. In this case the ligand activates a receptor, possibly associated with the membrane; either it is a classical ER (ER in green), an ER isoform (ER in red), or a distinct receptor (question symbol) or, alternatively, a signal activates a classical ER located in the cytoplasm. After this not very clear step, signaling cascades are initiated through second messengers (SM) that affect ion channels (++) or increase nitric oxide (NO) levels in the cytoplasm, and this leads to a rapid physiological response without gene modulation involvement. The ligand-independent pathway consists of activation through other signaling pathways, like growth factor signaling. In this model, activated kinases phosphorylate ERs and thereby activate them to dimerize, bind DNA, and regulate genes (Heldring et al. 2007).

2.4.1.1 Steroid signaling outside of the nucleus

It was observed that steroids also induce responses through rapid signaling which is not via transcriptional regulation (Roberts 1950; Levin and Hammes 2016). This can be explained by the binding of the steroid hormones at the receptors localized at the plasma membrane rather than in cytoplasm or nucleus. Extranuclear signaling by steroid receptors was observed first in plants. Even if there are no steroid receptors in the nucleus of plants, brassinosteroids localize to the cell membrane and bind tyrosine kinase receptor to signal and activate responses needed for plant flowering and fertility (Levin and Hammes 2016; Levin 2008; Belkhadir and Chory 2006). Then, work on glucocorticoid and estrogen rapid effect revealed a new aspect of steroid signaling (Roberts 1950).

Membrane steroid signaling might have evolved to the nuclear pathway as the genome became more complex in some organisms (Levin and Hammes 2016). Steroid receptors do not localize only in the nucleus, the cytoplasm or the plasma membrane, but also in the mitochondria, endoplasmic reticulum and other organelles (Hammes and Levin 2007). These extranuclear receptor pools work independently or together with the classic nuclear pools, in a cell-specific manner, to modulate cell functions (Levin and Hammes 2016).

2.4.2 Estrogen related receptor (ERR) group

As mentioned before, the nuclear receptors are a distinctive group of transcription factors which mediates the activity of steroid hormones, thyroid hormones, fat-soluble vitamins A and D, but the family includes also many orphan nuclear receptors (Evans and Mangelsdorf 2014).

A group of orphan receptors is represented by three estrogen related receptors - ERR α (NR3B1, *ESRR4* gene), ERR β (NR3B2, *ESRRB* gene) and ERR γ (NR3B3, *ESRRG* gene). They belong to the NR3 subfamily of steroid receptors which also includes ERs, PR, MR, GR, AR (Nuclear Receptors Nomenclature Committee 1999).

ERR α and ERR β were discovered using the DNA-binding domain of the human ER α cDNA as a hybridization probe in the search for gene products related to steroid hormone receptors (Giguere et al. 1988). ERR γ was discovered first during the investigation of a critical genomic locus for Usher's syndrome locus (Eudy et al. 1998) and subsequently by yeast two-hybrid screen (Hong, Yang, and Stallcup 1999).

ERRs have only 30-40% homology with the LBD of ER α and they do not bind or respond to estrogens or their derivatives, their transcriptional activity being constitutive (Giguere et al. 1988; Tremblay 2007).

The ERRs have the structure of a typical NR, which includes a NTD, a DNA-binding domain (DBD), a hinge region and a ligand-binding domain (LBD) (see Figure 18).

As in most NRs, the NTD contains the activation function domain AF-1 which provides a weak transcriptional activity, independent of the ligand. All three ERRs, have conserved motifs in their NTD, susceptible to post-translational modifications, such as phosphorylation and sumoylation, which can regulate transcriptional activity (Tremblay et al. 2008; Vu, Kraus, and Mertz 2007; Huss, Garbacz, and Xie 2015).

The ERRs paralogs show high sequence identity, especially ERR β and ERR γ (Figure 18) (Hong, Yang, and Stallcup 1999). The highly similar DBD of ERRs recognize a specific DNA sequence, named the ERR response element (ERRE), defined as 5'-TNAAGGTCA-3', a slightly extended form of the canonical half site 5'-AGGTCA-3' to

which most NRs bind (Sladek, Bader, and Giguere 1997). Thus, many genes are target of several ERRs.

ERRs can bind to their DNA motif as monomers, homodimers or heterodimers of two different ERRs (Huppunen, Wohlfahrt, and Aarnisalo 2004; Gearhart et al. 2003). The considerable similarity between the DBD of ERRs and ER α does not result in strong binding to ER response elements, even if they share common target genes (Giguère 2002).

Generally, ERRs occupy half sites of a multi-site module that mediates the response to estrogens (Huss, Garbacz, and Xie 2015), where ERRs and ERs can either cooperate or antagonize one another (Nengyu Yang et al. 1996; Huss, Garbacz, and Xie 2015).

For ERR α , the binding to ERREs and thus the control of ERR α transcriptional activity is also modulated by the acetylation of four lysine residues residing in the DBD (see Figure 18.A). The acetylation represses the function of ERR α by altering its DNA-binding activity and forms a dynamic acetylation/deacetylation switch for the transcriptional activity of ERR α (Huss, Garbacz, and Xie 2015; B. J. Wilson et al. 2010).

The C-terminal part of the LBD of the ERRs includes a well-conserved AF-2 domain, which is essential for co-regulator interactions (Figure 18.A). As stated before, ERRs are ligand-independent and constitutively active due to the conformation of their LBD, which allows co-regulator binding (S. Chen et al. 2001).

The resolved structure of ERR α and ERR γ shows that the LBP is occupied by side chains of amino acids which mimics the conformation adopted when the ligand is bound, to attract coactivator binding (Huss, Garbacz, and Xie 2015). For example, the crystal structure of ERR α , revealed that a part of the LBP is occupied by Phe328. As a consequence, the LBD can adopt an agonist-like conformation and can bind peroxisome proliferator-activated receptor γ (PPAR γ) coactivator-1 α (PGC-1 α) (Kallen et al. 2004).

ERRs activity depends on the presence of co-regulators proteins. Important coactivators are PGC1- α and PGC1- β , and members of the steroid receptor activator (SRC) family (Huss, Kopp, and Kelly 2002; Gaillard, Dwyer, and McDonnell 2007; W. Xie et al. 1999; Hong, Yang, and Stallcup 1999; Deblois and Giguère 2011).

Regarding the repression of ERRs transcriptional activity, a strong suppressor is the nuclear receptor interacting protein 140 (RIP140). The interaction between RIP140 and

ERR γ allows conformational constraints on the ERR γ complex, influencing its binding pattern and target gene recognition (Castet et al. 2006; Huss, Garbacz, and Xie 2015). A genetic controller of ERR α is the homeodomain-containing protein PROX1 inhibits the activity of the ERR α /PGC-1 α interacting with ERR α solely through its DBD (Charest-Marcotte et al. 2010). In sum, ERRs can function either as activators or repressors of gene expression, depending on the presence of and association with specific coregulators in a given cell type or tissue.

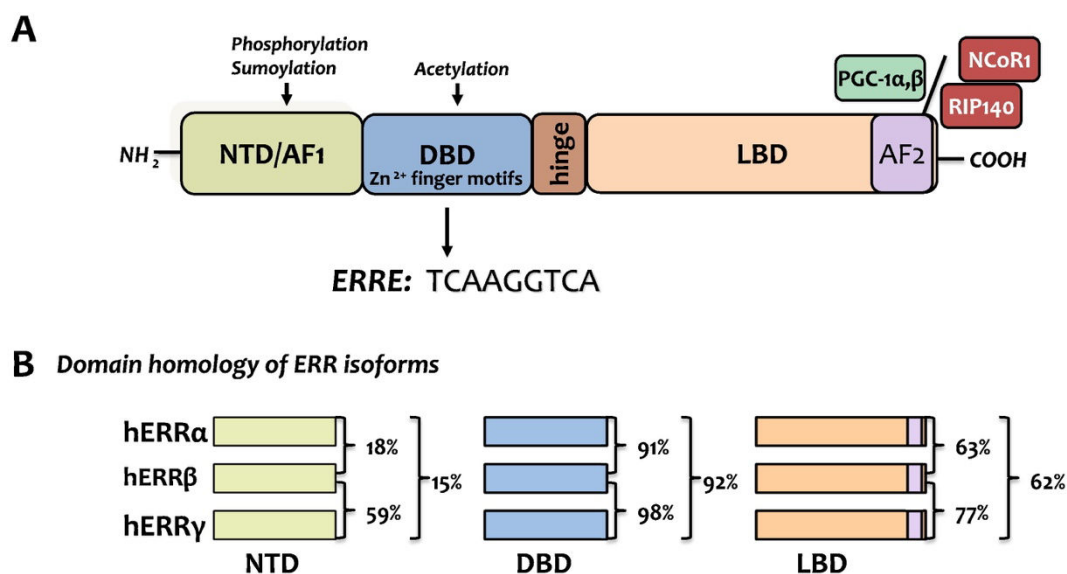


Figure 18. Structural organization and domain homology of Estrogen Related Receptors (ERRs).

A) The structure of ERRs is similar to that of other nuclear receptors and consists of the following regions: a poorly conserved N-terminal domain (NTD) with its activation function-1 (AF-1) which regulates transcription in a ligand independent manner and is subject to posttranslational modifications; a DNA Binding Domain (DBD), almost identical between the three ERR paralogs, which contains two zinc finger domains; a hinge region which links the DBD and LBD, and provides protein flexibility; a constitutively active Ligand Binding Domain (LBD) which recruits coregulators: coactivators as PGC-1 α/β or corepressors as NCoR1, RIP140. B) Amino-acid conservation between the different domains of each ERR. The three ERR isoforms present a high level of similarity, especially in their DBD (Huss, Garbacz, and Xie 2015).

2.4.2.1 ERRs are ubiquitously expressed, but at different levels depending on the cell type

Expression profiles obtained in mice, showed that $ERR\alpha$ is found in each cell and tissue analyzed and $ERR\gamma$ and $ERR\beta$ are widely expressed, with some exceptions, as listed below (Bookout et al. 2006; Deblois and Giguère 2011; Huss, Garbacz, and Xie 2015).

$ERR\alpha$ and $ERR\gamma$ present higher expression levels in tissues where the energy demand is elevated as heart, intestine, kidney, brown fat, skeletal muscle, and cytokine-activated macrophages (Giguere et al. 1988; Sladek, Bader, and Giguere 1997; Deblois and Giguère 2011; Huss, Garbacz, and Xie 2015).

$ERR\beta$ has a wide distribution, with increased presence in the eye, inner ear, extra-embryonic ectoderm of the developing placenta and in mouse embryonic stem cells. $ERR\beta$ is absent in the immune system and both $ERR\alpha$ and $ERR\gamma$ did not show detectable levels of expression in adult bone and skin (Tremblay 2007; Bookout et al. 2006; Pettersson et al. 1996; C.-Q. Xie et al. 2009; Deblois and Giguère 2011).

$ERR\gamma$ is also wide-spread, with a higher profile in developing heart, the spinal cord and some areas of the brain (Alaynick et al. 2007; Bookout et al. 2006).

2.4.2.2 Short overview on ERR-driven biological functions

Several studies of functional genomics, gene expression profiles, location and phenotypic analyses, revealed the functions of the ERRs in normal physiology and disease (Deblois and Giguère 2011).

NRs have been found to have a role in circadian clock, by linking biological timing to metabolic physiology. In particular, the ERRs contribute to the diurnal rhythm in several tissues: kidney, bones, liver, skeletal muscle and uterus (Xiaoyong Yang et al. 2006; Horard et al. 2004; Tremblay et al. 2010; Deblois and Giguère 2011).

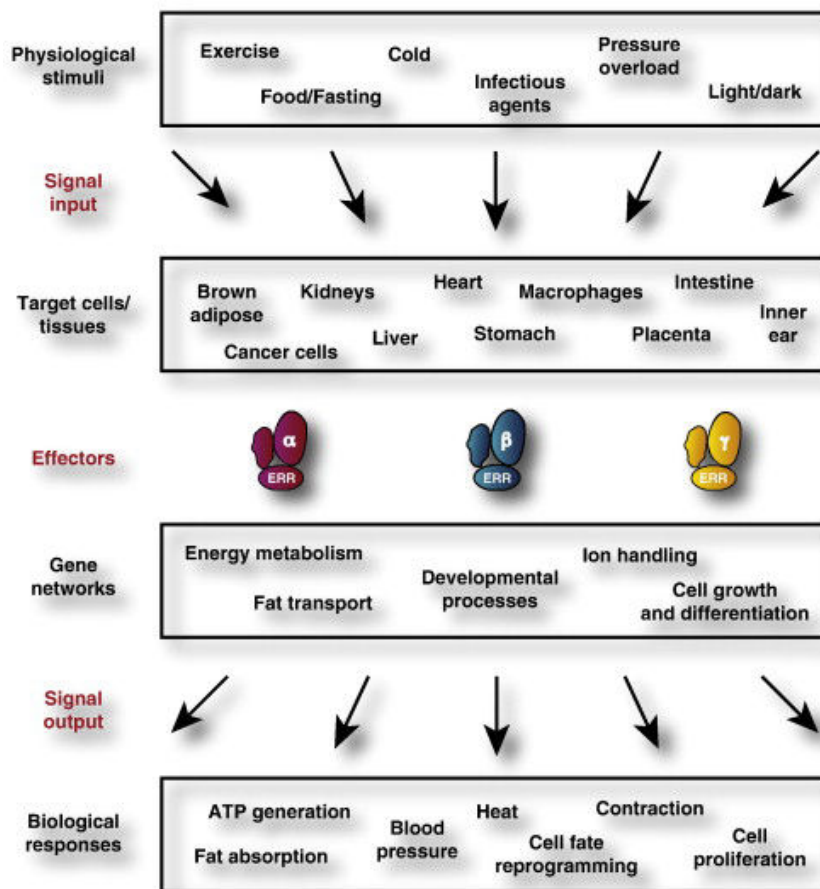


Figure 19. A schematic representation of the physiological responses of ERRs.

External/physiological stimuli induce a response from target tissues. The transcriptional activity of ERRs regulate the expression of a wide panel of gene networks. The output signal is a cell-type-specific biological response (Deblois and Giguère 2011).

2.4.2.3 **ERRα and ERRγ modulate mitochondrial biogenesis and energy metabolism**

The interaction between ERRα and ERRγ with PGC1-α and PGC1-β, factors important in metabolic regulation, and the discovery that genes implicated in mitochondrial β-oxidation of fatty acids and the control of energy metabolism are ERR target genes, have highlighted the role of ERRs in mitochondrial biogenesis and metabolism (Rangwala et al. 2007; Sladek, Bader, and Giguere 1997; J. Lin, Handschin, and Spiegelman 2005).

ChIP studies revealed that ERR paralogs bind mostly to their ERRE motif on genomic DNA and have been identified at the promoter regions of approximately 700 genes encoding mitochondrial proteins and at other genes involved in translation, glucosamine

pathway, energy sensing and growth factor/insulin signaling (Deblois et al. 2009; Fullwood et al. 2009; Sonoda et al. 2007; Deblois and Giguère 2011).

ERR α -deficient mice display an impaired adaptative thermogenesis, being incapable of maintaining a normal body temperature when exposed to moderate cold (Villena et al. 2007). This does not involve ERR α in the induction of thermogenic genes, but ERR α absence is connected with a reduced mitochondrial biogenesis and oxidative ability needed to generate heat (Deblois and Giguère 2011).

ERR α null-mice are lean and resistant to high-fat diet-induced obesity. In addition, these mice exhibit altered expression of target genes involved in fatty acid oxidation, ATP synthesis and translocation. Overall, this phenotype is characteristic of a dysfunctional heart and validates ERR α as essential for the functional adaptation of the heart to pressure overload, and loss of ERR α function can result in heart failure (Dufour et al. 2007; Huss et al. 2007).

Moreover, ERR α nul-mice retain Na⁺, showing deficiency in the regulation of Na⁺ and K⁺ homeostasis. Indeed, ERR α also regulates blood pressure and the renin-angiotensin pathway (Deblois and Giguère 2011; Tremblay et al. 2010).

Notably, ERR α and ERR γ have the same metabolic target genes, but accomplish distinct functions. In contrast with the mild renal phenotype showed by ERR α null-mice, the ERR γ null mice die after birth because of cardiac arrhythmia, decreased gastric acid and high levels of K⁺ (Alaynick et al. 2010).

2.4.2.4 ERR β regulates the body energy balance via ERR γ

The physiologic role of ERR β is less understood, because the homozygous deletion of the *ESRRB* gene is embryonic lethal due to placental dysfunction (Luo et al. 1997). However, the successful generation of ERR β $-/-$ mouse models, unveiled a role for ERR β in hypothalamic regulation of feeding behavior, satiety, whole body energy balance (Byerly et al. 2013; Huss, Garbacz, and Xie 2015). In these mice, the developing embryos were lean and with a high metabolic activity. The mice eat more, have increased insulin sensitivity and carbohydrate metabolism, as demonstrated by a higher respiratory exchange ratio. These observations suggest that their phenotype originates in the central nervous system, consistent with an increased hypothalamic expression of neuropeptide Y

(NPY) and agouti-related peptide (Agrp), neuropeptides that control feeding and energy consumption (Byerly et al. 2013; Huss, Garbacz, and Xie 2015).

In their study, Byerly and colleagues, observed that $ERR\gamma$ expression was increased in $ERR\beta^{-/-}$ mice, indicating that $ERR\gamma$ activation can stimulate some of the metabolic changes. The pharmacological activation of $ERR\gamma$ with a specific $ERR\beta/\gamma$ agonist, DY131, when $ERR\beta$ is deleted or not, decreased the satiety ratio.

The involvement of $ERR\gamma$ in metabolism was confirmed in $ERR\gamma^{-/-}$ mice, which have impaired cardiac activity associated with metabolic changes ($ERR\gamma$ directs and maintains the transition to oxidative metabolism in the postnatal heart). However, $ERR\beta^{-/-}$ mice do not show any heart defects, but their preference for carbohydrate metabolism, suggests that $ERR\beta$ and $ERR\gamma$ homodimers could regulate expression of overlapping target genes. As heterodimers, $ERR\beta$ and $ERR\gamma$ might differentially regulate their target genes depending on their expression levels. Collectively, the two receptors might balance the food behavior, regulating satiety and whole-body energy consumption (Huss, Garbacz, and Xie 2015).

$ERR\beta$ plays also a role in genetic reprogramming and this will be presented in the following section, in the context of functional redundancy between $ERR\beta$ and $ERR\gamma$.

2.5 Functional redundancy and cross-talk between nuclear receptors

In a physiological context, cells are continuously exposed to various combinations of hormones, and thus, multiple ligand-dependent nuclear receptors can be activated at once. As a consequence, studying the mechanism of action and function of a single factor can be challenging.

There is a high interest on unveiling how the co-activation of different nuclear receptors, influence each other's activity. I will discuss below some important aspects regarding the functional redundancy and the cross-talk of nuclear receptors.

2.5.1 Functional redundancy of nuclear receptors

Functional redundancy among genes has frequently been noticed, even if its significance is not always obvious (Thomas JH 1993). The observation that single knockouts of different nuclear receptors had no effect or a mild one, led to a hypothesis in which different NRs could be functionally redundant (Kastner, Mark, Chambon 1995). However, the lack of a defect in a single mutant does not mean that the function of that gene was not debilitated even at low level, but it highlights that the compensation by another receptor is good enough.

Seen in this way, the functional redundancy might assure the survival in less optimal conditions of life.

2.5.1.1 Functional redundancy in genetic reprogramming between $ERR\beta$ and $ERR\gamma$

$ERR\beta$ is one of the factors used in genetic reprogramming of differentiated cells into induced pluripotent stem cells (X. Chen et al. 2008; B. Feng et al. 2009). This receptor can substitute for the original reprogramming factor, KLF4, and can reprogram mouse fibroblasts when it is used together with SOX2 and POU5F1. Consistently, KLF factors bind to $ERR\beta$ genomic binding sites, and approximately 60% of $ERR\beta$ binding sites are found in NANOG/OCT4/SOX2 multiple transcription factor binding loci (X. Chen et al. 2008; Deblois and Giguère 2011).

Moreover, $ERR\beta$ knock-out mice embryo die at mid-gestation because of placental abnormalities (Luo et al. 1997).

Importantly, it was shown that besides $ERR\beta$, $ERR\gamma$ - but not $ERR\alpha$ - also holds reprogramming potential. $ERR\gamma$ -reprogrammed cells have the ability to differentiate into cells and tissues of the main embryonic lineages – mesoderm, ectoderm and endoderm. Given also the high structural similarities between $ERR\beta$ and $ERR\gamma$, $ERR\gamma$ can substitute for $ERR\beta$ in genetic reprogramming of mouse fibroblast. These are proofs of the functional redundancy between these two NRs in reprogramming (B. Feng et al. 2009). Other studies showed that TFs involved in reprogramming can be substituted by members of the same family (Nakagawa et al. 2007; Bluelloch et al. 2007; Jiang et al. 2008).

2.5.1.2 Example of functional redundancy between the RARs

Retinoids are non-steroid hormones which contribute to a variety of biological processes through two classes of nuclear receptors: the retinoic acid receptors (RARs) and the retinoid X receptors (RXRs), which work as RXR/RAR heterodimers (Mangelsdorf and Evans 1995, Kastner 1997). These receptors are activated by binding of retinoic acid produced from dietary vitamin A (retinol). Specifically, RARs are activated by all-trans retinoic acid (tRA) and by 9-*cis* retinoic acid (9C-RA) and RXRs are activated only by 9C-RA only (Hideki Chiba 1997).

There are multiple RAR and RXR isotypes and isoforms, but each has a particular expression pattern, depending on the cell type and developmental stage (Leid 1992 Chambon 1996). Distinct genes encode RAR ($RAR\alpha$, β , and γ) and RXR ($RXR\alpha$, β , and γ) receptors and additionally the different N-terminal parts of their isoforms result through alternative splicing and differential promoter usage (Hideki Chiba 1997, Bastien Rochette-Egly 2003).

The degree of functional redundancy for the RARs was studied in mice deficient for each RAR gene. Interestingly the double homozygote mutants show strong or lethal defects ($RAR\alpha^{-/-}/RAR\beta^{-/-}$ and $RAR\beta^{-/-}RAR\gamma^{-/-}$ mutants die very soon after birth) compared with single mutants which show few defects (Lohnes 1993, Lohnes 1994, Philippe Kastner and Susan Chan 2001).

Important studies on delineating the functions of multiple RARs were done using F9 embryonic teratocarcinoma stem cells which differentiate through the action of RARs and RXRs, when treated with retinoic acid RA (Boylan 1993, Boylan 1995). Separated $RAR\alpha$

or RAR γ knockout revealed that the loss of each receptor results in a specific pattern of metabolic changes, indicating that each RAR regulates a different subset of genes (Boylan 1993, Boylan 1995).

This reveals and highlights that different NR isoforms are not functionally equivalent, even if they can substitute for one another under certain conditions (Kastner, Mark, Chambon 1995).

However, the loss of function of RAR γ was partially rescued by overexpression of either RAR α or RAR β (Taneja 1995). Thus, the functional redundancy might assure the survival in less optimal conditions of life.

2.5.1.3 Functional redundancy between Nur77 and Nor-1 orphan steroid receptors

Another example of functional redundancy comes from Nur77 and Nor-1 orphan steroid receptors. Nur77 knockout mice do not show unusual phenotype or dysfunction suggesting a possible functional redundancy between NR4A members.

Nur77 is a constitutively active NR with no ligand identified so far (Evans 1988, LEE 1993). Its expression is highly induced during T-cell receptor (TCR) mediated apoptosis in immature thymocytes, T-cell hybridomas and is involved in the hypothalamic-pituitary axis signaling (Liu 1994, Woronicz 1994, Cheng 1997).

Nur77 binds to a NGFI-B DNA responsive element (NBRE), at a consensus sequence formed by the half site of the estrogen responsive element extended by two more adenine nucleotides at its 5' end (5'-AAAGGTCA-3') (Wilson et al. 1991). Nur77 binds to NBRE as a monomer or as a heterodimer with retinoid-X receptor in the presence of retinoic acid (Wilson et al. 1993b, Forman et al. 1995, Perlmann & Jansson 1995).

TCR-mediated apoptosis is not disturbed in Nur77 deficient mice (Lee 1995) and Nur77^{-/-} mice show a normal phenotype (Crawford 1995) implying the presence of a protein with redundant function to Nur77 (Chen LE 1997).

Based on DBD homology, Nur77 has two closely related proteins, Nurr1 and Nor-1 (Mangelsdorf 1995) which might have similar roles in TCR-mediated apoptosis. This hypothesis was explored and the results show that both Nor-1 and Nurr1 can transactivate NBRE-containing genes. Interestingly, transactivation activity of all the Nur77 family is repressed by the expression of a Nur77 dominant-negative protein. This could explain the

different phenotypes observed between Nur77-deficient mice and Nur77 dominant-negative transgenic mice, since the dominant-negative protein can suppress all the Nur77 family members, including Nor-1 protein activity. This suggests also that Nur77, Nurr1 and Nor-1 have similar transactivation activity, even if the efficiency of transactivation is different, Nur77 having the highest activity (Chen LE 1997).

It was observed, that in stimulated thymocytes, Nor-1 expression can be induced at the same level as Nur77, while Nurr1 is only transiently induced in T-cell hybridomas, but not in thymocytes. This makes Nor-1 a candidate for the redundancy function in apoptosis.

The constitutive expression of Nur77 induces apoptosis in thymocytes (Calnan et al., 1995; Weih, 1996). To test if Nor-1 or Nurr1 can also induce apoptosis in thymocytes, transgenic mice constitutively expressing full-length Nor-1 protein and Nurr1, respectively, have been generated, the expression of the protein was targeted to the thymocytes and changes in T-cell development and thymocyte number was evaluated. Transgenic Nor-1 showed the same level of apoptosis in thymocytes, as Nur77 expression. In contrast, Nurr1 transgenic mice did not show any significant thymocytes reduction, but it can be due to the weak level of Nurr1 expression (Chen LE 1997).

Collectively, the Nur77 and Nor-1 signaling pathways seem to converge. The overexpression of a dominant-negative Nur77 abolishes TCR mediated apoptosis and the Nur77 deficient mice lack a phenotype. In this case, Nor-1 function compensates the lack of Nur77 highlighting a functional redundancy of two closely related nuclear receptors. (Chen LE 1997).

2.5.2 Cross-talk between nuclear receptors

In the canonical mode of action of nuclear receptors, gene regulation is mediated by NR binding to their cognate regulatory DNA motif. The transcriptional cross-talk model proposes that two NRs can interact on the promoter of target genes with only a binding site for one of the two factors. Additional cross-talks can involve nuclear receptors and transcription factors from other families.

2.5.2.1 A working model for the cross-talk between ER α and ER β

The two distinct estrogen receptors, ER α and ER β , mediate the biological effects of the steroid hormone 17 β -estradiol (E2), with critical functions in differentiation, growth, male and female reproductive systems, skeletal and cardiovascular systems, or mammary gland.

ER α and ER β steroid receptors, are encoded by separated genes, but show high levels of amino-acid homology in their DBD and LBD (97% and 60% homology) (Giguere 1998). In contrast, their N-terminal domains are less conserved (18% homology) (Giguere 1998, Hall Couse Korach 2001).

ER α and ER β differ in tissue distribution, but they bind to the same DNA response elements and, at least *in vitro*, show a similar ligand binding affinity for most existent estrogenic substances or estrogenic antagonists (Kuiper, Gustafsson 1997). Each receptor displays a distinct expression pattern as revealed by tissue localization studies and the biological functions of the ER β may be dependent on the presence of ER α in certain cell types and tissues (Couse, Gustafsson, Korach 1997).

The high homology between the two receptors suggests that these receptors could be functionally redundant in estrogen signaling. Both receptors transactivate ERE reporter constructs in mammalian cells, even if ER α levels are higher (Cowley et al. 1997).

Importantly, ER β was shown to compensate for the loss of ER α in some pathways. Notably, in ER α KO mice, ER β sustained ER α -mediated E2 actions, as the induction of PR expression.

ER α and ER β can bind as heterodimers and ER β can bind to ERE of target genes also in a ligand-independent manner. This will attenuate the ligand-activated transcriptional activity of ER α , indicating a point of convergence for the signaling pathways of these two

receptors. These observations can explain how the transcriptional activity of ER α can be regulated by ER β in cells where the two NRs are co-expressed (Hall and McDonnell 1999).

It was shown that at subsaturated levels of estrogen, inactive ER β competes with ER α , binding first to target response elements and thus impeding ER α binding (Hall and McDonnell 1999). ER β has the ability to modulate cellular response to agonists, by switching from a transcriptional repressor to an activator as estradiol levels rise. This might control the various sensitivities to estrogens. Therefore, when hormone levels increase, activated ER α and ER β levels also increase, which leads to competition between active and inactive, unliganded ER β and transcription can progress (Hall and McDonnell 1999).

If it would be only about competitive interaction, then increased levels of ER β would decrease the hormone efficiency, as when only ER β is expressed in cells. Nevertheless, the interaction between ER α and ER β is more complex, given that when the hormone is in excess, ER β overexpression does not lower estradiol efficacy. It was proposed that when saturating levels of estradiol are present in the cell, ER α and ER β can work as heterodimers, with a transcriptional activity equal to that of ER α homodimers. In this scenario, ER β regulates ER α transcriptional activity at low hormone levels (Hall and McDonnell 1999).

2.5.2.2 Functional cross-talks between ERRs and ERs assist cell-type-specific estrogenic responses

Different transcription factors may bind to identical response elements in various cell types. Transcriptional crosstalk between nuclear receptors implies also competition for binding motifs and coregulators. As mentioned before, the steroid NR subfamily binds DNA as dimers to response elements, using one of the two known consensus hexamers: the AR, PR, GR, MR recognize the half-site AGAACA, while ERs bind to the hexamer AGGTCA, part of the estrogen response element (ERE). The orphan ERRs can bind both to the EREs, or as monomers to the extended consensus hexamer TNAAGGTCA, named ERR response element (ERRE) (Giguère 2002; Johnston et al. 1997). Moreover, the monomeric steroidogenic factor-1 (SF-1) recognizes also this class of binding sites (T. E. Wilson, Fahrner, and Milbrandt 1993). This overlap in binding specificity is complex,

given that not only ERR dimers can bind to ERE, but also ER α dimers (and not ER β) can recognize ERRE (Vanacker 1999). Moreover, it was shown that, *in vitro*, ERR α interacts with ER α , through protein-protein contacts, even if functional heterodimers were not observed (Nengyu Yang et al. 1996).

For example, the transcription of the human lactoferrin gene is mediated by ER α , through an ERE positioned in the promoter of the gene. However, it was demonstrated that this transcriptional activity is modulated by ERR α . Mutations made in the promoter of the lactoferrin gene, 26 bp upstream from ERE which houses the ERRE, considerably reduced the transcriptional activity mediated by ER α (Nengyu Yang et al. 1996; Teng 1992).

Coregulators are crucial for the transcriptional activity of the nuclear receptors and represent limiting factors within cells. Thus, NRs compete with each other and with other transcription factors to associate with coregulators (Meyer et al. 1989; Giguère 2002). The ERs recruit several coregulators, but mostly members of the steroid receptor coactivator (SRC) group (Moggs and Orphanides 2001). Also, ERR α interacts with SRC1 and competes with ER α for binding this common coactivator in transfected cells (Z. Zhang and Teng 2001). ERR α behaves as a significant modulator in many signaling pathways that require the same coactivator and similar DNA response elements (Z. Zhang and Teng 2000).

Yeast two-hybrid and biochemical assays showed that the orphan nuclear receptor small heterodimer partner (SHP) physically interacts with all three members of ERR subfamily through their AF-2 coactivator-binding site. Moreover, SHP promoter is regulated by ERR γ , but not ERR α and ERR β (Sanyal et al. 2002). However, SHP inhibits the activity of ERR family, suggesting another level of crosstalk between the two NR subfamilies (Giguère 2002). Notably, SHP also interacts specifically with estrogen receptor-alpha (ER α) and, in transient cotransfection assays, SHP inhibits estradiol-dependent activation by ER α (Seol et al. 1998).

Collectively, the ERRs and ERs modulate both common target genes and the overall response to estrogen in cell types where they are coexpressed. They are able to interfere or collaborate with each other to accomplish their functions (Giguère 2002).

3. Results

Contents

3.1	Thesis objectives	83
3.2	Research article - in preparation for submission: A subset of steroid receptors interacts with LINE-1 ORF2p and regulate retrotransposition	84

3.1 Thesis objectives

Despite the important role of L1 in generating new insertions in the human genome, how retrotransposition is achieved in a chromatin context and coordinated with cellular activities remains largely unknown and few host regulators are described. Thus, this study started with the aim of discovering cell host factors involved in the regulation of the LINE-1 retrotransposon.

To achieve this goal, the laboratory has performed yeast 2-hybrid screens using a collection of L1-derived fragments as baits and identified $ERR\alpha$, a transcription factor belonging to the steroid nuclear receptor subfamily, as a cellular partner of L1 ORF2p, essential component of the retrotransposition machinery.

For my PhD, I focused on finding other possible interaction within the steroid nuclear receptors superfamily and on unveiling the consequence of this interaction on L1 activity. Considering the existence of several steroid receptors, we explored other possible interactions and found that several members interact with ORF2p, highlighting a possible functional redundancy between $ERR\alpha$ and other steroid receptors. We evaluated the effects of $ERR\alpha$ overexpression and knock-down, as well as the artificial tethering of $ERR\alpha$ to a specific locus and the recruitment of the L1 retrotransposition machinery at the same locus.

All the results of this study are shown in the form of a research article, in preparation for submission. Our work shows how steroid nuclear receptors can influence L1 activity and link environmental and physiological signals with genomic plasticity.

3.2 Research article - in preparation for submission: A subset of steroid receptors interacts with LINE-1 ORF2p and regulate retrotransposition

A subset of steroid receptors interacts with LINE-1 ORF2p and regulate retrotransposition

Ramona Galantonu^{1,3}, Javier Garcia-Pizarro^{1,3}, Clément Monot¹, Margo Montandon¹, Nadira Lagha¹, Aurore-Cécile Valfort¹, Serdar Kasakyan¹, Pierre-Olivier Vidalain², Gael Cristofari^{1,4,*}

¹Université Côte d'Azur, Inserm, CNRS, IRCAN, Nice, France

²Université Paris Descartes, CNRS UMR 8601, Laboratoire de Chimie et Biochimie Pharmacologiques et Toxicologiques, Equipe Chimie and Biologie, Modélisation et Immunologie pour la Thérapie, 75006 Paris, France

³These authors contributed equally

⁴Lead contact

*Correspondence: Gael.Cristofari@unice.fr

Abstract

LINE-1 (L1) retrotransposons are major drivers of mammalian genome plasticity and evolution. They replicate by a copy-and-paste mechanism through an RNA intermediate and a reverse transcription step occurring directly at chromosomal target sites, a process termed target-primed reverse transcription. However, how this is achieved in a chromatin context and coordinated with cellular activities remains largely unknown. In addition, L1 replication is tightly controlled by cellular pathways limiting its mutagenic activity. To gain insights into the interplay between the L1 machinery and its cellular host, we performed yeast 2-hybrid screens using a collection of L1-derived fragments as baits. Here, we identify the estrogen-related receptor alpha (ERR α), a transcription factor belonging to the nuclear receptor family, as a cellular partner of L1 ORF2p, the catalytic component of the L1 replicative complex. SiRNA-mediated ERR α depletion marginally affects L1 retrotransposition. However, overexpression of an ERR α dominant-negative form inhibits retrotransposition. In addition, we observe that several close paralogs of

ERR α , including ERR β and ERR γ , and the estrogen and glucocorticoid receptors, also interact *in vivo* with ORF2p, highlighting a high level of redundancy. Interestingly, tethering of the ERR α domain interacting with ORF2p to a heterochromatic repeated locus specifically inhibits L1 retrotransposition, suggesting that the L1 machinery is sequestered in a chromosomal region refractory to L1 integration. Based on these observations, we propose that ERR α and several other related nuclear receptors might represent tethering factors that promote L1 access to particular genomic region. Collectively, these results link hormonal signaling pathways with the regulation of a major endogenous mutagen in mammals.

Keywords

transposable element, nuclear receptor, host factor, tethering, integration

Introduction

Repetitive DNA accounts for at least half of our genome^{1,2}. Most of these repeats are retrotransposons, *i.e.* mobile genetic elements, which proliferate through an RNA-mediated copy-and-paste mechanism, called retrotransposition. The Long INterspersed Element-1 (LINE-1 or L1) is the sole family of retrotransposons able to autonomously generate new copies in the genome of modern humans³⁻⁵. This activity originates from a very limited subset of transcriptionally active L1 loci, among approximately 100 copies, which are still replication-competent in our genome⁶⁻¹⁰. In addition, the L1 retrotransposon machinery can also mobilize *in trans* non-autonomous retrotransposons (*Alu*, SVA) or cellular RNAs (U6, mRNA), causing retropseudogene formation (reviewed in¹¹). Altogether, L1 elements are directly or indirectly responsible for at least one fourth of all structural variants in the human population, contributing to our genetic diversity^{5,12,13} and occasionally to genetic diseases¹⁴. Moreover, L1 is not only able to mobilize in the germline or in the early embryo^{15,16} but it can also transpose in some somatic tissues such as brain, and in many cancers contributing to tumor genome dynamics¹⁷⁻²⁰.

A replication-competent L1 element is ~6.0 kb in length and is transcribed from an internal promoter located in the 5'-untranslated region (UTR) (reviewed in¹¹). The full length L1 mRNA is a bicistronic polyadenylated RNA, which encodes two proteins, ORF1p and ORF2p, both required for L1 retrotransposition²¹⁻²³. ORF1p is a trimeric RNA-binding protein with nucleic acid chaperone properties^{24,25}. ORF2p exhibits endonuclease (EN) and reverse transcriptase (RT) activities^{23,26,27}. These proteins associate *in cis* with the L1 mRNA and the poly(A)-binding protein C1 (PABPC1) to form a ribonucleoprotein particle (RNP), the presumed replication intermediate of L1 replication²⁸⁻³⁵. The incorporation of ORF2p into L1 RNPs requires L1 mRNA poly(A) tail³⁶. The L1 RNP is imported into the nucleus where it can catalyze target-site reverse transcription (TPRT)³⁷. In this process, DNA cleavage and reverse transcription of the L1 RNA at genomic target sites are coordinated. TPRT is initiated by an EN-mediated nick at degenerate consensus sequences related to 5'-TTTT/A-3', followed by annealing of the L1 mRNA poly(A) tail to the T-rich tract of the target sequence, and extension of the free 3'-hydroxyl group by L1 RT activity using L1 mRNA as a template^{30,38-40}. However, many aspects of this reaction remain unsolved. More specifically, we do not know how

the L1 RNPs enter into the nucleus and access their target DNA in the context of chromatin and nuclear organization, or how second strand synthesis and ligation of the final DNA product are achieved. It is likely that these steps involve cellular host factors. For example, association of the sliding DNA clamp protein PCNA with ORF2p might help recruiting cellular factors involved in TPRT resolution³⁵. Inversely, several cellular pathways limit L1 retrotransposition (reviewed in^{41,42}). Many of them function by reducing L1 mRNA or RNP accumulation and often involve restriction factors enriched in cytoplasmic bodies related to stress granules where L1 RNPs also concentrate^{32,34,43}. In the nucleus, APOBEC3A is a potent inhibitor L1 retrotransposition, which can deaminate transiently exposed single-stranded DNA intermediates during TPRT⁴⁴. L1 replication intermediates can also be sensed and processed by DNA repair pathways, competing with TPRT and eventually leading to the insertion 5' truncated L1 copies⁴⁵⁻⁴⁸. These defective insertions form the majority of *de novo* insertions and represent dead-ends of L1 replication.

To get additional insights into the mechanisms of L1 replication and its interplay with cellular host factors, we performed yeast 2-hybrid screens to identify cellular partners of L1 proteins. This led us to discover that estrogen-related receptor alpha (ERR α), as well as other members of the nuclear receptor (NR) family, interact with ORF2p. The functional characterization of this interaction suggest that several NRs might regulate L1 retrotransposition by tethering the L1 machinery to particular genomic regions.

Results

Identifying L1 cellular partners by yeast 2-hybrid screens

To identify potential cellular partners of the L1 retrotransposition machinery, we screened mouse and human cDNA libraries by yeast 2-hybrid (Y2H) with a variety of ORF1p or ORF2p fragments as baits. These fragments originated either from a natural human L1 element (L1.3)⁴⁹ or from a codon-optimized mouse L1 element (mouse ORFeus)⁵⁰. Most L1 fragments recovered very few hits, and these hits were isolated as single clones, and therefore were considered as false positives (**Supplementary Table 1**). However, a fragment of mouse ORF2p allowed us to isolate - in two independent screens - several overlapping clones corresponding to mouse Estrogen-Related Receptor alpha (ERR α), a

transcription factor of the nuclear receptor (NR) family (**Fig. 1a**). Nuclear receptors have a modular organization, with a N-terminal DNA-binding domain (DBD) and a C-terminal ligand-binding domain (LBD), which also contains an activation domain, responsible for transcriptional coregulator recruitment⁵¹⁻⁵³. Hormone-dependent NR are often sequestered in the cytoplasm in the absence of their cognate hormone. In a classical NR-signaling pathway, ligand binding leads to receptor conformational change, nuclear translocation, DNA binding, transcriptional coregulator recruitment and eventually to gene activation or repression⁵¹. All isolated ERR α clones contain a C-terminal fragment spanning the last 232 amino-acids of ERR α , corresponding to its ligand-binding domain (LBD, **Fig. 1b**). ERR α is broadly expressed in human tissues⁵² and is predominantly nuclear⁵⁴. Of note, ERR α functions as a ligand-independent transcription factor, since its ligand-binding pocket is constitutively occupied by one of its own amino-acid, mimicking an activated state⁵⁵. Instead, ERR α activity seems to be mostly regulated by post-translational modifications, particularly in its N-terminal domain (NTD), which can repress ERR α transcriptional activity⁵⁴.

ERR α interacts with human L1 ORF2p

Human and mouse ERR α are 98% identical, however human and mouse ORF2p Cter domains only share 63% identity. To validate the interaction found by Y2H and to test whether it is functionally conserved from mouse to human, we performed co-immunoprecipitation experiments in human cells, using epitope-tagged human ERR α LBD and human L1 ORF2p Cter domain, which confirmed the Y2H results (**Fig. 1c**). Since the ORF2p Cter domain can interact with RNA⁵⁶, we also examined whether its interaction with ERR α might depend on RNA. RNase treatment did not abolish ERR α -ORF2p interaction (**Sup. Fig. 1a-b**), although, under similar experimental conditions, bulk RNA in cell extracts is completely degraded (**Sup. Fig. 1c**) and the well-established and RNA-dependent interaction between ORF1p and Mov10 is abolished⁵⁷ (**Sup. Fig. 1d**). Finally, we tested the ability of several variants of untagged ERR α to bind ORF2p Cter domain. These included full length ERR α , the LBD alone, or a deletion variant, missing the N-terminal domain, named Δ N (**Fig. 1d**). While the ORF2p Cter domain co-immunoprecipitates with both LBD and Δ N ERR α truncations, it barely interacts with full length ERR α (**Fig. 1e**), consistent with a regulatory role of the NTD. Altogether these

data indicate that ERR α binds L1 ORF2p Cter domain in human cells through its LBD, and independently of RNA, and suggest that this interaction could be modulated by the N-terminal domain of ERR α .

ERR α depletion does not affect L1 retrotransposition efficiency

To explore the possible impact of ERR α on L1 retrotransposition, we performed siRNA-mediated knock down of ERR α in U2OS cells, followed by transfection of a genetically marked L1 element to estimate L1 retrotransposition efficiency (**Fig. 2a,d**). Two days after L1 transfection, ERR α siRNA treatments resulted in an ~70-80% reduction of endogenous ERR α protein levels as compared to mock-transfected cells, without significantly affecting L1 expression as evidenced by ORF1p immunoblotting, at least for siRNA A and B (**Fig. 2b-c**). The plasmid-borne L1 element is a retrotransposition-competent L1 clone carrying a blasticidin-resistant marker, which becomes functional only after retrotransposition (JJ101/L1.3, **Fig. 2d**). As a control for cytotoxicity and for the capacity to form blasticidin-resistant colonies upon siRNA treatment, we transfected cells with pcDNA6, a plasmid containing a blasticidin selection cassette identical to that present in JJ101/L1.3 but without the intron. None of the specific siRNAs directed against ERR α significantly affects L1 retrotransposition efficiency under these experimental conditions. Thus, we conclude that ERR α is not an essential cellular cofactor for L1 retrotransposition, although we cannot exclude that the remaining ERR α could be sufficient to achieve such functions.

ERR α LBD overexpression acts as dominant negative on L1 retrotransposition

To further explore the potential impact of ERR α on L1 life cycle, we overexpressed ERR α and its truncated variants (**Fig. 1d**) together with a replication-competent L1 element containing a Luciferase-based retrotransposition marker (pYX14)⁵⁸. This marker is similar in principle to that described in **Fig. 2d**, but contains a Firefly Luciferase (FLuc) reporter interrupted with an intron instead of the blasticidin-resistance gene. It also contains a regular Renilla Luciferase (RLuc) expression cassette in the plasmid backbone which can be used as an internal normalizer in dual-Luciferase assays. Finally, L1 expression is only driven by its natural 5' UTR promoter (no CMV promoter). As controls, we used a defective L1 element (pYX15). In contrast to the Blast^R-based assay,

the Luc-based assay allows measuring retrotransposition in a short window of time. Therefore, we reasoned that it could mitigate possible direct or indirect transcriptional effects of $ERR\alpha$ and its variants on the different components of the assay and on cell physiology. First, we tested the ability of each $ERR\alpha$ construct to activate the transcription of a reporter gene under the control of a minimal SV40 promoter with or without estrogen-related receptor responsive elements (ERRE). In contrast to the LBD domain alone, both full length $ERR\alpha$ and the ΔN variant can activate ERRE-dependent transcription (**Fig. 3a**). Surprisingly, they also moderately stimulate transcription in the absence of ERRE sites (although not statistically significant), suggesting that they partly exert their effect on the SV40 promoter indirectly. Thus, as expected, both FL and ΔN are transcriptionally active, while the LBD, which cannot bind DNA, is not. Then, we tested whether these $ERR\alpha$ constructs could impact mobilization of L1 in the Luciferase-based retrotransposition assay. Neither full length $ERR\alpha$, nor its ΔN truncated form, affect L1 retrotransposition. In contrast, overexpressing the LBD domain alone reduces L1 mobilization by ~40% (**Fig. 3b**). Of note, the reduction of L1 retrotransposition upon LBD overexpression, does not coincide with a decreased activity of the L1 promoter or of the promoter of the reporter gene (5'UTR and SV40, respectively, **Fig. 3c**). Altogether, these results indicate that the LBD can act as a dominant-negative to suppress L1 retrotransposition.

Functional redundancy between $ERR\alpha$ and other steroid receptors

One possible explanation for the lack of effect of $ERR\alpha$ depletion on L1 retrotransposition is the redundancy of NRs. The NR family contains 48 members in humans. $ERR\alpha$ belongs to the steroid receptor superfamily, which includes two other close paralogs of $ERR\alpha$, namely $ERR\beta$ and $ERR\gamma$ (**Fig. 4a**). To address the possibility that ORF2p interacts with other members of the NR family, we used a fluorescent 2-hybrid (F2H) experimental scheme (**Fig. 4b**)⁵⁹. For this purpose, we took advantage of an established U2OS-derived cell line, which contains a large LacO repeat array in the euchromatic region of chromosome 1p36⁶⁰, in which we stably expressed our protein of interest, ORF2p-Cter fused to GFP, by retroviral transduction. In this cell line, the transfection of mCherry-LacI fused to various baits (X) leads to a nuclear red spot. If the protein of interest (Y) interacts with X, the GFP-Y fusion protein also accumulates at the

LacO array and forms a green spot co-localizing with the red spot. In the absence of interaction, the GFP-Y protein remains diffuse in the nucleus. First, we verified that the mCherry-LacI fusion proteins correctly fold by testing their ability to activate transcription in a Luciferase assay. Each of the mCherry-LacI fusion constructs were co-transfected with a plasmid containing a Firefly Luciferase gene under the control of a minimal SV40 promoter, with or without 3xLacO repeats, and with a plasmid expressing Renilla Luciferase under the control of a constitutive TK promoter, as a normalizer (**Fig. 4c**). As expected, none of the LacI fusion activates transcription in the absence of LacO repeat (**Fig. 4d**). However, when LacO repeats are present, all three ERR paralogs stimulates transcription ~3 to 10-fold as compared to the mCherry-LacI alone. As a positive control for transcriptional activation, mCherry-LacI fused to the strong viral VP16 activating domain stimulates transcription >100 fold. These results indicate that all these mCherry-LacI fusion proteins correctly fold and are transcriptionally active. Strikingly, the LBDs of all three ERR paralogs can interact with ORF2p-Cter by F2H in contrast to mCherry-LacI alone or fused to VP16 activation domain (**Fig. 4e,f**).

Next, we tested whether other related nuclear receptors belonging to the steroid receptor superfamily, could also interact with ORF2p (**Fig. 4a**). Notably, and in contrast to ERRs, the activity of these transcription factors is hormone-dependent. In a first set of control experiments, as previously achieved for ERRs, we tested whether the mCherry-LacI fusion proteins were transcriptionally active, as an indirect readout of proper folding. The fusion proteins with estrogen receptor (ER α and ER β) or glucocorticoid receptor (GR) LBDs are transcriptionally active and hormone-dependent (**Supplementary Fig. S3a,b**). Remarkably, they also all interact with ORF2p-Cter, but only in presence of their respective ligands (**Supplementary Fig. S3f,g**). In contrast, androgen-, mineralocorticoid-, and progesterone receptor fusion proteins were totally unable to stimulate transcription of the reporter gene with the LacO repeats (**Supplementary Fig. S3c-e**), indicating that these constructs do not properly fold *in vivo* or that their LBD regions are not sufficient to stimulate transcription. For this reason, we did not explore further the potential interaction of MR and PR with ORF2p. Since a previous report suggested that AR and ORF2p might cooperate to cause chromosome fusions⁶¹, we tested by F2H whether this inactive form of AR was nevertheless able to interact with ORF2p-Cter, but this was not the case under our experimental conditions (**Supplementary Fig. S3f,g**).

Thus, L1 ORF2p can interact with at least two thirds of the steroid receptors (all ERRs and ERs, as well as GR) and we cannot conclude for the remaining third (AR, MR and PR).

Tethering ERR α LBD to a heterochromatic repeated array inhibits L1 retrotransposition

Dissecting the possible role of ERR α in L1 retrotransposition cycle is rendered difficult by the pleiotropic effects of NR signaling pathways on cellular physiology, through their direct or indirect transcriptional activity, their high degree of interconnections, and their potential functional redundancy, at least with regard to ORF2p interaction. We reasoned that we could mitigate some of these issues by testing L1 retrotransposition in a cellular context in which ERR α LBD is tethered to a specific locus independently of its DNA-binding domain. For this purpose, we took advantage of the U2OS cells containing the LacO array (U2OS 2-6-3 clone), already used for the F2H experiments (**Fig. 4b**), in which we transfected a replication-competent L1 element containing the blasticidin-resistance retrotransposition marker or an RT-defective L1 as negative control (**Fig. 2d and 4a**). Parallel experiments were performed in the parental U2OS cell line, which does not contain the LacO array, as control. Four days after transfection, L1 expression, measured by ORF1p immunoblotting, was unaffected or slightly increased by the coexpression of mCherry-LacI-LBD as compared to mCherry-LacI (**Fig. 2b,c**). Of note, the overall levels of L1 overexpression were always less important in the parental cell line than in U2OS 2-6-3. For this reason, we only compared conditions within a particular cell line rather than between cell lines. In the parental U2OS cells (- LacO), overexpressing mCherry-LacI-LBD does not affect L1 retrotransposition as compared to mCherry-LacI alone (**Fig. 5d,e**). In contrast, when the LacO array is present (+ LacO), we observed a decreased number of blasticidin-resistant colonies upon L1 transfection when mCherry-LacI-LBD is expressed as compared to mCherry-LacI. Notably, mCherry-LacI-LBD overexpression in U2OS 2-6-3 cells also results in a slight reduction of pCDNA6-mediated colony formation, suggestive of diminished cell growth or slight cytotoxicity. However, quantification of the retrotransposition assay, and normalization by the pCDNA6 controls, indicate that tethering ERR α LBD to the LacO array specifically inhibits L1 retrotransposition, consistent with a model in which ERR α LBD attracts the L1 machinery to the LacO array and potentially prevents efficient retrotransposition.

Although the LacO array in U2OS 2-6-3 is inserted in a euchromatic region, the body of the array itself is heterochromatic⁶⁰. Therefore, our observations could also be the consequence of the silencing of L1 copies (and blasticidin marker) inserted into the LacO array. To test this possibility, we specifically measured the number of copies of the spliced blasticidin-resistance gene, which reflects the number of L1 cDNA formed, in unselected cells, by a multiplex droplet digital PCR (ddPCR) assay. As expected, the spliced product is undetectable in untransfected cells or cells transfected with an RT-deficient L1, but readily detected in cells transfected with a replication-competent L1 element, both in blasticidin-selected and unselected cells, demonstrating the specificity of the amplification (**Supplementary Fig. S4**). Strikingly, L1 cDNA formation is strongly diminished when ERR α LBD is tethered to the LacO array, but not when the array or the LBD is absent (**Fig. 5f**). Thus, we conclude that tethering the LBD of ERR α to the LacO array attracts and sequester the L1 machinery in a region refractory to retrotransposition, which reduces its overall efficiency.

Discussion

In summary, we show here that a subset of NRs, including all three ERRs, the two ER paralogs and GR, can physically interact with ORF2p, a key component of the L1 RNP, in human cultured cells (**Fig. 1 and 4, Supplementary Figures S1 and S3**). NR binding to ORF2p is mediated by their LBD. Knocking down a single NR (ERR α) is not sufficient to alter retrotransposition efficiency (**Fig. 2**). However, a truncated form of ERR α containing solely the LBD can act as a dominant-negative and inhibits L1 retrotransposition, indicating that ERR α can functionally impact L1 retrotransposition (**Fig. 3**). Finally, tethering the LBD of ERR α to a repeated heterochromatic array also reduces L1 retrotransposition efficiency, likely by sequestering the L1 RNP in a chromatin context poorly permissive to L1 integration (**Fig. 5**). Altogether, our data are consistent with a model in which several NR receptors act as cellular partners of L1 retrotransposition, and cooperate to guide the L1 RNP to particular euchromatic regions of the genome (**Fig. 6**).

Interestingly the integration of transposable elements and retroviruses is often guided by tethering factors that promote the integration of these mobile genetic elements into

specific chromosomal regions (see ⁶² for review). Tethering factors have been identified for many long-terminal repeat (LTR)-containing retroelements. Among others, the BET family of proteins favor murine leukemia virus (MLV) integration in promoters and enhancers ⁶³⁻⁶⁵, LEDGF/p75 guides human immunodeficiency virus (HIV) insertions into highly spliced and transcribed genes ^{66,67}, TFIIB and AC40 direct yeast Ty3 and Ty1 integration close to tRNA genes, respectively ^{68,69}, and finally yeast Ty5 preferentially targets subtelomeres through an interaction between Ty5 integrase and Sir4p ^{70,71}. Interestingly, Ty5 targeting to subtelomeres depends on the phosphorylation of Ty5 integrase. Under metabolic stress conditions, integrase phosphorylation is reduced and insertions are relocated throughout the genome, indicating that Ty5-mediated insertional mutagenesis is a regulated process ⁷². With regards to non-LTR-retrotransposons, evidence of a tethering mechanism is scarce ⁶². The most documented example is TRE5-A, a non-LTR retrotransposon of the L1 clade found in *Dictyostelium discoideum*, which inserts specifically upstream of tRNA genes through an ORF1p-TFIIB interaction ^{73,74}. There is currently no known integration site preference for mammalian L1 elements beside the EN target site of ORF2p, a degenerate consensus sequence related to 5'-TTTT/A-3' ^{23,75-77}. Our results suggest that some steroid receptors could function as tethering factors for L1 in mammals, by virtue of their interaction with the C-terminal end of ORF2p, and raise the intriguing possibility that *de novo* L1 insertions could disperse less randomly in the genome than previously anticipated. Consistently, a previous report implicated ORF2p in androgen receptor (AR)-mediated chromosome breaks and fusions in prostate cancer cells ⁶¹, although we could not conclusively test the interaction of ORF2p with AR under our experimental conclusions.

We originally identified ERR α as interacting with ORF2p and we extended this observation to a number of other steroid receptors. However, we cannot exclude that ORF2p interacts with several other NRs, outside of the steroid superfamily. Future studies will be necessary to delineate the precise set of NRs which can associate with ORF2p and whether some subsets could function predominantly in particular cell-types. This will be necessary to systematically test the genomic association between L1 integration sites and NR binding sites. Of note, several NR agonists or antagonists have been identified in reporter assays as regulating positively or negatively L1 promoter

activity^{78,79}. Thus, NR might influence the replication cycle of L1 elements both at transcriptional and post-transcriptional levels.

In conclusion, given the importance of steroid receptors in the physiology of mammals and their adaptation to environmental changes, we speculate that NR-mediated tethering of ORF2p might modulate the landscape of L1 integration in the genome and link environmental and physiological signals with genomic plasticity.

Methods

Yeast 2-hybrid screens

Our yeast two-hybrid (Y2H) protocol was previously described in details⁸⁰. Briefly, fragments from human or mouse L1 coding regions (Supplementary Table S1) were used as baits to screen human or mouse prey cDNA libraries, respectively. L1 fragments from the human L1.3 element (Genbank Acc No: L19088)⁴⁹ or from the mouse and codon-optimized ORFeus construct⁵⁰ were introduced by PCR amplification and *in vitro* recombination (BP cloning, Gateway system, Invitrogen) into pDONR207, a Gateway entry vector. After validation by sequencing, these L1 fragments were then transferred by recombinational cloning (LR reaction, Gateway system, Invitrogen) from pDONR207 into the yeast expression vector pPC97-GW (provided by Dr. Marc Vidal) to be expressed as fusion proteins with GAL4 DNA-Binding Domain (GAL4-BD) at the N terminus. Plasmids were transformed into the AH109 yeast strain (MAT α), which includes a His3 reporter gene to select for two-hybrid interactions. Then, GAL4-BD bait fusion proteins were tested for autonomous transactivation of the His3 reporter gene, but none of them showed some self-transactivation activity (no background in the absence of interaction). Commercial human spleen and mouse brain cDNA libraries cloned in the pPC86 vector backbone (Invitrogen), which contains a GAL4-Activating Domain (GAL4-AD), were introduced into the Y187 yeast strain (MATa) using a high-efficiency transformation protocol to conserve the clonal diversity of the libraries. We used a mating strategy between bait- and prey-containing strains to screen the cDNA libraries, and at least 4.10^7 diploids were generated per bait. After growth for 6-7 days on a selective culture medium (–L-W-H + 10 mM 3-aminotriazole), His⁺ colonies were picked. The GAL4-AD-cDNA fusions expressed in these clones were amplified by PCR and

sequenced to identify interactors by BLAST analysis (**Supplementary Table S1**). ERR α was the only hit found independently in several screens and represented by distinct overlapping fragments.

Cell culture and cell transfection

U2OS osteosarcoma cells (#92022711, ECACC, distributed by Sigma-Aldrich) and their derivatives, HeLa cervical cancer cells (CCL-2, ATCC, provided by ENS-Lyon cell bank), and HEK-293T human embryonic kidney cells (CRL-11268, ATCC, a kind gift from A. Cimorelli, ENS-Lyon) were grown in DMEM medium (Gibco) containing 4.5 g/L D-Glucose, 110 mg/L Sodium Pyruvate, and 862 mg/mL L-Alanyl-L-Glutamine (Glutamax), supplemented with 10% fetal bovine serum (FBS, Gibco), 100 U/mL penicillin and 100 μ g/mL streptomycin (Gibco). Cell lines were maintained at 37°C with 5% CO₂ in humidified incubators (Sanyo). The U2OS 2-6-3 clone is an U2OS-derived clone containing ~200 copies of a plasmid with 256 *Lac* operator (*LacO*) repeats⁶⁰ (kindly provided by D. Spector). The U2OS-263-ORF2Cter-GFP is a clonal derivative of U2OS 2-6-3 transduced with pVan571, a retroviral vector containing ORF2-Cter N-terminally fused to NLS-GFP in the pLNCX2 backbone (Clontech) and selected with G418, to facilitate F2H screens. The cell cultures were tested negative for mycoplasma infection using the MycoAlert Mycoplasma Detection Kit as routine test (Lonza).

Hormonal treatments

Prior to hormonal treatment, cells were incubated overnight in growth medium containing charcoal-treated FBS instead of complete FBS. The day after, we added either ethanol for mock-treated cells, or the one of the following ligands dissolved in ethanol at a final concentration of 10 nM: β -estradiol (E8875, Sigma), dexamethasone (D1756, Sigma), 5 α -androstan-17 β -ol-3-one (also known as dihydrotestosterone, A8380, Sigma), progesterone (P0130, Sigma), aldosterone (A9477, Sigma).

Plasmid constructions and preparation

Plasmid constructs are detailed bellow. Oligonucleotide sequences and cloning strategies used are available upon request. When indicated, plasmid constructs were generated by SLiCE cloning, a method based on *in vitro* homologous recombination⁸¹. LBD of the different NR were amplified by PCR from the Open Biosystems Human Full-Length

cDNA Set of Nuclear Receptors (MHS4911, Thermo Scientific) and corrected by PCR. All plasmid DNA was prepared with a Midiprep Plasmid DNA Kit (NucleoBond, Macherey-Nagel) and verified by Sanger sequencing (GATC Biotech).

pVan327 - pCMV-myc-Gateway/hERRalpha LBD contains a MYC-tag on the ligand binding domain of human ERRalpha (obtained through Gateway cloning).

pVan297 - pCI-neo-3xFlag/hORF2p opt C-term contains a triple FLAG tag on the C-terminal sequence of ORF2p.

pVan 576 - hERRα FL is derived from pF2H vector and contains full length human ERRα, cloned from the NR set, gene accession number BC092470 (Slice cloning).

pVan551 - hERRα ΔN is identical to pVan576, but contains a deletion variant of human ERRα, missing the N-terminal domain (Slice cloning).

pVan552 - hERRα LBD is identical to pVan576, but contains solely the ligand binding domain of human ERRα (Slice cloning).

pCEP4 (Life Technologies) is an episomal mammalian expression vector carrying the hygromycin B resistance gene.

pJJ101/L1.3 is a pCEP4-based plasmid containing an active human L1 (L1.3) (accession no. L19088)⁸² carrying a *mblastI* retrotransposition indicator cassette⁸³. L1 expression is augmented by a CMV promoter located upstream of the L1 5' UTR and an SV40 polyadenylation signal that is located downstream of the native L1 polyadenylation signal. This vector and the following pJJ105/L1.3 were a generous gift from Dr. J.L. Garcia-Perez (Univ. of Edinburgh, UK).

pJJ105/L1.3 is similar to pJJ101/L1.3, but contains a D702A missense mutation in the RT active site of L1.3 ORF2⁸³.

pcDNA6 (Life Technologies) is a mammalian expression vector with a blasticidin resistance selection cassette.

pJM101/L1.3 is similar to pJJ101/L1.3, but contains a *mneoI* retrotransposition reporter cassette.

pVan571 is a retroviral vector containing ORF2-Cter N-terminally fused to NLS-GFP in the pLNCX2 backbone (Clontech), used to transduce the U2OS 2-6-3 cell line, and to obtain the U2OS 263-ORF2Cter-GFP clone.

pVan 396 - mCherry-LacI is derived from pF2H vector and contains the fusion protein mCherry-LacI (Slice cloning). The F2H plasmids used for cloning were a generous gift from H. Leonhardt (Ludwig Maximilians Univ., Munich, Germany).

pVan386 - mCherry-LacI-LBD-hERR α is identical to pVan396 and contains the fusion protein mCherry-LacI and the ligand binding domain (LBD) of human ERR α . Other plasmids are identical to this one, but contain the ligand binding domain of other human nuclear receptors - ERR β , ERRY, ER α , ER β , GR, MR, AR, PR (Slice cloning).

pYX014 is dual luciferase vector that expresses a human L1 (L1RP) equipped with a Firefly Luciferase retrotransposition indicator cassette and an intact RLuc expression cassette on the vector backbone for normalizing transfection efficiency. L1 expression is driven by its own promoter, located in the L1 5' UTR⁵⁸; this vector and the following pYX015, were a generous gift of Wengfen An (South Dakota State Univ., SD, USA).

pYX015 is identical to pYX014 except that it carries two missense mutations in ORF1p and is thus retrotransposition-incompetent and used as negative control⁵⁸.

pGL3prom (FLuc) - commercial vector from Promega, with Firefly Luciferase (FLuc) cassette under the control of the SV40 promoter.

LacO-FLuc is identical to pGL3prom (Fluc), except that the SV40 promoter is preceded by 3 LacO repeats.

pYX013 (RLuc) expresses Renilla Luciferase (RLuc) from a constitutive TK promoter, used for transfection control.

pGL2prom contains a Firefly Luciferase (FLuc) cassette under the control of the SV40 promoter.

ERRE-luc is identical to pGL2prom vector, but has 3 binding sites for ERR (ERRE) upstream of SV40 promoter and Firefly Luciferase (FLuc) cassette.

pVan575 is identical to ERRE-luc, but it lacks ERRE sites.

pVan601 is identical to pGL2prom vector, but it does not contain the SV40 promoter or any other promoter.

pVan604 - L1 5'UTR is a pGL2prom vector with FLuc cassette under the control of the L1 5'UTR promoter (cloned from JM101/L1.3).

Antibodies

Primary antibodies used for immunoblots were directed against ERR α (C-terminal epitope, clone E1G1J, Cell Signaling Technology, 1:3,000); or N-terminal epitope, ab137489, 1:1,000), Flag (clone M2, 1:1,000, Sigma-Aldrich), hORF1p (rabbit polyclonal antibody, serum SE-6798, 1:1,000)^{6,39}, β -Tubulin (clone BT7R, Pierce Biotechnology, 1:10,000), mCherry (ab183628, 1:1000), HA (clone HA.C5, ab18181, 1:1,000), MOV10 (ab80613, 1:3000) or Myc (clone 9E10, Sigma, 1:2000). As secondary

antibodies (LI-COR Biosciences), we used IRDye 680RD Goat anti-Rabbit IgG, IRDye 800CW Goat anti-Rabbit IgG, IRDye 680RD Goat anti-Mouse IgG or IRDye 800CW Goat anti-Mouse IgG, diluted at 1:10,000.

For IP, antibodies were used as follow: 2 µg of the M2 Flag antibody, 1:50 diluted for the ERRα (C-terminal epitope, clone E1G1J, Cell Signaling Technology) or 2 µg of the Myc antibody.

Co-immunoprecipitation

HEK-293T cells were seeded in 6-well plate at 400,000 cells/well, and transfected on the next day with 4 µg of total DNA and 10 µl of Lipofectamine 2000. Forty-eight hours post-transfection, cell pellets were resuspended in 300 µL of lysis buffer (20 mM Tris-HCl pH 8, 137mM NaCl, 1% Triton X-100, 2mM EDTA, supplemented with protease inhibitor (Complete Mini, EDTA-free, Roche) and phosphatase inhibitor (Sigma) cocktails for 30 min at 4 °C.

Protein A or Protein G-sepharose beads (Sigma) were blocked in 1 mg/mL BSA for 2h at 4°C. Cell lysates were precleared with blocked beads for 1h at 4°C. Upon removal of the beads, we incubated the precleared lysates with the specific antibody for 1h at 4°C, and then we added fresh blocked beads for 1h at 4°C. Bound beads were washed 3 times for 10 min in lysis buffer (no BSA, no protease and phosphatase inhibitor) and beads were boiled in 2x Laemmli Buffer for 10 minutes at 98°C. To test whether interactions were RNA-dependent, cell lysates were incubated with a mixture of RNase A and RNase If (New England Biolabs) at the indicated concentration for 5 min at 37°C.

Immunoblots

Cells were trypsinized, counted, and directly boiled for 5 min at 95°C in 2x Laemmli Buffer (1x10⁶ cells / 100 µL). Samples were resolved by 10% polyacrylamide gel electrophoresis in 1x Tris-glycine, and wet-transferred onto a PVDF FL membrane (Millipore). The membranes were hydrated in methanol and incubated in blocking solution (Phosphate-buffered saline with 0.1% Tween 20 (PBS-T), containing 5% (w/v) fat-free milk) for 1h at room temperature. Then, membranes were incubated with a primary antibody diluted in blocking solution overnight at 4°C or for 1h at room temperature. After 4 washes in PBS-T, membranes were incubated for 45 min at room temperature with a secondary antibody coupled to infrared fluorochromes diluted in

Odyssey blocking buffer (LI-COR Biosciences), followed by 4 additional washes in PBS-T and one last wash in PBS for 5 min. Fluorescent signal was detected and quantified with a dual-channel Odyssey infrared imaging system (LI-COR Biosciences). When necessary, membranes were stripped for 30 min in stripping buffer (ST010, Euromedex), washed 5 times in distilled water, reblocked in blocking buffer and immunoblotting was repeated as described above.

Luciferase assays

HeLa cells were seeded in 12-well plates at 200,000 cells/well. The next day, cells were transfected with 3 μ L of Lipofectamine 2000 diluted in 100 μ L of Opti-MEM (Life Technologies) premixed with 0.5 μ g of each of the following plasmids diluted in 100 μ L of Opti-MEM total: (i) pGL3prom or LacO-Fluc; (ii) pYX013 (RLuc); and (iii) a plasmid expressing one of the mCherry-LacI fusion proteins (**Fig. 3c**). Four hours post-transfection, medium was replaced with fresh growth medium. For ligand-independent nuclear receptors, cells were lysed 24h post-transfection. For ligand-dependent nuclear receptors, cells were seeded and grown in media containing charcoal-treated FBS, transfected with plasmid DNA, treated with the indicated hormone 24h post-transfection, and lysed after 24h of hormonal treatment. RLuc and FLuc activities in cell extracts were measured with the Dual-Luciferase reporter assay system (Promega) following manufacturer's instructions using a Centro XS3 LB960 plate reader (Berthold technologies) and the Microwin 2000 software.

siRNA-mediated depletion

Pre-designed siRNAs were obtained from Sigma-Aldrich and include two control siRNA (SIRNA universal negative control #1- SIC001, SIRNA universal negative control #2- SIC002) and three siRNA directed against $ERR\alpha$ (MISSION PDSIRNA5D SASI_Hs01_00193458, MISSION PDSIRNA5D SASI_Hs01_00193459, MISSION PDSIRNA5D SASI_Hs01_00193460). 2×10^5 U2OS cells were plated in six-well plates. The next day, cells were transfected with 5 nM of each siRNA using the Mission siRNA transfection reagent (Sigma-Aldrich). Twenty-four hours after siRNA treatment, cells were transfected with L1 retrotransposition assay plasmid (JJ101/L1.3) or the pCDNA6 control, using 2 μ g of plasmid DNA and 6 μ L of jetPEI (Ozyme) diluted in 150 mM NaCl. Five hours post-transfection, medium was replaced with fresh growth medium.

Two days after DNA transfection, cells were collected and counted. A fraction was used for immunoblot analysis, and another for colony assay formation. Notably, we consistently observed increased colony formation in the retrotransposition assay with the control siRNA SIC001, consistent with previous observations reporting unexpected effects of some commercial non-targeting control siRNAs on retrotransposition assay plasmids ⁸⁴. None of the other targeting or non-targeting siRNAs had similar effects. Since the sequences of these controls are not provided by the manufacturers, we could not explore further this phenomenon.

Retrotransposition assays in cultured cells

Blasticidin-based assay. U2OS cells were seeded in six-well plates at 2×10^5 cells per well. The next day, cells were transfected with 2 μ g of plasmid DNA and 6 μ L of jetPEI (Ozyme) diluted in 150 mM NaCl. Five hours post-transfection, medium was replaced with fresh growth medium. Two or four days post-transfection, cells were trypsinized, counted and replated in medium supplemented with Blasticidine (Gibco, Life Technologies) at 10 μ g/ml final concentration. After 10 days of selection, colonies were fixed and stained with a solution containing 30% methanol (v/v), 10% acetic acid (v/v) and 0.2% Coomassie blue (m/v). The plates were scanned using an Odyssey imaging system (LI-COR Biosciences). The images were analyzed using the Colony Area Image J plugin ⁸⁵ using the area percentage.

Fluorescent 2-hybrid assay

U2OS-263-ORF2Cter-GFP cells were seeded in 24-well plates on cover slips at 100,000 cells/well. The following day, cells were transfected with 0.5 μ g of plasmid DNA using 1 μ L Lipofectamine 2000 diluted in 100 μ L Opti-MEM (Life Technologies), and medium was replaced with fresh medium after 4h. For ligand-independent nuclear receptors, 24h post-transfection, cells were washed twice with PBS, fixed with 4% formaldehyde diluted in PBS, for 10 minutes, washed 3 times for 10 min with PBS, and, finally, mounted with DAPI-containing Vectashield (Vector Laboratories). For ligand-dependent nuclear receptors, cells were seeded and grown in media containing charcoal-treated FBS, and transfected as described above. However, 24h post-transfection, cells were first treated with the indicated hormone for 24h, before applying the same protocol of fixation as described above. Cells from F2H experiments were imaged using a Zeiss Epi-

fluorescence Cytogenetic Microscope, using 63x oil immersion objective and Isis acquisition software (Metasystems).

The pictures were analyzed with CellProfiler (available at <http://cellprofiler.org/>). The pipeline to identify red and green spots in the nuclei and to measure their intensities modules is illustrated in **Supplementary Fig. S2**, and is annotated in detail in **Supplementary Table S3**. The script itself is provided as Supplementary file (pipeline_F2H-V7.cppipe).

Droplet digital PCR (ddPCR)

Genomic DNA was extracted using QiaAmp DNA Blood mini kit (Qiagen), following manufacturer's instructions and was quantified by spectrophotometry with a Nanodrop 2000 (ThermoFisher Scientific). PCR mixes were assembled in 20 μ L reactions containing 50 ng of genomic DNA in 1x ddPCR Supermix for Probes no dUTP (Bio-Rad) supplemented with 0.9 μ M of each LOU2258 and LOU2259 primers and 250 nM of the CALLI012 FAM-labelled probe to specifically amplifies and detects the spliced Blast^R gene exon-exon junction (**Supplementary Fig. S4 and Table S2**), as well as 1x premixed RNase P-TAMRA assay as an internal reference (Life Technologies). Then, 70 μ L of droplet generation oil for probes (Bio-Rad) were added to each reaction and droplets were generated according to QX200 Droplet generator's instruction manual (Bio-Rad). After droplet generation, 40 μ L of each reaction were transferred into a clean 96-well plate. Amplifications were performed with a C1000 Touch Thermal Cycler (Bio-Rad) under the following cycling conditions: 95°C for 10 min; [94°C for 30 s, 60°C for 1 min] x 39 cycles; 98°C for 10 min; 4°C until analysis. Ramping between each step was 2°C/s. Finally, droplets were analyzed using a QX200 Droplet Reader and the QuantaSoft software (Bio-Rad).

Supplementary information

Supplementary information includes 4 figures and 3 tables and can be found with this article online.

Acknowledgements

We thank IRCAN microscopy, cytometry and genomics facilities. We are grateful to J.-M. Vanacker (ENS-Lyon, Lyon, France) for providing ERR α cDNA and for useful

advices in the early steps of the project, to J. V. Moran (Univ. Michigan, MI, USA), J. D. Boeke (New-York Univ., NY, USA), and W. An (South Dakota State Univ., SD, USA) for sharing L1 plasmids, to D. L. Spector (CSHL, NY, USA) for providing the U2OS 2-6-3 cell line, and to H. Leonhardt (Ludwig Maximilians Univ., Munich, Germany) for sharing F2H plasmids. We also thank J.L. Garcia-Perez (Univ. of Edinburgh, UK) for L1 constructs and useful discussions. This work was supported by grants to GC from the European Research Council (ERC-2010-StG 243312, Retrogenomics), the French Government (National Research Agency, ANR) through the "Investments for the Future" program (LABEX SIGNALIFE, ANR-11-LABX-0028-01), and CNRS (GDR 3546). CM was supported by a PhD fellowship from the French Ministry of Research and from the Association pour la Recherche contre le Cancer.

Authors contributions

RG, JGP and CM designed, performed and analyzed the experiments; MM, ACV, SK contributed to preliminary experiments or technical development; POV performed the Y2H screen; GC supervised the project, provided funding, designed and analyzed experiments and wrote the F2H analysis pipeline. GC and RG drafted the manuscript and all authors revised it.

Competing financial interests

The authors declare no competing financial interests.

References

1. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
2. de Koning, A. P. J., Gu, W., Castoe, T. A., Batzer, M. A. & Pollock, D. D. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet.* **7**, e1002384 (2011).
3. Mills, R. E. *et al.* Recently mobilized transposons in the human and chimpanzee genomes. *Am J Hum Genet* **78**, 671–679 (2006).
4. Mills, R. E., Bennett, E. A., Iskow, R. C. & Devine, S. E. Which transposable elements are active in the human genome? *Trends Genet* **23**, 183–191 (2007).
5. Sudmant, P. H. *et al.* An integrated map of structural variation in 2,504 human genomes. *Nature* **526**, 75–81 (2015).
6. Philippe, C. *et al.* Activation of individual L1 retrotransposon instances is restricted to cell-type dependent permissive loci. *Elife* **5**, 166 (2016).
7. Scott, E. C. *et al.* A hot L1 retrotransposon evades somatic repression and initiates human colorectal cancer. *Genome Res* **26**, 745–755 (2016).
8. Deininger, P. *et al.* A comprehensive approach to expression of L1 loci. *Nucleic*

- Acids Res* **45**, e31–e31 (2016).
9. Brouha, B. *et al.* Hot L1s account for the bulk of retrotransposition in the human population. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5280–5285 (2003).
 10. Beck, C. R. *et al.* LINE-1 retrotransposition activity in human genomes. *Cell* **141**, 1159–1170 (2010).
 11. Richardson, S. R. *et al.* The Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes. *Microbiol Spectr* **3**, MDNA3–0061–2014 (2015).
 12. Kidd, J. M. *et al.* A human genome structural variation sequencing resource reveals insights into mutational mechanisms. *Cell* **143**, 837–847 (2010).
 13. Mir, A. A., Philippe, C. & Cristofari, G. euL1db: the European database of L1HS retrotransposon insertions in humans. *Nucleic Acids Research* **43**, D43–7 (2015).
 14. Hancks, D. C. & Kazazian, H. H. Roles for retrotransposon insertions in human disease. *Mobile DNA* **7**, 9 (2016).
 15. Richardson, S. R. *et al.* Heritable L1 retrotransposition in the mouse primordial germline and early embryo. *Genome Res* gr.219022.116 (2017). doi:10.1101/gr.219022.116
 16. Garcia-Perez, J. L., Widmann, T. J. & Adams, I. R. The impact of transposable elements on mammalian development. *Development* **143**, 4101–4114 (2016).
 17. Richardson, S. R., Morell, S. & Faulkner, G. J. L1 Retrotransposons and Somatic Mosaicism in the Brain. *Annu Rev Genet* **48**, 1–27 (2014).
 18. Goodier, J. L. Retrotransposition in tumors and brains. *Mob DNA* **5**, 11 (2014).
 19. Burns, K. H. Transposable elements in cancer. *Nat Rev Cancer* **17**, 415–424 (2017).
 20. Macia, A. *et al.* Engineered LINE-1 retrotransposition in nondividing human neurons. *Genome Res* **27**, 335–348 (2017).
 21. Alisch, R. S., Garcia-Perez, J. L., Muotri, A. R., Gage, F. H. & Moran, J. V. Unconventional translation of mammalian LINE-1 retrotransposons. *Genes Dev.* **20**, 210–224 (2006).
 22. Moran, J. V. *et al.* High frequency retrotransposition in cultured mammalian cells. *Cell* **87**, 917–927 (1996).
 23. Feng, Q., Moran, J. V., Kazazian, H. H. & Boeke, J. D. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87**, 905–916 (1996).
 24. Martin, S. L. & Bushman, F. D. Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. *Molecular and Cellular Biology* **21**, 467–475 (2001).
 25. Khazina, E. *et al.* Trimeric structure and flexibility of the L1ORF1 protein in human L1 retrotransposition. *Nat. Struct. Mol. Biol.* **18**, 1006–1014 (2011).
 26. Mathias, S. L., Scott, A. F., Kazazian, H. H., Boeke, J. D. & Gabriel, A. Reverse transcriptase encoded by a human transposable element. *Science* **254**, 1808–1810 (1991).
 27. Weichenrieder, O., Repanas, K. & Perrakis, A. Crystal structure of the targeting endonuclease of the human LINE-1 retrotransposon. *Structure* **12**, 975–986 (2004).
 28. Esnault, C., Maestre, J. & Heidmann, T. Human LINE retrotransposons generate processed pseudogenes. *Nat Genet* **24**, 363–367 (2000).
 29. Wei, W. *et al.* Human L1 retrotransposition: cis preference versus trans complementation. *Mol. Cell. Biol.* **21**, 1429–1439 (2001).
 30. Kulpa, D. A. & Moran, J. V. Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. *Nat. Struct. Mol. Biol.* **13**, 655–660 (2006).

31. Kulpa, D. A. & Moran, J. V. Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition. *Hum. Mol. Genet.* **14**, 3237–3248 (2005).
32. Doucet, A. J. *et al.* Characterization of LINE-1 ribonucleoprotein particles. *PLoS Genet* **6**, e1001150 (2010).
33. Dai, L., Taylor, M. S., O'Donnell, K. A. & Boeke, J. D. Poly(A) binding protein C1 is essential for efficient L1 retrotransposition and affects L1 RNP formation. *Molecular and Cellular Biology* **32**, 4323–4336 (2012).
34. Goodier, J. L., Cheung, L. E. & Kazazian, H. H. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. *Nucleic Acids Research* **41**, 7401–7419 (2013).
35. Taylor, M. S. *et al.* Affinity Proteomics Reveals Human Host Factors Implicated in Discrete Stages of LINE-1 Retrotransposition. *Cell* **155**, 1034–1048 (2013).
36. Doucet, A. J., Wilusz, J. E., Miyoshi, T., Liu, Y. & Moran, J. V. A 3' Poly(A) Tract Is Required for LINE-1 Retrotransposition. *Mol. Cell* **60**, 728–741 (2015).
37. Luan, D. D., Korman, M. H., Jakubczak, J. L. & Eickbush, T. H. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* **72**, 595–605 (1993).
38. Cost, G. J., Feng, Q., Jacquier, A. & Boeke, J. D. Human L1 element target-primed reverse transcription in vitro. *EMBO J.* **21**, 5899–5910 (2002).
39. Monot, C. E. M. *et al.* The specificity and flexibility of L1 reverse transcription priming at imperfect T-tracts. *PLoS Genet.* **9**, e1003499 (2013).
40. Viollet, S. E. B., Monot, C. E. M. & Cristofari, G. E. L. L1 retrotransposition: The snap-velcro model and its consequences. *Mob Genet Elements* **4**, e28907 (2014).
41. Goodier, J. L. Restricting retrotransposons: a review. *Mobile DNA* **7**, 16 (2016).
42. Pizarro, J. G. & Cristofari, G. Post-Transcriptional Control of LINE-1 Retrotransposition by Cellular Host Factors in Somatic Cells. *Front Cell Dev Biol* **4**, 210–9 (2016).
43. Goodier, J. L., Zhang, L., Vetter, M. R. & Kazazian, H. H. LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex. *Molecular and Cellular Biology* **27**, 6469–6483 (2007).
44. Richardson, S. R., Narvaiza, I. N. I., Planegger, R. A., Weitzman, M. D. & Moran, J. V. APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition. *Elife* **3**, e02008 (2014).
45. Gasior, S. L., Roy-Engel, A. M. & Deininger, P. L. ERCC1/XPF limits L1 retrotransposition. *DNA Repair (Amst)* **7**, 983–989 (2008).
46. Servant, G. *et al.* The Nucleotide Excision Repair Pathway Limits L1 Retrotransposition. *Genetics* **205**, 139–153 (2017).
47. Coufal, N. G. *et al.* Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20382–20387 (2011).
48. Suzuki, J. *et al.* Genetic evidence that the non-homologous end-joining repair pathway is involved in LINE retrotransposition. *PLoS Genet.* **5**, e1000461 (2009).
49. Dombroski, B. A., Scott, A. F. & Kazazian, H. H. Two additional potential retrotransposons isolated from a human L1 subfamily that contains an active retrotransposable element. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6513–6517 (1993).
50. Han, J. S. & Boeke, J. D. A highly active synthetic mammalian retrotransposon. *Nature* **429**, 314–318 (2004).

51. Levin, E. R. & Hammes, S. R. Nuclear receptors outside the nucleus: extranuclear signalling by steroid receptors. *Nat. Rev. Mol. Cell Biol.* **17**, 783–797 (2016).
52. Huss, J. M., Garbacz, W. G. & Xie, W. Constitutive activities of estrogen-related receptors: Transcriptional regulation of metabolism by the ERR pathways in health and disease. *Biochim. Biophys. Acta* **1852**, 1912–1927 (2015).
53. Evans, R. M. & Mangelsdorf, D. J. Nuclear Receptors, RXR, and the Big Bang. *Cell* **157**, 255–266 (2014).
54. Tremblay, A. M., Wilson, B. J., Yang, X.-J. & Giguère, V. Phosphorylation-Dependent Sumoylation Regulates Estrogen-Related Receptor- α and - γ Transcriptional Activity through a Synergy Control Motif. *Molecular Endocrinology* **22**, 570–584 (2008).
55. Kallen, J. *et al.* Evidence for ligand-independent transcriptional activation of the human estrogen-related receptor alpha (ERRalpha): crystal structure of ERRalpha ligand binding domain in complex with peroxisome proliferator-activated receptor coactivator-1alpha. *J. Biol. Chem.* **279**, 49330–49337 (2004).
56. Piskareva, O., Ernst, C., Higgins, N. & Schmatchenko, V. The carboxy-terminal segment of the human LINE-1 ORF2 protein is involved in RNA binding. *FEBS Open Bio* **3**, 433–437 (2013).
57. Goodier, J. L., Cheung, L. E. & Kazazian, H. H. MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. *PLoS Genet* **8**, e1002941 (2012).
58. Xie, Y., Rosser, J. M., Thompson, T. L., Boeke, J. D. & An, W. Characterization of L1 retrotransposition with high-throughput dual-luciferase assays. *Nucleic Acids Res* **39**, e16–e16 (2011).
59. Zolghadr, K. *et al.* A fluorescent two-hybrid assay for direct visualization of protein interactions in living cells. *Mol. Cell Proteomics* **7**, 2279–2287 (2008).
60. Janicki, S. M. *et al.* From silencing to gene expression: real-time analysis in single cells. *Cell* **116**, 683–698 (2004).
61. Lin, C. *et al.* Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell* **139**, 1069–1083 (2009).
62. Sultana, T., Zamborlini, A., Cristofari, G. & Lesage, P. Integration site selection by retroviruses and transposable elements in eukaryotes. *Nat Rev Genet* **18**, 292–308 (2017).
63. Sharma, A. *et al.* BET proteins promote efficient murine leukemia virus integration at transcription start sites. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 12036–12041 (2013).
64. De Rijck, J. *et al.* The BET family of proteins targets moloney murine leukemia virus integration near transcription start sites. *Cell Rep* **5**, 886–894 (2013).
65. Gupta, S. S. *et al.* Bromo- and extraterminal domain chromatin regulators serve as cofactors for murine leukemia virus integration. *J Virol* **87**, 12721–12736 (2013).
66. Ciuffi, A. *et al.* A role for LEDGF/p75 in targeting HIV DNA integration. *Nat. Med.* **11**, 1287–1289 (2005).
67. Singh, P. K. *et al.* LEDGF/p75 interacts with mRNA splicing factors and targets HIV-1 integration to highly spliced genes. *Genes Dev.* **29**, 2287–2297 (2015).
68. Kirchner, J., Connolly, C. M. & Sandmeyer, S. B. Requirement of RNA polymerase III transcription factors for in vitro position-specific integration of a retroviruslike element. *Science* **267**, 1488–1491 (1995).
69. Bridier-Nahmias, A. *et al.* Retrotransposons. An RNA polymerase III subunit determines sites of retrotransposon integration. *Science* **348**, 585–588 (2015).
70. Zhu, Y., Dai, J., Fuerst, P. G. & Voytas, D. F. Controlling integration specificity of a yeast retrotransposon. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5891–5895 (2003).

71. Xie, W. *et al.* Targeting of the yeast Ty5 retrotransposon to silent chromatin is mediated by interactions between integrase and Sir4p. *Molecular and Cellular Biology* **21**, 6606–6614 (2001).
72. Dai, J., Xie, W., Brady, T. L., Gao, J. & Voytas, D. F. Phosphorylation regulates integration of the yeast Ty5 retrotransposon into heterochromatin. *Mol. Cell* **27**, 289–299 (2007).
73. Siol, O. *et al.* Role of RNA polymerase III transcription factors in the selection of integration sites by the dictyostelium non-long terminal repeat retrotransposon TRE5-A. *Molecular and Cellular Biology* **26**, 8242–8251 (2006).
74. Chung, T., Siol, O., Dingermann, T. & Winckler, T. Protein interactions involved in tRNA gene-specific integration of Dictyostelium discoideum non-long terminal repeat retrotransposon TRE5-A. *Molecular and Cellular Biology* **27**, 8492–8501 (2007).
75. Gilbert, N., Lutz-Prigge, S. & Moran, J. V. Genomic deletions created upon LINE-1 retrotransposition. *Cell* **110**, 315–325 (2002).
76. Symer, D. E. *et al.* Human l1 retrotransposition is associated with genetic instability in vivo. *Cell* **110**, 327–338 (2002).
77. Jurka, J. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1872–1877 (1997).
78. Morales, J. F., Snow, E. T. & Murnane, J. P. Environmental factors affecting transcription of the human L1 retrotransposon. I. Steroid hormone-like agents. *Mutagenesis* **17**, 193–200 (2002).
79. Terasaki, N. *et al.* In vitro screening for compounds that enhance human l1 mobilization. *PLoS One* **8**, e74629 (2013).
80. Vidalain, P.-O. *et al.* A field-proven yeast two-hybrid protocol used to identify coronavirus-host protein-protein interactions. *Methods Mol. Biol.* **1282**, 213–229 (2015).
81. Zhang, Y., Werling, U. & Edelmann, W. SLiCE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Research* **40**, e55–e55 (2012).
82. Sassaman, D. M. *et al.* Many human L1 elements are capable of retrotransposition. *Nat Genet* **16**, 37–43 (1997).
83. Kopera, H. C., Moldovan, J. B., Morrish, T. A., Garcia-Perez, J. L. & Moran, J. V. Similarities between long interspersed element-1 (LINE-1) reverse transcriptase and telomerase. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20345–20350 (2011).
84. Cook, P. R. & Tabor, G. T. Deciphering fact from artifact when using reporter assays to investigate the roles of host factors on L1 retrotransposition. *Mobile DNA* **7**, 23 (2016).
85. Guzmán, C., Bagga, M., Kaur, A., Westermarck, J. & Abankwa, D. ColonyArea: an ImageJ plugin to automatically quantify colony formation in clonogenic assays. *PLoS One* **9**, e92444 (2014).

Figures and figure legends

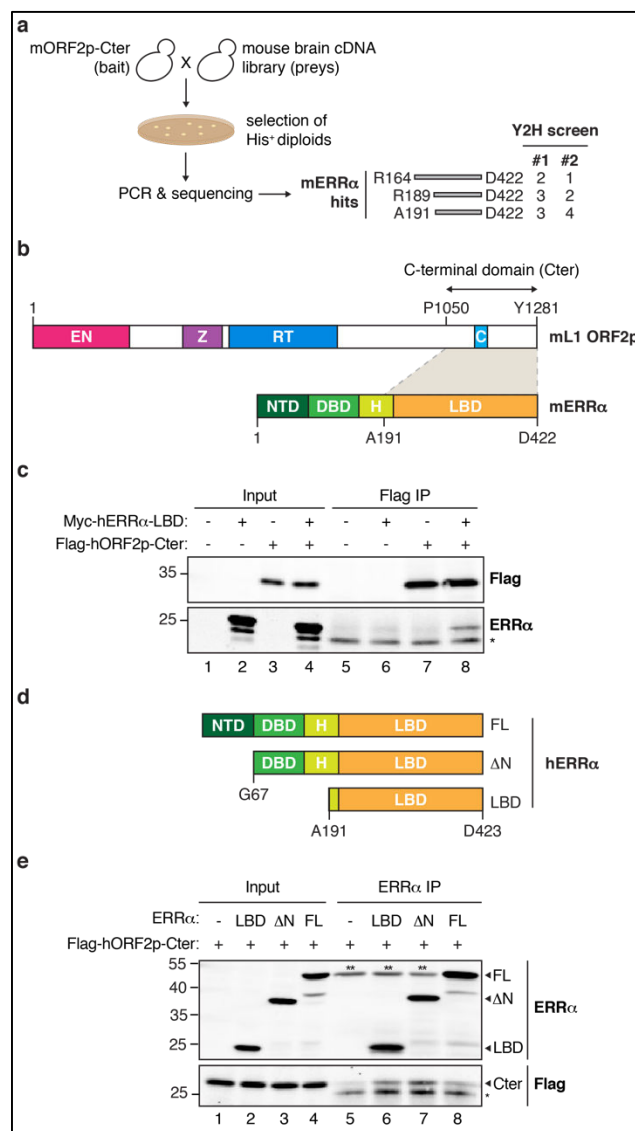


Figure 1. ERRα is a cellular partner of L1 ORF2p.

(a) Yeast 2-hybrid (Y2H) screens to identify cellular partners of L1 retrotransposition machinery. Several regions of human and mouse ORF2p (mORF2p) were used as baits in Y2H screens (**Supplementary Table 1**). Mouse ERRα (mERRα) was the most robust hit, independently identified multiple times as interacting with mORF2p-Cter domain (mORF2p-Cter). The table indicates the number of recovered clones in each of the two independent screens performed with mORF2p-Cter and for each distinct mERRα overlapping fragment found. (b) Schematic representation of mORF2p and mERRα domain structures. The minimal interacting regions identified by Y2H are indicated by a grey area. ORF2p: EN, endonuclease; Z, Z-domain; RT, reverse transcriptase; C, Cystein-rich domain. ERRα: NTD, N-terminal domain; DBD, DNA-binding domain; H, Hinge region; LBD, Ligand-binding domain. (c) Co-immunoprecipitation (coIP) of Myc-tagged human ERRα LBD domain (Myc-hERRα-LBD) with Flag-tagged ORF2p-C-terminal domain (Flag-ORF2p-Cter) in HEK293T cells. Immunoprecipitation (IP) was performed with a Flag antibody. Left, molecular weight (kDa) markers. Right (bold), antibodies used for immunoblotting. IP, immunoprecipitation. The band marked with a star (*) corresponds to the IgG light chain. (d) Structure of deletion

variants of hERR α used in this study. FL, full length. Note that hERR α is 98% identical to mERR α and is 1 amino-acid longer. (e) CoIP of Flag-tagged ORF2p-C-terminal domain (Flag-ORF2p-Cter) with various deletion variants of untagged hERR α in HEK293T cells. Immunoprecipitation (IP) was performed with an antibody directed against hERR α . Left, molecular weight (kDa) markers. Right (bold), antibodies used for immunoblotting. IP, immunoprecipitation. The bands marked with a star (*) or two stars (**) correspond to the IgG light and heavy chains, respectively.

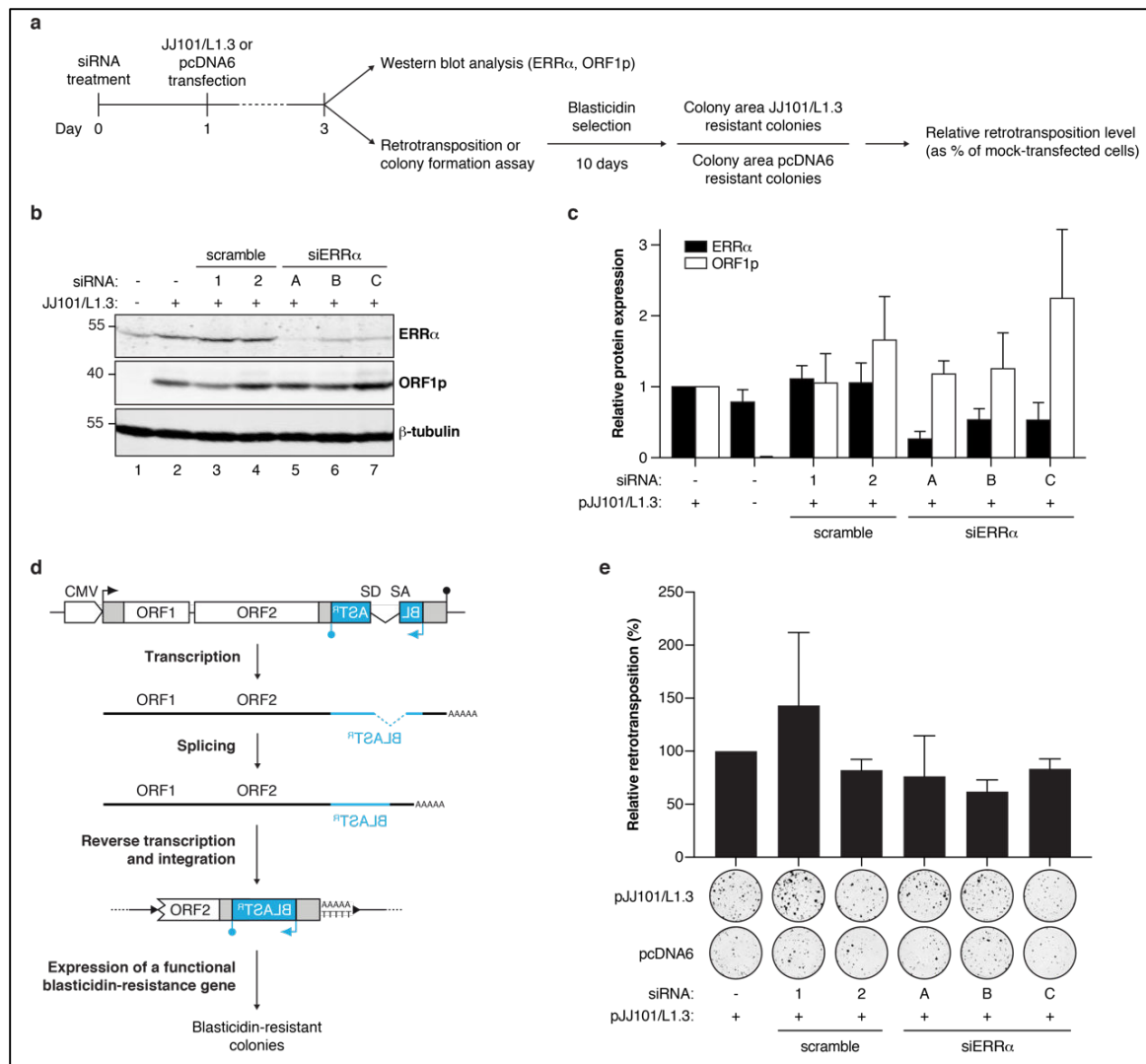


Figure 2. ERRα knock-down does not affect retrotransposition efficiency.

(a) Flow chart summarizing the experimental procedure. Twenty-four hours after siRNA treatment, U2OS cells were transfected with a retrotransposition-competent L1 element carrying a blasticidin-resistant marker, which becomes functional only after retrotransposition (JJ101/L1.3, see panel d). Two days after L1 transfection, a fraction of the cells was lysed for immunoblot analysis and another fraction was re-plated and selected for 10 days with blasticidin. As a control for cytotoxicity and for the capacity to form blasticidin-resistant colonies upon siRNA treatment, we replaced, in parallel transfections, the L1-containing plasmid by pcDNA6, a plasmid containing an identical blasticidin selection cassette as in JJ101/L1.3 but without the intron. (b-c) siRNA-mediated depletion of ERRα does not impact ORF1p protein levels. Representative immunoblot (b) and quantification (c) showing that several siRNA against ERRα (named A, B, C) efficiently reduce endogenous ERRα levels, as compared to two scramble siRNA (named 1,2). Under these experimental conditions, L1 expression measured through ORF1p protein levels is unchanged or only slightly increased. Expression levels were normalized to β-tubulin (b, bottom panel) and then to the mock-treated sample (b, lane 2). (b) Left, molecular weight (kDa) markers; right (bold), antibodies used for immunoblotting. (c) bars represent the mean ± s.d. of 3 biological replicates. (d) Principle of the retrotransposition cellular assay. The blasticidin-resistance (Blast^R)-based retrotransposition marker inserted in the 3' UTR of L1 becomes functional only upon transcription, splicing, reverse transcription and integration into the genome. In contrast, the plasmid-borne version does not express a

functional gene. (e) Retrotransposition assays in siRNA-treated U2OS cells. For each siRNA condition, the ratio of colonies obtained with JJ101/L1.3 and pcDNA6 was normalized by its matching mock-transfected control (arbitrary defined as 100%). Bars represent the mean \pm s.d. (n=3). Variations between siERR α -treated- and mock-transfected samples are not statistically significant (ratio paired t-test). Under each bar of the graph a picture of a representative well with stained colonies is displayed for illustrative purposes.

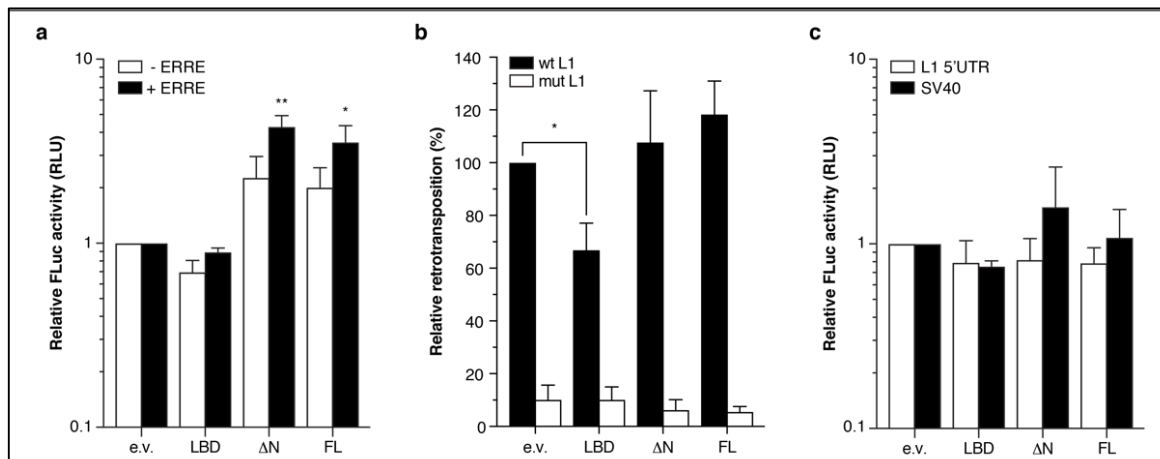


Figure 3. ERRα LBD acts as a dominant-negative on L1 retrotransposition.

(a) ERRE-transcriptional activity of ERRα variant constructs. Relative luciferase activity (RLU) was measured by co-transfecting in HeLa cells a Firefly Luciferase (FLuc) cassette under the control of a minimal SV40 promoter preceded or not by 3 estrogen-related receptor responsive element (ERRE) sequences, a Renilla Luciferase (RLuc) cassette expressed from a constitutive TK promoter as a transfection control, and the ERRα expression constructs. RLU was calculated by dividing FLuc values by RLuc values, and then normalized by its matching empty vector (e.v.) condition. Bars represent mean \pm s.d. (n=3, *, $p<0.05$; **, $p<0.01$; ratio paired t-test). White bars, FLuc construct with no ERRE; black bars, ERRE-containing FLuc construct. (b) Retrotransposition assays in HeLa cells overexpressing ERRα variants. For each condition, the FLuc/RLuc ratio was normalized by the empty vector (e.v.) control (arbitrary defined as 100%). Bars represent the mean \pm s.d. (n=3). Only the LBD condition shows a significant difference with the control (* $p<0.05$, ratio paired t-test). (c) Transcriptional activity of ERRα variant constructs on the L1 and SV40 promoters. Legend is as in (a). None of constructs shows a significant difference with the control.

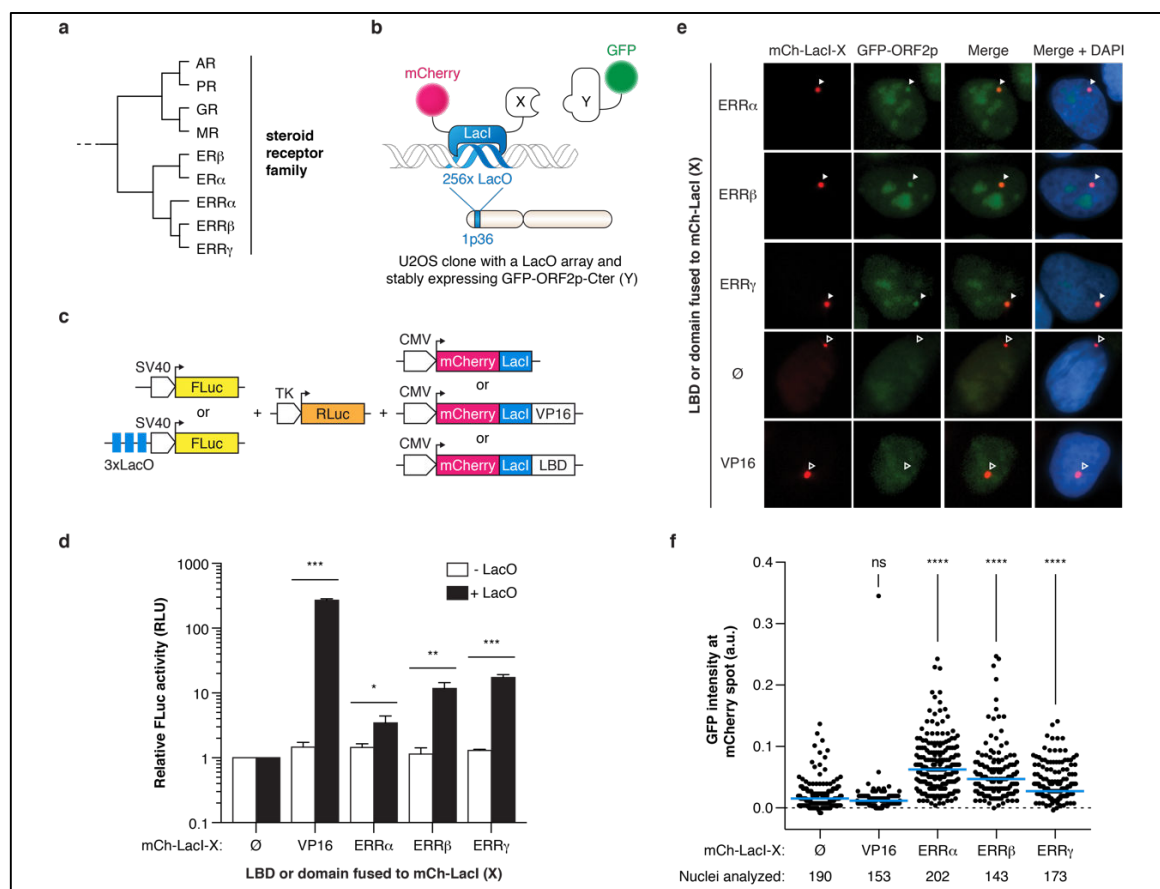


Figure 4. ORF2p interacts with all estrogen-related receptor (ERRs) paralogs.

(a) Position of hERR α among other nuclear receptors. Humans possess 48 nuclear receptor family members. For the sake of simplicity, only the steroid receptor superfamily is represented in this tree. (b) Principle of the fluorescent 2-hybrid (F2H) approach. An U2OS cell line containing ~200 copies of a plasmid with 256 *Lac* operator (LacO) tandem repeats in the chromosomal region 1p36 was transduced with a retroviral vector expressing a green fluorescent protein (GFP) N-terminally fused to hORF2p C-terminal domain (GFP-ORF2p-Cter). These cells are then transfected with constructs expressing mCherry-LacI fused or not to the LBD domain of each ERR analog or to VP16 activating domain as an additional negative control (shown in panel (c), constructs on the right). mCherry-LacI fusion proteins (mCherry-LacI-X) bind to the LacO array and form an intense fluorescent red spot in the nucleus. If the protein of interest (Y) interacts with X, the GFP-Y fusion protein also accumulates at the LacO array and forms a green spot co-localizing with the red spot. In the absence of interaction, the GFP-Y protein remains diffuse in the nucleus. Both constructs contain nuclear localization signals. (c-d) Transcriptional activity of mCherry-LacI fusion proteins. Relative luciferase activity (RLU) was measured by co-transfecting in HeLa cells a Firefly Luciferase (FLuc) cassette under the control of a minimal SV40 promoter preceded or not by 3 LacO repeats, a Renilla Luciferase (RLuc) cassette expressed from a constitutive TK promoter as a transfection control, and the mCherry-LacI fusion expression cassettes. RLU was calculated by dividing FLuc values by RLuc values, and then normalized by its matching mCherry-LacI condition. Bars represent mean \pm s.d. (n=3, *, p<0.05; **, p<0.01; ***, p<0.001, multiple t-test). White bars, FLuc construct with no LacO; black bars, LacO-containing FLuc construct. Note that only the LBD of each ERR protein was fused to mCherry-LacI. The condition labelled (-) corresponds to mCherry-LacI without any additional domain. (e-f) hORF2p can interact with ERR α , ERR β and ERR γ . Representative fluorescent microscopy images (e) and quantification (f) of F2H assays performed with GFP-ORF2p-Cter and varying mCherry-LacI fusion constructs. (e) All three ERR paralogs

Results

show a green spot co-localizing with the red spot (plain arrowheads) in contrast to the VP16 activating domain or to mCherry-LacI construct alone used as negative control (empty arrowheads). Nuclei were stained with DAPI (Blue, right panels). **(f)** Quantification of F2H assays. The median green fluorescence intensity at the red spot was calculated and the local background was subtracted. Local background was defined as the median green fluorescence intensity in a donut-shaped region around the red spot (**Supplementary Figure 2**). Ø indicates no LBD (i.e. mCherry-LacI alone). In the graph, dots correspond to individual values of 3 independent experiments pooled together, and blue horizontal lines to median values (each condition was compared to the mCherry-LacI-Ø construct; ns, not significant; ****, $p < 0.0001$; one-way ANOVA with Tukey's multiple test correction).

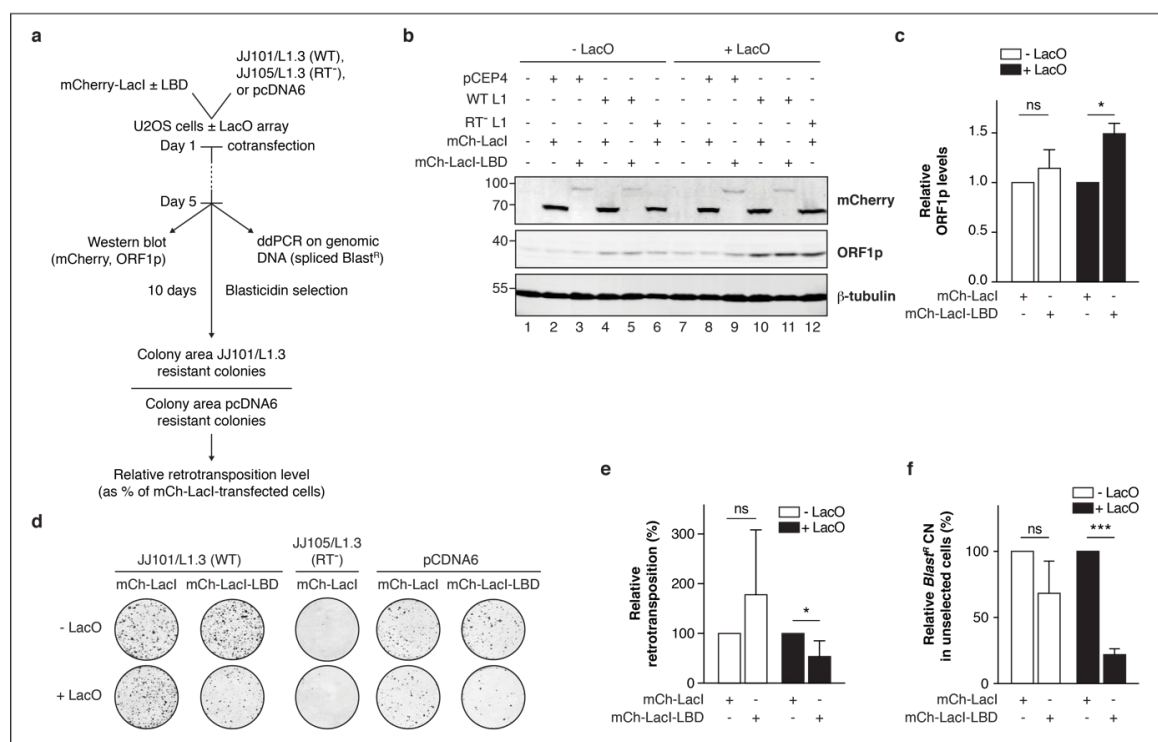


Figure 5. Tethering the LBD of ERRα to a LacO array inhibits L1 retrotransposition.

(a) Flow chart summarizing the experimental procedure. In each experiment, two cell lines are used in parallel: the U2OS 2.6.3 clone containing the LacO array (+ LacO) and the parental U2OS cell line without the LacO array (- LacO). Cells are co-transfected with JJ101/L1.3 (or its RT-minus derivative JJ105/L1.3) and either mCherry-LacI or mCherry-LacI fused to ERRα LBD domain (mCherry-LacI/+LBD). Four days after transfection, a fraction of the cells was subsequently analyzed by immunoblotting and droplet digital PCR (ddPCR), while another fraction was re-plated and selected with blasticidin for 10 days. As a control for cytotoxicity and for the capacity to form blasticidin-resistant colonies in the different experimental conditions, we transfected cells in parallel with the blasticidin-resistance-containing plasmid pCDNA6. (b-c) Expression of mCherry-LacI-LBD fusion proteins does not affect L1 expression measured by ORF1p immunoblotting. (b) Representative immunoblot. Left, molecular weight (kDa) markers; right (bold), antibodies used for immunoblotting. β-tubulin is used as a loading control. (c) Quantification of ORF1p expression, normalized by β-tubulin levels and by the mCherry-LacI condition in each cell line. Note that L1 expression was always slightly higher in U2OS 2.6.3 cells (+ LacO), but within a particular cell line, ORF1p levels were equivalent or slightly increased when comparing mCherry-LacI vs mCherry-LacI-LBD (lanes 4 vs 5, and lanes 10 vs 11). Bars represent the mean ± s.d. (n=3; ns, not significant, *, p<0.05, ratio paired t-test). (d-e) Retrotransposition assays upon ERRα-LBD tethering to a LacO array in U2OS cells. (d) Representative results of retrotransposition assay. Blasticidin-resistant foci were stained and pictures of wells corresponding to each construct used for co-transfection are displayed. Cells transfected with an L1 RT mutant (pJJ105/L1.3) serve as negative control. (e) Quantification of cellular retrotransposition assays. For each experimental condition, the ratio of colonies obtained with JJ101/L1.3 and pcDNA6 was normalized by its matching mCherry-LacI-transfected control (arbitrary defined as 100%). Bars represent the mean ± s.d. (n=5; p<0.05, ratio paired t-test). (f) Tethering of ERRα-LBD to the LacO array inhibits L1 RNA reverse transcription. Multiplex Taqman ddPCR assay to quantify spliced Blast^R DNA in unselected U2OS cells containing or not a LacO array and transfected with L1 and either mCherry-LacI or mCherry-LacI-LBD. Values were normalized to

Results

the matched mCherry-LacI control of each cell line. Bars indicate the mean \pm s.d. (n=4; ***, p<0.001; ns, not significant; ratio paired t test).

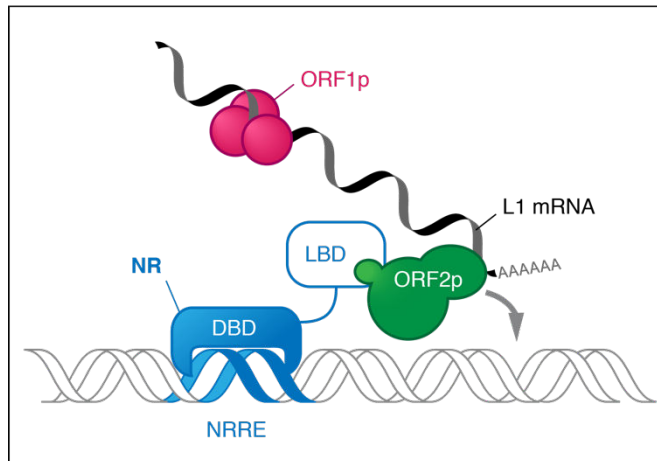
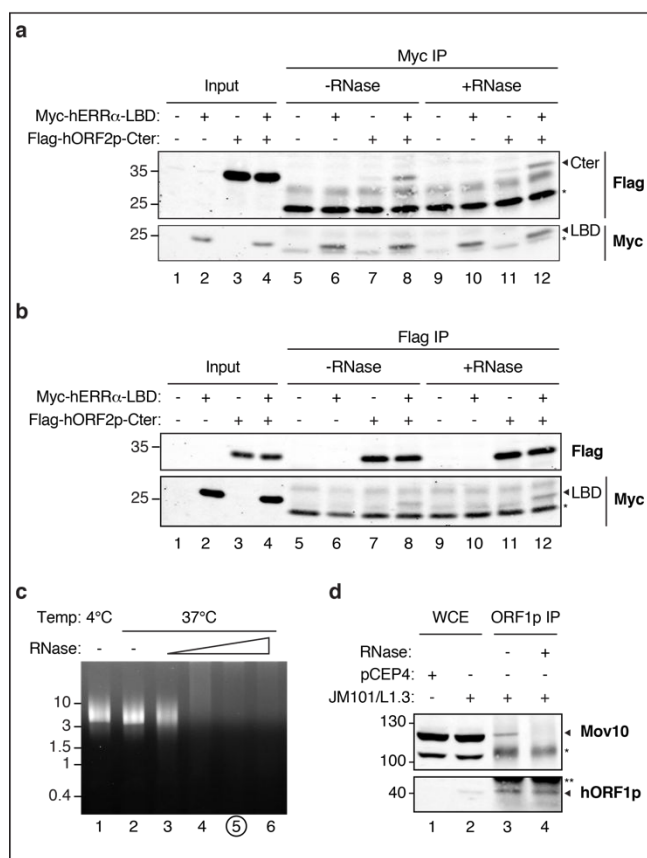


Figure 6. A model for NR-mediated tethering of the L1 RNP to chromosomes.

NR have the required structure for a tethering factor, with one domain binding to DNA, while another can recruit ORF2p through its C-terminal region. Note that NRs do not necessarily require direct binding to their DNA-binding site motif to exert their transcriptional activity, since they can bind indirectly to chromatin through protein-protein interactions with other transcription factors (Levin and Hammes 2016). For the sake of simplicity, only

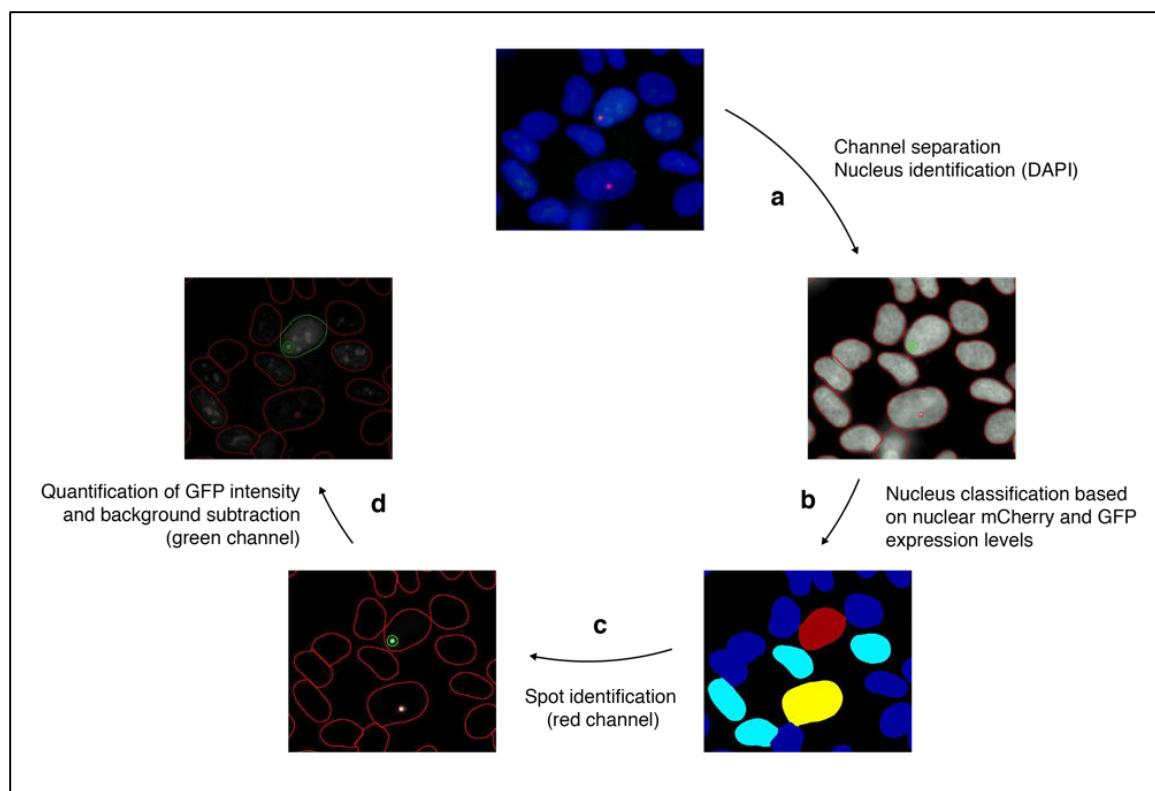
the direct mode of tethering is illustrated. NR, nuclear receptor; NRRE, nuclear receptor responsive element; LBD, ligand-binding domain; DBD, DNA-binding domain.

Supplementary data



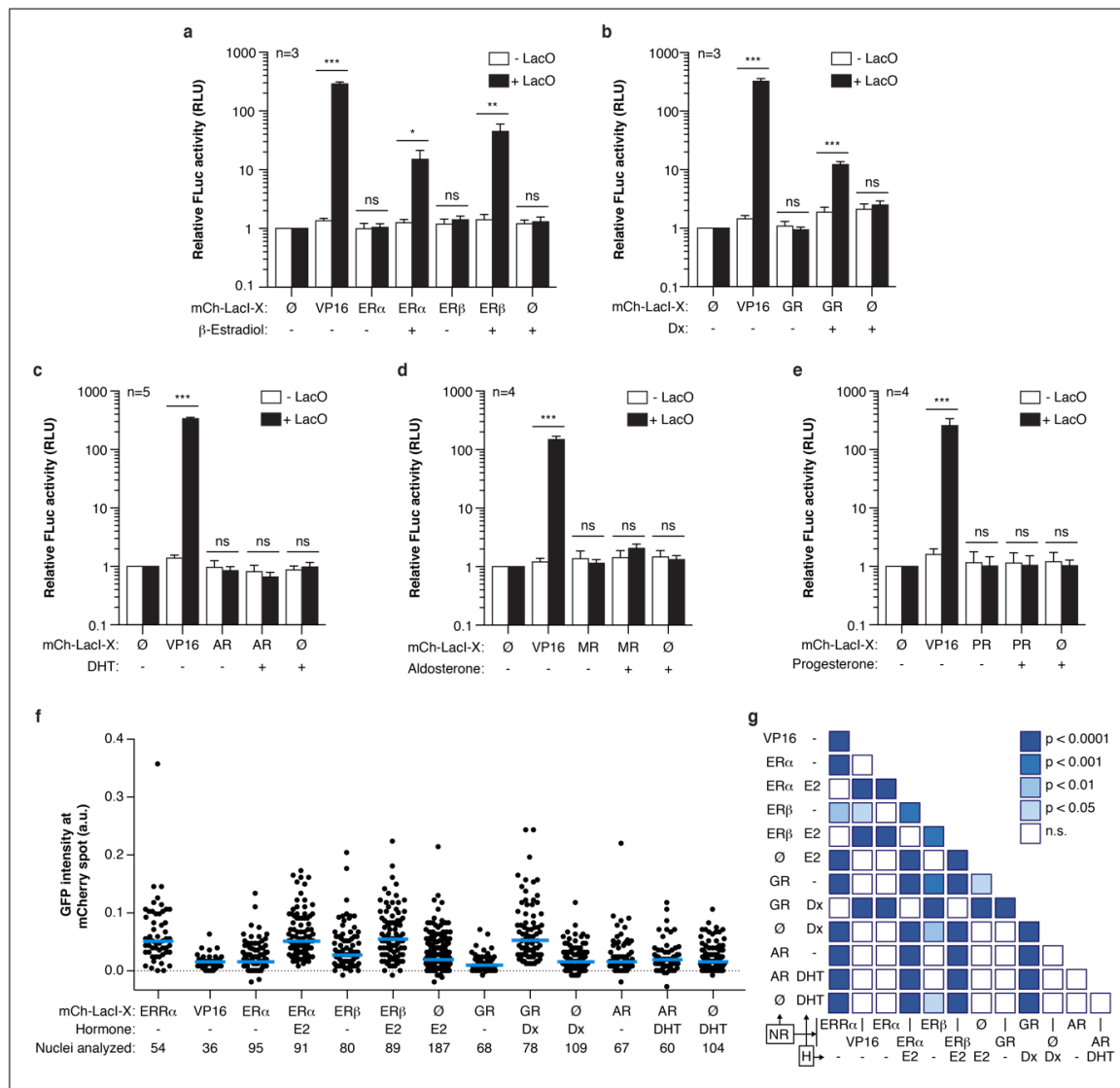
Supplementary Figure 1. The interaction between ERRα and ORF2p is RNA-independent. (a-b) coIP of Flag-tagged ORF2p-C-terminal domain (Flag-ORF2p-Cter) with Myc-tagged human ERRα LBD domain (Myc-hERRα-LBD) in HEK293T cell extracts pre-incubated or not for 5 min at 37°C with an RNase A/RNase I cocktail. Immunoprecipitation (IP) was performed with a Myc (a) or a Flag (b) antibody. Left, molecular weight (kDa) markers. Right (bold), antibodies used for immunoblotting. IP, immunoprecipitation. The bands marked with a star (*) correspond to the IgG light chain. (c) Activity of the RNase A/RNase I cocktail on total RNA present in cell extracts. Cell extracts were incubated for 5 min at 4°C or at 37°C with or without dilutions of an RNase A (5U/mL) and RNase I (200 U/mL) cocktail. RNA was purified and analyzed by agarose gel electrophoresis and ethidium bromide staining. Lane 3, 1:500; lane 4, 1:50; lane 5, 1:5; lane 6, undiluted. Conditions corresponding to lane 5 (circled) were used for all experiments with an RNase treatment step. (d) Control co-immunoprecipitation assay showing that under our RNase treatment conditions, the well-established and RNA-dependent interaction between ORF1p and Mov10 proteins is disrupted. HEK 293T cells were transfected either with a complete L1 expressing plasmid (JM101/L1.3) or the empty vector (pCEP4). RNase treatment was performed as in (a-c). Immunoprecipitation (IP) was performed with an antibody directed against hORF1p. Left, molecular weight (kDa) markers. Right (bold), antibodies used for immunoblotting. IP, immunoprecipitation; WCE, whole cell extracts. Arrowheads indicate the specific signals. Top, the bands marked with a star (*) are non-specific. Bottom, the cropped signal marked with a double star (**) corresponds to the IgG heavy chain. Note that the top and bottom blots correspond to the same samples, separated on different gels of 8% and 10% acrylamide for optimal resolution.

corresponding to lane 5 (circled) were used for all experiments with an RNase treatment step. (d) Control co-immunoprecipitation assay showing that under our RNase treatment conditions, the well-established and RNA-dependent interaction between ORF1p and Mov10 proteins is disrupted. HEK 293T cells were transfected either with a complete L1 expressing plasmid (JM101/L1.3) or the empty vector (pCEP4). RNase treatment was performed as in (a-c). Immunoprecipitation (IP) was performed with an antibody directed against hORF1p. Left, molecular weight (kDa) markers. Right (bold), antibodies used for immunoblotting. IP, immunoprecipitation; WCE, whole cell extracts. Arrowheads indicate the specific signals. Top, the bands marked with a star (*) are non-specific. Bottom, the cropped signal marked with a double star (**) corresponds to the IgG heavy chain. Note that the top and bottom blots correspond to the same samples, separated on different gels of 8% and 10% acrylamide for optimal resolution.



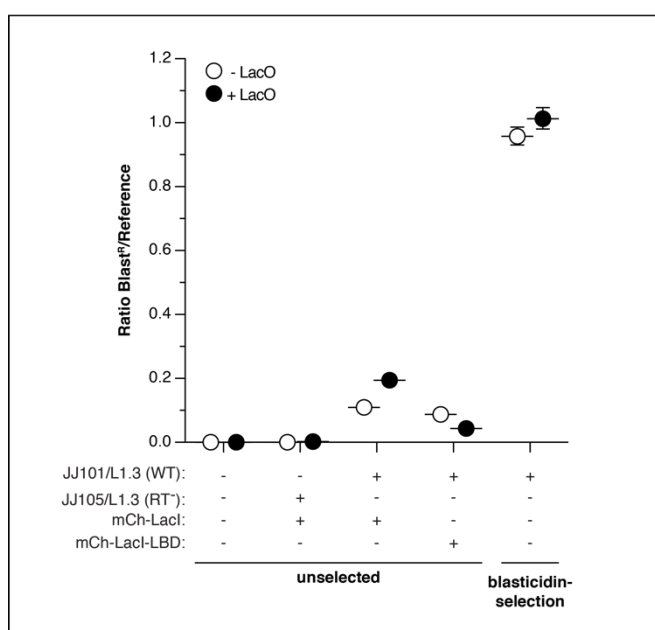
Supplementary Figure 2. An automated CellProfiler pipeline to process and analyze F2H microscopy images.

(a) Nuclei are identified with the blue channel (DAPI). (b) Red and green signal intensities are measured in each nucleus, and used to classify cells into 4 groups: mCherry⁻/GFP⁻ (dark blue), mCherry⁺/GFP⁻ (yellow), mCherry⁻/GFP⁺ (light blue), and mCherry⁺/GFP⁺ (red). Only mCherry⁺/GFP⁺ cells were considered. (c) Spots are identified in the red channel. (d) Median GFP intensity is measured in the spot and in a donut-shaped area around the spot to subtract local background (green channel). In all black and white pictures, the spots are circled (single circle for a dismissed mCherry⁺/GFP⁻ cell, and double-circled for a measured mCherry⁺/GFP⁺ cell with its local background).



Supplementary Figure 3. ORF2p interacts with several hormone-dependent steroid nuclear receptors.

(a-e) Transcriptional activity of mCherry-LacI fusion proteins. Legend and assays conditions were as described in Fig. 3c and 3d, except that growth medium was depleted from endogenous steroids by charcoal treatment and supplemented either with ethanol (-) or with a specific nuclear receptor agonist: β -estradiol (E2) for estrogen receptors (panel a, ER α and ER β), dexamethasone (Dx) for the glucocorticoid receptor (panel b, GR), dihydrotestosterone (DHT) for the androgen receptor (panel c, AR), aldosterone for the mineralocorticoid receptor (panel d, MR), and progesterone for the progesterone receptor (panel e, PR). Bars represent the mean \pm s.d. (n as indicated in each panel; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; multiple t-tests). Ø indicates no LBD (i.e. mCherry-LacI alone). Under our experimental conditions, the LBD domains of ER α , ER β , and GR fused to mCherry-LacI are transcriptionally active but their counterparts based on AR-, MR- or PR-LBD domains are not. (f) Quantification of F2H assays. Legend is similar to Fig. 3f. In the graph, dots correspond to individual values of 2 independent experiments pooled together, and blue horizontal lines represent median values. (g) One-way ANOVA analysis of data shown in (f), using Tukey's multiple test correction. NR, nuclear receptor; H, hormone.



Supplementary Figure 4. Specificity of the multiplex Taqman-based ddPCR assay to measure spliced Blast^R copy number.

The multiplex assay contains two sets of primer pairs and Taqman probes, one specific for the Blast^R exon-exon junction, and the other for a reference chromosomal locus used as normalizer (the RNase P gene, *RPPH1*). Signal is detected only with a retrotransposition-competent L1 (JJ101/L1.3), but neither with an RT-defective L1 (JJ105/L1.3), nor in untransfected cells. As a positive control, DNA from blasticidin-selected cells show strong enrichment of Blast^R copy number. Dots represent mean \pm Poisson confidence limits reported by QuantaSoft for n=20,000 droplets.

Supplementary Table S1. Results of Y2H screens performed with human and mouse L1 fragments as baits.

Human and mouse L1 fragments were screened against human thymus and mouse brain cDNA libraries, respectively. For each hit, the number corresponds to the number of clones sequenced. Hits highlighted in red are common Y2H false positives. Hits highlighted in grey were below the cut-off of ≥ 3 clones generally used to select candidate interactors. ERR α is highlighted in green.

Supplementary Table S3. Detailed CellProfiler pipeline to analyze F2H images.

No.	Analysis module	Module description
1	Load images	
2	Crop	Crop the picture to eliminate black borders and annotations included by microscope acquisition software *** Eventually adjust the left, right, top and bottom positions to your specific needs/pictures
3	Color to gray	Split the RGB picture into gray-scaled images for each color (Red-mCherry, Green-GFP, Blue-DAPI)
4	Soften	Soften the blue image to facilitate the automatique detection of the nuclei.
5	Identify primary objects	Nuclei identification step (based on DAPI, ie Blue, picture).
6	Enhance edges	Red spot identification.
7	Enhance or suppress features	Association of the found dots to their corresponding nuclei.
8	Enhance or suppress features	If there are secondary spots in the nucleus, will be kept only the biggest one (Red_Enhanced).
9	Mask image	Mask the red channel to keep only nuclear signal before spot identification
10	Identify primary objects	Spot identification step (based on mCherry, ie Red, picture). Note that the lower bound on threshold (0.002) is critical to remove artefactual « granular » hits.
11	Relate objects	Associate Spots with their Nucleus.
12	Measure Object Size Shape	Measure Area of Spots. This step is needed to keep only the largest spot if several spots per nucleus are found.
13	Measure Object Intensity	Filter all the nuclei and keep only the GFP+/mCherry+.
14	Filter objects	Filter Spots for each Nucleus to keep only the biggest (if several were found).
15	Relate objects	Associate biggest spot of each nucleus with their Nucleus.
16	Measure Object Intensity	Measure the intensity of GFP in the nucleus (to filter out GFP- cells).

Results

17	Classify objects	Classify each cell to know if it was doubly transfected (mCherry and GFP).
18	Filter objects	Keep only the co-transfected nuclei, but eliminate the super-bright red cells.
19	Mask objects	Keep only the spots from the co-transfected nuclei.
20	Relate objects	Associate filtered spots with their Nucleus.
21	Color to Gray	Create a binary mask with nuclei (1/2)
22	Apply treshold	Create a binary mask with nuclei (2/2)
23	Measure Object Intensity	Measure intensity-based values in each Nucleus and associated Spot.
24	Overlay outlines	Overlay the Blue (DAPI) original picture with outlines of all Nuclei and Spots found (red outlines) and those selected for analysis (green outlines). For QC purpose (used for SaveImages).
25	Tile	Create a picture with all overlays (each color with outlined Nuclei and Spots). Only green outlined objects are analyzed. For QC purpose (used in SaveImages).
26	Calculate Math	For the GFP: Calculate the ratio of the median intensity in the Spot vs the median intensity in the Nuclei. This corresponds to the enrichment of the GFP signal in the Spot as compared to the Nucleus (taken as background).
27	Export to spreadsheet	Export quantitative analyses to .csv file that can be imported in Excel.
28	Save images	For each analysis, export the overlays of each color with the outlines of objects (generated in the Tile module). For QC purpose.

4. General discussion and perspectives

Contents

4.1 In the tethering model, the integration of transposable elements is guided to specific chromosomal regions by cellular factors	126
4.1.1 A broad diversity of tethering factors contributes to the diversity of integration site preferences among TEs	126
4.1.2 Properties to be considered when evaluating a putative tethering factor	127
4.2 Association between steroid receptor and retrotransposon activities	131
4.2.1 Possible impact of nuclear receptors on L1 retrotransposon	134

“We think in generalities, but live in detail.” Alfred North Whitehead

Because L1 retrotransposon encodes the catalytic activities required for DNA cleavage and its reverse transcription, it is considered as an 'autonomous' transposable element. However, it is clear that host factors contribute to L1 expression, replication and regulation, particularly in a cellular context. Thus, to discover cellular pathways and factors involved in the regulation of L1 retrotransposon mobility, our laboratory has performed yeast 2-hybrid (Y2H) screens to identify cellular partners of the human L1 retrotransposition machinery and identified several potential hits. Among them, we identified ERR α , a steroid nuclear receptor, as a cellular partner of ORF2p, an essential component of the L1 machinery. The existence of ERR α paralogs and other steroid receptors prompted us to test whether ORF2p could also interact with other members of this nuclear receptor subfamily. Our results show that a subset of NRs, including all three ERRs, the two ER paralogs and GR, and possibly others, can physically interact with ORF2p in human cultured cells, linking retrotransposition with steroid signaling pathways.

4.1 In the tethering model, the integration of transposable elements is guided to specific chromosomal regions by cellular factors

Chromatin conformation has a significant impact on the integration process by blocking or opening the access of the integration machinery to the target DNA. However, the preference for nucleosome-free regions in euchromatic or heterochromatic environments, or for DNase-sensitive sites or nucleosome-bound DNA, influences target site selection locally and cannot explain all the TE insertion in a genome.

The integration of transposable elements and retroviruses is often also determined by tethering factors that guide the integration of these mobile genetic elements into specific chromosomal regions (for review see (Sultana et al. 2017)). This mechanism, known as the “tethering model”, originally proposed that integration is guided by a cellular protein that binds to the site of integration where it can recruit the preintegration complex of LTR-retrotransposons and retroviruses (Bushman 2003; Kirchner, Connolly, and Sandmeyer 1995).

4.1.1 A broad diversity of tethering factors contributes to the diversity of integration site preferences among TEs

Specific tethering factors have been identified mostly for long-terminal repeat (LTR) retroelements. For instance, the bromodomain and extra terminal (BET) family of proteins interacts with the integrase of murine leukemia virus (MLV) and favors its integration in promoters and enhancers (Sharma et al. 2013; De Rijck et al. 2013; Gupta et al. 2013). LEDGF/p75 interacts with the IN of human immunodeficiency virus (HIV) and guides its activity to highly spliced, active and transcribed genes (Ciuffi et al. 2005; Singh et al. 2015). TFIIII and AC40 direct yeast Ty3 and Ty1 integration close to tRNA genes, respectively (Kirchner, Connolly, and Sandmeyer 1995; Bridier-Nahmias et al. 2015), and yeast Ty5 preferentially targets subtelomeres through an interaction between Ty5 integrase and Sir4p (Zhu et al. 2003; W. Xie et al. 2001). Notably, the phosphorylation of Ty5 integrase is a major determinant of Ty5 targeting to subtelomeres. The phosphorylation of the IN is reduced under metabolic stress conditions and insertions are relocated throughout the genome, suggesting that Ty5-mediated insertional mutagenesis is a controlled process (Dai et al. 2007).

For non-LTR-retrotransposons, examples of tethering are less documented (Sultana et al. 2017). The most detailed report is TRE5-A, a non-LTR retrotransposon of the L1 clade found in *Dictyostelium discoideum*, which integrates specifically upstream of tRNA genes through an ORF1p-TFIIIB interaction (Siol et al. 2006; Chung et al. 2007).

We hypothesized that some steroid receptors could function as tethering factors for L1 in mammals, through their interaction with the C-terminal end of ORF2p. This suggests also the intriguing possibility that *de novo* L1 insertions could disperse less randomly in the genome than previously anticipated.

4.1.2 Properties to be considered when evaluating a putative tethering factor

To be a tethering factor, a cellular protein should be evaluated taking in consideration the following properties (Sultana et al. 2017):

- i) **Interaction:** To interact with one of the components of the integration complex.
- ii) **Endogenous distribution:** Its distribution should correspond with the integration preference of the mobile elements (Sharma et al. 2013; Bridier-Nahmias et al. 2015; Zou and Voytas 1997; De Rijck et al. 2013; De Rijck et al. 2010; Qi et al. 2012).
- iii) **Relocation:** Its artificial genomic relocation should relocate insertions in the same genomic sites (Zhu et al. 2003; Ferris et al. 2010; Silvers et al. 2010; Gijssbers et al. 2010).
- iv) **Loss-of-function:** Abolishing the interaction should result in changes in the distribution insertion (Bridier-Nahmias et al. 2015).

Here we discuss whether $ERR\alpha$ and other NRs do fulfill these experimental criteria.

(i) Interaction

We originally **identified $ERR\alpha$ as interacting with ORF2p** and this observation was extended to more members of the steroid nuclear receptors subfamily. However, 48 nuclear receptors are encoded in the human genome and we cannot exclude that ORF2p might interact with other NRs, outside of the steroid group. Thus, we currently do not

know which is the complete set of NRs interacting. Additionally, even in one subfamily, as the steroid one, the regulatory and signaling pathways depend very much on the cell type used for the study. The specificity and level of expression of nuclear receptors in different cell lines might reveal that the interaction between a given NR and L1 is specific to a certain tissue and regulatory process.

Future studies will be necessary to precisely define which NRs can associate with ORF2p and whether some of them could function in a cell type-dependent manner. This will be necessary to systematically test the genomic association between L1 integration sites and NR binding sites.

(ii) Endogenous distribution

There is **currently no known integration site preference for mammalian L1 elements** beside the endonuclease (EN) target site of ORF2p, a degenerate consensus sequence related to 5'-TTTT/A-3' (Khazina et al. 2011; Gilbert, Lutz-Prigge, and Moran 2002; Symer et al. 2002; Jurka 1997). No published study has explored how *de novo* L1 insertions are distributed in the genome, as it has been achieved for several LTR-retroelements (Sultana et al. 2017). **For possible future directions**, the global distribution of new L1 insertions could be obtained by next-generation sequencing (Philippe et al. 2016). To assess whether L1 preferentially integrates nearby interacting NRs, we would need first to precise the full set of interacting NR. Then, we could propose to perform ChIP-seq on these factors in the same cell type used for mapping L1 insertion sites, and to correlate NR peaks with L1 insertions. A previous report implicated androgen receptor (AR) binding to DNA in triggering ORF2p-mediated chromosome breaks and fusions in prostate cancer cells. This observation is consistent with a colocalization of AR with ORF2p at particular genomic locations (C. Lin et al. 2009). However, we were not able to demonstrate ORF2p binding to AR LBD under our experimental conditions. It is possible that other AR domains are involved.

(iii) Relocation

Artificial tethering of ERR α to a specific locus has been chosen to address a possible causal link between ERR α DNA binding and the recruitment of the L1 retrotransposition machinery at the same locus. The LacO-LacI bacterial system is used in mammalian cells and provides a strong tether by virtue of the large number of LacO sequences and to the

high affinity of binding of the LacI protein to the LacO sequence (Czapiewski, Robson, and Schirmer 2016; Kumaran and Spector 2008).

The cell line U2OS 2-6-3 used in this study, was created as a cell system to study the dynamics of gene expression in vivo (Janicki et al. 2004). The LacO repeats are integrated in a single integration site in an euchromatic region of chromosome 1p36. However, the integrated LacO array itself forms a heterochromatic structure. Generally, heterochromatin is viewed as a highly condensed state, which contributes to the repression of transcription by impeding transcription factors from accessing their binding elements.

The number of **L1 retrotransposition events diminishes upon tethering to the LacO array**, using a retrotransposition genetic assay, implying that $ERR\alpha$ relocation to the LacO array functionally impact L1 retrotransposition. The selection of the insertions by blasticidine might prevent detection of insertions occurring in a heterochromatic LacO array. However, quantification of insertions in unselected cells by droplet digital PCR shows that it actually directly reduces the number of L1 retrotransposition events.

Thus, a possible explanation for the reduced L1 retrotransposition would be that $ERR\alpha$ indeed tethers L1 to the LacO array, but the latter would not be a favorable site of integration. Alternatively, tethering could simply limit the number of possible target sites in the genome and thus could reduce the overall retrotransposition efficiency.

An immediate future direction is to investigate the location of L1 insertions upon $ERR\alpha$ tethering to the LacO array. We ask if the L1 insertions still occurring in this context are enriched in the LacO array or are rather inserted somewhere else in the genome. To achieve this, we designed a ChIP experiment in which we pull down mCherry-LacI fusion proteins and we quantify the retrotransposed BlastR marker.

We took advantage of the LacO-LacI system to directly tether $ERR\alpha$ LBD independently of NR DNA-binding sequence. This allowed us to circumvent the redundancy between nuclear receptors. However, even if this tethering system mimics the binding of the NR to DNA, the LacO array is not the natural environment where a NR or L1 would act. More specifically, the LacO array has some intrinsic limitations due to its heterochromatic and repetitive state. Hence, it can behave as a fragile site (Jacome and Fernandez-Capetillo

2011), particularly upon LacI binding, since this can promote replication blocks within the array (Payne et al. 2006). **In spite of these limitations**, the LacO system is widely used to study the biology of the nucleus and is a powerful tool to bring factors to a fixed chromatin location, without disrupting chromosome organization and recruitment of the tagged loci to silent nuclear domains (Jacome and Fernandez-Capetillo 2011; Finlan et al. 2008; Kumaran and Spector 2008; Reddy et al. 2008; Soutoglou and Misteli 2008).

(iv) Loss-of-function

Abolishing the interaction between a retroelement and its predominant tethering factor can result in diverse integration profiles (lost, partially conserved, more random or alternative distribution). Often the new distribution is not random. For instance, in the case of Ty1, *de novo* insertions are redistributed from tRNA genes to subtelomeric regions, when the interaction between AC40 and IN is disrupted (Bridier-Nahmias et al. 2015). Predicting the outcome of a loss-of-interaction might be even more challenging when several tethering factors are functionally redundant, whether they belong or not to the same protein family. The pattern of HIV integration into active genes does not change after depletion of LEDGF/p75, because another related protein, HRP2, is compensating. Even the concomitant depletion of both LEDGF/p75 and HRP2 does not randomize HIV insertions, which remain frequent in active genes (Schrijvers et al. 2012; Hao Wang et al. 2012).

In the present study, **knocking down a single nuclear receptor (ERR α) is not sufficient to alter retrotransposition** efficiency. The functional redundancy between nuclear receptors could explain such a result. However, we cannot exclude that siRNA-mediated knockdown does not completely deplete ERR α from the system, and that even very low ERR α levels could be sufficient to tether the L1 machinery.

As a possible future direction, the RNAi library approach, used for functional gene discovery within a predefined protein family, might offer more answers than single receptor siRNA-mediated knock-down. Also, ERR α gene could be knockout by CRISPR/Cas9 mediated genome editing.

In conclusion, ERR α and other NRs interacting with ORF2p fulfil several criteria supporting their role as L1 tethering factors. So far, their functional redundancy prevented us to conclusively test some of these criteria. Future experiments defining the major set of factors in a given cell-type will be necessary to definitely address this question.

4.2 Association between steroid receptor and retrotransposon activities

The steroid receptors are part of an adaptive response important to maintain homeostasis when facing environmental and physiological stresses. For example, estrogen and glucocorticoid receptors regulate response to oxidative stress (S.-R. Lee et al. 2013; Simoes et al. 2012). The role of ERs in glucose metabolism is highlighted in breast cancer cells where they support the adaptation to glucose availability. When glucose is abundant, membrane ERs, promote glucose uptake and use for glycolysis, which is the main source of energy for tumor growth and it keeps glutathione levels low to limit oxidative stress. When glucose levels are low, membrane ERs promotes glucose uptake and glycolysis, enhancing cell viability under nutrient stress (O'Mahony et al. 2012). The physiological consequences of glucocorticoid production are an increased glucose production and immunosuppression, which are responses to low blood glucose and inflammation, respectively, the most common insults the human body confronts (Revollo and Cidlowski 2009; Chrousos 1995; Barnes 1998). In *C. elegans*, nutrient deficiency enhances steroid signalling which is necessary for the lifespan extension observed upon caloric restriction and TOR signalling, a nutrient-sensing kinase (Thondamal et al. 2014).

Given the importance of steroid receptors in the physiology of multicellular organisms and their adaptation to environmental changes, we speculate that NR-mediated tethering of ORF2p might modulate the landscape of L1 integration in the genome and link environmental and physiological signals with genomic plasticity. However, additional links between steroid receptor and retrotransposon activities have been previously highlighted.

L1 promoter contains binding sites for different transcription factors which could modulate L1 transcription, depending on the availability of those factors in a specific cell

or tissue. Thus, exposure to agents which alter the function or expression of those factors, could also influence L1 transcription (Ade et al. 2017).

Some studies focused directly on testing the influence of various agents or environmental factors on the promoter of L1. Notably, a group of researchers used beta galactosidase assay and L1 promoter fused to the LacZ reporter gene and observed that treatment with steroid hormone-like agents (serum, testosterone, dihydrotestosterone, pesticides) increases slightly the activity of the L1 promoter in cultured cells (Morales, Snow, and Murnane 2002). However, this study lacks measurement of RNA levels to know if the influence of those agents on the L1 promoter is indeed affecting transcription. Also, controls of cytotoxicity which are a must, are missing, as factors which affect cell viability and efficiency to express a certain reporter will distort the results (Ade et al. 2017).

Terasaki and colleagues tested 95 compounds for their potential to enhance transcription of human L1 (Terasaki et al. 2013). They used genotoxic agents, commercially available drugs and compounds which induce cellular stress. Strikingly, out of 95 compounds, 15 increased L1 promoter activity, including 8 drugs available on the pharmaceutical market. Among these, have been tested agonists or antagonists of nuclear receptors and their effect on L1 promoter activity was evaluated by luciferase reporter gene assay. After 6 and 24 hours of treatment, a mineralocorticoid-receptor antagonist and the estrogen hormone (17 β -estradiol or E2) showed no significant change on L1 promoter activity. In contrast, the thyroid hormone (triiodothyronine) and the progesterone decreased L1 promoter activity after 24 hours treatment. An increase in L1 promoter activity was observed after exposure to peroxisome proliferator-activated receptor α (PPAR α) agonist, drug used for the treatment of the metabolic syndrome symptoms, mainly for lowering triglycerides and blood sugar. An increased L1 promoter activity was detected also after treatment with anti-steroid drugs, which are steroidogenesis inhibitors used in adrenocortical carcinoma (Terasaki et al. 2013).

Similarly, Alu expression can be stimulated at the transcriptional level by glucocorticoids (Sun and Frankel 1986). Notably, the highest level of retrotransposition is found in steroidogenic organs as brain, placenta and gonads (Dupressoir, Lavialle, and Heidmann 2012; Hunter, McEwen, and Pfaff 2014; Pillai and Chuma 2012; Ross, Weiner, and Lin 2014).

NR agonists and antagonists form an important class of pharmacological drugs. We can cite for example, glucocorticoids in oncology, where they are used for their pro-apoptotic action to treat lymphoproliferative disorders and also to relieve the side effects of chemotherapy and radiotherapy (K.-T. Lin and Wang 2016; Gennari et al. 1996). Another famous example is mifepristone, known as RU486, a potent antagonist of progesterone receptor, used for emergency contraception, termination of pregnancy and menstrual regulation (Ho, Yu Ng, and Tang 2002).

Even if there is not yet a clear image about which compounds have the potential to alter L1 retrotransposition and at which doses, a risk assessment of induced retrotransposition during drug development and drug administration, should be considered. This is also true for other substances, such as the pesticides or herbicides, which are often hormonal dysregulators (Ade et al. 2017).

Inversely, transposable elements have significantly contributed to the evolution of our genome, by dispersing transcription factor binding sites and ready-made promoter elements (Hunter et al. 2015; Rebollo, Romanish, and Mager 2012; Faulkner and Carninci 2014), in particular steroid receptor responsive elements (Cotnoir-White, Laperrière, and Mader 2011). For example, Alu sequences are rich in responsive elements for progesterone, glucocorticoid and vitamin D (Jacobsen et al. 2009; Gombart, Saito, and Koeffler 2009).

Recent observations have connected circadian rhythm and L1 retrotransposition (deHaro et al. 2014). Environmental light exposure at night is recognized by World Health Organization as a disruptor of the circadian system and as a possible carcinogen, due to an increased rate of cancers in shift workers. During light exposure, melatonin secretion is suppressed. Interestingly, overexpression of the melatonin receptor (MT1) reduces L1 retrotransposition in cultured cells, decreasing both the levels of L1 mRNA and ORF1p protein. In addition, use of human blood rich in melatonin suppressed endogenous L1 mRNA expression during *in situ* perfusion of tissue-isolated xenografts of human cancer (deHaro et al. 2014). These experiments suggest a model in which the disruption of circadian rhythm activates L1 elements, which might contribute to an increased cancer

risk in shift workers. However, causative relationships between these phenomena still need to be demonstrated (Ade et al. 2017).

Interestingly, nuclear receptors have also been linked to the circadian rhythm, given their role in metabolic regulation. Several of them (e.g. Rev-erba, RORα, PPARs) are integral components of the molecular clock machinery and regulate downstream target genes in a circadian manner, coupling peripheral circadian clocks with metabolic outputs, and coordinating biological timing with metabolic physiology (Duez and Staels 2010; Xiaoyong Yang et al. 2006; X Yang, Lamia, and Evans 2007).

Collectively, these observations suggest that nuclear receptor signalling pathways can regulate L1 retrotransposition at several levels, from transcription to insertion site selection. In a broader perspective, they connect environmental, physiological or pathological signalling to genome (in)stability, a driving mechanism of cancer and aging.

4.2.1 Possible impact of nuclear receptors on L1 retrotransposon

The impact transposons have on the function of genomes depends on the position of their insertion sites. This work could provide the first evidence that nuclear receptors contribute to L1 distribution in the genome by tethering L1 machinery to specific sites. These sites could be chosen depending on the different expression of nuclear receptors in different cell types or in response to diverse external or internal signals, which impact both the nuclear receptors and L1 retrotransposon.

Moreover, considering the hypothesis of Barbara McClintock, that transposons alter gene expression in ways that allow cells to respond to stress and if integration is not only random, then the targeting and tethering mechanisms evolved with the role to protect and help the host.

At the environmental level, as mentioned before, stress has been shown to increase L1 mobilization (Terasaki et al. 2013; deHaro et al. 2014). Also, it is well known that stressful events, result in secretion of glucocorticoid hormones which bind and activate the glucocorticoid (GR) receptor, considered the main mediator of the stress response, maintaining the cell health and wellbeing (Mifsud and Reul 2016).

In a very speculative manner, we could imagine a scenario where upon stress, L1 increased mobilization could be controlled, by restriction or insertion site selection, via interaction with the glucocorticoid receptor.

In addition, recent studies highlight the role of GR in the proliferation, differentiation, migration and functional integration of newborn neurons in the hippocampus (Saaltink and Vreugdenhil 2014; Mifsud and Reul 2016). Except the epithelial cancers, most of known somatic retrotransposition occurs in brain and in particular the hippocampus, a neurogenic niche, supports pronounced L1 activity (Coufal et al. 2009; Evrony et al. 2012; Upton et al. 2015; Macia et al. 2017) .

In cancers, the mechanisms which keep under control the mobile DNA are dysregulated. This allows L1 to be expressed and be a dynamic component of cancer genomes (Burns 2017). L1 ORF1 protein is overexpressed in human cancers and constitutes a hallmark, being detected by immunohistochemistry in half of all epithelial tumors (Rodić et al. 2014).

Given their involvement in the regulation of many and diverse human developmental and physiological functions, nuclear receptors have been implicated also in several disease types, as cancer, obesity, diabetes, infertility.

In this case, steroid nuclear receptors could again play a protective role, by limiting the overall efficiency of L1 retrotransposition and by mediating the insertions to specific genomic sites, which will not add more mutagenesis to an already destabilized genome.

Hormonal therapy is used in the treatment of several cancers, contributing to reduced recurrence and longer survival rates. The duration and doses of hormonal therapy is not very well understood yet. More data is needed to avoid appearance of steroid-resistance over time and to develop new steroid-based treatments (Ahmad and Kumar 2011). On the same speculative note, it could be possible that the hormones received as therapy for specific diseases, would help the overall stabilization of the genome, by regulating also L1, the only active retrotransposon in humans.

5. References

- Ade, Catherine M, Geraldine Servant, Maria E Morales, and Astrid M Roy-Engel. 2017. "Environment, Cellular Signaling, and L1 Activity." In *Human Retrotransposons in Health and Disease*, 20:157–94. Cham: Springer International Publishing. doi:10.1007/978-3-319-48344-3_7.
- Ahmad, Nihal, and Raj Kumar. 2011. "Steroid Hormone Receptors in Cancer Development: a Target for Cancer Therapeutics.." *Cancer Letters* 300 (1): 1–9. doi:10.1016/j.canlet.2010.09.008.
- Alaynick, William A, James M Way, Stephanie A Wilson, William G Benson, Liming Pei, Michael Downes, Ruth Yu, et al. 2010. "ERRgamma Regulates Cardiac, Gastric, and Renal Potassium Homeostasis.." *Molecular Endocrinology (Baltimore, Md.)* 24 (2): 299–309. doi:10.1210/me.2009-0114.
- Alaynick, William A, Richard P Kondo, Wen Xie, Weimin He, Catherine R Dufour, Michael Downes, Johan W Jonker, et al. 2007. "ERRγ Directs and Maintains the Transition to Oxidative Metabolism in the Postnatal Heart." *Cell Metabolism* 6 (1): 13–24. doi:10.1016/j.cmet.2007.06.007.
- Alce, Timothy M, and Waldemar Popik. 2004. "APOBEC3G Is Incorporated Into Virus-Like Particles by a Direct Interaction with HIV-1 Gag Nucleocapsid Protein.." *The Journal of Biological Chemistry* 279 (33): 34083–86. doi:10.1074/jbc.C400235200.
- Alisch, Reid S, Jose L Garcia-Perez, Alysson R Muotri, Fred H Gage, and John V Moran. 2006. "Unconventional Translation of Mammalian LINE-1 Retrotransposons.." *Genes & Development* 20 (2): 210–24. doi:10.1101/gad.1380406.
- Alves, G, A Tatro, and T Fanning. 1996. "Differential Methylation of Human LINE-1 Retrotransposons in Malignant Cells.." *Gene* 176 (1-2): 39–44. doi:10.1016/0378-1119(96)00205-3.
- Athanikar, Jyoti N, Richard M Badge, and John V Moran. 2004. "A YY1-Binding Site Is Required for Accurate Human LINE-1 Transcription Initiation.." *Nucleic Acids Research* 32 (13). England: 3846–55. doi:10.1093/nar/gkh698.
- Austen, M, B Lüscher, and J M Lüscher-Firzlaff. 1997. "Characterization of the Transcriptional Regulator YY1. the Bipartite Transactivation Domain Is Independent of Interaction with the TATA Box-Binding Protein, Transcription Factor IIB, TAFII55, or cAMP-Responsive Element-Binding Protein (CPB)-Binding Protein.." *The Journal of Biological Chemistry* 272 (3): 1709–17.
- Bannert, Norbert, and Reinhard Kurth. 2006. "The Evolutionary Dynamics of Human Endogenous Retroviral Families.." *Annu Rev Genomics Hum Genet* 7 (1). United States: 149–73. doi:10.1146/annurev.genom.7.080505.115700.
- Barkhem, Tomas, Stefan Nilsson, and Jan-Åke Gustafsson. 2004. "Molecular Mechanisms, Physiological Consequences and Pharmacological Implications of Estrogen Receptor Action." *American Journal of Pharmacogenomics* 4 (1). Springer International Publishing: 19–28. doi:10.2165/00129785-200404010-00003.
- Barlow, C, S Hirotsune, R Paylor, M Liyanage, M Eckhaus, F Collins, Y Shiloh, et al. 1996. "Atm-Deficient Mice: a Paradigm of Ataxia Telangiectasia.." *Cell* 86 (1): 159–71.
- Barnes, P J. 1998. "Anti-Inflammatory Actions of Glucocorticoids: Molecular Mechanisms.." *Clinical Science* 94 (6). Portland Press Limited: 557–72.

- doi:10.1042/cs0940557.
- Batzer, Mark A, and Prescott L Deininger. 2002. "Alu Repeats and Human Genomic Diversity.." *Nat Rev Genet* 3 (5). England: 370–79. doi:10.1038/nrg798.
- Beck, Christine R, Beck, C R, Pamela Collier, P Collier, C Macfarlane, Catriona Macfarlane, Maika Malig, et al. 2010. "LINE-1 Retrotransposition Activity in Human Genomes.." *Cell* 141 (7): 1159–70. doi:10.1016/j.cell.2010.05.021.
- Beck, Christine R, José Luis Garcia-Perez, Richard M Badge, and John V Moran. 2011. "LINE-1 Elements in Structural Variation and Disease.." *Annual Review of Genomics and Human Genetics* 12: 187–215. doi:10.1146/annurev-genom-082509-141802.
- Beck-Engeser, Gabriele B, Dan Eilat, and Matthias Wabl. 2011. "An Autoimmune Disease Prevented by Anti-Retroviral Drugs.." *Retrovirology* 8 (1): 91. doi:10.1186/1742-4690-8-91.
- Becker, K G, G D Swergold, K Ozato, and R E Thayer. 1993. "Binding of the Ubiquitous Nuclear Transcription Factor YY1 to a Cis Regulatory Sequence in the Human LINE-1 Transposable Element.." *Human Molecular Genetics* 2 (10): 1697–1702.
- Belancio, Victoria P, Dale J Hedges, and Prescott Deininger. 2008. "Mammalian Non-LTR Retrotransposons: for Better or Worse, in Sickness and in Health.." *Genome Res* 18 (3). United States: 343–58. doi:10.1101/gr.5558208.
- Belfort, Marlene, M Joan Curcio, and Neal F Lue. 2011. "Telomerase and Retrotransposons: Reverse Transcriptases That Shaped Genomes.." In, 108:20304–10. National Acad Sciences. doi:10.1073/pnas.1100269109.
- Belkhadir, Youssef, and Joanne Chory. 2006. "Brassinosteroid Signaling: a Paradigm for Steroid Hormone Signaling From the Cell Surface.." *Science* 314 (5804). American Association for the Advancement of Science: 1410–11. doi:10.1126/science.1134040.
- Belshaw, R, A L A Dawson, J Woolven-Allen, J Redding, A Burt, and M Tristem. 2005. "Genomewide Screening Reveals High Levels of Insertional Polymorphism in the Human Endogenous Retrovirus Family HERV-K(HML2): Implications for Present-Day Activity.." *Journal of Virology* 79 (19): 12507–14. doi:10.1128/JVI.79.19.12507-12514.2005.
- Bennett, E Andrew, Heiko Keller, Ryan E Mills, Steffen Schmidt, John V Moran, Oliver Weichenrieder, and Scott E Devine. 2008. "Active Alu Retrotransposons in the Human Genome.." *Genome Res* 18 (12). United States: 1875–83. doi:10.1101/gr.081737.108.
- Bennett, Sara M, Derek S Woods, Katherine S Pawelczak, and John J Turchi. 2012. "Multiple Protein-Protein Interactions Within the DNA-PK Complex Are Mediated by the C-Terminus of Ku 80.." *International Journal of Biochemistry and Molecular Biology* 3 (1): 36–45.
- Bishop, Kate N, Mohit Verma, Eun-Young Kim, Steven M Wolinsky, and Michael H Malim. 2008. "APOBEC3G Inhibits Elongation of HIV-1 Reverse Transcripts.." Edited by Thomas J Hope. *PLoS Pathog* 4 (12): e1000231. doi:10.1371/journal.ppat.1000231.
- Black, Ben E, James M Holaska, Fraydoon Rastinejad, and Bryce M Paschal. 2001. "DNA Binding Domains in Diverse Nuclear Receptors Function as Nuclear Export Signals." *Current Biology* 11 (22): 1749–58. doi:10.1016/S0960-9822(01)00537-1.
- Blelloch, Robert, Monica Venere, Jonathan Yen, and Miguel Ramalho-Santos. 2007. "Generation of Induced Pluripotent Stem Cells in the Absence of Drug Selection.." *Cell Stem Cell* 1 (3): 245–47. doi:10.1016/j.stem.2007.08.008.
- Boeke, J D, D J Garfinkel, C A Styles, and G R Fink. 1985. "Ty Elements Transpose Through an RNA Intermediate.." *Cell* 40 (3): 491–500.

- Bogerd, Hal P, Heather L Wiegand, Amy E Hulme, Jose L Garcia-Perez, K Sue O'Shea, John V Moran, and Bryan R Cullen. 2006. "Cellular Inhibitors of Long Interspersed Element 1 and Alu Retrotransposition.." *Proceedings of the National Academy of Sciences of the United States of America* 103 (23): 8780–85. doi:10.1073/pnas.0603313103.
- Boissinot, S, P Chevret, and A V Furano. 2000. "L1 (LINE-1) Retrotransposon Evolution and Amplification in Recent Human History.." *Mol Biol Evol* 17 (6): 915–28. doi:10.1093/oxfordjournals.molbev.a026372.
- Bommer, Martin, Arndt Benecke, Hinrich Gronemeyer, and Cécile Rochette-Egly. 2002. "TIF2 Mediates the Synergy Between RARalpha 1 Activation Functions AF-1 and AF-2.." *The Journal of Biological Chemistry* 277 (40). American Society for Biochemistry and Molecular Biology: 37961–66. doi:10.1074/jbc.M206001200.
- Bookout, Angie L, Yangsik Jeong, Michael Downes, Ruth T Yu, Ronald M Evans, and David J Mangelsdorf. 2006. "Anatomical Profiling of Nuclear Receptor Expression Reveals a Hierarchical Transcriptional Network.." *Cell* 126 (4): 789–99. doi:10.1016/j.cell.2006.06.049.
- Bourguet, W, M Ruff, P Chambon, H Gronemeyer, and D Moras. 1995. "Crystal Structure of the Ligand-Binding Domain of the Human Nuclear Receptor RXR-Alpha.." *Nature* 375 (6530). Nature Publishing Group: 377–82. doi:10.1038/375377a0.
- Brélivet, Yann, Natacha Rochel, and Dino Moras. 2012. "Structural Analysis of Nuclear Receptors: From Isolated Domains to Integral Proteins.." *Molecular and Cellular Endocrinology* 348 (2): 466–73. doi:10.1016/j.mce.2011.08.015.
- Bridier-Nahmias, Antoine, Aurélie Tchalikian-Cosson, Joshua A Baller, Rachid Menouni, Hélène Fayol, Amando Flores, Ali Saïb, Michel Werner, Daniel F Voytas, and Pascale Lesage. 2015. "Retrotransposons. an RNA Polymerase III Subunit Determines Sites of Retrotransposon Integration.." *Science* 348 (6234): 585–88. doi:10.1126/science.1259114.
- Brouha, B, J Schustak, R M Badge, S Lutz-Prigge, A H Farley, J V Moran, and H H Jr Kazazian. 2003. "Hot L1s Account for the Bulk of Retrotransposition in the Human Population.." *Proceedings of the National Academy of Sciences of the United States of America* 100 (9). Department of Genetics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA.: 5280–85. doi:10.1073/pnas.0831042100.
- Bubeck, Doryen, Martin A M Reijns, Stephen C Graham, Katy R Astell, E Yvonne Jones, and Andrew P Jackson. 2011. "PCNA Directs Type 2 RNase H Activity on DNA Replication and Repair Substrates.." *Nucleic Acids Research* 39 (9): 3652–66. doi:10.1093/nar/gkq980.
- Burns, Kathleen H. 2017. "Transposable Elements in Cancer." *Nat Rev Cancer* 17 (7): 415–24. doi:10.1038/nrc.2017.35.
- Bushman, Frederic D. 2003. "Targeting Survival: Integration Site Selection by Retroviruses and LTR-Retrotransposons.." *Cell* 115 (2): 135–38.
- Byerly, Mardi S, Muhannad Al Salayta, Roy D Swanson, Kiwook Kwon, Jonathan M Peterson, Zhikui Wei, Susan Aja, Timothy H Moran, Seth Blackshaw, and G William Wong. 2013. "Estrogen-Related Receptor B Deletion Modulates Whole-Body Energy Balance via Estrogen-Related Receptor Γ and Attenuates Neuropeptide Y Gene Expression.." *The European Journal of Neuroscience* 37 (7): 1033–47. doi:10.1111/ejn.12122.
- Callahan, Kathryn E, Alison B Hickman, Charles E Jones, Rodolfo Ghirlando, and Anthony V Furano. 2012. "Polymerization and Nucleic Acid-Binding Properties of

- Human L1 ORF1 Protein..” *Nucleic Acids Research* 40 (2). England: Oxford University Press: 813–27. doi:10.1093/nar/gkr728.
- Castet, Audrey, Adrien Herledan, Sandrine Bonnet, Stéphan Jalaguier, Jean-Marc Vanacker, and Vincent Cavaillès. 2006. “Receptor-Interacting Protein 140 Differentially Regulates Estrogen Receptor-Related Receptor Transactivation Depending on Target Genes..” *Molecular Endocrinology* 20 (5): 1035–47. doi:10.1210/me.2005-0227.
- Castro-Diaz, Nathaly, Gabriela Ecco, Andrea Coluccio, Adamandia Kapopoulou, Benyamin Yazdanpanah, Marc Friedli, Julien Duc, Suk Min Jang, Priscilla Turelli, and Didier Trono. 2014. “Evolutionally Dynamic L1 Regulation in Embryonic Stem Cells..” *Genes & Development* 28 (13): 1397–1409. doi:10.1101/gad.241661.114.
- Chandra, Vikas, Pengxiang Huang, Yoshitomo Hamuro, Srilatha Raghuram, Yongjun Wang, Thomas P Burris, and Fraydoon Rastinejad. 2008. “Structure of the Intact PPAR-Gamma-RXR- Nuclear Receptor Complex on DNA..” *Nature* 456 (7220): 350–56. doi:10.1038/nature07413.
- Charest-Marcotte, Alexis, Catherine R Dufour, Brian J Wilson, Annie M Tremblay, Lillian J Eichner, Daniel H Arlow, Vamsi K Mootha, and Vincent Giguère. 2010. “The Homeobox Protein Prox1 Is a Negative Modulator of ERR{Alpha}/PGC-1{Alpha} Bioenergetic Functions..” *Genes & Development* 24 (6). Cold Spring Harbor Lab: 537–42. doi:10.1101/gad.1871610.
- Chen, H, C E Lilley, Q Yu, D V Lee, J Chou, I Narvaiza, N R Landau, and M D Weitzman. 2006. “APOBEC3A Is a Potent Inhibitor of Adeno-Associated Virus and Retrotransposons.” *Curr Biol* 16 (5). Infectious Disease Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA.: 480–85. doi:10.1016/j.cub.2006.01.031.
- Chen, Shiuan, Dujin Zhou, Chun Yang, and Mark Sherman. 2001. “Molecular Basis for the Constitutive Activity of Estrogen-Related Receptor A-1.” *The Journal of Biological Chemistry* 276 (30): 28465–70. doi:10.1074/jbc.M102638200.
- Chen, Xi, Han Xu, Ping Yuan, Fang Fang, Mikael Huss, Vinsensius B Vega, Eleanor Wong, et al. 2008. “Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells.” *Cell* 133 (6): 1106–17. doi:10.1016/j.cell.2008.04.043.
- Chiruvella, Kishore K, Zhuobin Liang, and Thomas E Wilson. 2013. “Repair of Double-Strand Breaks by End Joining..” *Cold Spring Harb Perspect Biol* 5 (5): a012757–57. doi:10.1101/cshperspect.a012757.
- Chiu, Ya-Lin L, and Warner C Greene. 2008. “The APOBEC3 Cytidine Deaminases: an Innate Defensive Network Opposing Exogenous Retroviruses and Endogenous Retroelements..” *Annu Rev Immunol* 26 (1). United States: 317–53. doi:10.1146/annurev.immunol.26.021607.090350.
- Chrousos, George P. 1995. “The Hypothalamic–Pituitary–Adrenal Axis and Immune-Mediated Inflammation.” Edited by Jeffrey S Flier and Lisa H Underhill. *N Engl J Med* 332 (20): 1351–63. doi:10.1056/NEJM199505183322008.
- Chu, W M, W M Liu, and C W Schmid. 1995. “RNA Polymerase III Promoter and Terminator Elements Affect Alu RNA Expression..” *Nucleic Acids Research* 23 (10): 1750–57.
- Chung, Thanh, Oliver Siol, Theodor Dingermann, and Thomas Winckler. 2007. “Protein Interactions Involved in tRNA Gene-Specific Integration of Dictyostelium Discoideum Non-Long Terminal Repeat Retrotransposon TRE5-a..” *Molecular and Cellular Biology* 27 (24): 8492–8501. doi:10.1128/MCB.01173-07.

- Ciuffi, Angela, Manuel Llano, Eric Poeschla, Christian Hoffmann, Jeremy Leipzig, Paul Shinn, Joseph R Ecker, and Frederic Bushman. 2005. "A Role for LEDGF/P75 in Targeting HIV DNA Integration.." *Nature Medicine* 11 (12): 1287–89. doi:10.1038/nm1329.
- Claessens, Frank, and Daniel T Gewirth. 2004. "DNA Recognition by Nuclear Receptors.." *Essays in Biochemistry* 40. Portland Press Limited: 59–72. doi:10.1042/bse0400059.
- Coffin, John M, Stephen H Hughes, Harold E Varmus, J M Coffin, S H Hughes, and H E Varmus. 1997. "The Interactions of Retroviruses and Their Hosts." Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.
- Comeaux, Matthew S, Astrid M Roy-Engel, Dale J Hedges, and Prescott L Deininger. 2009. "Diverse Cis Factors Controlling Alu Retrotransposition: What Causes Alu Elements to Die?." *Genome Res* 19 (4). United States: 545–55. doi:10.1101/gr.089789.108.
- Cook, Pamela R, Charles E Jones, and Anthony V Furano. 2015. "Phosphorylation of ORF1p Is Required for L1 Retrotransposition.." *Proceedings of the National Academy of Sciences of the United States of America* 112 (14): 4298–4303. doi:10.1073/pnas.1416869112.
- Cordaux, Richard, and Mark A Batzer. 2009. "The Impact of Retrotransposons on Human Genome Evolution.." *Nat Rev Genet* 10 (10). England: 691–703. doi:10.1038/nrg2640.
- Cost, Gregory J, Qinghua Feng, Alain Jacquier, and Jef D Boeke. 2002. "Human L1 Element Target-Primed Reverse Transcription in Vitro.." *The EMBO Journal* 21 (21). England: 5899–5910. doi:10.1093/emboj/cdf592.
- Cotnoir-White, David, David Laperrière, and Sylvie Mader. 2011. "Evolution of the Repertoire of Nuclear Receptor Binding Sites in Genomes." *Molecular and Cellular Endocrinology* 334 (1-2). Elsevier Ireland Ltd: 76–82. doi:10.1016/j.mce.2010.10.021.
- Coufal, Nicole G, Jose L Garcia-Perez, Grace E Peng, Gene W Yeo, Yangling Mu, Michael T Lovci, Maria Morell, K Sue O'Shea, John V Moran, and Fred H Gage. 2009. "L1 Retrotransposition in Human Neural Progenitor Cells.." *Nature* 460 (7259): 1127–31. doi:10.1038/nature08248.
- Coufal, Nicole G, José Luis Garcia-Perez, Grace E Peng, Maria C N Marchetto, Alysso R Muotri, Yangling Mu, Christian T Carson, Angela Macia, John V Moran, and Fred H Gage. 2011. "Ataxia Telangiectasia Mutated (ATM) Modulates Long Interspersed Element-1 (L1) Retrotransposition in Human Neural Stem Cells.." *Proceedings of the National Academy of Sciences of the United States of America* 108 (51): 20382–87. doi:10.1073/pnas.1100273108.
- Cowley, S M, S Hoare, S Mosselman, and M G Parker. 1997. "Estrogen Receptors Alpha and Beta Form Heterodimers on DNA.." *The Journal of Biological Chemistry* 272 (32): 19858–62.
- Craig, Nancy L. 2002. "Mobile DNA: an Introduction." In *Mobile DNA II*, 3–11. American Society of Microbiology. doi:10.1128/9781555817954.ch1.
- Crow, Y J, D N Black, M Ali, J Bond, A P Jackson, M Lefson, J Michaud, et al. 2003. "Cree Encephalitis Is Allelic with Aicardi-Goutières Syndrome: Implications for the Pathogenesis of Disorders of Interferon Alpha Metabolism.." *J Med Genet* 40 (3): 183–87.
- Crow, Yanick J, Bruce E Hayward, Rekha Parmar, Peter Robins, Andrea Leitch, Manir Ali, Deborah N Black, et al. 2006. "Mutations in the Gene Encoding the 3'-5' DNA

- Exonuclease TREX1 Cause Aicardi-Goutières Syndrome at the AGS1 Locus.” *Nat Genet* 38 (8): 917–20. doi:10.1038/ng1845.
- Czapiewski, Rafal, Michael I Robson, and Eric C Schirmer. 2016. “Anchoring a Leviathan: How the Nuclear Membrane Tethers the Genome.” *Front Genet* 7 (e157). Frontiers: 82. doi:10.3389/fgene.2016.00082.
- Dai, Junbiao, Weiwu Xie, Troy L Brady, Jiquan Gao, and Daniel F Voytas. 2007. “Phosphorylation Regulates Integration of the Yeast Ty5 Retrotransposon Into Heterochromatin.” *Molecular Cell* 27 (2): 289–99. doi:10.1016/j.molcel.2007.06.010.
- Darimont, B D, R L Wagner, J W Apreletti, M R Stallcup, P J Kushner, J D Baxter, R J Fletterick, and K R Yamamoto. 1998. “Structure and Specificity of Nuclear Receptor-Coactivator Interactions.” *Genes & Development* 12 (21). Cold Spring Harbor Laboratory Press: 3343–56.
- Daskalos, Alexandros, Georgios Nikolaidis, George Xinarianos, Paraskevi Savvari, Adrian Cassidy, Roubini Zakopoulou, Athanasios Kotsinas, Vassilis Gorgoulis, John K Field, and Triantafillos Liloglou. 2009. “Hypomethylation of Retrotransposable Elements Correlates with Genomic Instability in Non-Small Cell Lung Cancer.” *Int J Cancer* 124 (1). United States: 81–87. doi:10.1002/ijc.23849.
- De Luca, Chiara, Fiorella Guadagni, Paola Sinibaldi-Vallebona, Steno Sentinelli, Michele Gallucci, Andreas Hoffmann, Gerald G Schumann, Corrado Spadafora, and Ilaria Sciamanna. 2016. “Enhanced Expression of LINE-1-Encoded ORF2 Protein in Early Stages of Colon and Prostate Transformation.” *Oncotarget* 7 (4): 4048–61. doi:10.18632/oncotarget.6767.
- De Rijck, Jan, Christine de Kogel, Jonas Demeulemeester, Sofie Vets, Sara El Ashkar, Nirav Malani, Frederic D Bushman, et al. 2013. “The BET Family of Proteins Targets Moloney Murine Leukemia Virus Integration Near Transcription Start Sites.” *Cell Rep* 5 (4): 886–94. doi:10.1016/j.celrep.2013.09.040.
- De Rijck, Jan, Koen Bartholomeeusen, Hugo Ceulemans, Zeger Debyser, and Rik Gijssbers. 2010. “High-Resolution Profiling of the LEDGF/P75 Chromatin Interaction in the ENCODE Region.” *Nucleic Acids Research* 38 (18): 6135–47. doi:10.1093/nar/gkq410.
- Deblois, Geneviève, and Vincent Giguère. 2011. “Functional and Physiological Genomics of Estrogen-Related Receptors (ERRs) in Health and Disease.” *Biochimica Et Biophysica Acta* 1812 (8): 1032–40. doi:10.1016/j.bbadis.2010.12.009.
- Deblois, Geneviève, Jacqueline A Hall, Marie-Claude Perry, Josée Laganière, Majid Ghahremani, Morag Park, Michael Hallett, and Vincent Giguère. 2009. “Genome-Wide Identification of Direct Target Genes Implicates Estrogen-Related Receptor Alpha as a Determinant of Breast Cancer Heterogeneity.” *Cancer Res* 69 (15). American Association for Cancer Research: 6149–57. doi:10.1158/0008-5472.CAN-09-1251.
- deHaro, Dawn, Kristine J Kines, Mark Sokolowski, Robert T Dauchy, Vincent A Streva, Steven M Hill, John P Hanifin, George C Brainard, David E Blask, and Victoria P Belancio. 2014. “Regulation of L1 Expression and Retrotransposition by Melatonin and Its Receptor: Implications for Cancer Risk Associated with Light Exposure at Night.” *Nucleic Acids Res* 42 (12): 7694–7707. doi:10.1093/nar/gku503.
- Denli, Ahmet M, Inigo Narvaiza, Bilal E Kerman, Monique Pena, Christopher Benner, Maria C N Marchetto, Jolene K Diedrich, et al. 2015. “Primate-Specific ORF0 Contributes to Retrotransposon-Mediated Diversity.” *Cell* 163 (3). United States: 583–93. doi:10.1016/j.cell.2015.09.025.

- Dewannieux, Marie, and Thierry Heidmann. 2005. "Role of Poly(a) Tail Length in Alu Retrotransposition.." *Genomics* 86 (3): 378–81. doi:10.1016/j.ygeno.2005.05.009.
- Dewannieux, Marie, Cécile Esnault, and Thierry Heidmann. 2003. "LINE-Mediated Retrotransposition of Marked Alu Sequences.." *Nat Genet* 35 (1). United States: 41–48. doi:10.1038/ng1223.
- Dewannieux, Marie, Francis Harper, Aurélien Richaud, Claire Letzelter, David Ribet, Gérard Pierron, and Thierry Heidmann. 2006. "Identification of an Infectious Progenitor for the Multiple-Copy HERV-K Human Endogenous Retroelements.." *Genome Research* 16 (12): 1548–56. doi:10.1101/gr.5565706.
- Dombroski, B A, S L Mathias, E Nanthakumar, A F Scott, and H H Kazazian. 1991. "Isolation of an Active Human Transposable Element.." *Science* 254 (5039). UNITED STATES: 1805–8. doi:10.1126/science.1662412.
- Dong, B, and R H Silverman. 1995. "2-5A-Dependent RNase Molecules Dimerize During Activation by 2-5A.." *The Journal of Biological Chemistry* 270 (8): 4133–37.
- Doucet, Aurélien J, Amy E Hulme, Elodie Sahinovic, Deanna A Kulpa, John B Moldovan, Hui C Kopera, Jyoti N Athanikar, et al. 2010. "Characterization of LINE-1 Ribonucleoprotein Particles.." *PLOS Genetics* 6 (10): e1001150. doi:10.1371/journal.pgen.1001150.
- Doucet, Aurélien J, Gaëtan Droc, Oliver Siol, Jérôme Audoux, and Nicolas Gilbert. 2015. "U6 snRNA Pseudogenes: Markers of Retrotransposition Dynamics in Mammals.." *Mol Biol Evol* 32 (7): 1815–32. doi:10.1093/molbev/msv062.
- Doucet, Aurélien J, Jeremy E Wilusz, Tomoichiro Miyoshi, Ying Liu, and John V Moran. 2015. "A 3' Poly(a) Tract Is Required for LINE-1 Retrotransposition.." *Molecular Cell* 60 (5): 728–41. doi:10.1016/j.molcel.2015.10.012.
- Douglas, Pauline, Shikha Gupta, Nick Morrice, Katheryn Meek, and Susan P Lees-Miller. 2005. "DNA-PK-Dependent Phosphorylation of Ku70/80 Is Not Required for Non-Homologous End Joining." *DNA Repair* 4 (9): 1006–18. doi:10.1016/j.dnarep.2005.05.003.
- Duez, Hélène, and Bart Staels. 2010. "Nuclear Receptors Linking Circadian Rhythms and Cardiometabolic Control.." *Arteriosclerosis, Thrombosis, and Vascular Biology* 30 (8). American Heart Association, Inc.: 1529–34. doi:10.1161/ATVBAHA.110.209098.
- Dufour, Catherine R, Brian J Wilson, Janice M Huss, Daniel P Kelly, William A Alaynick, Michael Downes, Ronald M Evans, Mathieu Blanchette, and Vincent Giguère. 2007. "Genome-Wide Orchestration of Cardiac Functions by the Orphan Nuclear Receptors ERRα and γ." *Cell Metabolism* 5 (5): 345–56. doi:10.1016/j.cmet.2007.03.007.
- Dupressoir, A, C Lavialle, and T Heidmann. 2012. "From Ancestral Infectious Retroviruses to Bona Fide Cellular Genes: Role of the Captured Syncytins in Placentation." *Placenta* 33 (9): 663–71. doi:10.1016/j.placenta.2012.05.005.
- Ehrlund, Anna, and Eckardt Treuter. 2012. "Ligand-Independent Actions of the Orphan Receptors/Corepressors DAX-1 and SHP in Metabolism, Reproduction and Disease.." *The Journal of Steroid Biochemistry and Molecular Biology* 130 (3-5): 169–79. doi:10.1016/j.jsbmb.2011.04.007.
- Eickbush, T H, and V K Jamburuthugoda. 2008. "The Diversity of Retrotransposons and the Properties of Their Reverse Transcriptases.." *Virus Res* 134 (1-2). Department of Biology, University of Rochester, Rochester, NY 14627, USA.: 221–34. doi:10.1016/j.virusres.2007.12.010.
- Esnault, C, J Maestre, and T Heidmann. 2000. "Human LINE Retrotransposons Generate

- Processed Pseudogenes.." *Nat Genet* 24 (4): 363–67. doi:10.1038/74184.
- Eudy, J D, S Yao, M D Weston, M Ma-Edmonds, C B Talmadge, J J Cheng, W J Kimberling, and J Sumegi. 1998. "Isolation of a Gene Encoding a Novel Member of the Nuclear Receptor Superfamily From the Critical Region of Usher Syndrome Type IIa at 1q41.." *Genomics* 50 (3): 382–84. doi:10.1006/geno.1998.5345.
- Evans, Ronald M. 2005. "The Nuclear Receptor Superfamily: a Rosetta Stone for Physiology.." *Molecular Endocrinology* 19 (6): 1429–38. doi:10.1210/me.2005-0046.
- Evans, Ronald M, and David J Mangelsdorf. 2014. "Nuclear Receptors, RXR, and the Big Bang." *Cell* 157 (1). Elsevier Inc.: 255–66. doi:10.1016/j.cell.2014.03.012.
- Evrony, Gilad D, Xuyu Cai, Eunjung Lee, L Benjamin Hills, Princess C Elhosary, Hillel S Lehmann, J J Parker, et al. 2012. "Single-Neuron Sequencing Analysis of L1 Retrotransposition and Somatic Mutation in the Human Brain.." *Cell* 151 (3). United States: 483–96. doi:10.1016/j.cell.2012.09.035.
- Ewing, Adam D, and Haig H Kazazian. 2010. "High-Throughput Sequencing Reveals Extensive Variation in Human-Specific L1 Content in Individual Human Genomes.." *Genome Res* 20 (9): 1262–70. doi:10.1101/gr.106419.110.
- Ewing, Adam D, Anthony Gacita, Laura D Wood, Florence Ma, Dongmei Xing, Min-Sik Kim, Srikanth S Manda, et al. 2015. "Widespread Somatic L1 Retrotransposition Occurs Early During Gastrointestinal Cancer Evolution.." *Genome Res* 25 (10). United States: Cold Spring Harbor Lab: 1536–45. doi:10.1101/gr.196238.115.
- Ewing, Adam D, Tracy J Ballinger, Dent Earl, Christopher C Harris, Li Ding, Richard K Wilson, and David Haussler. 2013. "Retrotransposition of Gene Transcripts Leads to Structural Variation in Mammalian Genomes.." *Genome Biology* 14 (3). England: R22. doi:10.1186/gb-2013-14-3-r22.
- Fanning, Thomas, and Maxine Singer. 1987. "The LINE-1 DNA Sequences in Four Mammalian Orders Predict Proteins That Conserve Homologies to Retrovirus Proteins." *Nucleic Acids Research* 15 (5). Oxford Univ Press: 2251–60. doi:10.1093/nar/15.5.2251.
- Farley, Alexander H, Eline T Luning Prak, and Haig H Kazazian. 2004. "More Active Human L1 Retrotransposons Produce Longer Insertions.." *Nucleic Acids Research* 32 (2): 502–10. doi:10.1093/nar/gkh202.
- Faulkner, Geoffrey J, and Piero Carninci. 2014. "Altruistic Functions for Selfish DNA." *Cell Cycle* 8 (18): 2895–2900. doi:10.4161/cc.8.18.9536.
- Fedoroff, N V. 2012. "McClintock's Challenge in the 21st Century." *Proceedings of the National Academy of Sciences of the United States of America* 109 (50): 20200–20203. doi:10.1073/pnas.1215482109.
- Feng, Bo, Jianming Jiang, Petra Kraus, Jia-Hui Ng, Jian-Chien Dominic Heng, Yun-Shen Chan, Lai-Ping Yaw, et al. 2009. "Reprogramming of Fibroblasts Into Induced Pluripotent Stem Cells with Orphan Nuclear Receptor Esrrb." *Nat Cell Biol* 11 (2). Nature Publishing Group: 197–203. doi:10.1038/ncb1827.
- Feng, Q, J V Moran, H H Kazazian, and J D Boeke. 1996. "Human L1 Retrotransposon Encodes a Conserved Endonuclease Required for Retrotransposition.." *Cell* 87 (5): 905–16. doi:10.1016/s0092-8674(00)81997-2.
- Ferris, Andrea L, Xiaolin Wu, Christina M Hughes, Claudia Stewart, Steven J Smith, Thomas A Milne, Gang G Wang, et al. 2010. "Lens Epithelium-Derived Growth Factor Fusion Proteins Redirect HIV-1 DNA Integration.." *Proceedings of the National Academy of Sciences of the United States of America* 107 (7): 3135–40. doi:10.1073/pnas.0914142107.
- Finlan, Lee E, Duncan Sproul, Inga Thomson, Shelagh Boyle, Elizabeth Kerr, Paul Perry,

- Bauke Ylstra, Jonathan R Chubb, and Wendy A Bickmore. 2008. "Recruitment to the Nuclear Periphery Can Alter Expression of Genes in Human Cells.." *PLOS Genetics* 4 (3): e1000039–13. doi:10.1371/journal.pgen.1000039.
- Fischer, Katharina, Sharon M Kelly, Kate Watt, Nicholas C Price, and Iain J McEwan. 2010. "Conformation of the Mineralocorticoid Receptor N-Terminal Domain: Evidence for Induced and Stable Structure.." *Molecular Endocrinology (Baltimore, Md.)* 24 (10): 1935–48. doi:10.1210/me.2010-0005.
- Fullwood, Melissa J, Mei Hui Liu, You Fu Pan, Jun Liu, Han Xu, Yusoff Bin Mohamed, Yuriy L Orlov, et al. 2009. "An Oestrogen-Receptor-Alpha-Bound Human Chromatin Interactome.." *Nature* 462 (7269). Nature Publishing Group: 58–64. doi:10.1038/nature08497.
- Furano, A V. 2000. "The Biological Properties and Evolutionary Dynamics of Mammalian LINE-1 Retrotransposons.." *Progress in Nucleic Acid Research and Molecular Biology* 64: 255–94.
- Furtak, Vyacheslav, Alok Mulky, Stephen A Rawlings, Lina Kozhaya, KyeongEun Lee, Vineet N Kewalramani, and Derya Unutmaz. 2010. "Perturbation of the P-Body Component Mov10 Inhibits HIV-1 Infectivity.." Edited by Christopher Arendt. *PLoS One* 5 (2): e9081. doi:10.1371/journal.pone.0009081.
- Gaillard, Stéphanie, Mary A Dwyer, and Donald P McDonnell. 2007. "Definition of the Molecular Basis for Estrogen Receptor-Related Receptor-Alpha-Cofactor Interactions.." *Molecular Endocrinology* 21 (1): 62–76. doi:10.1210/me.2006-0179.
- Gallastegui, Nerea, Jonathan A G Mackinnon, Robert J Fletterick, and Eva Estébanez-Perpiñá. 2015. "Advances in Our Structural Understanding of Orphan Nuclear Receptors.." *Trends in Biochemical Sciences* 40 (1). Elsevier Ltd: 25–35. doi:10.1016/j.tibs.2014.11.002.
- Garcia-Perez, Jose L, Aurélien J Doucet, Alain Bucheton, John V Moran, and Nicolas Gilbert. 2007. "Distinct Mechanisms for Trans-Mediated Mobilization of Cellular RNAs by the LINE-1 Reverse Transcriptase.." *Genome Res* 17 (5): 602–11. doi:10.1101/gr.5870107.
- Garcia-Perez, Jose L, Thomas J Widmann, and Ian R Adams. 2016. "The Impact of Transposable Elements on Mammalian Development.." *Development (Cambridge, England)* 143 (22). Oxford University Press for The Company of Biologists Limited: 4101–14. doi:10.1242/dev.132639.
- Garza, Anna M S, Shagufta H Khan, and Raj Kumar. 2010. "Site-Specific Phosphorylation Induces Functionally Active Conformation in the Intrinsically Disordered N-Terminal Activation Function (AF1) Domain of the Glucocorticoid Receptor.." *Molecular and Cellular Biology* 30 (1): 220–30. doi:10.1128/MCB.00552-09.
- Gasior, Stephen L, Astrid M Roy-Engel, and Prescott L Deininger. 2008. "ERCC1/XPF Limits L1 Retrotransposition.." *DNA Repair (Amst)* 7 (6): 983–89. doi:10.1016/j.dnarep.2008.02.006.
- Gasior, Stephen L, Timothy P Wakeman, Bo Xu, and Prescott L Deininger. 2006. "The Human LINE-1 Retrotransposon Creates DNA Double-Strand Breaks.." *J Mol Biol* 357 (5). England: 1383–93. doi:10.1016/j.jmb.2006.01.089.
- Gearhart, Micah D, Signe M A Holmbeck, Ronald M Evans, H Jane Dyson, and Peter E Wright. 2003. "Monomeric Complex of Human Orphan Estrogen Related Receptor-2 with DNA: a Pseudo-Dimer Interface Mediates Extended Half-Site Recognition.." *J Mol Biol* 327 (4): 819–32.
- Gennari, A, B Salvadori, A Tognoni, and P F Conte. 1996. "Rapid Intravenous

- Premedication with Dexamethasone Prevents Hypersensitivity Reactions to Paclitaxel..” *Annals of Oncology : Official Journal of the European Society for Medical Oncology* 7 (9): 978–79.
- Giguere, V, N Yang, P Segui, and R M Evans. 1988. “Identification of a New Class of Steroid Hormone Receptors..” *Nature* 331 (6151). Nature Publishing Group: 91–94. doi:10.1038/331091a0.
- Giguère, Vincent. 2002. “To ERR in the Estrogen Pathway..” *Trends Endocrinol Metab* 13 (5): 220–25.
- Gijbbers, Rik, Keshet Ronen, Sofie Vets, Nirav Malani, Jan De Rijck, Melissa McNeely, Frederic D Bushman, and Zeger Debyser. 2010. “LEDGF Hybrids Efficiently Retarget Lentiviral Integration Into Heterochromatin..” *Molecular Therapy : the Journal of the American Society of Gene Therapy* 18 (3): 552–60. doi:10.1038/mt.2010.36.
- Gilbert, N, S Lutz-Prigge, and J V Moran. 2002. “Genomic Deletions Created Upon LINE-1 Retrotransposition.” *Cell* 110 (3). Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109, USA. gilbertn\@umich.edu: 315–25. doi:10.1016/s0092-8674(02)00828-0.
- Gilbert, Nicolas, Sheila Lutz, Tammy A Morrish, and John V Moran. 2005. “Multiple Fates of L1 Retrotransposition Intermediates in Cultured Human Cells..” *Molecular and Cellular Biology* 25 (17). United States: 7780–95. doi:10.1128/MCB.25.17.7780-7795.2005.
- Gillet, Ludovic C J, and Orlando D Schärer. 2006. “Molecular Mechanisms of Mammalian Global Genome Nucleotide Excision Repair..” *Chemical Reviews* 106 (2). American Chemical Society: 253–76. doi:10.1021/cr040483f.
- Goldstone, David C, Valerie Ennis-Adeniran, Joseph J Hedden, Harriet C T Groom, Gillian I Rice, Evangelos Christodoulou, Philip A Walker, et al. 2011. “HIV-1 Restriction Factor SAMHD1 Is a Deoxynucleoside Triphosphate Triphosphohydrolase..” *Nature* 480 (7377). England: 379–82. doi:10.1038/nature10623.
- Gombart, Adrian F, Tsuyako Saito, and H Phillip Koeffler. 2009. “Exaptation of an Ancient Alu Short Interspersed Element Provides a Highly Conserved Vitamin D-Mediated Innate Immune Response in Humans and Primates..” *BMC Genomics* 10 (1). BioMed Central: 321. doi:10.1186/1471-2164-10-321.
- Goodier, John L. 2016. “Restricting Retrotransposons: a Review..” *Mobile DNA* 7 (1): 16. doi:10.1186/s13100-016-0070-z.
- Goodier, John L, and Haig H Kazazian. 2008. “Retrotransposons Revisited: the Restraint and Rehabilitation of Parasites..” *Cell* 135 (1): 23–35. doi:10.1016/j.cell.2008.09.022.
- Goodier, John L, Gavin C Pereira, Ling E Cheung, Rebecca J Rose, and Haig H Kazazian. 2015. “The Broad-Spectrum Antiviral Protein ZAP Restricts Human Retrotransposition..” *PLOS Genetics* 11 (5): e1005252. doi:10.1371/journal.pgen.1005252.
- Goodier, John L, Lili Zhang, Melissa R Vetter, and Haig H Kazazian. 2007. “LINE-1 ORF1 Protein Localizes in Stress Granules with Other RNA-Binding Proteins, Including Components of RNA Interference RNA-Induced Silencing Complex..” *Molecular and Cellular Biology* 27 (18). United States: 6469–83. doi:10.1128/MCB.00332-07.
- Goodier, John L, Ling E Cheung, and Haig H Kazazian. 2012. “MOV10 RNA Helicase Is a Potent Inhibitor of Retrotransposition in Cells..” *PLOS Genetics* 8 (10). United States: e1002941. doi:10.1371/journal.pgen.1002941.

- Goodier, John L, Prabhat K Mandal, Lili Zhang, and Haig H Kazazian. 2010. "Discrete Subcellular Partitioning of Human Retrotransposon RNAs Despite a Common Mechanism of Genome Insertion.." *Human Molecular Genetics* 19 (9). England: 1712–25. doi:10.1093/hmg/ddq048.
- Govindan, M V, and N Warriar. 1998. "Reconstitution of the N-Terminal Transcription Activation Function of Human Mineralocorticoid Receptor in a Defective Human Glucocorticoid Receptor.." *The Journal of Biological Chemistry* 273 (38): 24439–47.
- Görlich, Dirk, and Ulrike Kutay. 1999. "Transport Between the Cell Nucleus and the Cytoplasm." *Annu Rev Cell Dev Biol* 15 (1): 607–60. doi:10.1146/annurev.cellbio.15.1.607.
- Grawunder, U, D Zimmer, P Kulesza, and M R Lieber. 1998. "Requirement for an Interaction of XRCC4 with DNA Ligase IV for Wild-Type v(D)J Recombination and DNA Double-Strand Break Repair in Vivo.." *The Journal of Biological Chemistry* 273 (38): 24708–14.
- Grieves, Jessica L, Jason M Fye, Scott Harvey, Jason M Grayson, Thomas Hollis, and Fred W Perrino. 2015. "Exonuclease TREX1 Degrades Double-Stranded DNA to Prevent Spontaneous Lupus-Like Inflammatory Disease.." *Proceedings of the National Academy of Sciences of the United States of America* 112 (16): 5117–22. doi:10.1073/pnas.1423804112.
- Gronemeyer, Hinrich, and Dino Moras. 1995. "How to Finger DNA." *Nature* 375 (6528): 190–91. doi:10.1038/375190a0.
- Gupta, Saumya Shree, Tobias Maetzig, Goedele N Maertens, Azar Sharif, Michael Rothe, Magdalena Weidner-Glunde, Melanie Galla, Axel Schambach, Peter Cherepanov, and Thomas F Schulz. 2013. "Bromo- and Extraterminal Domain Chromatin Regulators Serve as Cofactors for Murine Leukemia Virus Integration.." *J Virol* 87 (23): 12721–36. doi:10.1128/JVI.01942-13.
- Hall, J M, and D P McDonnell. 1999. "The Estrogen Receptor Beta-Isoform (ERbeta) of the Human Estrogen Receptor Modulates ERalpha Transcriptional Activity and Is a Key Regulator of the Cellular Response to Estrogens and Antiestrogens.." *Endocrinology* 140 (12): 5566–78. doi:10.1210/endo.140.12.7179.
- Hammarsten, O, L G DeFazio, and G Chu. 2000. "Activation of DNA-Dependent Protein Kinase by Single-Stranded DNA Ends.." *The Journal of Biological Chemistry* 275 (3): 1541–50.
- Hammel, Michal, Yaping Yu, Brandi L Mahaney, Brandon Cai, Ruiqiong Ye, Barry M Phipps, Robert P Rambo, et al. 2010. "Ku and DNA-Dependent Protein Kinase Dynamic Conformations and Assembly Regulate DNA Binding and the Initial Non-Homologous End Joining Complex.." *The Journal of Biological Chemistry* 285 (2): 1414–23. doi:10.1074/jbc.M109.065615.
- Hammes, Stephen R, and Ellis R Levin. 2007. "Extranuclear Steroid Receptors: Nature and Actions.." *Endocrine Reviews* 28 (7): 726–41. doi:10.1210/er.2007-0022.
- Han, Jeffrey S, and Jef D Boeke. 2004. "A Highly Active Synthetic Mammalian Retrotransposon.." *Nature* 429 (6989): 314–18. doi:10.1038/nature02535.
- Hancks, Dustin C, and Haig H Kazazian. 2012. "Active Human Retrotransposons: Variation and Disease.." *Curr Opin Genet Dev* 22 (3). England: 191–203. doi:10.1016/j.gde.2012.02.006.
- Hancks, Dustin C, and Haig H Kazazian. 2016. "Roles for Retrotransposon Insertions in Human Disease.." *Mobile DNA* 7 (1): 9. doi:10.1186/s13100-016-0065-9.
- Hancks, Dustin C, John L Goodier, Prabhat K Mandal, Ling E Cheung, and Haig H Kazazian. 2011. "Retrotransposition of Marked SVA Elements by Human L1s in

- Cultured Cells.” *Hum Mol Genet* 20 (17): 3386–3400. doi:10.1093/hmg/ddr245.
- Hard, T, E Kellenbach, R Boelens, B A Maler, K Dahlman, L P Freedman, J Carlstedt-Duke, K R Yamamoto, J A Gustafsson, and R Kaptein. 1990. “Solution Structure of the Glucocorticoid Receptor DNA-Binding Domain.” *Science* 249 (4965): 157–60. doi:10.1126/science.2115209.
- Harris, C R, A Dewan, A Zupnick, R Normart, A Gabriel, C Prives, A J Levine, and J Hoh. 2009. “P53 Responsive Elements in Human Retrotransposons.” *Oncogene* 28 (44): 3857–65. doi:10.1038/onc.2009.246.
- Heery, D M, E Kalkhoven, S Hoare, and M G Parker. 1997. “A Signature Motif in Transcriptional Co-Activators Mediates Binding to Nuclear Receptors.” *Nature* 387 (6634). Nature Publishing Group: 733–36. doi:10.1038/42750.
- Heldring, Nina, Ashley Pike, Sandra Andersson, Jason Matthews, Guojun Cheng, Johan Hartman, Michel Tujague, et al. 2007. “Estrogen Receptors: How Do They Signal and What Are Their Targets.” *Physiological Reviews* 87 (3): 905–31. doi:10.1152/physrev.00026.2006.
- Henry, Michel, Denise Guétard, Rodolphe Suspène, Christophe Rusniok, Simon Wain-Hobson, and Jean-Pierre Vartanian. 2009. “Genetic Editing of HBV DNA by Monodomain Human APOBEC3 Cytidine Deaminases and the Recombinant Nature of APOBEC3G.” Edited by Peter Sommer. *PLoS One* 4 (1): e4277. doi:10.1371/journal.pone.0004277.
- Heras, Sara R, Sara Macias, Mireya Plass, Noem i Fernandez, David Cano, Eduardo Eyra, Jos e L Garcia-Perez, and Javier F C a ceres. 2013. “The Microprocessor Controls the Activity of Mammalian Retrotransposons.” *Nature Structural & Molecular Biology* 20 (10): 1173–81. doi:10.1038/nsmb.2658.
- Ho, Pak Chung, Ernest Hung Yu Ng, and Oi Shan Tang. 2002. “Mifepristone: Contraceptive and Non-Contraceptive Uses.” *Current Opinion in Obstetrics & Gynecology* 14 (3): 325–30.
- Hogg, J Robert, and Stephen P Goff. 2010. “Upf1 Senses 3'UTR Length to Potentiate mRNA Decay.” *Cell* 143 (3): 379–89. doi:10.1016/j.cell.2010.10.005.
- Hohjoh, H, and M F Singer. 1996. “Cytoplasmic Ribonucleoprotein Complexes Containing Human LINE-1 Protein and RNA.” *The EMBO Journal* 15 (3). Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.: 630–39.
- Hohjoh, H, and M F Singer. 1997. “Sequence-Specific Single-Strand RNA Binding Protein Encoded by the Human LINE-1 Retrotransposon.” *The EMBO Journal* 16 (19). ENGLAND: 6034–43. doi:10.1093/emboj/16.19.6034.
- Hollstein, M, D Sidransky, B Vogelstein, and C C Harris. 1991. “P53 Mutations in Human Cancers.” *Science* 253 (5015): 49–53.
- Holmes, S E, M F Singer, and G D Swergold. 1992. “Studies on P40, the Leucine Zipper Motif-Containing Protein Encoded by the First Open Reading Frame of an Active Human LINE-1 Transposable Element.” *The Journal of Biological Chemistry* 267 (28). UNITED STATES: 19765–68.
- Hong, H, L Yang, and M R Stallcup. 1999. “Hormone-Independent Transcriptional Activation and Coactivator Binding by Novel Orphan Nuclear Receptor ERR3.” *The Journal of Biological Chemistry* 274 (32): 22618–26.
- Horard, B, B Rayet, G Triqueneaux, V Laudet, F Delaunay, and J M Vanacker. 2004. “Expression of the Orphan Nuclear Receptor ERRalpha Is Under Circadian Regulation in Estrogen-Responsive Tissues.” *Journal of Molecular Endocrinology* 33 (1): 87–97.

- Höss, M, P Robins, T J Naven, D J Pappin, J Sgouros, and T Lindahl. 1999. "A Human DNA Editing Enzyme Homologous to the Escherichia Coli DnaQ/MutD Protein.." *The EMBO Journal* 18 (13): 3868–75. doi:10.1093/emboj/18.13.3868.
- Hu, Siqi, Jian Li, Fengwen Xu, Shan Mei, Yann Le Duff, Lijuan Yin, Xiaojing Pang, et al. 2015. "SAMHD1 Inhibits LINE-1 Retrotransposition by Promoting Stress Granule Formation.." *PLOS Genetics* 11 (7). United States: e1005367. doi:10.1371/journal.pgen.1005367.
- Hu, X, and M A Lazar. 1999. "The CoRNR Motif Controls the Recruitment of Corepressors by Nuclear Hormone Receptors.." *Nature* 402 (6757). ENGLAND: 93–96. doi:10.1038/47069.
- Huang, Feng, Junsong Zhang, Yijun Zhang, Guannan Geng, Juanran Liang, Yingniang Li, Jingliang Chen, Chao Liu, and Hui Zhang. 2015. "RNA Helicase MOV10 Functions as a Co-Factor of HIV-1 Rev to Facilitate Rev/RRE-Dependent Nuclear Export of Viral mRNAs.." *Virology* 486 (December): 15–26. doi:10.1016/j.virol.2015.08.026.
- Hughes, Jennifer F, and John M Coffin. 2004. "Human Endogenous Retrovirus K Solo-LTR Formation and Insertional Polymorphisms: Implications for Human and Viral Evolution.." *Proceedings of the National Academy of Sciences of the United States of America* 101 (6): 1668–72. doi:10.1073/pnas.0307885100.
- Hunter, Richard G, Bruce S McEwen, and Donald W Pfaff. 2014. "Environmental Stress and Transposon Transcription in the Mammalian Brain." *Mobile Genetic Elements* 3 (2): e24555–4. doi:10.4161/mge.24555.
- Hunter, Richard G, Khatuna Gagnidze, Bruce S McEwen, and Donald W Pfaff. 2015. "Stress and the Dynamic Genome: Steroids, Epigenetics, and the Transposome.." *Proceedings of the National Academy of Sciences of the United States of America* 112 (22): 6828–33. doi:10.1073/pnas.1411260111.
- Huppunen, Johanna, Gerd Wohlfahrt, and Piia Aarnisalo. 2004. "Requirements for Transcriptional Regulation by the Orphan Nuclear Receptor ERRgamma." *Molecular and Cellular Endocrinology* 219 (1-2): 151–60. doi:10.1016/j.mce.2004.01.002.
- Huss, Janice M, Ken-ichi Imahashi, Catherine R Dufour, Carla J Weinheimer, Michael Courtois, Atilla Kovacs, Vincent Giguère, Elizabeth Murphy, and Daniel P Kelly. 2007. "The Nuclear Receptor ERRα Is Required for the Bioenergetic and Functional Adaptation to Cardiac Pressure Overload." *Cell Metabolism* 6 (1): 25–37. doi:10.1016/j.cmet.2007.06.005.
- Huss, Janice M, Ryan P Kopp, and Daniel P Kelly. 2002. "Peroxisome Proliferator-Activated Receptor Coactivator-1α (PGC-1α) Coactivates the Cardiac-Enriched Nuclear Receptors Estrogen-Related Receptor-A and -Γ." *The Journal of Biological Chemistry* 277 (43): 40265–74. doi:10.1074/jbc.M206324200.
- Huss, Janice M, Wojciech G Garbacz, and Wen Xie. 2015. "Constitutive Activities of Estrogen-Related Receptors: Transcriptional Regulation of Metabolism by the ERR Pathways in Health and Disease.." *Biochimica Et Biophysica Acta* 1852 (9): 1912–27. doi:10.1016/j.bbadis.2015.06.016.
- Ingraham, Holly A, and Matthew R Redinbo. 2005. "Orphan Nuclear Receptors Adopted by Crystallography.." *Current Opinion in Structural Biology* 15 (6): 708–15. doi:10.1016/j.sbi.2005.10.009.
- Ishii, Ken J, Cevayir Coban, Hiroki Kato, Ken Takahashi, Yuichi Torii, Fumihiko Takeshita, Holger Ludwig, et al. 2006. "A Toll-Like Receptor-Independent Antiviral Response Induced by Double-Stranded B-Form DNA.." *Nature Immunology* 7 (1): 40–48. doi:10.1038/ni1282.

- Iskow, Rebecca C, Michael T McCabe, Ryan E Mills, Spencer Torene, W Stephen Pittard, Andrew F Neuwald, Erwin G Van Meir, Paula M Vertino, and Scott E Devine. 2010. "Natural Mutagenesis of Human Genomes by Endogenous Retrotransposons.." *Cell* 141 (7). United States: 1253–61. doi:10.1016/j.cell.2010.05.020.
- Ivics, Z, P B Hackett, R H Plasterk, and Z Izsvak. 1997. "Molecular Reconstruction of Sleeping Beauty, a Tc1-Like Transposon From Fish, and Its Transposition in Human Cells." *Cell* 91 (4). Department of Genetics and Cell Biology, University of Minnesota, St. Paul 55108-1095, USA.: 501–10. doi:10.1016/s0092-8674(00)80436-5.
- Ivics, Zoltán, and Zsuzsanna Izsvák. 2010. "The Expanding Universe of Transposon Technologies for Gene and Cell Engineering." *Mobile DNA* 1 (1). BioMed Central Ltd: 25. doi:10.1186/1759-8753-1-25.
- Jacobs, Frank M J, David Greenberg, Ngan Nguyen, Maximilian Haeussler, Adam D Ewing, Sol Katzman, Benedict Paten, Sofie R Salama, and David Haussler. 2015. "An Evolutionary Arms Race Between KRAB Zinc-Finger Genes ZNF91/93 and SVA/L1 Retrotransposons." *Nature* 516 (7530). Nature Publishing Group: 242–45. doi:10.1038/nature13760.
- Jacobsen, Britta M, Purevsuren Jambal, Stephanie A Schittone, and Kathryn B Horwitz. 2009. "ALU Repeats in Promoters Are Position-Dependent Co-Response Elements (coRE) That Enhance or Repress Transcription by Dimeric and Monomeric Progesterone Receptors.." *Molecular Endocrinology (Baltimore, Md.)* 23 (7): 989–1000. doi:10.1210/me.2009-0048.
- Jacome, Ariana, and Oscar Fernandez-Capetillo. 2011. "Lac Operator Repeats Generate a Traceable Fragile Site in Mammalian Cells.." *EMBO Rep* 12 (10). England: 1032–38. doi:10.1038/embor.2011.158.
- Janicki, Susan M, Toshiro Tsukamoto, Simone E Salghetti, William P Tansey, Ravi Sachidanandam, Kannanganattu V Prasanth, Thomas Ried, et al. 2004. "From Silencing to Gene Expression: Real-Time Analysis in Single Cells.." *Cell* 116 (5): 683–98. doi:10.1016/s0092-8674(04)00171-0.
- Jarmuz, Adam, Ann Chester, Jayne Bayliss, Jane Gisbourne, Ian Dunham, James Scott, and Naveenan Navaratnam. 2002. "An Anthropoid-Specific Locus of Orphan C to U RNA-Editing Enzymes on Chromosome 22.." *Genomics* 79 (3): 285–96. doi:10.1006/geno.2002.6718.
- Jenster, G, H A van der Korput, J Trapman, and A O Brinkmann. 1995. "Identification of Two Transcription Activation Units in the N-Terminal Domain of the Human Androgen Receptor.." *The Journal of Biological Chemistry* 270 (13): 7341–46.
- Jiang, Jianming, Yun-Shen Chan, Yui-Han Loh, Jun Cai, Guo-Qing Tong, Ching-Aeng Lim, Paul Robson, Sheng Zhong, and Huck Hui Ng. 2008. "A Core Klf Circuitry Regulates Self-Renewal of Embryonic Stem Cells.." *Nat Cell Biol* 10 (3). Nature Publishing Group: 353–60. doi:10.1038/ncb1698.
- Johnston, Stephen D, Xuedong Liu, Fengrong Zuo, Theresa L Eisenbraun, Steven R Wiley, Richard J Kraus, and Janet E Mertz. 1997. "Estrogen-Related Receptor A1 Functionally Binds as a Monomer to Extended Half-Site Sequences Including Ones Contained Within Estrogen-Response Elements." *Molecular Endocrinology* 11 (3): 342–52. doi:10.1210/mend.11.3.9897.
- Jurka, J. 1997. "Sequence Patterns Indicate an Enzymatic Involvement in Integration of Mammalian Retrotransposons.." *Proceedings of the National Academy of Sciences of the United States of America* 94 (5): 1872–77.

- Kajikawa, Masaki, and Norihiro Okada. 2002. "LINEs Mobilize SINEs in the Eel Through a Shared 3' Sequence.." *Cell* 111 (3). United States: 433–44. doi:10.1016/s0092-8674(02)01041-3.
- Kallen, Joerg, Jean-Marc Schlaeppli, Francis Bitsch, Ireos Filipuzzi, Alain Schilb, Virginie Riou, Alexander Graham, Andre Strauss, Martin Geiser, and Brigitte Fournier. 2004. "Evidence for Ligand-Independent Transcriptional Activation of the Human Estrogen-Related Receptor Alpha (ERRalpha): Crystal Structure of ERRalpha Ligand Binding Domain in Complex with Peroxisome Proliferator-Activated Receptor Coactivator-1alpha.." *The Journal of Biological Chemistry* 279 (47): 49330–37. doi:10.1074/jbc.M407999200.
- Kano, H, I Godoy, C Courtney, M R Vetter, G L Gerton, E M Ostertag, and H H Jr Kazazian. 2009. "L1 Retrotransposition Occurs Mainly in Embryogenesis and Creates Somatic Mosaicism." *Genes & Development* 23 (11). Department of Genetics, University of Pennsylvania, School of Medicine, Philadelphia, Pennsylvania 19104, USA.: 1303–12. doi:10.1101/gad.1803909.
- Kazazian, H H, C Wong, H Youssoufian, A F Scott, D G Phillips, and S E Antonarakis. 1988. "Haemophilia a Resulting From De Novo Insertion of L1 Sequences Represents a Novel Mechanism for Mutation in Man.." *Nature* 332 (6160). ENGLAND: 164–66. doi:10.1038/332164a0.
- Kazazian, Haig H. 2014. "Processed Pseudogene Insertions in Somatic Cells.." *Mobile DNA* 5 (1). England: 20. doi:10.1186/1759-8753-5-20.
- Keay, June, and Joseph W Thornton. 2009. "Hormone-Activated Estrogen Receptors in Annelid Invertebrates: Implications for Evolution and Endocrine Disruption.." *Endocrinology* 150 (4): 1731–38. doi:10.1210/en.2008-1338.
- Khan, Hameed, Arian Smit, and St e phane Boissinot. 2006. "Molecular Evolution and Tempo of Amplification of Human LINE-1 Retrotransposons Since the Origin of Primates.." *Genome Res* 16 (1). United States: 78–87. doi:10.1101/gr.4001406.
- Khazina, Elena, and Oliver Weichenrieder. 2009. "Non-LTR Retrotransposons Encode Noncanonical RRM Domains in Their First Open Reading Frame.." *Proceedings of the National Academy of Sciences of the United States of America* 106 (3). United States: 731–36. doi:10.1073/pnas.0809964106.
- Khazina, Elena, Vincent Truffault, Regina B u ttner, Steffen Schmidt, Murray Coles, and Oliver Weichenrieder. 2011. "Trimeric Structure and Flexibility of the L1ORF1 Protein in Human L1 Retrotransposition.." *Nature Structural & Molecular Biology* 18 (9). United States: 1006–14. doi:10.1038/nsmb.2097.
- Kim, Baek, Laura A Nguyen, Waaqo Daddacha, and Joseph A Hollenbaugh. 2012. "Tight Interplay Among SAMHD1 Protein Level, Cellular dNTP Levels, and HIV-1 Proviral DNA Synthesis Kinetics in Human Primary Monocyte-Derived Macrophages.." *The Journal of Biological Chemistry* 287 (26): 21570–74. doi:10.1074/jbc.C112.374843.
- Kinomoto, M, T Kanno, M Shimura, Y Ishizaka, A Kojima, T Kurata, T Sata, and K Tokunaga. 2007. "All APOBEC3 Family Proteins Differentially Inhibit LINE-1 Retrotransposition." *Nucleic Acids Research* 35 (9). Department of Pathology, National Institute of Infectious Diseases, Tokyo 162-8640, Japan.: 2955–64. doi:10.1093/nar/gkm181.
- Kirchner, J, C M Connolly, and S B Sandmeyer. 1995. "Requirement of RNA Polymerase III Transcription Factors for in Vitro Position-Specific Integration of a Retroviruslike Element.." *Science* 267 (5203): 1488–91.
- Kobayashi, K, Y Nakahori, M Miyake, K Matsumura, E Kondo-Iida, Y Nomura, M

- Segawa, et al. 1998. "An Ancient Retrotransposal Insertion Causes Fukuyama-Type Congenital Muscular Dystrophy.." *Nature* 394 (6691): 388–92. doi:10.1038/28653.
- Koyama, Takayoshi, Juan Fernando Arias, Yukie Iwabu, Masaru Yokoyama, Hideaki Fujita, Hironori Sato, and Kenzo Tokunaga. 2013. "APOBEC3G Oligomerization Is Associated with the Inhibition of Both Alu and LINE-1 Retrotransposition.." Edited by Chen Liang. *PLoS One* 8 (12): e84228. doi:10.1371/journal.pone.0084228.
- Krishna, T S, X P Kong, S Gary, P M Burgers, and J Kuriyan. 1994. "Crystal Structure of the Eukaryotic DNA Polymerase Processivity Factor PCNA.." *Cell* 79 (7): 1233–43.
- Krokan, Hans E, Finn Drabløs, and Geir Slupphaug. 2002. "Uracil in DNA--Occurrence, Consequences and Repair.." *Oncogene* 21 (58): 8935–48. doi:10.1038/sj.onc.1205996.
- Kubo, Shuji, Maria Del Carmen Selemé, Harris S Soifer, José Luis García Pérez, John V Moran, Haig H Kazazian, and Noriyuki Kasahara. 2006. "L1 Retrotransposition in Nondividing and Primary Human Somatic Cells.." *Proceedings of the National Academy of Sciences of the United States of America* 103 (21): 8036–41. doi:10.1073/pnas.0601954103.
- Kulpa, Deanna A, and John V Moran. 2006. "Cis-Preferential LINE-1 Reverse Transcriptase Activity in Ribonucleoprotein Particles.." *Nature Structural & Molecular Biology* 13 (7). United States: 655–60. doi:10.1038/nsmb1107.
- Kumar, Raj, and Iain J McEwan. 2012. "Allosteric Modulators of Steroid Hormone Receptors: Structural Dynamics and Gene Regulation.." *Endocrine Reviews* 33 (2): 271–99. doi:10.1210/er.2011-1033.
- Kumaran, R Ileng, and David L Spector. 2008. "A Genetic Locus Targeted to the Nuclear Periphery in Living Cells Maintains Its Transcriptional Competence.." *J Cell Biol* 180 (1). Rockefeller University Press: 51–65. doi:10.1083/jcb.200706060.
- Kurokawa, R, J DiRenzo, M Boehm, J Sugarman, B Gloss, M G Rosenfeld, R A Heyman, and C K Glass. 1994. "Regulation of Retinoid Signalling by Receptor Polarity and Allosteric Control of Ligand Binding.." *Nature* 371 (6497): 528–31. doi:10.1038/371528a0.
- Kuwabara, Tomoko, Jenny Hsieh, Alysson Muotri, Gene Yeo, Masaki Warashina, Dieter Chichung Lie, Lynne Moore, Kinichi Nakashima, Makoto Asashima, and Fred H Gage. 2009. "Wnt-Mediated Activation of NeuroD1 and Retro-Elements During Adult Neurogenesis.." *Nat Neurosci* 12 (9). United States: 1097–1105. doi:10.1038/nn.2360.
- Lackey, Lela, Zachary L Demorest, Allison M Land, Judd F Hultquist, William L Brown, and Reuben S Harris. 2012. "APOBEC3B and AID Have Similar Nuclear Import Mechanisms.." *J Mol Biol* 419 (5): 301–14. doi:10.1016/j.jmb.2012.03.011.
- Laguette, Nadine, Bijan Sobhian, Nicoletta Casartelli, Mathieu Ringard, Christine Chable-Bessia, Emmanuel S e g e r a l, Ahmad Yatim, St e phane Emiliani, Olivier Schwartz, and Monsef Benkirane. 2011. "SAMHD1 Is the Dendritic- and Myeloid-Cell-Specific HIV-1 Restriction Factor Counteracted by Vpx.." *Nature* 474 (7353). England: 654–57. doi:10.1038/nature10117.
- Lahouassa, Hichem, Waaqo Daddacha, Henning Hofmann, Diana Ayinde, Eric C Logue, Lo ic Dragin, Nicolin Bloch, et al. 2012. "SAMHD1 Restricts the Replication of Human Immunodeficiency Virus Type 1 by Depleting the Intracellular Pool of Deoxynucleoside Triphosphates.." *Nature Immunology* 13 (3). United States: 223–28. doi:10.1038/ni.2236.
- Lander, E S, L M Linton, B Birren, C Nusbaum, M C Zody, J Baldwin, K Devon, et al. 2001. "Initial Sequencing and Analysis of the Human Genome.." *Nature* 409 (6822):

- 860–921. doi:10.1038/35057062.
- Lavery, Derek N, and Iain J McEwan. 2005. “Structure and Function of Steroid Receptor AF1 Transactivation Domains: Induction of Active Conformations..” *The Biochemical Journal* 391 (Pt 3). Portland Press Limited: 449–64. doi:10.1042/BJ20050872.
- Lavie, Laurence, Esther Maldener, Brook Brouha, Eckart U Meese, and Jens Mayer. 2004. “The Human L1 Promoter: Variable Transcription Initiation Sites and a Major Impact of Upstream Flanking Sequence on Promoter Activity..” *Genome Res* 14 (11). United States: 2253–60. doi:10.1101/gr.2745804.
- Lee, Eunjung, Rebecca Iskow, Lixing Yang, Omer Gokcumen, Psalm Haseley, Lovelace J Luquette, Jens G Lohr, et al. 2012. “Landscape of Somatic Retrotransposition in Human Cancers..” *Science* 337 (6097). United States: 967–71. doi:10.1126/science.1222077.
- Lee, Sung-Ryul, Hyoung-Kyu Kim, In-Sung Song, Jaeboum Youm, Louise Anne Dizon, Seung-Hun Jeong, Tae-Hee Ko, et al. 2013. “Glucocorticoids and Their Receptors: Insights Into Specific Roles in Mitochondria..” *Progress in Biophysics and Molecular Biology* 112 (1-2): 44–54. doi:10.1016/j.pbiomolbio.2013.04.001.
- Lee, Young Nam, and Paul D Bieniasz. 2007. “Reconstitution of an Infectious Human Endogenous Retrovirus..” *PLoS Pathog* 3 (1): e10. doi:10.1371/journal.ppat.0030010.
- Leibold, D M, G D Swergold, M F Singer, R E Thayer, B A Dombroski, and T G Fanning. 1990. “Translation of LINE-1 DNA Elements in Vitro and in Human Cells.” *Proceedings of the National Academy of Sciences of the United States of America* 87 (18). Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892.: 6990–94. doi:10.1073/pnas.87.18.6990.
- Lempiäinen, Harri, and Thanos D Halazonetis. 2009. “Emerging Common Themes in Regulation of PIKKs and PI3Ks..” *The EMBO Journal* 28 (20): 3067–73. doi:10.1038/emboj.2009.281.
- Levin, Ellis R. 2008. “Rapid Signaling by Steroid Receptors..” *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology* 295 (5). American Physiological Society: R1425–30. doi:10.1152/ajpregu.90605.2008.
- Levin, Ellis R, and Stephen R Hammes. 2016. “Nuclear Receptors Outside the Nucleus: Extranuclear Signalling by Steroid Receptors..” *Nature Reviews. Molecular Cell Biology* 17 (12): 783–97. doi:10.1038/nrm.2016.122.
- Li, N, W Zhang, and X Cao. 2000. “Identification of Human Homologue of Mouse IFN-Gamma Induced Protein From Human Dendritic Cells..” *Immunology Letters* 74 (3): 221–24.
- Li, Peng, Juan Du, John L Goodier, Jingwei Hou, Jian Kang, Haig H Kazazian, Ke Zhao, and Xiao-Fang Yu. 2017. “Aicardi-Goutières Syndrome Protein TREX1 Suppresses L1 and Maintains Genome Integrity Through Exonuclease-Independent ORF1p Depletion..” *Nucleic Acids Res* 45 (8): 4619–31. doi:10.1093/nar/gkx178.
- Li, Xiaoyu, Jianyong Zhang, Rui Jia, Vicky Cheng, Xin Xu, Wentao Qiao, Fei Guo, Chen Liang, and Shan Cen. 2013. “The MOV10 Helicase Inhibits LINE-1 Mobility..” *The Journal of Biological Chemistry* 288 (29). United States: 21148–60. doi:10.1074/jbc.M113.465856.
- Lin, Chunru, Liuqing Yang, Bogdan Tanasa, Kasey Hutt, Bong-gun Ju, Kenny Ohgi, Jie Zhang, et al. 2009. “Nuclear Receptor-Induced Chromosomal Proximity and DNA Breaks Underlie Specific Translocations in Cancer..” *Cell* 139 (6): 1069–83. doi:10.1016/j.cell.2009.11.030.

- Lin, Jiandie, Christoph Handschin, and Bruce M Spiegelman. 2005. "Metabolic Control Through the PGC-1 Family of Transcription Coactivators.." *Cell Metabolism* 1 (6): 361–70. doi:10.1016/j.cmet.2005.05.004.
- Lin, Kai-Ti, and Lu-Hai Wang. 2016. "New Dimension of Glucocorticoids in Cancer Treatment." *Steroids* 111 (July): 84–88. doi:10.1016/j.steroids.2016.02.019.
- Lindahl, T, J A Gally, and G M Edelman. 1969. "Properties of Deoxyribonuclease 3 From Mammalian Tissues.." *The Journal of Biological Chemistry* 244 (18): 5014–19.
- Liu, H S, M S Jan, C K Chou, P H Chen, and N J Ke. 1999. "Is Green Fluorescent Protein Toxic to the Living Cells?." *Biochem Biophys Res Commun* 260 (3): 712–17. doi:10.1006/bbrc.1999.0954.
- Luan, D D, M H Korman, J L Jakubczak, and T H Eickbush. 1993. "Reverse Transcription of R2Bm RNA Is Primed by a Nick at the Chromosomal Target Site: a Mechanism for Non-LTR Retrotransposition.." *Cell* 72 (4). Department of Biology, University of Rochester, New York 14627.: 595–605. doi:10.1016/0092-8674(93)90078-5.
- Luisi, B F, W X Xu, Z Otwinowski, L P Freedman, K R Yamamoto, and P B Sigler. 1991. "Crystallographic Analysis of the Interaction of the Glucocorticoid Receptor with DNA.." *Nature* 352 (6335). Nature Publishing Group: 497–505. doi:10.1038/352497a0.
- Luo, Jiangming, Robert Sladek, Jo-Ann Bader, Annie Matthyssen, Janet Rossant, and Vincent Giguère. 1997. "Placental Abnormalities in Mouse Embryos Lacking the Orphan Nuclear Receptor ERR-B." *Nature* 388 (6644). Nature Publishing Group: 778–82. doi:10.1038/42022.
- Macfarlane, Catriona, and Peter Simmonds. 2004. "Allelic Variation of HERV-K(HML-2) Endogenous Retroviral Elements in Human Populations." *J Mol Evol* 59 (5): 642–56. doi:10.1007/s00239-004-2656-1.
- Macia, Angela, Thomas J Widmann, Sara R Heras, Veronica Ayllon, Laura Sanchez, Meriem Benkaddour-Boumzaouad, Martin Muñoz-Lopez, et al. 2017. "Engineered LINE-1 Retrotransposition in Nondividing Human Neurons.." *Genome Res* 27 (3): 335–48. doi:10.1101/gr.206805.116.
- Mader, S, J Y Chen, Z Chen, J White, P Chambon, and H Gronemeyer. 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 (13). European Molecular Biology Organization: 5029–41.
- Maksakova, I A, M T Romanish, L Gagnier, C A Dunn, L N van de Lagemaat, and D L Mager. 2006. "Retroviral Elements and Their Hosts: Insertional Mutagenesis in the Mouse Germ Line." *PLOS Genetics* 2 (1). Terry Fox Laboratory, BC Cancer Agency, Vancouver, British Columbia, Canada.: e2. doi:10.1371/journal.pgen.0020002.
- Malfavon-Borja, Ray, Sara L Sawyer, Lily I Wu, Michael Emerman, and Harmit S Malik. 2013. "An Evolutionary Screen Highlights Canonical and Noncanonical Candidate Antiviral Genes Within the Primate TRIM Gene Family.." *Genome Biology and Evolution* 5 (11): 2141–54. doi:10.1093/gbe/evt163.
- Mangelsdorf, D J, C Thummel, M Beato, P Herrlich, G Schütz, K Umesono, B Blumberg, et al. 1995. "The Nuclear Receptor Superfamily: the Second Decade.." *Cell* 83 (6): 835–39.
- Marchetto, Maria C N, Inigo Narvaiza, Ahmet M Denli, Christopher Benner, Thomas A Lazzarini, Jason L Nathanson, Apu a C M Paquola, et al. 2013. "Differential L1 Regulation in Pluripotent Stem Cells of Humans and Apes.." *Nature* 503 (7477). England: 525–29. doi:10.1038/nature12686.

- Margolis, Ronald N, and Sylvia Christakos. 2010. "The Nuclear Receptor Superfamily of Steroid Hormones and Vitamin D Gene Regulation. an Update.." *Annals of the New York Academy of Sciences* 1192 (1). Blackwell Publishing Inc: 208–14. doi:10.1111/j.1749-6632.2009.05227.x.
- Martin, S L. 1991. "Ribonucleoprotein Particles with LINE-1 RNA in Mouse Embryonal Carcinoma Cells.." *Molecular and Cellular Biology* 11 (9). UNITED STATES: 4804–7. doi:10.1128/mcb.11.9.4804.
- Martin, S L, and F D Bushman. 2001. "Nucleic Acid Chaperone Activity of the ORF1 Protein From the Mouse LINE-1 Retrotransposon." *Molecular and Cellular Biology* 21 (2). Department of Cellular and Structural Biology, University of Colorado School of Medicine, Denver, Colorado 80262, USA. sandy.martin@uchsc.edu: 467–75. doi:10.1128/mcb.21.2.467-475.2001.
- Martin, S L, D Branciforte, D Keller, and D L Bain. 2003. "Trimeric Structure for an Essential Protein in L1 Retrotransposition." *Proceedings of the National Academy of Sciences of the United States of America* 100 (24). Department of Cell and Developmental Biology and Program in Molecular Biology, University of Colorado School of Medicine, 4200 East Ninth Avenue, Denver, CO 80262, USA. sandy.martin@uchsc.edu: 13815–20. doi:10.1073/pnas.2336221100.
- Mathias, S L, A F Scott, H H Kazazian, J D Boeke, and A Gabriel. 1991. "Reverse Transcriptase Encoded by a Human Transposable Element.." *Science* 254 (5039). UNITED STATES: 1808–10. doi:10.1126/science.1722352.
- Mazur, Alexey K. 2001. "Molecular Dynamics of Minimal B-DNA." *Journal of Computational Chemistry* 22 (4): 457–67. doi:10.1002/1096-987X(200103)22:4<457::AID-JCC1016>3.0.CO;2-H.
- McCLINTOCK, B. 1950. "The Origin and Behavior of Mutable Loci in Maize.." *Proceedings of the National Academy of Sciences of the United States of America* 36 (6): 344–55.
- McKenna, Neil J, Rainer B Lanz, and Bert W O'Malley. 1999. "Nuclear Receptor Coregulators: Cellular and Molecular Biology 1." *Endocrine Reviews* 20 (3): 321–44. doi:10.1210/edrv.20.3.0366.
- McNally, J G, W G Müller, D Walker, R Wolford, and G L Hager. 2000. "The Glucocorticoid Receptor: Rapid Exchange with Regulatory Sites in Living Cells.." *Science* 287 (5456): 1262–65.
- Mehta, Anuja, and James E Haber. 2014. "Sources of DNA Double-Strand Breaks and Models of Recombinational DNA Repair.." *Cold Spring Harb Perspect Biol* 6 (9): a016428–28. doi:10.1101/cshperspect.a016428.
- Meijsing, Sebastiaan H, Miles A Pufall, Alex Y So, Darren L Bates, Lin Chen, and Keith R Yamamoto. 2009. "DNA Binding Site Sequence Directs Glucocorticoid Receptor Structure and Activity.." *Science* 324 (5925): 407–10. doi:10.1126/science.1164265.
- Meister, Gunter, Markus Landthaler, Lasse Peters, Po Yu Chen, Henning Urlaub, Reinhard Lührmann, and Thomas Tuschl. 2005. "Identification of Novel Argonaute-Associated Proteins.." *Curr Biol* 15 (23): 2149–55. doi:10.1016/j.cub.2005.10.048.
- Meyer, Marc-Etienne, Hinrich Gronemeyer, Bernard Turcotte, Marie-Thérèse Bocquel, Diane Tasset, and Pierre Chambon. 1989. "Steroid Hormone Receptors Compete for Factors That Mediate Their Enhancer Function." *Cell* 57 (3): 433–42. doi:10.1016/0092-8674(89)90918-5.
- Mifsud, Karen R, and Johannes M H M Reul. 2016. "Acute Stress Enhances Heterodimerization and Binding of Corticosteroid Receptors at Glucocorticoid Target Genes in the Hippocampus.." *Proceedings of the National Academy of Sciences of the*

- United States of America* 113 (40): 11336–41. doi:10.1073/pnas.1605246113.
- Miki, Y, I Nishisho, A Horii, Y Miyoshi, J Utsunomiya, K W Kinzler, B Vogelstein, and Y Nakamura. 1992. “Disruption of the APC Gene by a Retrotransposal Insertion of L1 Sequence in a Colon Cancer..” *Cancer Res* 52 (3). Department of Biochemistry, Cancer Institute, Tokyo, Japan.: 643–45.
- Mills, Ryan E, E Andrew Bennett, Rebecca C Iskow, and Scott E Devine. 2007. “Which Transposable Elements Are Active in the Human Genome?.” *Trends Genet* 23 (4). England: 183–91. doi:10.1016/j.tig.2007.02.006.
- Minakami, R, K Kurose, K Etoh, Y Furuhashi, M Hattori, and Y Sakaki. 1992. “Identification of an Internal Cis-Element Essential for the Human L1 Transcription and a Nuclear Factor(S) Binding to the Element..” *Nucleic Acids Research* 20 (12): 3139–45. doi:10.1093/nar/20.12.3139.
- Mir, Ashfaq A, Claude Philippe, and Gael Cristofari. 2015. “euL1db: the European Database of L1HS Retrotransposon Insertions in Humans..” *Nucleic Acids Research* 43 (Database issue): D43–D47. doi:10.1093/nar/gku1043.
- Mitra, Rupak, Xianghong Li, Aurélie Kapusta, David Mayhew, Robi D Mitra, Cédric Feschotte, and Nancy L Craig. 2013. “Functional Characterization of piggyBat From the Bat *Myotis Lucifugus* Unveils an Active Mammalian DNA Transposon..” *Proceedings of the National Academy of Sciences of the United States of America* 110 (1): 234–39. doi:10.1073/pnas.1217548110.
- Moggs, Jonathan G, and George Orphanides. 2001. “Estrogen Receptors: Orchestrators of Pleiotropic Cellular Responses.” *EMBO Reports* 2 (9). EMBO Press: 775–81. doi:10.1093/embo-reports/kve185.
- Moldovan, George-Lucian L, Boris Pfander, and Stefan Jentsch. 2007. “PCNA, the Maestro of the Replication Fork..” *Cell* 129 (4). United States: 665–79. doi:10.1016/j.cell.2007.05.003.
- Moldovan, John B, and John V Moran. 2015. “The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and Alu Retrotransposition.” Edited by Harmit S Malik. *PLOS Genetics* 11 (5): e1005121–34. doi:10.1371/journal.pgen.1005121.
- Monot, Cl e ment, Monika Kuciak, S e bastien Viollet, Ashfaq Ali Mir, Caroline Gabus, Jean-Luc L Darlix, and Ga e l Cristofari. 2013. “The Specificity and Flexibility of L1 Reverse Transcription Priming at Imperfect T-Tracts..” *PLoS Genetics* 9 (5). United States: e1003499. doi:10.1371/journal.pgen.1003499.
- Mooslehner, K, U Müller, U Karls, L Hamann, and K Harbers. 1991. “Structure and Expression of a Gene Encoding a Putative GTP-Binding Protein Identified by Provirus Integration in a Transgenic Mouse Strain..” *Molecular and Cellular Biology* 11 (2): 886–93.
- Morales, Jos e F, Elizabeth T Snow, and John P Murnane. 2002. “Environmental Factors Affecting Transcription of the Human L1 Retrotransposon. I. Steroid Hormone-Like Agents..” *Mutagenesis* 17 (3). England: 193–200. doi:10.1093/mutage/17.3.193.
- Moran, J V, S E Holmes, T P Naas, R J DeBerardinis, J D Boeke, and H H Kazazian. 1996. “High Frequency Retrotransposition in Cultured Mammalian Cells..” *Cell* 87 (5). UNITED STATES: 917–27. doi:10.1016/s0092-8674(00)81998-4.
- Morrish, Tammy A, José Luis Garcia-Perez, Thomas D Stamato, Guillermo E Taccioli, JoAnn Sekiguchi, and John V Moran. 2007. “Endonuclease-Independent LINE-1 Retrotransposition at Mammalian Telomeres..” *Nature* 446 (7132): 208–12. doi:10.1038/nature05560.
- Morrish, Tammy A, Nicolas Gilbert, Jeremy S Myers, Bethaney J Vincent, Thomas D Stamato, Guillermo E Taccioli, Mark A Batzer, and John V Moran. 2002. “DNA

- Repair Mediated by Endonuclease-Independent LINE-1 Retrotransposition.." *Nat Genet* 31 (2). United States: 159–65. doi:10.1038/ng898.
- Muckenfuss, H, M Hamdorf, U Held, M Perkovic, J Lower, K Cichutek, E Flory, G G Schumann, and C Munk. 2006. "APOBEC3 Proteins Inhibit Human LINE-1 Retrotransposition." *The Journal of Biological Chemistry* 281 (31). Division of Medical Biotechnology, Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, D-63225 Langen, Germany.: 22161–72. doi:10.1074/jbc.m601716200.
- Mullican, S E, J R DiSpirito, and M A Lazar. 2013. "The Orphan Nuclear Receptors at Their 25-Year Reunion." *Journal of Molecular Endocrinology* 51 (3): T115–40. doi:10.1530/JME-13-0212.
- Muotri, Alysso R, Maria C N Marchetto, Nicole G Coufal, Ruth Oefner, Gene Yeo, Kinichi Nakashima, and Fred H Gage. 2010. "L1 Retrotransposition in Neurons Is Modulated by MeCP2.." *Nature* 468 (7322): 443–46. doi:10.1038/nature09544.
- Nagy, L, H Y Kao, J D Love, C Li, E Banayo, J T Gooch, V Krishna, K Chatterjee, R M Evans, and J W Schwabe. 1999. "Mechanism of Corepressor Binding and Release From Nuclear Hormone Receptors.." *Genes & Development* 13 (24). Cold Spring Harbor Laboratory Press: 3209–16.
- Nagy, Laszlo, and John W R Schwabe. 2004. "Mechanism of the Nuclear Receptor Molecular Switch.." *Trends in Biochemical Sciences* 29 (6): 317–24. doi:10.1016/j.tibs.2004.04.006.
- Nakagawa, Masato, Michiyo Koyanagi, Koji Tanabe, Kazutoshi Takahashi, Tomoko Ichisaka, Takashi Aoi, Keisuke Okita, Yuji Mochiduki, Nanako Takizawa, and Shinya Yamanaka. 2007. "Generation of Induced Pluripotent Stem Cells Without Myc From Mouse and Human Fibroblasts." *Nat Biotechnol* 26 (1). Nature Publishing Group: 101–6. doi:10.1038/nbt1374.
- Nelson, C C, S C Hendy, R J Shukin, H Cheng, N Bruchovsky, B F Koop, and P S Rennie. 1999. "Determinants of DNA Sequence Specificity of the Androgen, Progesterone, and Glucocorticoid Receptors: Evidence for Differential Steroid Receptor Response Elements.." *Molecular Endocrinology* 13 (12): 2090–2107. doi:10.1210/mend.13.12.0396.
- Niewiadomska, Anna Maria, Chunjuan Tian, Lindi Tan, Tao Wang, Phuong Thi Nguyen Sarkis, and Xiao-Fang F Yu. 2007. "Differential Inhibition of Long Interspersed Element 1 by APOBEC3 Does Not Correlate with High-Molecular-Mass-Complex Formation or P-Body Association.." *J Virol* 81 (17). United States: 9577–83. doi:10.1128/JVI.02800-06.
- Nigumann, Pilvi, Kaja Redik, Kert M a tlik, and Mart Speek. 2002. "Many Human Genes Are Transcribed From the Antisense Promoter of L1 Retrotransposon.." *Genomics* 79 (5). United States: 628–34. doi:10.1006/geno.2002.6758.
- Nisole, Sébastien, Clare Lynch, Jonathan P Stoye, and Melvyn W Yap. 2004. "A Trim5-Cyclophilin a Fusion Protein Found in Owl Monkey Kidney Cells Can Restrict HIV-1.." *Proceedings of the National Academy of Sciences of the United States of America* 101 (36): 13324–28. doi:10.1073/pnas.0404640101.
- Nuclear Receptors Nomenclature Committee. 1999. "A Unified Nomenclature System for the Nuclear Receptor Superfamily.." *Cell* 97 (2): 161–63.
- O'Mahony, Fiona, Mahnaz Razandi, Ali Pedram, Brian J Harvey, and Ellis R Levin. 2012. "Estrogen Modulates Metabolic Pathway Adaptation to Available Glucose in Breast Cancer Cells.." *Molecular Endocrinology (Baltimore, Md.)* 26 (12): 2058–70. doi:10.1210/me.2012-1191.
- Ogino, Shuji, Katsuhiko Noshio, Gregory J Kirkner, Takako Kawasaki, Andrew T Chan,

- Eva S Schernhammer, Edward L Giovannucci, and Charles S Fuchs. 2008. "A Cohort Study of Tumoral LINE-1 Hypomethylation and Prognosis in Colon Cancer.." *J Natl Cancer Inst* 100 (23): 1734–38. doi:10.1093/jnci/djn359.
- Ohshima, Takayuki, Hiroshi Koga, and Kunitada Shimotohno. 2004. "Transcriptional Activity of Peroxisome Proliferator-Activated Receptor Gamma Is Modulated by SUMO-1 Modification.." *The Journal of Biological Chemistry* 279 (28): 29551–57. doi:10.1074/jbc.M403866200.
- Okabe, Yasutaka, Kohki Kawane, Shizuo Akira, Tadatsugu Taniguchi, and Shigekazu Nagata. 2005. "Toll-Like Receptor-Independent Gene Induction Program Activated by Mammalian DNA Escaped From Apoptotic DNA Degradation.." *J Exp Med* 202 (10): 1333–39. doi:10.1084/jem.20051654.
- Ostertag, E M, and H H Kazazian. 2001. "Twin Priming: a Proposed Mechanism for the Creation of Inversions in L1 Retrotransposition.." *Genome Res* 11 (12). United States: 2059–65. doi:10.1101/gr.205701.
- Ostertag, Eric M, John L Goodier, Yue Zhang, and Haig H Kazazian. 2003. "SVA Elements Are Nonautonomous Retrotransposons That Cause Disease in Humans.." *Am J Hum Genet* 73 (6): 1444–51. doi:10.1086/380207.
- Pace, John K, and Cédric Feschotte. 2007. "The Evolutionary History of Human DNA Transposons: Evidence for Intense Activity in the Primate Lineage.." *Genome Research* 17 (4): 422–32. doi:10.1101/gr.5826307.
- Pak, Vladimir, Gisela Heidecker, Vinay K Pathak, and David Derse. 2011. "The Role of Amino-Terminal Sequences in Cellular Localization and Antiviral Activity of APOBEC3B.." *J Virol* 85 (17): 8538–47. doi:10.1128/JVI.02645-10.
- Pawlak, Michal, Philippe Lefebvre, and Bart Staels. 2012. "General Molecular Biology and Architecture of Nuclear Receptors.." *Current Topics in Medicinal Chemistry* 12 (6). Insm: 486–504.
- Payne, Bryony T I, Ingeborg C van Knippenberg, Hazel Bell, Sergio R Filipe, David J Sherratt, and Peter McGlynn. 2006. "Replication Fork Blockage by Transcription Factor-DNA Complexes in Escherichia Coli." *Nucleic Acids Research* 34 (18): 5194–5202. doi:10.1093/nar/gkl682.
- Perepelitsa-Belancio, Victoria, and Prescott Deininger. 2003. "RNA Truncation by Premature Polyadenylation Attenuates Human Mobile Element Activity.." *Nat Genet* 35 (4): 363–66. doi:10.1038/ng1269.
- Perissi, V, L M Staszewski, E M McInerney, R Kurokawa, A Krones, D W Rose, M H Lambert, M V Milburn, C K Glass, and M G Rosenfeld. 1999. "Molecular Determinants of Nuclear Receptor-Corepressor Interaction.." *Genes & Development* 13 (24). Cold Spring Harbor Laboratory Press: 3198–3208.
- Pettersson, K, K Svensson, R Mattsson, B Carlsson, R Ohlsson, and A Berkenstam. 1996. "Expression of a Novel Member of Estrogen Response Element-Binding Nuclear Receptors Is Restricted to the Early Stages of Chorion Formation During Mouse Embryogenesis.." *Mechanisms of Development* 54 (2): 211–23.
- Philippe, Claude, Dulce B Vargas-Landin, Aurélien J Doucet, Dominic van Essen, Jorge Vera-Otarola, Monika Kuciak, Antoine Corbin, Pilvi Nigumann, and Gael Cristofari. 2016. "Activation of Individual L1 Retrotransposon Instances Is Restricted to Cell-Type Dependent Permissive Loci.." *eLife* 5 (March): 166. doi:10.7554/eLife.13926.
- Pillai, Ramesh S, and Shinichiro Chuma. 2012. "piRNAs and Their Involvement in Male Germline Development in Mice.." *Development, Growth & Differentiation* 54 (1). Blackwell Publishing Ltd: 78–92. doi:10.1111/j.1440-169X.2011.01320.x.
- Poukka, H, U Karvonen, O A Janne, and J J Palvimo. 2000. "Covalent Modification of

- the Androgen Receptor by Small Ubiquitin-Like Modifier 1 (SUMO-1).” *Proceedings of the National Academy of Sciences of the United States of America* 97 (26): 14145–50. doi:10.1073/pnas.97.26.14145.
- Pourcet, Benoit, Inés Pineda-Torra, Bruno Derudas, Bart Staels, and Corine Glineur. 2010. “SUMOylation of Human Peroxisome Proliferator-Activated Receptor Alpha Inhibits Its Trans-Activity Through the Recruitment of the Nuclear Corepressor NCoR.” *The Journal of Biological Chemistry* 285 (9): 5983–92. doi:10.1074/jbc.M109.078311.
- Powell, Rebecca D, Paul J Holland, Thomas Hollis, and Fred W Perrino. 2011. “Aicardi-Goutieres Syndrome Gene and HIV-1 Restriction Factor SAMHD1 Is a dGTP-Regulated Deoxynucleotide Triphosphohydrolase.” *The Journal of Biological Chemistry* 286 (51): 43596–600. doi:10.1074/jbc.C111.317628.
- Qi, Xiaojie, Kenneth Daily, Kim Nguyen, Haoyi Wang, David Mayhew, Paul Rigor, Sholeh Forouzan, et al. 2012. “Retrotransposon Profiling of RNA Polymerase III Initiation Sites.” *Genome Res* 22 (4): 681–92. doi:10.1101/gr.131219.111.
- Raiz, Julija, Annette Damert, Sergiu Chira, Ulrike Held, Sabine Klawitter, Matthias Hamdorf, Johannes L o wer, Wolf H Str a tling, Roswitha L o wer, and Gerald G Schumann. 2012. “The Non-Autonomous Retrotransposon SVA Is Trans-Mobilized by the Human LINE-1 Protein Machinery.” *Nucleic Acids Research* 40 (4). England: 1666–83. doi:10.1093/nar/gkr863.
- Rangwala, Shamina M, Xiaoyan Li, Loren Lindsley, Xiaomei Wang, Stacey Shaughnessy, Thomas G Daniels, Joseph Szustakowski, N R Nirmala, Zhidan Wu, and Susan C Stevenson. 2007. “Estrogen-Related Receptor Alpha Is Essential for the Expression of Antioxidant Protection Genes and Mitochondrial Function.” *Biochem Biophys Res Commun* 357 (1): 231–36. doi:10.1016/j.bbrc.2007.03.126.
- Rastinejad, F, T Perlmann, R M Evans, and P B Sigler. 1995. “Structural Determinants of Nuclear Receptor Assembly on DNA Direct Repeats.” *Nature* 375 (6528): 203–11. doi:10.1038/375203a0.
- Rastinejad, Fraydoon, Pengxiang Huang, Vikas Chandra, and Sepideh Khorasanizadeh. 2013. “Understanding Nuclear Receptor Form and Function Using Structural Biology.” *Journal of Molecular Endocrinology* 51 (3). BioScientifica: T1–T21. doi:10.1530/JME-13-0173.
- Rebollo, Rita, Mark T Romanish, and Dixie L Mager. 2012. “Transposable Elements: an Abundant and Natural Source of Regulatory Sequences for Host Genes.” *Annual Review of Genetics* 46 (1): 21–42. doi:10.1146/annurev-genet-110711-155621.
- Reddy, K L, J M Zullo, E Bertolino, and H Singh. 2008. “Transcriptional Repression Mediated by Repositioning of Genes to the Nuclear Lamina.” *Nature* 452 (7184): 243–47. doi:10.1038/nature06727.
- Renaud, J P, N Rochel, M Ruff, V Vivat, P Chambon, H Gronemeyer, and D Moras. 1995. “Crystal Structure of the RAR-Gamma Ligand-Binding Domain Bound to All-Trans Retinoic Acid.” *Nature* 378 (6558): 681–89. doi:10.1038/378681a0.
- Revollo, Javier R, and John A Cidlowski. 2009. “Mechanisms Generating Diversity in Glucocorticoid Receptor Signaling.” *Annals of the New York Academy of Sciences* 1179 (1). Blackwell Publishing Inc: 167–78. doi:10.1111/j.1749-6632.2009.04986.x.
- Rice, Gillian I, Jacquelyn Bond, Aruna Asipu, Rebecca L Brunette, Iain W Manfield, Ian M Carr, Jonathan C Fuller, et al. 2009. “Mutations Involved in Aicardi-Goutières Syndrome Implicate SAMHD1 as Regulator of the Innate Immune Response.” *Nature Genetics* 41 (7): 829–32. doi:10.1038/ng.373.
- Richardson, Sandra R, Aurélien J Doucet, Huira C Kopera, John B Moldovan, José Luis

- Garcia-Perez, and John V Moran. 2015. "The Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes.." *Microbiology Spectrum* 3 (2): MDNA3-0061-2014. doi:10.1128/microbiolspec.MDNA3-0061-2014.
- Richardson, Sandra R, Inigo Narvaiza, Randy A Planegger, Matthew D Weitzman, and John V Moran. 2014. "APOBEC3A Deaminates Transiently Exposed Single-Strand DNA During LINE-1 Retrotransposition.." *Elife* 3. United States: e02008. doi:10.7554/elife.02008.
- Roberts, Ff. 1950. "Stress and the General Adaptation Syndrome." *Bmj* 2 (4670): 104–5. doi:10.1136/bmj.2.4670.104-a.
- Rodić, Nemanja, Reema Sharma, Rajni Sharma, John Zampella, Lixin Dai, Martin S Taylor, Ralph H Hruban, et al. 2014. "Long Interspersed Element-1 Protein Expression Is a Hallmark of Many Human Cancers.." *Am J Pathol* 184 (5): 1280–86. doi:10.1016/j.ajpath.2014.01.007.
- Roemer, Sarah C, Douglas C Donham, Lori Sherman, Vickie H Pon, Dean P Edwards, and Mair E A Churchill. 2006. "Structure of the Progesterone Receptor-Deoxyribonucleic Acid Complex: Novel Interactions Required for Binding to Half-Site Response Elements.." *Molecular Endocrinology* 20 (12): 3042–52. doi:10.1210/me.2005-0511.
- Rollins, R A. 2005. "Large-Scale Structure of Genomic Methylation Patterns." *Genome Research* 16 (2): 157–63. doi:10.1101/gr.4362006.
- Ross, Robert J, Molly M Weiner, and Haifan Lin. 2014. "PIWI Proteins and PIWI-Interacting RNAs in the Soma.." *Nature* 505 (7483). Nature Research: 353–59. doi:10.1038/nature12987.
- Rubin, Carol M, Catherine M Houck, Prescott L Deininger, Theodore Friedmann, and Carl W Schmid. 1980. "Partial Nucleotide Sequence of the 300-Nucleotide Interspersed Repeated Human DNA Sequences." *Nature* 284 (5754): 372–74. doi:10.1038/284372a0.
- Saaltink, Dirk-Jan, and Erno Vreugdenhil. 2014. "Stress, Glucocorticoid Receptors, and Adult Neurogenesis: a Balance Between Excitation and Inhibition?." *Cell Mol Life Sci* 71 (13): 2499–2515. doi:10.1007/s00018-014-1568-5.
- Saito, Kenichiro, Kazuyuki Kawakami, Isao Matsumoto, Makoto Oda, Go Watanabe, and Toshinari Minamoto. 2010. "Long Interspersed Nuclear Element 1 Hypomethylation Is a Marker of Poor Prognosis in Stage IA Non-Small Cell Lung Cancer.." *Clin Cancer Res* 16 (8). United States: 2418–26. doi:10.1158/1078-0432.CCR-09-2819.
- Sanyal, Sabyasachi, Joon-Young Kim, Han-Jong Kim, Jun Takeda, Yoon-Kwang Lee, David D Moore, and Hueng-Sik Choi. 2002. "Differential Regulation of the Orphan Nuclear Receptor Small Heterodimer Partner(SHP) Gene Promoter by Orphan Nuclear Receptor ERR Isoforms." *The Journal of Biological Chemistry* 277 (3). American Society for Biochemistry and Molecular Biology: 1739–48. doi:10.1074/jbc.M106140200.
- Sayah, David M, Elena Sokolskaja, Lionel Berthoux, and Jeremy Luban. 2004. "Cyclophilin a Retrotransposition Into TRIM5 Explains Owl Monkey Resistance to HIV-1.." *Nature* 430 (6999): 569–73. doi:10.1038/nature02777.
- Schrijvers, Rik, Sofie Vets, Jan De Rijck, Nirav Malani, Frederic D Bushman, Zeger Debyser, and Rik Gijssbers. 2012. "HRP-2 Determines HIV-1 Integration Site Selection in LEDGF/P75 Depleted Cells.." *Retrovirology* 9 (1). BioMed Central: 84. doi:10.1186/1742-4690-9-84.
- Schwabe, J W, L Chapman, J T Finch, and D Rhodes. 1993. "The Crystal Structure of the Estrogen Receptor DNA-Binding Domain Bound to DNA: How Receptors

- Discriminate Between Their Response Elements.." *Cell* 75 (3): 567–78.
- Scott, Emma C, and Scott E Devine. 2017. "The Role of Somatic L1 Retrotransposition in Human Cancers.." *Viruses* 9 (6): 131. doi:10.3390/v9060131.
- Scott, Emma C, Eugene J Gardner, Ashiq Masood, Nelson T Chuang, Paula M Vertino, and Scott E Devine. 2016. "A Hot L1 Retrotransposon Evades Somatic Repression and Initiates Human Colorectal Cancer.." *Genome Res* 26 (6): 745–55. doi:10.1101/gr.201814.115.
- Sen, Shurjo K, Charles T Huang, Kyudong Han, and Mark A Batzer. 2007. "Endonuclease-Independent Insertion Provides an Alternative Pathway for L1 Retrotransposition in the Human Genome.." *Nucleic Acids Research* 35 (11). England: 3741–51. doi:10.1093/nar/gkm317.
- Seol, W, B Hanstein, M Brown, and D D Moore. 1998. "Inhibition of Estrogen Receptor Action by the Orphan Receptor SHP (Short Heterodimer Partner).." *Molecular Endocrinology* 12 (10): 1551–57. doi:10.1210/mend.12.10.0184.
- Sharma, Amit, Ross C Larue, Matthew R Plumb, Nirav Malani, Frances Male, Alison Slaughter, Jacques J Kessl, et al. 2013. "BET Proteins Promote Efficient Murine Leukemia Virus Integration at Transcription Start Sites.." *Proceedings of the National Academy of Sciences of the United States of America* 110 (29): 12036–41. doi:10.1073/pnas.1307157110.
- Sheehy, Ann M, Nathan C Gaddis, Jonathan D Choi, and Michael H Malim. 2002. "Isolation of a Human Gene That Inhibits HIV-1 Infection and Is Suppressed by the Viral Vif Protein.." *Nature* 418 (6898): 646–50. doi:10.1038/nature00939.
- Shigeoka, Toshiaki, Sayaka Kato, Masashi Kawaichi, and Yasumasa Ishida. 2012. "Evidence That the Upf1-Related Molecular Motor Scans the 3'-UTR to Ensure mRNA Integrity.." *Nucleic Acids Research* 40 (14): 6887–97. doi:10.1093/nar/gks344.
- Shiloh, Y. 2001. "ATM and ATR: Networking Cellular Responses to DNA Damage.." *Curr Opin Genet Dev* 11 (1): 71–77.
- Shukla, Ruchi, Kyle R Upton, Martin Muñoz-Lopez, Daniel J Gerhardt, Malcolm E Fisher, Thu Nguyen, Paul M Brennan, et al. 2013. "Endogenous Retrotransposition Activates Oncogenic Pathways in Hepatocellular Carcinoma.." *Cell* 153 (1): 101–11. doi:10.1016/j.cell.2013.02.032.
- Sijbers, A M, W L de Laat, R R Ariza, M Biggerstaff, Y F Wei, J G Moggs, K C Carter, et al. 1996. "Xeroderma Pigmentosum Group F Caused by a Defect in a Structure-Specific DNA Repair Endonuclease.." *Cell* 86 (5): 811–22.
- Silverman, Robert H. 2007. "Viral Encounters with 2',5'-Oligoadenylate Synthetase and RNase L During the Interferon Antiviral Response.." *J Virol* 81 (23): 12720–29. doi:10.1128/JVI.01471-07.
- Silvers, Robert M, Johanna A Smith, Michael Schowalter, Samuel Litwin, Zhihui Liang, Kyla Geary, and René Daniel. 2010. "Modification of Integration Site Preferences of an HIV-1-Based Vector by Expression of a Novel Synthetic Protein.." *Human Gene Therapy* 21 (3). Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA: 337–49. doi:10.1089/hum.2009.134.
- Simental, J A, M Sar, M V Lane, F S French, and E M Wilson. 1991. "Transcriptional Activation and Nuclear Targeting Signals of the Human Androgen Receptor.." *The Journal of Biological Chemistry* 266 (1): 510–18.
- Simoes, Davina C M, Anna-Maria G Psarra, Thais Mauad, Ioanna Pantou, Charis Roussos, Constantine E Sekeris, and Christina Gratiou. 2012. "Glucocorticoid and Estrogen Receptors Are Reduced in Mitochondria of Lung Epithelial Cells in

- Asthma.” Edited by Annabel Valledor. *PLoS One* 7 (6). Public Library of Science: e39183. doi:10.1371/journal.pone.0039183.
- Singh, Parmit Kumar, Matthew R Plumb, Andrea L Ferris, James R Iben, Xiaolin Wu, Hind J Fadel, Brian T Luke, et al. 2015. “LEDGF/P75 Interacts with mRNA Splicing Factors and Targets HIV-1 Integration to Highly Spliced Genes.” *Genes & Development* 29 (21): 2287–97. doi:10.1101/gad.267609.115.
- Siol, Oliver, Moustapha Boutliliss, Thanh Chung, Gernot Glöckner, Theodor Dinger, and Thomas Winckler. 2006. “Role of RNA Polymerase III Transcription Factors in the Selection of Integration Sites by the Dictyostelium Non-Long Terminal Repeat Retrotransposon TRE5-a.” *Molecular and Cellular Biology* 26 (22): 8242–51. doi:10.1128/MCB.01348-06.
- Sladek, R, J A Bader, and V Giguere. 1997. “The Orphan Nuclear Receptor Estrogen-Related Receptor Alpha Is a Transcriptional Regulator of the Human Medium-Chain Acyl Coenzyme a Dehydrogenase Gene.” *Molecular and Cellular Biology* 17 (9). American Society for Microbiology (ASM): 5400–5409.
- Smit, A F, G T o th, A D Riggs, and J Jurka. 1995. “Ancestral, Mammalian-Wide Subfamilies of LINE-1 Repetitive Sequences.” *J Mol Biol* 246 (3). ENGLAND: 401–17. doi:10.1006/jmbi.1994.0095.
- Smit-McBride, Z. 1994. “DNA Sequence Specificity of the v-Erb a Oncoprotein/Thyroid Hormone Receptor: Role of the P-Box and Its Interaction with More N-Terminal Determinants of DNA Recognition.” *Molecular Endocrinology* 8 (7): 819–28. doi:10.1210/me.8.7.819.
- Solyom, Szilvia, Adam D Ewing, Eric P Rahrmann, Tara Doucet, Heather H Nelson, Michael B Burns, Reuben S Harris, et al. 2012. “Extensive Somatic L1 Retrotransposition in Colorectal Tumors.” *Genome Res* 22 (12): 2328–38. doi:10.1101/gr.145235.112.
- Sonoda, J, J Laganier, I R Mehl, G D Barish, L W Chong, X Li, I E Scheffler, et al. 2007. “Nuclear Receptor ERR and Coactivator PGC-1beta Are Effectors of IFN- - Induced Host Defense.” *Genes & Development* 21 (15): 1909–20. doi:10.1101/gad.1553007.
- Soutoglou, Evi, and Tom Misteli. 2008. “Activation of the Cellular DNA Damage Response in the Absence of DNA Lesions.” *Science* 320 (5882): 1507–10. doi:10.1126/science.1159051.
- Spagnolo, Laura, Angel Rivera-Calzada, Laurence H Pearl, and Oscar Llorca. 2006. “Three-Dimensional Structure of the Human DNA-PKcs/Ku70/Ku80 Complex Assembled on DNA and Its Implications for DNA DSB Repair.” *Molecular Cell* 22 (4): 511–19. doi:10.1016/j.molcel.2006.04.013.
- Speak, M. 2001. “Antisense Promoter of Human L1 Retrotransposon Drives Transcription of Adjacent Cellular Genes.” *Molecular and Cellular Biology* 21 (6): 1973–85. doi:10.1128/MCB.21.6.1973-1985.2001.
- Stark, George R, Ian M Kerr, Bryan R G Williams, Robert H Silverman, and Robert D Schreiber. 1998. “How Cells Respond to Interferons.” *Annu Rev Biochem* 67 (1): 227–64. doi:10.1146/annurev.biochem.67.1.227.
- Stenglein, Mark D, and Reuben S Harris. 2006. “APOBEC3B and APOBEC3F Inhibit L1 Retrotransposition by a DNA Deamination-Independent Mechanism.” *The Journal of Biological Chemistry* 281 (25). United States: 16837–41. doi:10.1074/jbc.M602367200.
- Stenglein, Mark D, Hiroshi Matsuo, and Reuben S Harris. 2008. “Two Regions Within the Amino-Terminal Half of APOBEC3G Cooperate to Determine Cytoplasmic

- Localization..” *J Virol* 82 (19): 9591–99. doi:10.1128/JVI.02471-07.
- Stetson, Daniel B, Joan S Ko, Thierry Heidmann, and Ruslan Medzhitov. 2008. “Trex1 Prevents Cell-Intrinsic Initiation of Autoimmunity..” *Cell* 134 (4). United States: 587–98. doi:10.1016/j.cell.2008.06.032.
- Streva, Vincent A, Vallmer E Jordan, Sara Linker, Dale J Hedges, Mark A Batzer, and Prescott L Deininger. 2015. “Sequencing, Identification and Mapping of Primed L1 Elements (SIMPLE) Reveals Significant Variation in Full Length L1 Elements Between Individuals..” *BMC Genomics* 16 (1): 220. doi:10.1186/s12864-015-1374-y.
- Sultana, Tania, Alessia Zamborlini, Gael Cristofari, and Pascale Lesage. 2017. “Integration Site Selection by Retroviruses and Transposable Elements in Eukaryotes..” *Nat Rev Genet* 18 (5): 292–308. doi:10.1038/nrg.2017.7.
- Sun, L H, and F R Frankel. 1986. “The Induction of Alu-Sequence Transcripts by Glucocorticoid in Rat Liver Cells..” *Journal of Steroid Biochemistry* 25 (2): 201–7.
- Suzuki, Jun, Katsumi Yamaguchi, Masaki Kajikawa, Kenji Ichiyangi, Noritaka Adachi, Hideki Koyama, Shunichi Takeda, and Norihiro Okada. 2009. “Genetic Evidence That the Non-Homologous End-Joining Repair Pathway Is Involved in LINE Retrotransposition..” *PLoS Genetics* 5 (4). United States: e1000461. doi:10.1371/journal.pgen.1000461.
- Suzuki, Sadako, Shigekazu Sasaki, Hiroshi Morita, Yutaka Oki, Daisuke Turiya, Takeshi Ito, Hiroko Misawa, Keiko Ishizuka, and Hirotohi Nakamura. 2010. “The Role of the Amino-Terminal Domain in the Interaction of Unliganded Peroxisome Proliferator-Activated Receptor Gamma-2 with Nuclear Receptor Co-Repressor..” *Journal of Molecular Endocrinology* 45 (3): 133–45. doi:10.1677/JME-10-0007.
- Swergold, G D. 1990. “Identification, Characterization, and Cell Specificity of a Human LINE-1 Promoter..” *Molecular and Cellular Biology* 10 (12). UNITED STATES: 6718–29. doi:10.1128/mcb.10.12.6718.
- Symer, David E, Carla Connelly, Suzanne T Szak, Emerita M Caputo, Gregory J Cost, Giovanni Parmigiani, and Jef D Boeke. 2002. “Human L1 Retrotransposition Is Associated with Genetic Instability in Vivo..” *Cell* 110 (3): 327–38. doi:10.1016/s0092-8674(02)00839-5.
- Takimoto, Glenn S, Lin Tung, Hany Abdel-Hafiz, Michael G Abel, Carol A Sartorius, Jennifer K Richer, Britta M Jacobsen, David L Bain, and Kathryn B Horwitz. 2003. “Functional Properties of the N-Terminal Region of Progesterone Receptors and Their Mechanistic Relationship to Structure..” *The Journal of Steroid Biochemistry and Molecular Biology* 85 (2-5): 209–19.
- Tanenbaum, D M, Y Wang, S P Williams, and P B Sigler. 1998. “Crystallographic Comparison of the Estrogen and Progesterone Receptor's Ligand Binding Domains..” *Proceedings of the National Academy of Sciences of the United States of America* 95 (11). National Academy of Sciences: 5998–6003.
- Taylor, Martin S, John LaCava, Paolo Mita, Kelly R Molloy, Cheng Ran Lisa Huang, Donghui Li, Emily M Adney, et al. 2013. “Affinity Proteomics Reveals Human Host Factors Implicated in Discrete Stages of LINE-1 Retrotransposition.” *Cell* 155 (5). Elsevier: 1034–48. doi:10.1016/j.cell.2013.10.021.
- Tchénio, T, J F Casella, and T Heidmann. 2000. “Members of the SRY Family Regulate the Human LINE Retrotransposons..” *Nucleic Acids Research* 28 (2): 411–15. doi:10.1093/nar/28.2.411.
- Teng, C T. 1992. “Differential Molecular Mechanism of the Estrogen Action That Regulates Lactoferrin Gene in Human and Mouse.” *Molecular Endocrinology* 6 (11): 1969–81. doi:10.1210/me.6.11.1969.

- Terasaki, Natsuko, John L Goodier, Ling E Cheung, Yue J Wang, Masaki Kajikawa, Haig H Kazazian, and Norihiro Okada. 2013. "In Vitro Screening for Compounds That Enhance Human L1 Mobilization.." *PLoS One* 8 (9). United States: e74629. doi:10.1371/journal.pone.0074629.
- Thomas, Charles A, Apu a C M Paquola, and Alysson R Muotri. 2012. "LINE-1 Retrotransposition in the Nervous System.." *Annu Rev Cell Dev Biol* 28 (1). United States: 555–73. doi:10.1146/annurev-cellbio-101011-155822.
- Thomas, Charles A, Leon Tejwani, Cleber A Trujillo, Priscilla D Negraes, Roberto H Herai, Pinar Mesci, Angela Macia, Yanick J Crow, and Alysson R Muotri. 2017. "Modeling of TREX1-Dependent Autoimmune Disease Using Human Stem Cells Highlights L1 Accumulation as a Source of Neuroinflammation.." *Cell Stem Cell* 21 (3): 319–331.e8. doi:10.1016/j.stem.2017.07.009.
- Thondamal, Manjunatha, Michael Witting, Philippe Schmitt-Kopplin, and Hugo Aguilaniu. 2014. "Steroid Hormone Signalling Links Reproduction to Lifespan in Dietary-Restricted *Caenorhabditis Elegans*.." *Nature Communications* 5 (September). Nature Publishing Group: 4879. doi:10.1038/ncomms5879.
- Togashi, Marie, Sabine Borngraeber, Ben Sandler, Robert J Fletterick, Paul Webb, and John D Baxter. 2005. "Conformational Adaptation of Nuclear Receptor Ligand Binding Domains to Agonists: Potential for Novel Approaches to Ligand Design.." *The Journal of Steroid Biochemistry and Molecular Biology* 93 (2-5): 127–37. doi:10.1016/j.jsbmb.2005.01.004.
- Torrents, David, Mikita Suyama, Evgeny Zdobnov, and Peer Bork. 2003. "A Genome-Wide Survey of Human Pseudogenes.." *Genome Research* 13 (12): 2559–67. doi:10.1101/gr.1455503.
- Tremblay. 2007. "The NR3B Subgroup: an overERRview.." *Nuclear Receptor Signaling* 4. doi:10.1621/nrs.05009.
- Tremblay, Annie M, Brian J Wilson, Xiang-Jiao Yang, and Vincent Giguère. 2008. "Phosphorylation-Dependent Sumoylation Regulates Estrogen-Related Receptor- α and - γ Transcriptional Activity Through a Synergy Control Motif.." *Molecular Endocrinology* 22 (3): 570–84. doi:10.1210/me.2007-0357.
- Tremblay, Annie M, Catherine R Dufour, Majid Ghahremani, Timothy L Reudelhuber, and Vincent Giguère. 2010. "Physiological Genomics Identifies Estrogen-Related Receptor α as a Regulator of Renal Sodium and Potassium Homeostasis and the Renin-Angiotensin Pathway.." *Molecular Endocrinology (Baltimore, Md.)* 24 (1): 22–32. doi:10.1210/me.2009-0254.
- Tubio, Jose M C, Yilong Li, Young Seok Ju, Inigo Martincorena, Susanna L Cooke, Marta Tojo, Gunes Gundem, et al. 2014. "Mobile DNA in Cancer. Extensive Transduction of Nonrepetitive DNA Mediated by L1 Retrotransposition in Cancer Genomes.." *Science* 345 (6196): 1251343–43. doi:10.1126/science.1251343.
- Ullu, E, and C Tschudi. 1984. "Alu Sequences Are Processed 7SL RNA Genes.." *Nature* 312 (5990): 171–72.
- Ulrich, Helle D. 2011. "Timing and Spacing of Ubiquitin-Dependent DNA Damage Bypass.." *FEBS Letters* 585 (18): 2861–67. doi:10.1016/j.febslet.2011.05.028.
- Umesono, K, K K Murakami, C C Thompson, and R M Evans. 1991. "Direct Repeats as Selective Response Elements for the Thyroid Hormone, Retinoic Acid, and Vitamin D3 Receptors.." *Cell* 65 (7): 1255–66.
- Upton, Kyle R, Daniel J Gerhardt, J Samuel Jesuadian, Sandra R Richardson, Francisco J Sánchez-Luque, Gabriela O Bodea, Adam D Ewing, et al. 2015. "Ubiquitous L1 Mosaicism in Hippocampal Neurons.." *Cell* 161 (2): 228–39.

- doi:10.1016/j.cell.2015.03.026.
- van den Hurk, J A, I C Meij, M C Seleme, H Kano, K Nikopoulos, L H Hoefsloot, E A Siermans, et al. 2007. "L1 Retrotransposition Can Occur Early in Human Embryonic Development." *Hum Mol Genet* 16 (13). Department of Human Genetics, Radboud University Nijmegen Medical Centre, 6500 HB Nijmegen, The Netherlands.: 1587–92. doi:10.1093/hmg/ddm108.
- van Hoesel, Anneke Q, Cornelis J H van de Velde, Peter J K Kuppen, Gerrit Jan Liefers, Hein Putter, Yusuke Sato, David A Elashoff, et al. 2012. "Hypomethylation of LINE-1 in Primary Tumor Has Poor Prognosis in Young Breast Cancer Patients: a Retrospective Cohort Study." *Breast Cancer Res Treat* 134 (3): 1103–14. doi:10.1007/s10549-012-2038-0.
- Vanacker, J M. 1999. "Transcriptional Targets Shared by Estrogen Receptor- Related Receptors (ERRs) and Estrogen Receptor (ER) Alpha , but Not by ERbeta." *The EMBO Journal* 18 (15). EMBO Press: 4270–79. doi:10.1093/emboj/18.15.4270.
- Vanin, E F. 1984. "Processed Pseudogenes. Characteristics and Evolution.." *Biochimica Et Biophysica Acta* 782 (3): 231–41.
- Villena, Josep A, M Benjamin Hock, William Y Chang, Joanalyn E Barcas, Vincent Giguère, and Anastasia Kralli. 2007. "Orphan Nuclear Receptor Estrogen-Related Receptor Alpha Is Essential for Adaptive Thermogenesis.." *Proceedings of the National Academy of Sciences of the United States of America* 104 (4). National Acad Sciences: 1418–23. doi:10.1073/pnas.0607696104.
- Viollot, S e bastien, Cl e ment Monot, and Ga e l Cristofari. 2014. "L1 Retrotransposition: the Snap-Velcro Model and Its Consequences.." *Mobile Genetic Elements* 4 (1): e28907. doi:10.4161/mge.28907.
- Voss, Ty C, and Gordon L Hager. 2013. "Dynamic Regulation of Transcriptional States by Chromatin and Transcription Factors." *Nature Publishing Group* 15 (2). Nature Publishing Group: 69–81. doi:10.1038/nrg3623.
- Vu, Elizabeth H, Richard J Kraus, and Janet E Mertz. 2007. "Phosphorylation-Dependent Sumoylation of Estrogen-Related Receptor Alpha1.." *Biochemistry* 46 (34): 9795–9804. doi:10.1021/bi700316g.
- Walker, J R, R A Corpina, and J Goldberg. 2001. "Structure of the Ku Heterodimer Bound to DNA and Its Implications for Double-Strand Break Repair.." *Nature* 412 (6847): 607–14. doi:10.1038/35088000.
- Wallace, Nicholas A, Stephen L Gasior, Zachary J Faber, Heather L Howie, Prescott L Deininger, and Denise A Galloway. 2013. "HPV 5 and 8 E6 Expression Reduces ATM Protein Levels and Attenuates LINE-1 Retrotransposition.." *Virology* 443 (1). United States: 69–79. doi:10.1016/j.virol.2013.04.022.
- Wang, Chenguang, Maofu Fu, Ruth H Angeletti, Linda Siconolfi-Baez, Anne T Reutens, Chris Albanese, Michael P Lisanti, et al. 2001. "Direct Acetylation of the Estrogen Receptor A Hinge Region by P300 Regulates Transactivation and Hormone Sensitivity." *The Journal of Biological Chemistry* 276 (21): 18375–83. doi:10.1074/jbc.M100800200.
- Wang, Hao, Kellie A Jurado, Xiaolin Wu, Ming-Chieh Shun, Xiang Li, Andrea L Ferris, Steven J Smith, et al. 2012. "HRP2 Determines the Efficiency and Specificity of HIV-1 Integration in LEDGF/P75 Knockout Cells but Does Not Contribute to the Antiviral Activity of a Potent LEDGF/P75-Binding Site Integrase Inhibitor.." *Nucleic Acids Research* 40 (22): 11518–30. doi:10.1093/nar/gks913.
- Wang, Hui, Jinchuan Xing, Deepak Grover, Dale J Hedges, Kyudong Han, Jerilyn A Walker, and Mark A Batzer. 2005. "SVA Elements: a Hominid-Specific Retroposon

- Family.” *Journal of Molecular Biology* 354 (4): 994–1007. doi:10.1016/j.jmb.2005.09.085.
- Wang, Xiaojun, Yanxing Han, Ying Dang, William Fu, Tao Zhou, Roger G Ptak, and Yong-Hui Zheng. 2010. “Moloney Leukemia Virus 10 (MOV10) Protein Inhibits Retrovirus Replication.” *The Journal of Biological Chemistry* 285 (19): 14346–55. doi:10.1074/jbc.M110.109314.
- Wang, Zhulun, Gérard Benoit, Jinsong Liu, Srividya Prasad, Piia Aarnisalo, Xiaohong Liu, Haoda Xu, Nigel P C Walker, and Thomas Perlmann. 2003. “Structure and Function of Nurrl Identifies a Class of Ligand-Independent Nuclear Receptors.” *Nature* 423 (6939): 555–60. doi:10.1038/nature01645.
- Webb, Paul, Ngoc-Ha Nguyen, Grazia Chiellini, Hikari A I Yoshihara, Suzana T Cunha Lima, James W Apriletti, Ralff C J Ribeiro, et al. 2002. “Design of Thyroid Hormone Receptor Antagonists From First Principles.” *The Journal of Steroid Biochemistry and Molecular Biology* 83 (1-5): 59–73.
- Wei, W, N Gilbert, S L Ooi, J F Lawler, E M Ostertag, H H Kazazian, J D Boeke, and J V Moran. 2001. “Human L1 Retrotransposition: Cis Preference Versus Trans Complementation.” *Molecular and Cellular Biology* 21 (4). Department of Human Genetics, The University of Michigan Medical School, Ann Arbor, Michigan 48109, USA.: 1429–39. doi:10.1128/MCB.21.4.1429-1439.2001.
- Wei, W, T A Morrish, R S Alisch, and J V Moran. 2000. “A Transient Assay Reveals That Cultured Human Cells Can Accommodate Multiple LINE-1 Retrotransposition Events.” *Anal Biochem* 284 (2): 435–38. doi:10.1006/abio.2000.4675.
- Wicker, Thomas, Fran c ois Sabot, Aur e lie Hua-Van, Jeffrey L Bennetzen, Pierre Capy, Boulos Chalhoub, Andrew Flavell, et al. 2007. “A Unified Classification System for Eukaryotic Transposable Elements.” *Nat Rev Genet* 8 (12). England: 973–82. doi:10.1038/nrg2165.
- Wilson, Brian J, Annie M Tremblay, Geneviève Deblois, Guillaume Sylvain-Drolet, and Vincent Giguère. 2010. “An Acetylation Switch Modulates the Transcriptional Activity of Estrogen-Related Receptor Alpha.” *Molecular Endocrinology (Baltimore, Md.)* 24 (7): 1349–58. doi:10.1210/me.2009-0441.
- Wilson, T E, T J Fahrner, and J Milbrandt. 1993. “The Orphan Receptors NGFI-B and Steroidogenic Factor 1 Establish Monomer Binding as a Third Paradigm of Nuclear Receptor-DNA Interaction.” *Molecular and Cellular Biology* 13 (9): 5794–5804. doi:10.1128/MCB.13.9.5794.
- Wissing, Silke, Mauricio Montano, José Luis Garcia-Perez, John V Moran, and Warner C Greene. 2011. “Endogenous APOBEC3B Restricts LINE-1 Retrotransposition in Transformed Cells and Human Embryonic Stem Cells.” *The Journal of Biological Chemistry* 286 (42). United States: 36427–37. doi:10.1074/jbc.M111.251058.
- Witherspoon, D J, E E Marchani, W S Watkins, C T Ostler, S P Wooding, B A Anders, J D Fowlkes, et al. 2006. “Human Population Genetic Structure and Diversity Inferred From Polymorphic L1(LINE-1) and Alu Insertions.” *Human Heredity* 62 (1): 30–46. doi:10.1159/000095851.
- Wylie, Annika, Amanda E Jones, Alejandro D'Brot, Wan-Jin Lu, Paula Kurtz, John V Moran, Dinesh Rakheja, et al. 2016. “P53 Genes Function to Restrain Mobile Elements.” *Genes & Development* 30 (1): 64–77. doi:10.1101/gad.266098.115.
- Xie, Chang-Qing, Yangsik Jeong, Mingui Fu, Angie L Bookout, Minerva T Garcia-Barrio, Tingwan Sun, Bong-Hyun Kim, et al. 2009. “Expression Profiling of Nuclear Receptors in Human and Mouse Embryonic Stem Cells.” *Molecular Endocrinology (Baltimore, Md.)* 23 (5): 724–33. doi:10.1210/me.2008-0465.

- Xie, W, H Hong, N N Yang, R J Lin, C M Simon, M R Stallcup, and R M Evans. 1999. "Constitutive Activation of Transcription and Binding of Coactivator by Estrogen-Related Receptors 1 and 2.." *Molecular Endocrinology* 13 (12): 2151–62. doi:10.1210/mend.13.12.0381.
- Xie, W, X Gai, Y Zhu, D C Zappulla, R Sternglanz, and D F Voytas. 2001. "Targeting of the Yeast Ty5 Retrotransposon to Silent Chromatin Is Mediated by Interactions Between Integrase and Sir4p.." *Molecular and Cellular Biology* 21 (19): 6606–14. doi:10.1128/MCB.21.19.6606-6614.2001.
- Xie, Yi, Lajos Mates, Zoltán Ivics, Zsuzsanna Izsvák, Sandra L Martin, and Wenfeng An. 2013. "Cell Division Promotes Efficient Retrotransposition in a Stable L1 Reporter Cell Line.." *Mobile DNA* 4 (1): 10. doi:10.1186/1759-8753-4-10.
- Xiong, Y, and T H Eickbush. 1990. "Origin and Evolution of Retroelements Based Upon Their Reverse Transcriptase Sequences.." *The EMBO Journal* 9 (10). European Molecular Biology Organization: 3353–62.
- Xu, H Eric, Thomas B Stanley, Valerie G Montana, Millard H Lambert, Barry G Shearer, Jeffery E Cobb, David D McKee, et al. 2002. "Structural Basis for Antagonist-Mediated Recruitment of Nuclear Co-Repressors by PPARalpha.." *Nature* 415 (6873). Nature Publishing Group: 813–17. doi:10.1038/415813a.
- Xu, H, P Zhang, L Liu, and M Y Lee. 2001. "A Novel PCNA-Binding Motif Identified by the Panning of a Random Peptide Display Library.." *Biochemistry* 40 (14): 4512–20.
- Yan, Catherine T, Cristian Boboila, Ellen Kris Souza, Sonia Franco, Thomas R Hickernell, Michael Murphy, Sunil Gumaste, et al. 2007. "IgH Class Switching and Translocations Use a Robust Non-Classical End-Joining Pathway.." *Nature* 449 (7161): 478–82. doi:10.1038/nature06020.
- Yang, Nengyu, Hiroyuki Shigeta, Huiping Shi, and Christina T Teng. 1996. "Estrogen-Related Receptor, hERR1, Modulates Estrogen Receptor-Mediated Response of Human Lactoferrin Gene Promoter.." *The Journal of Biological Chemistry* 271 (10): 5795–5804. doi:10.1074/jbc.271.10.5795.
- Yang, Nuo, and Haig H Kazazian. 2006. "L1 Retrotransposition Is Suppressed by Endogenously Encoded Small Interfering RNAs in Human Cultured Cells.." *Nature Structural & Molecular Biology* 13 (9). United States: 763–71. doi:10.1038/nsmb1141.
- Yang, Nuo, Lin Zhang, Yue Zhang, and Haig H Kazazian. 2003. "An Important Role for RUNX3 in Human L1 Transcription and Retrotransposition.." *Nucleic Acids Research* 31 (16). England: 4929–40. doi:10.1093/nar/gkg663.
- Yang, X, K A Lamia, and R M Evans. 2007. "Nuclear Receptors, Metabolism, and the Circadian Clock.." *Cold Spring Harbor Symposia on Quantitative Biology* 72 (1). Cold Spring Harbor Laboratory Press: 387–94. doi:10.1101/sqb.2007.72.058.
- Yang, Xiaoyong, Michael Downes, Ruth T Yu, Angie L Bookout, Weimin He, Marty Straume, David J Mangelsdorf, and Ronald M Evans. 2006. "Nuclear Receptor Expression Links the Circadian Clock to Metabolism.." *Cell* 126 (4): 801–10. doi:10.1016/j.cell.2006.06.050.
- Yoder, J A, C P Walsh, and T H Bestor. 1997. "Cytosine Methylation and the Ecology of Intragenomic Parasites.." *Trends Genet* 13 (8). ENGLAND: 335–40. doi:10.1016/s0168-9525(97)01181-5.
- Yohn, Chris T, Zhaoshi Jiang, Sean D McGrath, Karen E Hayden, Philipp Khaitovich, Matthew E Johnson, Marla Y Eichler, et al. 2005. "Lineage-Specific Expansions of Retroviral Insertions Within the Genomes of African Great Apes but Not Humans

- and Orangutans..” Edited by Mike Tristem. *PLoS Biol* 3 (4): e110. doi:10.1371/journal.pbio.0030110.
- Yu, Yaping, Brandi L Mahaney, Ken-Ichi Yano, Ruiqiong Ye, Shujuan Fang, Pauline Douglas, David J Chen, and Susan P Lees-Miller. 2008. “DNA-PK and ATM Phosphorylation Sites in XLF/Cernunnos Are Not Required for Repair of DNA Double Strand Breaks..” *DNA Repair (Amst)* 7 (10): 1680–92. doi:10.1016/j.dnarep.2008.06.015.
- Zhan, Yan-yan, Yan Chen, Qian Zhang, Jia-jia Zhuang, Min Tian, Hang-zi Chen, Lian-ru Zhang, et al. 2012. “The Orphan Nuclear Receptor Nur77 Regulates LKB1 Localization and Activates AMPK..” *Nature Chemical Biology* 8 (11): 897–904. doi:10.1038/nchembio.1069.
- Zhang, Ao, Beihua Dong, Aurélien J Doucet, John B Moldovan, John V Moran, and Robert H Silverman. 2014. “RNase L Restricts the Mobility of Engineered Retrotransposons in Cultured Human Cells..” *Nucleic Acids Research* 42 (6): 3803–20. doi:10.1093/nar/gkt1308.
- Zhang, Yanqiao, Heidi R Kast-Woelbern, and Peter A Edwards. 2003. “Natural Structural Variants of the Nuclear Receptor Farnesoid X Receptor Affect Transcriptional Activation..” *The Journal of Biological Chemistry* 278 (1): 104–10. doi:10.1074/jbc.M209505200.
- Zhang, Z. 2003. “Millions of Years of Evolution Preserved: a Comprehensive Catalog of the Processed Pseudogenes in the Human Genome.” *Genome Research* 13 (12): 2541–58. doi:10.1101/gr.1429003.
- Zhang, Zhiping, and Christina T Teng. 2000. “Estrogen Receptor-Related Receptor A1 Interacts with Coactivator and Constitutively Activates the Estrogen Response Elements of the Human Lactoferrin Gene.” *The Journal of Biological Chemistry* 275 (27): 20837–46. doi:10.1074/jbc.M001880200.
- Zhang, Zhiping, and Christina T Teng. 2001. “Estrogen Receptor A and Estrogen Receptor-Related Receptor A1 Compete for Binding and Coactivator.” *Molecular and Cellular Endocrinology* 172 (1-2): 223–33. doi:10.1016/S0303-7207(00)00372-5.
- Zhao, R, K Gish, M Murphy, Y Yin, D Notterman, W H Hoffman, E Tom, D H Mack, and A J Levine. 2000. “Analysis of P53-Regulated Gene Expression Patterns Using Oligonucleotide Arrays..” *Genes & Development* 14 (8): 981–93.
- Zhu, Yunxia, Junbiao Dai, Peter G Fuerst, and Daniel F Voytas. 2003. “Controlling Integration Specificity of a Yeast Retrotransposon..” *Proceedings of the National Academy of Sciences of the United States of America* 100 (10): 5891–95. doi:10.1073/pnas.1036705100.
- Zingler, Nora, Ute Willhoeft, Hans-Peter P Brose, Volker Schoder, Thomas Jahns, Kay-Martin O M Hanschmann, Tammy A Morrish, Johannes L o wer, and Gerald G Schumann. 2005. “Analysis of 5‘ Junctions of Human LINE-1 and Alu Retrotransposons Suggests an Alternative Model for 5’-End Attachment Requiring Microhomology-Mediated End-Joining..” *Genome Res* 15 (6). United States: 780–89. doi:10.1101/gr.3421505.
- Zou, S, and D F Voytas. 1997. “Silent Chromatin Determines Target Preference of the *Saccharomyces* Retrotransposon Ty5..” *Proceedings of the National Academy of Sciences of the United States of America* 94 (14). National Academy of Sciences: 7412–16.
- Zwart, Wilbert, Renée de Leeuw, Mariska Rondaij, Jacques Neefjes, Michael A Mancini, and Rob Michalides. 2010. “The Hinge Region of the Human Estrogen Receptor

References

Determines Functional Synergy Between AF-1 and AF-2 in the Quantitative Response to Estradiol and Tamoxifen.." *Journal of Cell Science* 123 (Pt 8): 1253–61. doi:10.1242/jcs.061135.