Genomic Analysis of the Nuclear Receptor Family: New Insights into Structure, Regulation, and Evolution

Zhengdong Zhang 1, Paula E. Burch 1, Austin J. Cooney 2, Rainer B. Lanz 2, Fred A. Pereira 2,3, Jiaqian Wu 1,
Richard A. Gibbs 1, George Weinstock 1,4, David A. Wheeler 1

1. Human Genome Sequencing Center, Department of Molecular and Human Genetics, 2. Department of Molecular and Cellular Biology
3. Huffington Center on Aging, Department of Otolaryngology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030
4. Department of Microbiology and Molecular Genetics, University of Texas Medical School, PO Box 20708, Houston, TX 77225

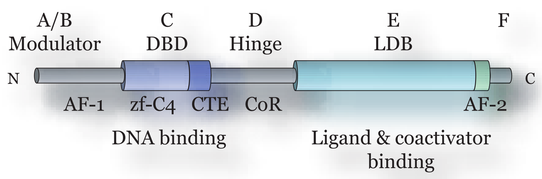
Summary

This study
- provided a global inventory of nuclear receptors in human, mouse, and rat.
- identified eleven syntenic NR gene blocks and three small NR gene clusters.
- revealed the gene structures of the DNA-binding domain and the ligand-binding domain of the NR genes.
- furnished separate accounts of the evolutionary attributes of the DBD and LBD.

Introduction

Nuclear receptors are
- a family of transcription factors, regulating a group of diverse and crucial biological functions, such as metabolism, homeostasis, development, and reproduction.
- structually similar and evolutionarily related.
- important pharmacological targets.

Domain organization



Methods

62 representative mammalian NRs

HMM-PFAM

NR domain Pfam profiles

Genomic Sequences

GENEWISE

NR domains in genome

62 representative mammalian NRs

Grouping, BLASTP

NR genes in genome

Poisson dist. & chi-square test

Seq. anal.

Evol. anal.

NR gene clusters in genome

Conserved blocks, gene structures

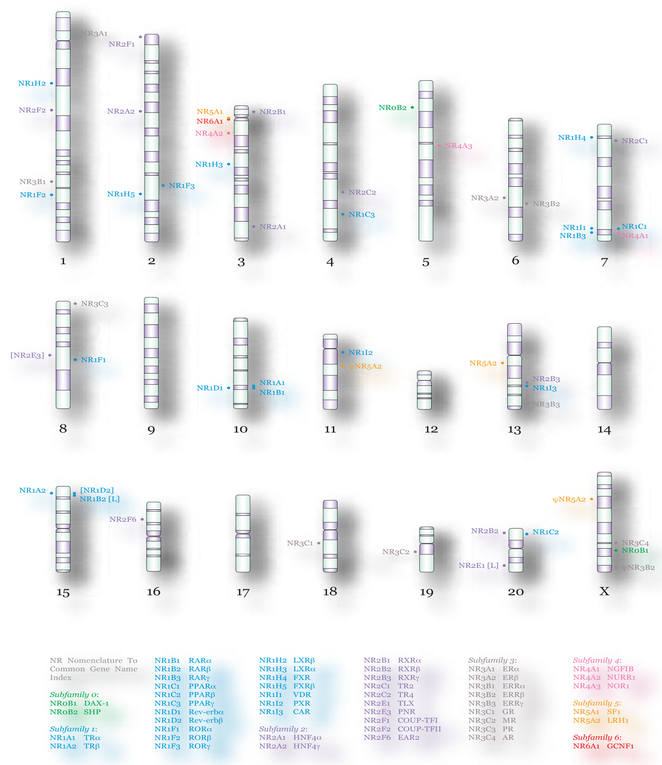
Trees, KA/KS

Results

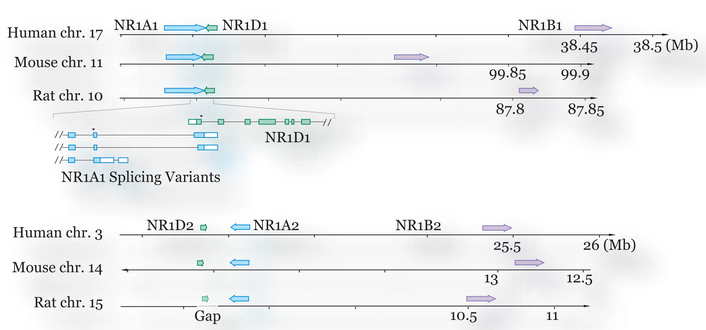
NR gene counts at a glance

Sequence	Rat	Mouse	Human
NR genes	47	49	48
NR genes, aberrant	0	1	1
NR pseudogenes	3	4	3
Domain singletons	1	2	1

Chromosomal landscape of rat NR genes

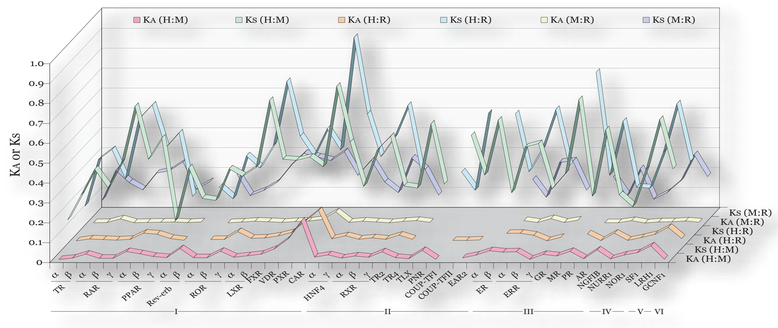


Two of three NR gene clusters



The most remarkable feature in this figure is the overlap of the 3'-most exons of one of the variants of TRα and Rev-erbα in the first cluster. The third cluster, not shown in this figure, is composed of NR5A1 and NR6A1. The current NR gene clusters, whose structures may be preserved by selective advantage, are probably the vestiges of an ancient large NR gene cluster.

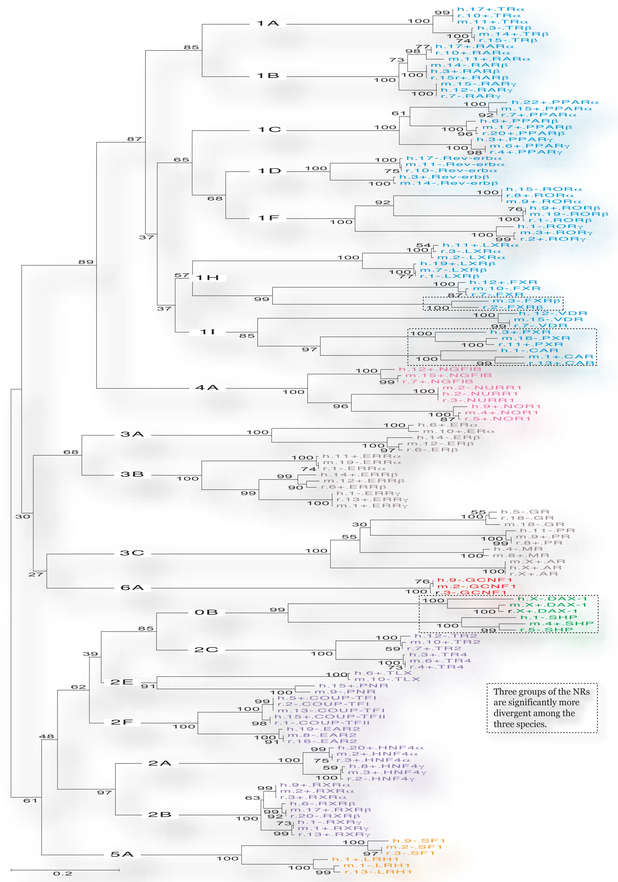
KA and KS in the LBDs



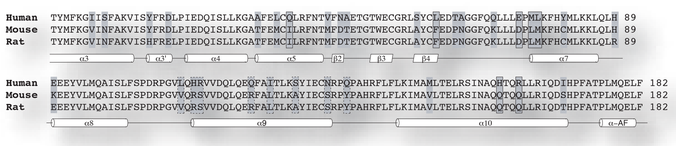
Nucleotide substitution analysis reveals the synonymous rates in the sequences encoding the LBDs of CAR and PXR are average whereas the nonsynonymous rates, 6.4 and 3.7 times higher than the average, have been substantially elevated. PXR and CAR, in group NR1I, share some ligands and regulate overlapping but distinct sets of genes involved in xenobiotic detoxification.

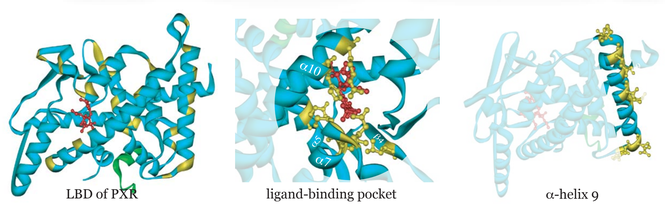
* This work was supported by grants 5 U54 HG02051 (Human/MB) from NHGRI and 5 U54 HG02345 (Rat) from NHGRI/NHLBI to R.A.G.

Unrooted LBD tree



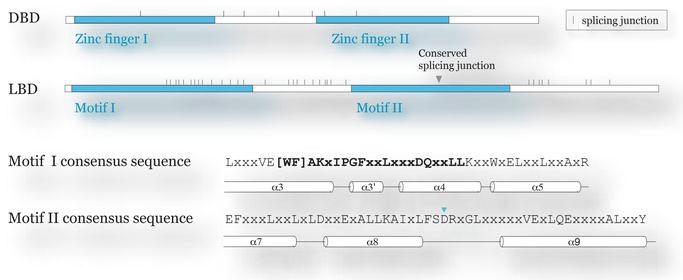
Variable sites in the LBD of PXR





Our mapping of the 33 variable sites to the tertiary structure of the LBD of the human PXR categorized them into three groups: seven sites line the inner surface of the ligand-binding pocket, and are so positioned that they could possibly form direct contacts with the bound ligand; eight variable sites were found to be almost evenly distributed in the α-helix 9 with their side chains protruding toward the surface; the remaining sites are distributed throughout the LBD.

Gene structure of the DBD and LBD



The conserved splice junction in the LBD was likely to have originated early in the family and have been subsequently conserved in evolution. It can serve as another signature of the NRs, in addition to the DBD and LBD, on the DNA level.