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. Huntington 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
  family of diverse and crucial biological functions, such as metabolism, homeostasis, development, and reproduction.
- structurally similar and evolutionarily related.
- important pharmacological targets.

Domain organization

Methods
62 representative mammalian NRs

HMM-PFAM

Genomic Sequences

NR domain
Pfam profiles

GENEWISE

NR domains in genome

Grouping, BLASTP

NR genes in genome

Evod.
anal.

Seq.
anal.

Trees, Ks/Ks

NR gene clusters in genome

Conserved blocks, gene structures

Results
NR gene counts at a glance

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Rat</th>
<th>Mouse</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>NR genes</td>
<td>47</td>
<td>49</td>
<td>48</td>
</tr>
<tr>
<td>NR genes, aberrant</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NR pseudogenes</td>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Domain singletons</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Chromosomal landscape of rat NR genes

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.

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.