Trichoplax, the simplest known animal, contains an estrogen-related receptor but no estrogen receptor: Implications for estrogen receptor evolution

Michael E. Baker*
Department of Medicine, 0693
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0693

*Corresponding author
E-mail: mbaker@ucsd.edu
Phone: 858-534-8317; Fax: 858-822-0873

Abstract. Although, as their names imply, estrogen receptors [ERs] and estrogen-related receptors [ERRs] are related transcription factors, their evolutionary relationships to each other are not fully understood. To elucidate the origins and evolution of ERs and ERRs, we searched for their orthologs in the recently sequenced genome of Trichoplax, the simplest known animal, and in the genomes of three lophotrochozoans: Capitella, an annelid worm, Helobdella robusta, a leech, and Lottia gigantea, a snail. BLAST searches found an ERR in Trichoplax, but no ER. BLAST searches also found ERRs in all three lophotrochozoans and invertebrate-like ERs in Capitella and Lottia, but not in Helobdella. Unexpectedly we find that the Capitella ER sequence is closest to ERβ, unlike the other invertebrate ER sequences, which are closest to ERα. Our database searches and phylogenetic analysis indicate that invertebrate ERs evolved in a lophotrochozoan and steroid-binding ERs evolved in a deuterostome.

Key words: estrogen receptor evolution, invertebrate estrogen receptors; estrogen related receptor, nuclear receptors, Trichoplax, lophotrochozoans,

Introduction.

From the beginning, when estrogen-related receptor α [ERRα] and ERRβ were first cloned [1], the ERR has been an enigma [2, 3]. As its name implies, ERR sequences are similar to that of vertebrate estrogen receptors [ER]. The ligand-binding domain of human ER and ERRα and ERRβ have about 35% sequence identity and 60% positive matches, when conservative replacements such as arginine/lysine and glutamic acid/aspartic acid are considered.
Yet these ERRs do not bind estradiol or other steroids [1-3]. Subsequently, ERRγ [4] was cloned and also found to lack steroid-binding activity. Indeed, a bona fide biological ligand for an ERR has not yet been identified. As a result, the ERR belongs to the orphan receptor group [5, 6] in the nuclear receptor family of transcription factors [7-10].

An explanation for the absence of steroid binding by ERRs came from analysis of the crystal structures of human ERRα [11, 12] and ERRγ [13], which showed that the ligand binding site is too small to accommodate a steroid [11-14]. The crystal structures also showed that in ERRα and ERRγ the activation function 2 (AF2) domain on α-helix 12 is in a conformation for productive interactions with coactivators [15, 16], which explains why the ERR does not require a ligand to become transcriptionally active in cell assays [1-5]. In the last few years, there has been progress in beginning to elucidate ERR functions, which include regulating bone formation [2, 3, 17, 18] and mitochondrial biogenesis [18-20].

Complicating understanding of the evolution of ERRs and vertebrate ERs was the cloning in the last five years of several invertebrate ERs from mollusks [21-25]. Invertebrate ERs, such as octopus ER, have about 34% sequence identity and 58% positive matches with the estrogen-binding domain in human ERRα, and 29% identity and 56% positive matches with human ERRγ. Similar to ERRs, invertebrate ERs do not bind estradiol with high affinity, in contrast to vertebrate ERs, which are activated by 0.2 nM estradiol [26]. Also, similar to ERRs, invertebrate ERs are constitutively active transcription factors in cell assays [21-25]. A biological function for invertebrate ERs has not been reported.

The phylogenetic relationships of vertebrate and invertebrate ERs to each other and to ERRs are still not fully understood [9, 17, 18, 21, 25]. When did the ancestral ER/ERR arise? Was this ancestor more like an ERR or an ER? How did the estrogen-binding vertebrate ER and the constitutively active invertebrate ERs evolve [8, 27-29]? That is, did vertebrate and invertebrate ERs evolve from a gene duplication of an ancestral ER, or did the vertebrate and invertebrate ERs evolve from separate ancestral genes? An opportunity to address these questions comes from recent sequencing by the Joint Genome Initiative [http://genome.jgi-psf.org] of genomes of Trichoplax, which is considered to be the simplest metazoan [30-32], and of three lophotrochozoans: Capitella, a segmented worm, Helobdella, a leech and Lottia, a snail.

As reported here, BLAST [33] searches found an ERR, but no ER in Trichoplax, indicating that ERRs are more ancient than ERs. BLAST searches of the three recently
sequenced lophotrochozoan genomes found ERRs in *Capitella, Helobdella* and *Lottia*, and invertebrate ERs in *Capitella* and *Lottia*. The current genome release of *Helobdella* does not contain an invertebrate ER. To our surprise, a BLAST search of GenBank with the *Capitella* ER sequence indicates that it is closest to ERβ, in contrast to the other invertebrate ER sequences, which are closest to ERα.

The evidence that invertebrate-like ERs are restricted to lophotrochozoans and our phylogenetic analysis of protostome and deuterostome ERRs and ERs indicates that invertebrate ERs share a common ancestor with protostome ERRs, and steroid-binding vertebrate ERs evolved from an ancestor in a deuterostome [28].

**Methods**

BLAST [33] was used to collect ERR and ER sequences from the JGI server [http://genome.jgi-psf.org] and GenBank. Two different methods, Clustal X 2.0 [34], which uses a neighbor-joining algorithm [35], and PHYML [36], which uses a maximum likelihood algorithm, were used to construct phylogenetic trees of various ERs, ERRs, human retinoid X receptor-α (RXRα) and amphioxus RXR.

For the Clustal X 2.0 phylogeny, the multiple alignment of ERs, ERRs and RXRs was done using the iteration option for each alignment step in the multiple alignment. This alignment was converted to a phylogenetic tree using the neighbor-joining algorithm [35] with a correction of branch lengths for rate heterogeneity between sites.

For PhyML, the Muscle algorithm [37] was used to construct a multiple alignment. PhyML was used with the WAG substitution model [38] and a gamma distribution of rates between sites (four categories, parameter α estimated by PhyML), and 100 bootstrap replicates. The phylogenetic trees with Clustal X 2.0 and PhyML gave similar topologies and bootstrap values.

clavigera (rock shell) ER [GenBank:BAC66480], Human ERβ [GenBank:6166154], Xenopus tropicalis ERα [GenBank:NP_988866] and ERβ [GenBank:NP_001035101], Drosophila melanogaster ERR [GenBank:NP_729340], Apis mellifera ERR [GenBank:110756963], Daphnia ERR [jgi|Dappu1|46682], Capitella ERR [jgi|Capca1|108381], Lottia ERR [jgi|Lotgi1|168715], Helobdella ERR [jgi|Helro1|106750], human RXRα [GenBank:NP_002948] and amphioxus RXR [GenBank:AAM46151]

Results and Discussion

Four Nuclear Receptors are present in a basal diploblast

The DNA and ligand-binding domains of human ERRγ, human ERα, octopus ER, Aplysia ER, Thais ER and oyster ER were used as queries for BLAST searches for orthologs in Trichoplax, Capitella, Helobdella, and Lottia on the JGI server. The BLAST search of Trichoplax with human ERRγ yielded four high scoring nuclear receptors. Searches with human ERα and invertebrate ERs found the same genes in Trichoplax. To classify the four Trichoplax genes, we used their sequences as queries for BLAST searches of GenBank. This identified Trichoplax jgi|Triad1|16711|gw1.23.179.1 as an ortholog of ERR; the other Trichoplax genes appear to be orthologs of COUP, RXR or HNF4 [Table 1]. Thus, genes with similarity to vertebrate ERR, COUP, RXR and HNF4 are found in a primitive multicellular animal belonging to the phylum Placozoa [30-32].

Our analysis does not exclude the possibility of other nuclear receptors in Trichoplax because our BLAST search focused on finding ancestors of ER and ERR. Also nuclear receptor genes may have been lost in Trichoplax during its evolution from an ancestral metazoan. Analyses of other simple metazoan genomes will provide a more definitive inventory of nuclear receptors in basal metazoans.

We focused the rest of our analyses on the relationship of the Trichoplax ERR-like gene to invertebrate and vertebrate ERRs and ERs. BLAST searches of the JGI server retrieved ERRs from Capitella, Helobdella and Lottia, and an invertebrate ER from Capitella and Lottia. BLAST did not find an invertebrate ER in Helobdella.
Table 1. Nuclear Receptor Genes in Trichoplax

<table>
<thead>
<tr>
<th>Gene ID in JGI Databank</th>
<th>Homolog in GenBank</th>
<th>BLAST score</th>
<th>% Identity and %Positives, and Gaps</th>
</tr>
</thead>
<tbody>
<tr>
<td>jgi</td>
<td>Triad1</td>
<td>16711</td>
<td>gw1.23.179.1</td>
</tr>
<tr>
<td>&gt;jgi</td>
<td>Triad1</td>
<td>49897</td>
<td>fgeneshTA2_pm_C_scaffold_2000050</td>
</tr>
<tr>
<td>jgi</td>
<td>Triad1</td>
<td>50786</td>
<td>fgeneshTA2_pm_C_scaffold_12000032</td>
</tr>
<tr>
<td>&gt;jgi</td>
<td>Triad1</td>
<td>9010</td>
<td>gw1.23.150.1</td>
</tr>
<tr>
<td>&gt;jgi</td>
<td>Triad1</td>
<td>21656</td>
<td>e_gw1.2.1246.1</td>
</tr>
</tbody>
</table>

The Trichoplax genome at JGI was searched with the amino acid sequence for human ERα. Column 1 lists the five entries that were retrieved, two of which correspond to the ligand-binding and DNA-binding domains of human ERRγ. Columns 2 and 3 show the highest scoring entry in GenBank and its BLAST score. Column 4 shows the % identities, positives, which include identities and conservative replacements, and the gaps in the BLAST alignment.

A BLAST search of the JGI server retrieved an ERR from Daphnia, a water flea. BLAST did not find an invertebrate ER in Daphnia. We also used BLAST to retrieve invertebrate ER sequences from A. californica, O. vulgaris, C. gigas, and two snails: T. clavigera [23] and M. cornuarietis [22] and ERR sequences from M. cornuarietis, A. mellifera, and D. melanogaster.

**Divergence of Capitella ER from other invertebrate ERs.**

To better understand the relationship of invertebrate ERs to steroid-binding ERs, we used BLAST to search GenBank with the domain on each invertebrate ER which corresponds to the steroid binding domain on vertebrate ERs, with the goal of determining how similar each invertebrate ER is to the steroid-binding vertebrate ERs. To our surprise, Capitella ER is closest to vertebrate ERβ, unlike the other invertebrate ER sequences, which are closest to ERα. For example, BLAST found minnow ERβ [GenBank: ABS84945] as the closest vertebrate protein to Capitella ER. Following in the BLAST output were eleven ERβ entries and then amphioxus ERR [GenBank: AAU88063]. Much later in the BLAST output was mouse ERα [GenBank:
BLAST searches showed that the other invertebrate ERs were closest to vertebrate ER\(\alpha\), and then closest to vertebrate ER\(\beta\) and then closest to vertebrate ERR.

To follow-up these BLAST analyses, we did pairwise BLAST comparisons of each invertebrate ER with human ER\(\alpha\), ER\(\beta\) and ERR\(\gamma\). As shown in Table 2, pairwise BLAST analyses show that *Capitella* ER is closer to human ER\(\beta\) than to human ER\(\alpha\). Interestingly, octopus ER has about equal sequence similarity to human ER\(\alpha\) and human ER\(\beta\). However, a BLAST search of GenBank found that octopus ER clearly was closest to ER\(\alpha\). BLAST found the closest vertebrate sequence to octopus ER was golden hamster ER\(\alpha\) [GenBank: AAD53956], which was followed by over twenty ER\(\alpha\) sequences. The pairwise BLAST analyses in Table 2 show that the other invertebrate ERs are closer to ER\(\alpha\) than to ER\(\beta\).

Table 2. Invertebrate ERs and ERRs in JGI and GenBank

<table>
<thead>
<tr>
<th>Animal</th>
<th>Characteristics</th>
<th>ER</th>
<th>ERR</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Trichoplax</em></td>
<td>Diploblast Placozoan</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Fruit Fly <em>Drosophila</em></td>
<td>Ecdysozoan Arthropod</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Water Flea <em>Daphnia</em></td>
<td>Ecdysozoan Arthropod</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Sea Slug <em>Aplysia</em> **</td>
<td>Lophotrochozoan Mollusk</td>
<td>Yes, *</td>
<td>None in GenBank</td>
</tr>
<tr>
<td>Snail <em>Thais</em> **</td>
<td>Lophotrochozoan Mollusk</td>
<td>Yes, *</td>
<td>None in GenBank</td>
</tr>
<tr>
<td>Snail <em>Marissa</em> **</td>
<td>Lophotrochozoan Mollusk</td>
<td>Yes, *</td>
<td>Yes</td>
</tr>
<tr>
<td>Oyster <em>Crassostrea</em> **</td>
<td>Lophotrochozoan Mollusk</td>
<td>Yes, *</td>
<td>None in GenBank</td>
</tr>
<tr>
<td>Octopus **</td>
<td>Lophotrochozoan Mollusk</td>
<td>Yes, *</td>
<td>None in GenBank</td>
</tr>
<tr>
<td>Snail <em>Lottia</em></td>
<td>Lophotrochozoan Mollusk</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Bristle Worm <em>Capitella</em></td>
<td>Lophotrochozoan Annelid</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Leech <em>Helobdella</em></td>
<td>Lophotrochozoan Annelid</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*Constitutive transcriptional activity. Receptor does not bind estradiol. **Complete genome has not been sequenced.
**Invertebrate ERs evolved in lophotrochozoans**

To clarify further the evolutionary relationships of various invertebrate and vertebrate ERRs and ERs, we constructed a phylogenetic tree of their ligand-binding domains as shown in Figure 1. The vertebrate ER and ERR part of the phylogeny is in agreement with previous analyses [8, 29, 39]. The phylogeny indicates that vertebrate ERs and invertebrate ERs diverged from a common ancestor at node A before the evolution of deuterostomes.

![Phylogenetic tree of invertebrate and vertebrate ERs and ERRs.](image)

**Figure 1. Phylogenetic analysis of invertebrate and vertebrate ERs and ERRs.**

The phylogenetic tree was constructed using PhyML [36] under the WAG substitution model [38], with a gamma distribution of rates between sites (four categories, parameter $\alpha$ estimated by PhyML), and 100 bootstrap replicates. Shown at the nodes are bootstrap values for each branch of the tree, which is the percent this cluster was found in the 100 bootstrap trials. Branches with bootstrap values that are greater than fifty percent are significant. Branch lengths are proportional to the distance between proteins. Due to space limitations, we show values for selected branches. Human RXR$\alpha$ and amphioxus RXR were used as outgroups for the phylogenetic tree.
Like our BLAST analyses [Table 2], the phylogeny shows that *Capitella* ER has diverged substantially from the other invertebrate ERs, which cluster together.

The phylogeny and the absence of invertebrate ERs in ecdyzoa suggests that invertebrate ERs arose in a lophotrochozoan from an ERR-like ancestor. A practical application of this phylogeny is to suggest that invertebrate ERs are likely to have function(s) that resemble ERR functions [2, 17-20].

The evolution of invertebrate ERs from an ERR-like ancestor is consistent with functional similarities between invertebrate ERs and vertebrate ERRs. Both vertebrate ERRs and invertebrate ERs are constitutively active and do not bind estradiol. The crystal structures of human ERRs [11-14] and a 3D model of octopus ER [40] indicates that their ligand-binding domains are too small to accommodate estradiol.

**Steroid-regulated vertebrate ERs evolved in a deuterostome**

The absence of an invertebrate ER outside of lophotrochozoans and the absence of an invertebrate ER in the recently completed sea urchin genome [41] suggests that a steroid-binding vertebrate ER evolved in a deuterostome [27, 28, 42, 43], in which case, vertebrate and invertebrate ERs evolved from different ancestors.

Which mutations led to the evolution of estrogen-dependent activation in vertebrate ERs? Analyses of the 3D structures of human ERRα [11] and ERRγ [13] reveals that the volume of their ligand-binding pockets are about 100 Å³ and 220 Å³, respectively, which is much less than 369 Å³ found in human ERα [14]. The more compact ligand-binding pocket in ERRs [13,14] is thought to explain why ERRs do not bind estradiol, which has a van der Waals volume of 251 Å³ [44] and, thus, easily fits into human ERα. The 3D structures of ERRα and ERRγ also reveal that the AF2 domain is in a position to have productive interactions with coactivators and regulate gene transcription.

Information clarifying the basis for the transcriptional properties of ERRγ comes from Greschik et al. [13], who modeled estradiol in ERRγ and compared it with estradiol in ERα [45]. They identified two residues, Leu-345 and Phe-435 in ERRγ that had steric clashes with the D ring of estradiol. Mutation of these residues to Ile and Leu, respectively, as found in ERα, reduced steric interference with estradiol. As a result, the mutant ERRγ bound estradiol, although with low affinity. However, as Greschik et al. [13] noted, unexpectedly, there was no
change in transcriptional activity of the ERRγ mutant, which suggests that in the estradiol-ERRγ mutant complex, the AF2 domain is in the proper configuration to bind coactivators. This contrasts with binding of estradiol to vertebrate ERs, which causes a conformational change in AF2 on α-helix 12, so that the ER can bind co-activators [45]. A similar conformational change occurs in other steroid receptors upon binding of their cognate steroid [15, 16].

If the ligand-activated vertebrate ER evolved from an ERR, then Greschik et al.’s studies indicate that additional mutation(s) in the ERR ancestor had to occur to increase the affinity for estradiol and also alter the configuration of AF2 in order for binding of a ligand to be required for transcriptional activity. Alignment ERRγ with human ERα [13] [40] reveals an insertion of a total of twelve amino acids distributed among three sites in the steroid binding domain of ERα compared to ERRγ. These insertions map to loops between α-helices in human ERα. One or more of these insertions may be important in altering the ligand-binding pocket and/or conformation of AF2 to yield a steroid-dependent ER.  

References


