Available online at www.sciencedirect.com
ScienceDirect
Gene 392 (2007) 164–173
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

The complete nuclear estrogen receptor family in the rainbow trout: Discovery of the novel *ERα2* and both *ERβ* isoforms

James J. Nagler [a,b,*], Tim Cavileer [a,b], Jack Sullivan [a], Daniel G. Cyr [c], Caird Rexroad III [d]

[a] *Department of Biological Sciences, University of Idaho, Moscow, ID, 83844-3051, USA*
[b] *Center for Reproductive Biology, University of Idaho, Moscow, ID, 83844-3051, USA*
[c] *INRS–Institut Armand-Frappier, Université du Québec, 245 Boulevard Hymus, Pointe-Claire, QC, Canada H9R 1G6*
[d] *USDA/ARS National Center for Cool and Cold Water Aquaculture, 11876 Leetown Road, Kearneysville, WV, 25430, USA*

Received 7 July 2006; received in revised form 7 November 2006; accepted 4 December 2006
Available online 20 January 2007
Received by M. D'Urso

Abstract

Estrogen hormones interact with cellular *ER*s to exert their biological effects in vertebrate animals. Similar to other animals, fishes have two distinct *ER* subtypes, *ERα* (NR3A1) and *ERβ* (NR3A2). The *ERβ* subtype is found as two different isoforms in several fish species because of a gene duplication event. Although predicted, two different isoforms of *ERα* have not been demonstrated in any fish species. In the rainbow trout (*Oncorhynchus mykiss*), the only *ER* described is an isoform of the *ERα* subtype (i.e. *ERα1*, NR3A1a). The purpose of this study was to determine whether the gene for the other *ERα* isoform, *ERα2* (i.e., NR3A1b), exists in the rainbow trout. A RT-PCR and cloning strategy, followed by screening a rainbow trout BAC library yielded a unique DNA sequence coding for 558 amino acids. The deduced amino acid sequence had a 75.4% overall similarity to *ERα1*. Both the rainbow trout *ERβ* subtypes, *ERβ1* [NR3A2a] and *ERβ2*, [NR3A2b] which were previously unknown in this species, were also sequenced as part of this study, and the amino acid sequences were found to be very different from the *ERα*s (~40% similarity). *ERβ1* and *ERβ2* had 594 and 604 amino acids, respectively, and had 57.6% sequence similarity when compared to one another. This information provides what we expect to be the first complete nuclear *ER* gene family in a fish. A comprehensive phylogenetic analysis with all other known fish *ER* gene sequences was undertaken to understand the evolution of fish *ER*s. The results show a single *ERα* subtype clade, with the closest relative to rainbow trout *ERα2* being rainbow trout *ERα1*, suggesting a recent, unique duplication event to create these two isoforms. For the *ERβ* subtype there are two distinct subclades, one represented by the *ERβ1* isoform and the other by the *ERβ2* isoform. The rainbow trout *ERβ1* and *ERβ2* are not closely associated with each other, but instead fall into their respective *ERβ* subclades with other known fish species. Real-time RT-PCR was used to measure the mRNA levels of all four *ER* isoforms (*ERα1*, *ERα2*, *ERβ1*, and *ERβ2*) in stomach, spleen, heart, brain, pituitary, muscle, anterior kidney, posterior kidney, liver, gill, testis and ovary samples from rainbow trout. The mRNAs for each of the four *ER*s were detected in every tissue examined. The liver tended to have the highest *ER* mRNA levels along with the testes, while the lowest levels were generally found in the stomach or heart. The nuclear *ER*s have a significant and ubiquitous distribution in the rainbow trout providing the potential for complex interactions that involve the functioning of many organ systems.

Keywords: Fish; Estradiol; Salmonid; Phylogeny; Real-time RT-PCR

Abbreviations: BAC, bacterial artificial chromosome; bp, base pair; cDNA, DNA complementary to RNA; Cm, chloramphenicol; cRNA, transcribed ribonucleic acid; *ER*, estrogen receptor; *ERα*, *ER* alpha; *ERβ*, *ER* beta; EtdBr, ethidium bromide; *eGFP*, enhanced green fluorescent protein; Km, kanamycin; LB, Luria–Bertani; ML, Maximum likelihood; MP, Maximum parsimony; ME, Minimum evolution; mRNA, messenger ribonucleic acid; nt, nucleotide; PCR, polymerase chain reaction; pmol, picomole; RACE, rapid amplification of complementary ends; RT, reverse transcription; TAE, Tris acetic acid EDTA; TBE, Tris borate EDTA; TE, Tris EDTA.

* Corresponding author. Department of Biological Sciences and Center for Reproductive Biology, University of Idaho, Moscow, ID, 83844-3051, USA. Tel.: +1 208 885 4382; fax: +1 208 885 7905.
E-mail address: jamesn@uidaho.edu (J.J. Nagler).

0378-1119/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2006.12.030

1. Introduction

Estrogen hormones have multi-faceted and wide-ranging effects in vertebrate animals (Tsai and O'Malley, 1994; Lange et al., 2003). In order for estrogens, such as estradiol-17β, to exert their biological effects they must interact with cellular *ER*s (Edwards, 2005). It is now understood that *ER*s are part of two distinct estrogenic transduction pathways (Edwards, 2005). One provides direct genomic (i.e., nuclear) control in which *ER*s act as transcription factors within the cell nucleus. The other involves a rapid, nongenomic pathway initiated by membrane bound *ER*s at the cell surface (Pappas et al., 1995; Loomis and Thomas, 2000; Simoncini and Genazzani, 2003). Recent evidence demonstrates that the *ER*s involved in the nongenomic pathway belong to the G protein-coupled receptor family, which is distinct and different from nuclear *ER*s (Revankar et al., 2005; Thomas et al., 2005). Nuclear *ER*s belong to a well-established superfamily of steroid hormone receptors that include androgen, progestin, and mineralocorticoid receptors (Laudet et al., 1992; Thornton, 2001). Within the nuclear *ER* group, there are two distinct subtypes, *ERα* and *ERβ*. Each *ER* subtype is encoded by separate genes with unique transcriptional activities (Cheung et al., 2003), and the proteins differ significantly in their amino acid sequence, size, and ligand binding characteristics (Pettersson and Gustafsson, 2001).

In ray-finned fishes (Actinopterygii) the *ER* story is more complex than in other vertebrate classes because many genes are duplicated. A current hypothesis, with considerable support, suggests that a whole genome duplication event occurred in ray-finned fishes after they diverged from the lobe-finned fishes (Sarcopterygii) (Amores et al., 1994). There is evidence for this duplication event in the genes from numerous ray-finned fish species (e.g. *IGF*-1, *IGF*-2; *GH*-1, *GH*-2) (Wittbrodt et al., 1998; Palti et al., 2004). With respect to fish *ER*s, two different isoforms of the *ERβ* subtype have been reported in several ray-finned fish species (Tchoudakova et al., 1999; Ma et al., 2000; Menuet et al., 2002; Halm et al., 2004; Hawkins and Thomas, 2004; Sabo-Attwood et al., 2004) presumably because of genome duplication. Since the fish *ERβ* subtype has been duplicated, it is predicted that the other *ER* subtype, *ERα*, should also be found as two isoforms. To date this has not been demonstrated in any fish species.

The first *ER* identified in fishes, *ERα* (i.e., *ERα1* [NR3A1a: Nuclear Receptors Nomenclature Committee, 1999]), was reported by Pakdel et al. (1989, 1990) for the rainbow trout (*Oncorhynchus mykiss*). Subsequently in this species, two different *ERα* splice variants from the ovary, a long and a short form, that differ by the addition of 45 amino acids at the N-terminal end of the long variant, were found (Pakdel et al., 2000). The remainder of the DNA sequence is identical for both the long and short variants. These kinds of *ER* splice variants have been found in mammals too (Hirata et al., 2003). In 2000, we reported that another *ERα* mRNA distinct from those reported by Pakdel et al. (2000) was present in rainbow trout ovary and testes (Nagler and Krisfalusi, 2000). It was theorized that this *ERα* was the other, genome-duplicated isoform belonging to the *ERα* subtype, similar to the situation for the *ERβ* subtype with two isoforms reported in other fish species (Menuet et al., 2002; Halm et al., 2004; Hawkins and Thomas, 2004; Sabo-Attwood et al., 2004). This paper reports on the novel DNA sequences of this *ERα* isoform (*ERα2* [NR3A1b] and both the *ERβ*s, *ERβ1* [NR3A2a] and *ERβ2* [NR3A2b]) from the rainbow trout. This completes the currently known family of nuclear *ER*s for the rainbow trout, with both isoforms for each of the *ERα* and *ERβ* subtypes, for the first time in any fish. Comprehensive phylogenetic analyses were undertaken to determine the relationship of the rainbow trout *ER* sequences to each other and within the context of known *ER* sequences from other fish species. Measurement of mRNA levels for all 4 *ER* isoforms by real-time RT-PCR in a variety of different tissues was done to establish the distribution and potential biological significance of these receptors in the rainbow trout.

2. Materials and methods

2.1. ERα2 DNA sequence

2.1.1. Reverse transcription-polymerase chain reaction and cDNA cloning

Samples of rainbow trout ovary were obtained, and total RNA and mRNA isolations performed as previously described (Nagler et al., 2000). RT reactions used 1 μg mRNA and Superscript II (GibcoBRL) combined with 50 pmol of ER4 primer (Nagler et al., 2000). A portion (4% by volume) of the completed RT reaction was added to the PCR reagents according to the GibcoBRL protocol, along with 50 pmol ER3 primer (Nagler et al., 2000). Typically, 5 μl of the final reaction mixture was loaded on a 1% agarose gel, containing EtdBr, and electrophoresed in 1× TBE buffer (Ausubel et al., 1998). DNA size markers (1 kb; GibcoBRL) were run in a well adjacent to the experimental samples.

RT-PCR products were ligated into the *pCR* II plasmid vector and used to transform INVαF' cells supplied in the Original TA Cloning Kit (Invitrogen). Plasmid DNA minipreps were made from single bacterial colonies grown overnight on LB-agar plates containing Km with the QIAprep kit (QIAGEN). Plasmid DNA from clones containing the appropriate sized insert, as determined by *Eco*R1 (GibcoBRL) restriction digest, was selected for sequencing. DNA sequencing, in both directions, was done using an automated system (Applied Biosystems). Clones from two separate RT-PCRs were analyzed to construct the final cDNA sequence from which the amino acid sequence was deduced.

RACE PCR was conducted with the Marathon kit (Clontech) using 1 μg mRNA. The 1st and 2nd strand DNA syntheses were conducted according to the Marathon protocol. The adaptor ligation was done at room temperature for 4 h, then the DNA diluted 250-fold with tricine-EDTA, denatured, and stored frozen at −20 °C. 5′ RACE PCR was performed according to the Marathon protocol using ER4 and AP-1 (kit supplied) primers with 5 μl of the DNA template. The thermal profile was: 94 °C for 10 s, then 30 cycles of 94 °C for 5 s, 64 °C for 5 min, and 68 °C for 4 min. The 5′ RACE PCR generated DNA was diluted 50-fold and used for touchdown PCR with the nested primers NER-4 (5′-GTTGCCTTGCTTGGGTCCTA-3′) and AP-2 (kit supplied). The touchdown PCR thermal profile was: 94 °C for

10 s, followed by 5 cycles of 94 °C for 5 s and 68 °C for 4 min, then 5 cycles of 94 °C for 5 s and 66 °C for 4 min, followed by 20 cycles of 94 °C for 5 s and 64 °C for 4 min. The resulting PCR products were cloned and sequenced as described above.

2.1.2. Bacterial artificial chromosome analysis

A custom primer set, (ER2-A 5′-CCTAGCCATCTCACCCCAGAAA-3′ and ER2-8 5′-CTGAGCTGGGGAGTGGAG-3′), designed to target the A/B-domain of the *ERα*-2 cDNA sequence, was used to produce a 255 bp DNA fragment by PCR amplification. The 255 bp probe was generated by PCR using 18.5 µl sterile distilled water, 2.5 µl 10× buffer (Invitrogen), 0.5 µl 10 mM dNTPs (Invitrogen), 0.75 µl 50 mM $MgCl_2$ (Invitrogen), 0.75 µl ER2-A at 20 µM, 0.75 µl ER2-8 at 20 µM, 0.25 µl *Taq* DNA Polymerase (Invitrogen) at 5 U/µl and 1.0 µl (~30 ng) of the 1:10 diluted template (the cloned 5′ RACE product described above). The thermocycling profile included 5 min denaturing at 94 °C, followed by 30 cycles of 30 s at 94 °C, 1 min at 59 °C and 1 min at 72 °C. This DNA fragment was used as a probe to screen a 10× BAC library for the rainbow trout (Palti et al., 2004). Thirteen positive BAC clones were identified and subsequently grown in liquid medium using Cm selection. Plasmid DNA was isolated from each clone by the alkaline-lysis method (Ausubel et al., 1998) and PCR analysis using primers ER2-A and ER2-8 showed that the correct target fragment could be amplified in 12 of the 13 clones. Two of these clones (ER#1 and ER#3) were selected for further DNA sequence analysis.

2.2. ERβ DNA sequences

2.2.1. ERβ1 RT-PCR and cDNA cloning

Samples of immature rainbow trout testis, ovary, liver and total RNA were obtained as previously described (Nagler et al., 2000). Individual 20 µl RT reactions were performed using 3–5 µg total RNA and Superscript II (Invitrogen) combined with 50 ng of oligo $d(T)_{12-18}$ primer as described by the manufacturer's protocol for 'First-Strand cDNA synthesis using SUPERSCRIPT II for RT-PCR'. A portion (4% by volume) of the completed RT reaction was added to standard PCR reagents, along with 10 pmol *ERβ* degenerate primers and *Taq* DNA polymerase (Invitrogen). Thermocycling conditions used were as described by Sabo-Attwood et al. (2004).

Degenerate custom primers were designed from a conserved region of 22 aligned ray-finned fish *ERβ* sequences, because four initial primer sets designed to a purported rainbow trout *ERβ* DNA sequence (GenBank accession no. AJ289883) failed to generate fragments from rainbow trout oligo d(T) primed cDNAs and genomic DNA. Custom primers *ERβ*-DF (5′-TGYGARGSVTGYAARGCYTTYTTCAA-3′) and *ERβ*-DR (5′-AANCCWGGDAYHTKYTTVGCCCAG-3′) were subsequently used to generate multiple clones as described above.

RT-PCR products were separated on 1% agarose gels in 1× TAE buffer (Ausubel et al., 1998). Bands were cut out, eluted through modified filtered pipet tips, and ligated into the *pCR* II TOPO TA plasmid vector and used to transform TOP10 cells (Invitrogen). Plasmid DNA from clones containing the appropriate sized insert, as determined by *Eco*R1 (Invitrogen) restriction digest, was selected for sequencing. DNA sequencing was done in both directions using an automated sequencer (Applied Biosystems). Sequences from several clones were used to search GenBank and TIGR databases. Matches indicated that the clones contained fragments from *ERβ*-like genes in the C, D, and E-domains. Primers designed to Atlantic salmon (*Salmo salar*) *ERβ* sequence (GenBank accession no. AY508959) (*ERβ*-As1 5′-CACGGATGGATTGCTACTCCA-3′ and *ERβ*-As2 5′-AGCTTCAGAAGGGTCGCAGA-3′) were used to generate additional clones from rainbow trout cDNA representing the 5′ region of an *ERβ*1-like DNA sequence. This DNA sequence was used to search the TIGR *O. mykiss* database and one match indicated EST BX860575 might contain the entire sequence. This clone, tcba0008c.h.15, obtained from the AGENAE Resource Centre (INRA, Jouy-en-Josas, France), was sequenced in its entirety.

2.2.2. ERβ2 RT-PCR and cDNA cloning

Several clones, generated with *ERβ* degenerative primers described above, had significant similarity with other ray-finned fish *ERβ2* sequences. Gene specific primers were designed to this sequence and used to generate 5′ and 3′ RACE products using the BD SMART RACE cDNA Amplification Kit (BD Biosciences Clontech).

Rainbow trout mRNA was purified from total RNA using the NucleoTrap mRNA Nucleic Acid Purification kit (BD Biosciences Clontech). 5′ RACE-Ready cDNA was made using 1 mg mRNA, 5′-CDS primer, and the BD SMART II Oligo A primer in a 10 µl reaction. 3′-RACE-Ready cDNA was similarly made using the kit supplied 3′-CDS primer A and 1 µg rainbow trout mRNA according to the manufacturer's protocol.

Gene specific primer *ERβ2*-214f (5′-CCAGAACAGCTGATCTCCTGCATCATGG-3′) and the Universal Primer Mix from the kit were used to obtain clones representing the 3′ end of the rainbow *ERβ2*. RACE products were amplified using the BD Advantage 2 PCR Enzyme System (BD Biosciences Clontech) according to the manufacturer's protocol. The resulting RACE PCR products were cloned into *pCR* II TA vector (Invitrogen) and sequenced as described earlier.

Gene specific primer ER2-363r (5′-GGCCCAGTTGATCATGAGGACCAACTGG-3′) and the Universal Primer Mix were used to generate clones representing the 5′ region of the rainbow *ERβ2*. RACE PCR, cloning and sequencing were done as described in Section 2.2.1.

2.3. Phylogenetic analyses

Nucleotide sequence representing the C, D, and E-domains of seventy-four ray-finned fish *ER* sequences found in the GenBank database, and that of rainbow trout *ERα2* (GenBank accession no. DQ177438), *ERβ1* (GenBank accession no. DQ177439), *ERβ2* (GenBank accession no. DQ248229), and coho salmon (*Oncorhynchus kisutch*) *ERα1* (GenBank accession no. DQ248228), were aligned using ClustalX. The hyper variable domains A/B and F were excluded because of alignment ambiguity. The *ER* sequence used for the blue tilapia (*Oreochromis aureus*) (listed as OauX93554) is a compilation

of 10 exons (GenBank accession nos. X93555-X93561, Z46665-Z46669). The sea lamprey (*Petromyzon marinus*) *ER* sequence (GenBank accession no. AY028456) was used as an outgroup.

Aligned nucleotide sequences generated from ClustalX were used for phylogeny estimation under maximum parsimony (MP), minimum evolution (ME), maximum-likelihood (ML), and Bayesian approaches. Identical sets of analyses were run with gaps treated as missing data and with gap sites deleted (this had no effect on results). Maximum parsimony analyses were conducted with equal weights using PAUP* (Swofford, 2002) and consisted of heuristic searches with starting trees generated with 100 replicate random addition sequences, and tree bifurcation replication branch swapping. This analysis generated 10 equally parsimonious trees that only differed by slight rearrangements at a few tips, and produced no evidence of multiple peaks across tree space. Nodal support in MP analyses was estimated using bootstrap analysis (Felsenstein, 1985). In addition, because the data exhibit a deviation from homogeneity because of frequencies (as detected using PAUP*; $P=0.0005$), we conducted a heuristic search under the ME criterion from LogDet distances (Lockhart et al., 1994).

DT-ModSel (Minin et al., 2003) was used to select a model for ML and Bayesian estimation. An iterative search strategy (Sullivan et al., 1996, 2005; Swofford et al., 1996) for phylogeny estimation under ML was employed. We used MrBayes (v.3.1.2; Ronquist and Huelsenbeck, 2005) to estimate nodal probabilities under a GTR+I+Γ model of nt sequence evolution. Duplicate runs were conducted, each of which consisted of four Metropolis coupled chains, with uniform priors across all parameters except branch lengths, where exponential priors were used. The chains were run for 5 million generations and eventually terminated because standard deviations of partition frequencies approached zero, suggesting that the runs had converged to the same posterior probability distribution.

2.4. Quantification of ER mRNA in rainbow trout tissues

The mRNAs for the four rainbow trout *ER*s (*ERα1*, *ERα2*, *ERβ1*, and *ERβ2*) were quantified in a number of different rainbow trout tissues using real-time RT-PCR. Samples (100 mg) from the stomach, spleen, heart, brain, pituitary, muscle, anterior kidney, posterior kidney, liver, gill and testis were collected from three juvenile male rainbow trout and flash frozen in liquid nitrogen. Ovary tissue (100 mg) from three similarly aged female rainbow trout was also collected. After freezing all tissues were stored at −80 °C. Total RNA was isolated from the frozen tissues and first-strand cDNA synthesized as described above in Section 2.2.1.

Primers for the four *ER* genes were designed using Primer Express v2.0 (Applied Biosystems). The following primer sets (*ERα1*: 5′-CCCTGCTGGTGACAGAGAGAA-3′, 5′-ATCCTCCACCACCATTGAGACT-3′; *ERα2*: 5′-GTGGCACTGCTGGTGACAAC-3′, 5′-ACCACCGAAGCTGCTGTTCT-3′; *ERβ1*: 5′-CCCAAGCGGGTCCTAGCT-3′, 5′-TCCTCATGTCCTTCTGGAGGAA-3′; *ERb2*: 5′-CTGACCCCAGAACAGCTGATC-3′, 5′-TCGGCCAGGTTGGTAAGTG-3′) were used in real-time RT-PCRs to measure *ER* mRNA levels in the twelve different tissues. Total RNA was isolated as described previously and SuperScript™ II RT (Invitrogen) was used to generate first stand cDNA from 400 ng of total RNA for each sample.

As an internal standard to account for amplification differences between samples, 60 pg of the enhanced green fluorescent protein (*eGFP*) *in vitro* transcribed RNA (cRNA) was added to the total RNA before cDNA synthesis. The *eGFP* was used because an endogenous gene could not be found that consistently and uniformly amplified in the real-time RT-PCRs across the various tissues examined. Plasmid *eGFP* was subcloned from pEGFP-N1 (U55762; Clontech) into *pCR* II (Invitrogen) and cRNA generated and optimized for use as an *in vitro* reference gene as described by Gilsbach et al. (2006). The resulting spiked cDNA samples were diluted 1:20 in TE buffer, pH 8.0. Primer sequences used for *eGFP* were 5′-CTGCTGCCCGACAACCA-3′ (forward) and 5′-TGTGATCGCGCTTCTCGTT-3′ (reverse).

For each candidate gene two 20 μl reactions consisting of 10 μl 2× SYBR Green PCR mix (this master mix contains SYBR Green 1 Dye and AmpliTaq Gold® DNA Polymerase; Applied Biosystems), 6.0 μl of H_2O, 2.0 μl of cDNA template, and 2.0 μl of either the *ER* primer set or *eGFP* primer set (2 pmol of each primer) were combined using a 96-well plate format. A no-primer control treatment for each template used 2.0 μl of TE in the reaction mixture instead of primers. Real-time RT-PCRs were performed using an ABI 7900HT Prism Sequence Detection System (Applied Biosystems) in conjunction with Applied Biosystems Prism SDS 2.2.2 software. The thermal program used was 50 °C for 2 min, 95 °C for 10 min, and 40 cycles of 95 °C for 15 s and 60 °C for 1 min. A representative RT-PCR sample for each gene was electrophoresed on a 3% high-resolution agarose gel to verify the presence of a single amplicon, and disassociation curves for each sample were analyzed on all plates.

Relative mRNA expression was determined using absolute quantities calculated from standard curves (Applied Biosystems, 1997). Standard curves for each *ER* and *eGFP* were generated by diluting 1 ng of the specific amplicon from 10^{-4} through 10^{-7} dilutions. For each sample, the *ER* gene was normalized to the *in vitro* reference gene (*eGFP*) by dividing the absolute value of the gene by the absolute value of the reference gene. Normalized quantities were therefore expressed as relative ER mRNA expression, without using a calibrator (Filby and Tyler, 2005).

3. Results and discussion

3.1. Sequence analysis

A unique cDNA sequence was obtained from rainbow trout ovary mRNA by RT-PCR using primers ER3 and ER4 that resembled a portion of the E-domain region of a previously documented rainbow trout *ERα* (GenBank accession nos. AJ242740 and AJ242741). This novel sequence (DQ177438), termed *ERα2* (see below), was extended in the 5′ direction by

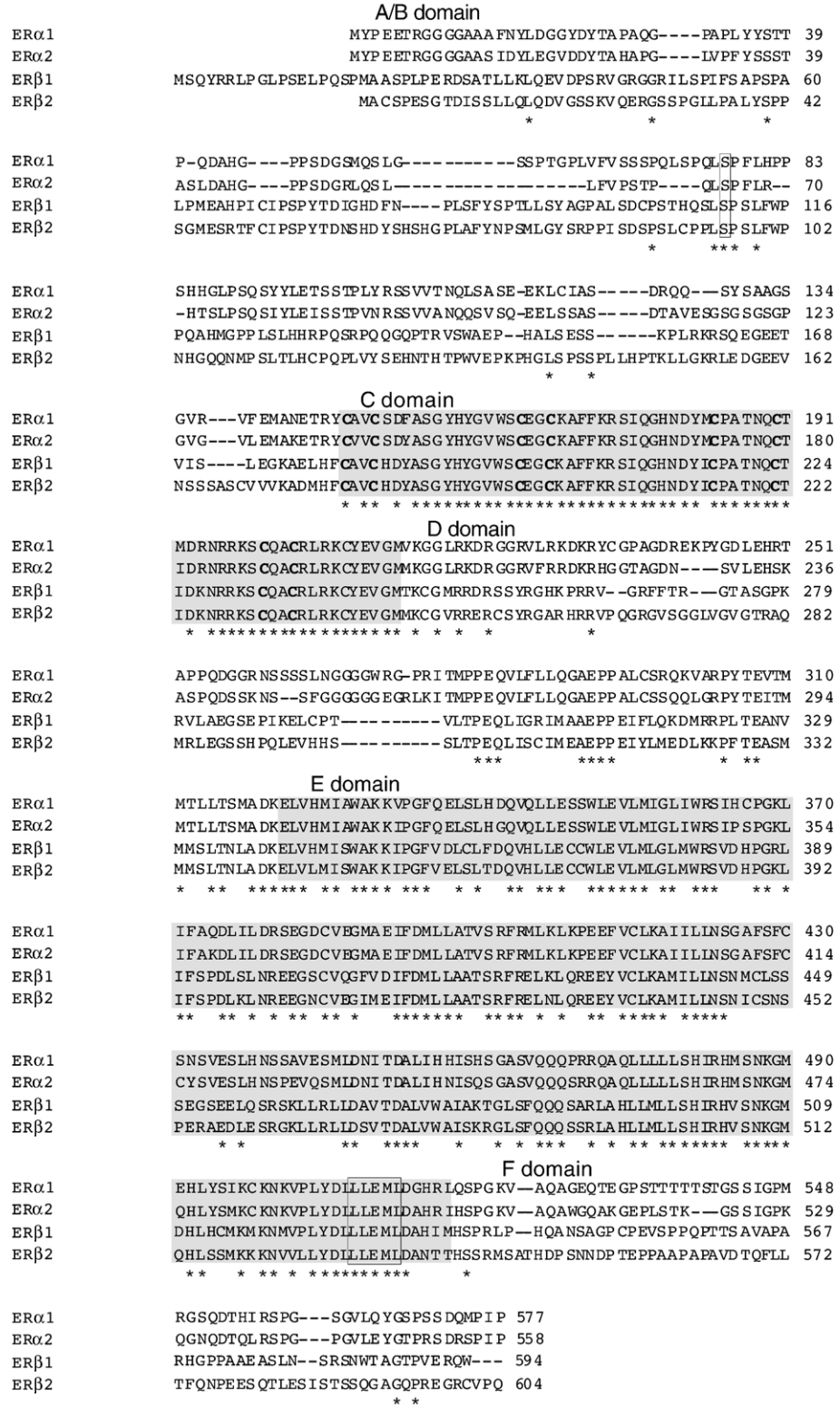

Fig. 1. Alignment comparison of the deduced amino acid sequences for rainbow trout *ERα1* (GenBank accession no. AJ242741), *ERα2* (GenBank accession no. DQ177438), *ERβ1* (GenBank accession no. DQ177439), and *ERβ2* (GenBank accession no. DQ248229). The '*' indicates positions which have a single, fully conserved residue. The different domains of the proteins, A–F, are indicated beginning at the N-terminus (top). The different domains are demarcated by the C- and E-domains, which are shaded in gray. Conserved residues within the transactivation factors AF-1 (A/B-domain) and AF-2 (F-domain) are boxed. The zinc-finger motif cysteines are indicated in bold within the C-domain.

RACE PCR, which terminated beyond a putative start site for an N-terminus (data not shown). The cDNA sequence of *ERα2* in the 3′ direction was obtained by direct sequencing of the BAC clones ER#1 and ER#3. Combining this information yielded a cDNA sequence of 2571 bases with a protein-coding region of 1674 bases and a deduced amino acid sequence for a protein containing 558 amino acids (Fig. 1). A BLAST similarity search of the GenBank database showed that the DNA sequences with the greatest similarity to this novel gene were from rainbow trout *ERα* sequences (GenBank accession nos. AJ242740 and AJ242741; Pakdel et al., 1989, 1990). An alignment comparison of the *ERα2* amino acid sequence reported here with the rainbow trout ERα amino acid sequence from GenBank accession no. AJ242741 corroborated the BLAST similarity search, indicating a member of the *ER* class from the steroid hormone superfamily. This novel rainbow trout *ER* showed high similarity with rainbow trout *ERα* (AJ242741) in the C- (95.5%) and E- (91.4%) domains, but less in the A/B- (57.0%), D- (69.2%), and F- (60.0%) domains (Table 1). This is typical amongst nuclear steroid hormone receptors where the C- and E-domains are the most conserved, whereas the others are less so (Tsai and O'Malley, 1994; Tan et al., 1996).

Table 1
Percent similarity across and between domains for the rainbow trout *ERα1*, *ERα2*, *ERβ1*, and *ERβ2* isoforms

	ERα1	*ERα2*	*ERβ1*	*ERβ2*	a.a.
Across all domains					
ERα1	100				577
ERα2	75.4	100			558
ERβ1	39.4	39.8	100		594
ERβ2	41.1	30.9	57.6	100	604
A/B-domain					
ERα1	100				147
ERα2	57.0	100			136
ERβ1	16.6	18.3	100		180
ERβ2	18.5	16.6	36.7	100	178
C-domain					
ERα1	100				66
ERα2	95.5	100			66
ERβ1	92.4	93.9	100		66
ERβ2	92.4	93.9	100	100	66
D-domain					
ERα1	100				107
ERα2	69.2	100			102
ERβ1	25.2	26.5	100		93
ERβ2	24.8	29.4	50.0	100	98
E-domain					
ERα1	100				197
ERα2	91.4	100			197
ERβ1	61.9	62.4	100		197
ERβ2	62.9	65.0	82.7	100	197
F-domain					
ERα1	100				60
ERα2	60.0	100			57
ERβ1	21.3	15.0	100		58
ERβ2	10.8	16.7	20.0	100	65

Comparison was made using MEGA 3.1.

Significantly, less similarity was noted in all domains when this sequence was compared with the other two rainbow trout nuclear receptor genes reported below (i.e., *ERβ*s) demonstrating that this novel *ER* is an *ERα* (Table 1).

The first *ER* in fishes, an *ERα*, was reported by Pakdel et al. (1989) for the rainbow trout. Because ray-finned fishes are thought to have undergone two whole genome expansions that resulted in the duplication of many (or all) genes (Wittbrodt et al., 1998; Thornton, 2001) it has been anticipated that two isoforms of each *ER* subtype should be present. In support of this, there is evidence in several species of fish for both isoforms of the *ERβ* subtype (Tchoudakova et al., 1999; Ma et al., 2000; Menuet et al., 2002; Halm et al., 2004; Hawkins and Thomas, 2004; Sabo-Attwood et al., 2004). However, a second isoform of *ERα* has never been reported in any fish until now. It is proposed that the first rainbow trout *ER* (Pakdel et al., 1989) be termed *ERα1* and the other *ERα* isoform reported in this study be denoted *ERα2*.

Two other unique cDNA sequences, believed to be the two isoforms of *ERβ* in the rainbow trout, were also discovered as part of this study. The cDNAs were obtained by degenerate primed PCR, followed by 3′- and 5′ RACE PCR to extend the cDNA sequence in both directions. The DNA products were cloned and sequenced to piece together the complete coding and non-coding regions. One cDNA (putative *ERβ1*) had a cDNA sequence of 2403 bp with a protein-coding region of 1782 bases and a deduced amino acid sequence for a protein containing 594 amino acids (Fig. 1). BLAST similarity searches of the GenBank database showed that this DNA sequence had the greatest similarities with other fish *ERβ1*s. The other cDNA (putative *ERβ2*) had a sequence of 2348 bp with a protein-coding region of 1812 bases and a deduced amino acid sequence for a protein containing 604 amino acids (Fig. 1). BLAST similarity searches of the GenBank database showed that this cDNA sequence had the greatest similarities with other fish *ERβ2*s. These two rainbow trout *ERβ*s when compared to one another had identical C-domains (100%) and very similar E-domains (82.7%) domains, but far less similarity in the A/B- (36.7%), D- (50.0%), and F- (20.0%) domains (Table 1). Therefore, similar to the situation in a number of other ray-finned fish species, in which two isoforms of the *ERβ* subtype are present (Tchoudakova et al., 1999; Ma et al., 2000; Menuet et al., 2002; Hawkins and Thomas, 2004; Sabo-Attwood et al., 2004; Halm et al., 2004), two different *ERβ* isoforms (i.e., *ERβ1* and *ERβ2*) are found in the rainbow trout also.

The amino acid alignment comparisons of the whole protein (Fig. 1) and discrete domains (Table 1) for all four rainbow trout *ER*s show key features that underline their relatedness. This is particularly evident in the C-domain with the eight highly conserved cysteine residues that make up the zinc-finger motifs. Similarly, the signature amino acids for the transactivation factors AF-1 and AF-2 found in the A/B- and F-domains, respectively, are identical across all four rainbow trout *ER*s isoforms. These hallmarks establish that these are all nuclear *ER*s.

3.2. Phylogenetic relationships

All the phylogenetic analyses conducted produced broadly congruent phylogenies. Thus, we present and discuss the

ML and Bayesian estimates of phylogeny for all currently documented fish *ER*s, which is very strongly supported (Fig. 2). This shows a clear separation of the *ERα* and *ERβ* subtypes into two distinct clades (Fig. 2). Within the *ERα* clade, the rainbow trout *ERα2* isoform that we have identified (OmyDQ177438) is most closely related to a clade containing the rainbow trout *ERα1* isoform and *ERα*s from other members of the Family Salmonidae. Were these two sequences to represent allelic forms of the same locus, we would expect them to be monophyletic (i.e., be sister taxa). The Bayesian test we conducted for this hypothesis indicates that it has a very low posterior probability ($P > 0.00001$), suggesting that the two *ERα* forms that have now been identified in rainbow trout represent distinct loci. The strongly supported position of the new *ERα*

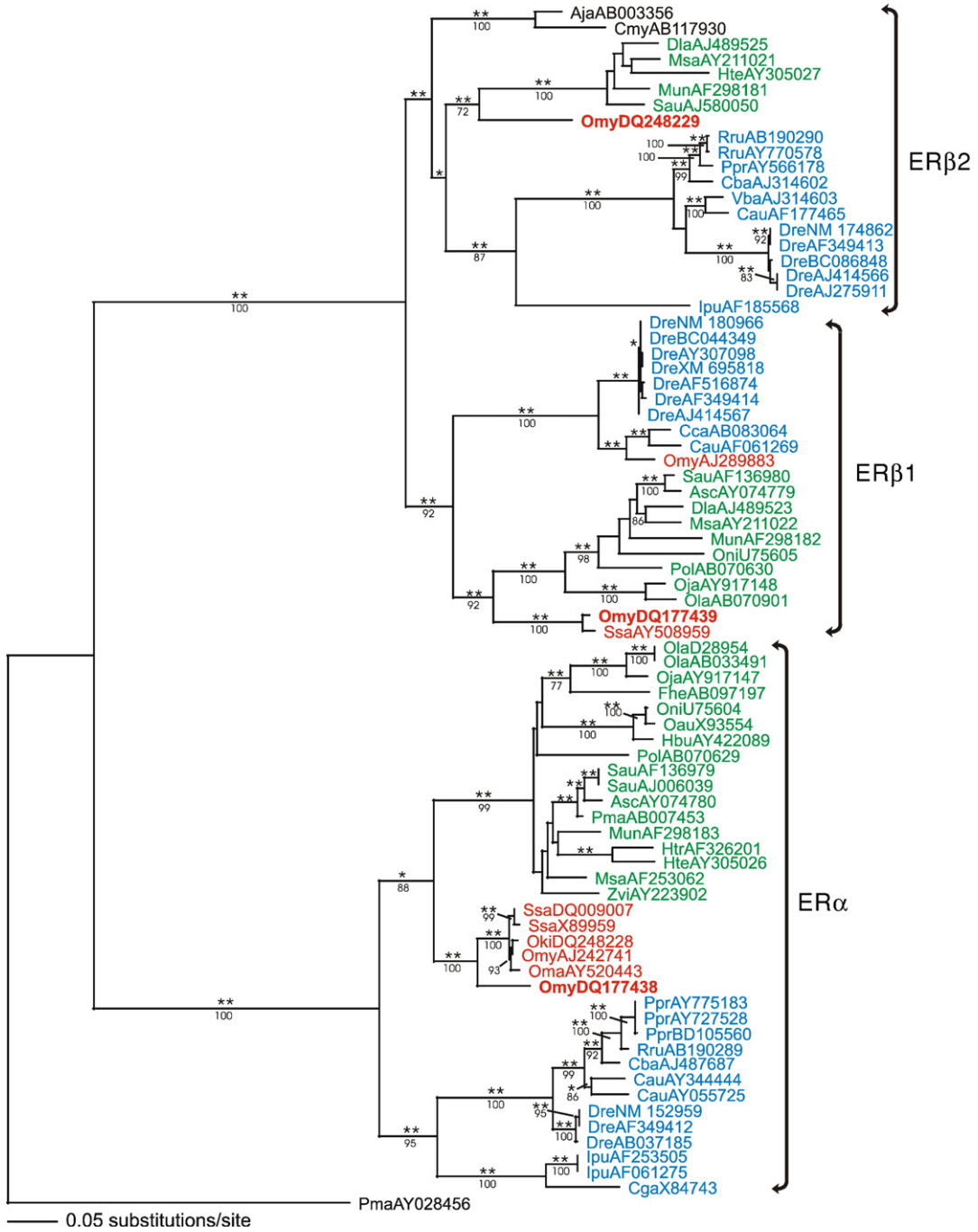

Fig. 2. Maximum-likelihood estimate of the gene-family phylogeny for all currently documented fish ERs. This estimate was derived following an iterative search strategy, under the model selected using decision theory. Taxon labels in green are sequences of Acanthopterygii, in blue are Ostariophysi, in red are Protacanthopterygii (i.e., salmonids), and the others in black. The rainbow trout sequences reported in this study are in bold red. Branches denoted by two asterisks are supported by Bayesian posterior probabilities of 100%, those with a single asterisk have probabilities >95%. Nodes with support between 50% and 95% are indicated. Parsimony bootstraps are indicated below the branches.

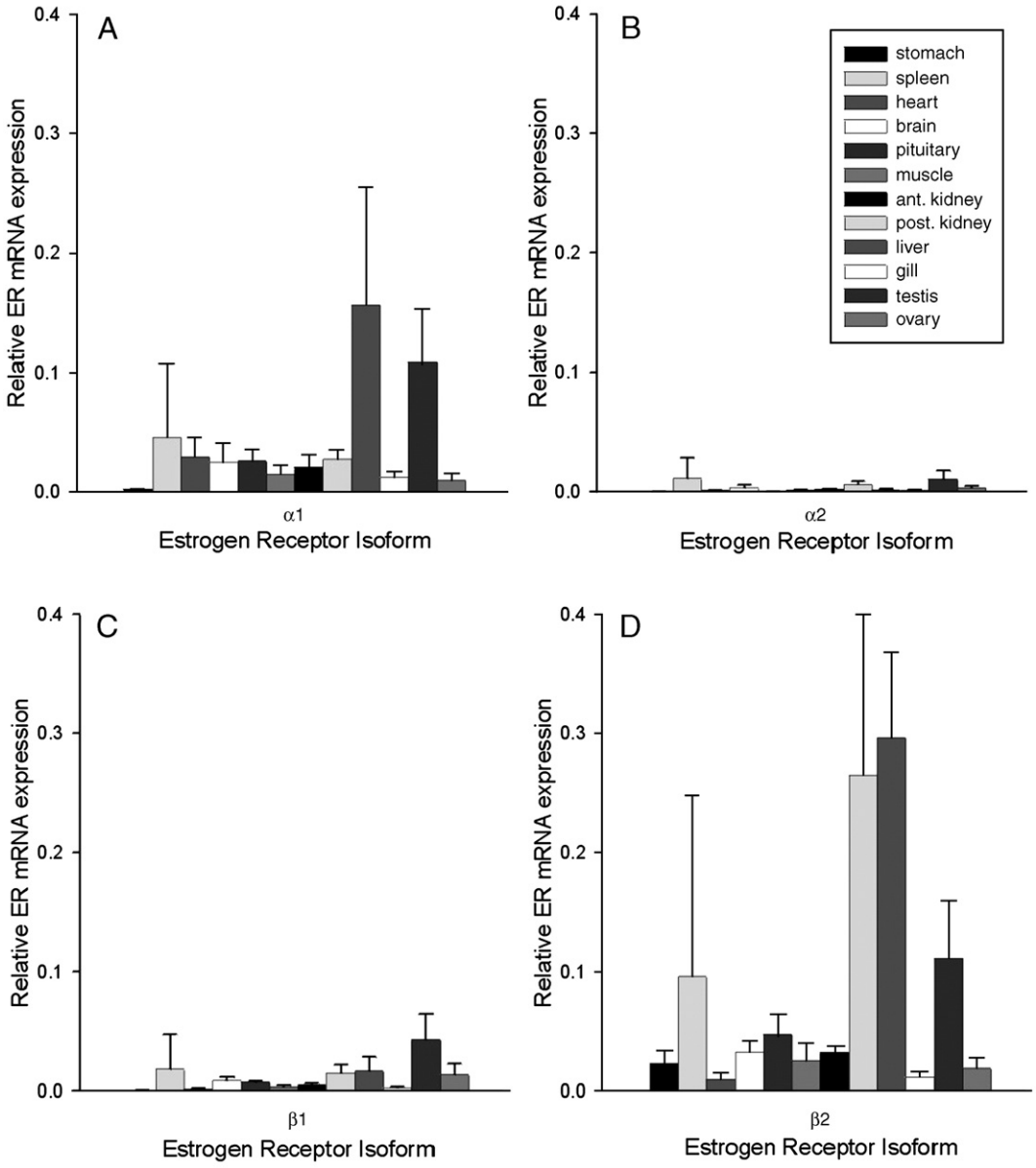

Fig. 3. Histograms showing the relative *ER* mRNA expression, measured by real-time RT-PCR, for each *ER* isoform (A—*ERα1*:*α1*; B—*ERα2*:*α2*; C—*ERβ1*:*β1*; D—*ERβ2*:*β2*) in various rainbow trout tissues (tissue legend, from top to bottom, corresponds with left to right on the x-axis). The data is presented as a mean+/− standard deviation (n=three fish). Note: all *ER* mRNAs were detected in all tissues studied, however due to low *ER* mRNA levels in some cases the relative expression is not observable due to the *y*-axis scale used.

sequence as sister to all other known salmonid *ERα1* sequences corroborates this view and suggests the possibility that there has been a salmonid-specific duplication of the *ERα* locus. The presence of *ERα2* in other salmonid fish species however remains to be determined. By contrast, two goldfish (*Carassius auratus*) *ERα* sequences (GenBank accession nos. AY344444 and AY055725) are sister to each other and may represent allelic variants of the same (*ERα1*) locus, similar to *ERα1* splice variants reported for rainbow trout (Pakdel et al., 2000) and channel catfish (Patino et al., 2000).

Collectively, these results demonstrate that a gene duplication of the *ERα* subtype to create the two isoforms, *ERα1* and *ERα2*, has occurred in the rainbow trout as predicted. However, in contrast to what was initially hypothesized the rainbow trout *ERα* duplication event appears to be a single, independent gene duplication and not part of a genome wide mechanism. This is based on the widely divergent phylogenetic patterns displayed by the two *ER* subtypes (Fig. 2). The duplication of the rainbow trout *ERα* subtype occurred much more recently than the duplication of teleost *ERβ* subtypes and there is no evidence yet for simultaneous duplication of *ERα* subtypes in other teleost species.

A very different pattern of relationships is exhibited by the *ERβ* subtype. There is a clear, strongly supported, early secondary split between the two *ERβ* isoforms (i.e. *ERβ1* and *ERβ2*) in this clade (Fig. 2), a pattern consistent with an early genome wide duplication event in ray-finned fishes. In general, the *ERβ1* and *ERβ2* topologies are largely congruent with well-corroborated fish phylogenies. Within the *ERβ2* subclade, the new rainbow

trout *ERβ2* sequence (OmyDQ248229) is sister to sequences of the acanthopterygians, whereas the ostariophysian sequences do not form a monophyletic group (although paraphyly of the ostariophysian *ERβ2* sequences is not strongly supported). Within the *ERβ1* subclade, the new rainbow trout *ERβ1* sequence we have identified (OmyDQ177439) is sister to the Atlantic salmon sequence (SsaAY508959) rather than to the rainbow trout *ERβ1* sequence that has been reported previously (OmyAJ289883). This previously reported sequence (i.e., OmyAJ289883) is sister to the ostariophysian *ERβ1* sequences. This phylogenetic placement, taken together with our inability to reproduce this sequence by PCR with primers that match the sequence perfectly (discussed above), suggests that OmyAJ289883 is spurious, and not actually a rainbow trout sequence.

3.3. ER mRNA amount and distribution in different rainbow trout tissues

Real-time PCR was used to measure the mRNA levels of all four *ER* isoforms (*ERα1*, *ERα2*, *ERβ1*, and *ERβ2*) in a number of different tissues in the rainbow trout. The mRNAs for each of the four *ER*s were detected in every tissue examined (Fig. 3), although the levels varied considerably. The liver tended to have the highest *ER* mRNA levels along with the testes, except for *ERα2*, in which case the spleen was the highest. The lowest levels were generally found in the stomach, the exception being *ERβ2* where the heart had the lowest levels measured. This data demonstrates a ubiquitous distribution of nuclear *ER*s in the rainbow trout, in line with the numerous functions estrogens are reported to have in diverse tissues and organs of vertebrate animals (Britt and Findlay, 2002; Goetz et al., 1999; Hess, 2003; Lindberg et al., 2001). The high *ER* mRNA levels in the liver, particularly for *ERα1* and *ERβ2*, are in keeping with the effect estrogen has on this organ to induce vitellogenesis in rainbow trout (Campbell and Idler, 1980; Sumpter, 1985). The relatively prominent levels of all *ER* mRNAs in the testes of these juvenile male fish suggest a reproductive role that may be more significant at this life history phase than in the female (the ovary *ER* mRNA levels being lower). The results of this study now dramatically increase the potential complexity for interaction between *ER* isoforms in these and other organs in this fish. More in-depth investigations of individual tissues at different points during development and throughout the life history of the rainbow trout will now be required to better understand the biological significance of each *ER* isoform.

Acknowledgments

This work was supported by the NSF-Idaho EPSCoR Program and by the National Science Foundation Cooperative Agreement number EPS-9720634 to JJN, and National Institute of Health 5-R01 ES 12446-2 to JJN.

References

Amores, A., et al., 1994. Zebrafish *hox* clusters and vertebrate genome evolution. Science 282, 1711–1714.

Applied Biosystems, 1997. ABI User Bulletin #2. Foster City, CA.

Ausubel, F.M., et al., 1998. Current Protocols in Molecular Biology. John Wiley and Sons, Inc, New York.

Britt, K.L., Findlay, J.K., 2002. Estrogen actions in the ovary revisited. J. Endocrinol. 175, 269–276.

Campbell, C.M., Idler, D.R., 1980. Characterization of an estradiol-induced protein from rainbow trout as vitellogenin by the composition and radioimmunological cross reactivity to ovarian yolk fractions. Biol. Reprod. 22, 605–617.

Cheung, E., Schwabish, M.A., Kraus, W.L., 2003. Chromatin exposes intrinsic differences in the transcriptional activities of estrogen receptors α and β. EMBO J. 22, 600–611.

Edwards, D.P., 2005. Regulation of signal transduction pathways by estrogen and progesterone. Annu. Rev. Physiol. 67, 23.1–23.42.

Felsenstein, J., 1985. Confidence limits on phylogeny: an approach using the bootstrap. Evolution 3, 783–791.

Filby, A.L., Tyler, C.R., 2005. Molecular characterization of estrogen receptors 1, 2a, and 2b and their tissue and ontogenic expression profiles in fathead minnow (*Pimephales promelas*). Biol. Reprod. 73, 648–662.

Gilsbach, R., Kouta, M., Bonisch, H., Bruss, M., 2006. Comparison of in vitro and in vivo reference genes for internal standardization of real-time PCR data. BioTechniques 40, 173–177.

Goetz, R.M., Thatte, H.S., Prabhakar, P., Cho, M.R., Michel, T., Golan, D.E., 1999. Estradiol induces the calcium-dependent translocation of endothelial nitric oxide synthase. Proc. Natl. Acad. Sci. U. S. A. 96, 2788–2793.

Halm, S., et al., 2004. Cloning, characterization, and expression of three oestrogen receptors (ERa, ERb1 and ERb2) in the European sea bass, *Dicentrarchus labrax*. Mol. Cell. Endocrinol. 223, 63–75.

Hawkins, M.B., Thomas, P., 2004. The unusual binding properties of the third distinct teleost estrogen receptor subtype ERβa are accompanied by highly conserved amino acid changes in the ligand binding domain. Endocrinology 145, 2968–2977.

Hess, R.A., 2003. Estrogen in the adult male reproductive tract: a review. Reprod. Biol. Endocrinol. 1, 52.

Hirata, S., Shoda, T., Kato, J., Hoshi, K., 2003. Isoform/variant mRNAs for sex steroid hormone receptors in humans. Trends Endocrinol. Metab. 14, 124–129.

Lange, I.G., Hartel, A., Meyer, H.H.D., 2003. Evolution of oestrogen functions in vertebrates. J. Steroid Biochem. Mol. Biol. 83, 219–226.

Laudet, V., Hänni, C., Coll, J., Catzeflis, C., Stéhelin, D., 1992. Evolution of the nuclear receptor gene family. EMBO J. 11, 1003–1013.

Lindberg, M.K., Alatalo, S.L., Halleen, J.M., Mohan, S., Gustafsson, J.-A., Ohlsson, C., 2001. Estrogen receptor specificity in the regulation of the skeleton in female mice. J. Endocrinol. 171, 229–236.

Lockhart, P.J., Steel, M.A., Hendy, M.D., Penny, D., 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. 12, 28–52.

Loomis, A.K., Thomas, P., 2000. Effects of estrogens and xenoestrogens on androgen production by Atlantic croaker testes in vitro: evidence for a nongenomic action mediated by an estrogen membrane receptor. Biol. Reprod. 62, 995–1004.

Ma, C.H., Dong, K.W., Yu, K.L., 2000. cDNA cloning and expression of a novel estrogen receptor beta-subtype in goldfish (*Carassius auratus*). Biochim. Biophys. Acta 1490, 145–152.

Menuet, A., et al., 2002. Molecular characterization of three estrogen receptor forms in zebrafish: binding characteristics, transactivation properties, and tissue distributions. Biol. Reprod. 66, 1881–1892.

Minin, V., Abdo, Z., Joyce, P., Sullivan, J., 2003. Performance-based selection of likelihood models for phylogeny estimation. Syst. Biol. 52, 674–683.

Nagler, J.J., Krisfalusi, M., 2000. Different gonadal estrogen receptor messenger RNAs in the rainbow trout: molecular characterization and measurement during reproductive development. Proc. 6th Int. Symp. Reprod. Physiol. Fish, July 4–9, 1999, Bergen, Norway, p. 183.

Nagler, J.J., Krisfalusi, M., Cyr, D.G., 2000. Quantification of rainbow trout (*Oncorhynchus mykiss*) estrogen receptor messenger RNA and its expression in the ovary during the reproductive cycle. J. Mol. Endocrinol. 25, 243–251.

Nuclear Receptors Nomenclature Committee, 1999. A unified nomenclature system for the nuclear receptor superfamily. Cell 97, 161–163.

Pakdel, F., Le Guellec, C., Vaillant, C., Le Roux, M.G., Valotaire, Y., 1989. Identification and estrogen induction of two estrogen receptors (ER)

messenger ribonucleic acids in the rainbow trout liver: sequence homology with other ERs. Mol. Endocrinol. 3, 44–51.

Pakdel, F., Le Gac, F., Le Goff, P., Valotaire, Y., 1990. Full-length sequence and in vitro expression of rainbow trout estrogen receptor cDNA. Mol. Cell. Endocrinol. 71, 195–204.

Pakdel, F., Metivier, R., Flouriot, G., Valotaire, Y., 2000. Two estrogen receptor (ER) isoforms with different estrogen dependencies are generated from the trout ER gene. Endocrinology 141, 571–580.

Palti, Y., Gahr, S.A., Hansen, J.D., Rexroad III, C.E., 2004. Characterization of a new BAC library for rainbow trout: evidence for multi-locus duplication. Anim. Genet. 35, 130–133.

Pappas, T.C., Gametchu, B., Watson, C.S., 1995. Membrane estrogen receptors identified by antibody and impeded ligand labeling. FASEB J. 9, 404–410.

Patino, R., Xia, Z., Gale, W.L., Wu, C., Maule, A.G., Chang, X., 2000. Novel transcripts of the estrogen receptor α gene in channel catfish. Gen. Comp. Endocrinol. 120, 314–325.

Pettersson, K., Gustafsson, J.-Å., 2001. Role of estrogen receptor beta in estrogen action. Annu. Rev. Physiol. 63, 165–192.

Revankar, C.M., Cimino, D.F., Sklar, L.A., Arterburn, J.B., Prossnitz, E.R., 2005. A transmembrane intracellular estrogen receptor mediates rapid cell signaling. Science 307, 1625–1630.

Ronquist, F., Huelsenbeck, J.P., 2005. Bayesian analysis of molecular evolution using MrBayes. In: Nielsen, R. (Ed.), Statistical Methods in Molecular Evolution. Springer, New York.

Sabo-Attwood, T., Kroll, K.J., Denslow, N.D., 2004. Differential expression of largemouth bass (*Micropterus salmoides*) estrogen receptor isotypes alpha, beta, and gamma by estradiol. Mol. Cell. Endocrinol. 218, 107–118.

Simoncini, T., Genazzani, A.R., 2003. Non-genomic actions of sex hormones. Eur. J. Endocrinol. 148, 281–292.

Sullivan, J., Holsinger, K.E., Simon, C., 1996. The effect of topology on estimates of among-site rate variation. J. Mol. Evol. 42, 308–312.

Sullivan, J., Abdo, Z., Joyce, P., Swofford, D.L., 2005. Evaluating the performance of a successive-approximations approach to parameter optimization in maximum-likelihood phylogeny estimation. Mol. Biol. Evol. 22, 1386–1392.

Sumpter, J.P., 1985. The purification, radioimmunoassay and plasma levels of vitellogenin from the rainbow trout, *Salmo gairdneri*. In: Lofts, B., Holmes, W.N. (Eds.), Current Trends in Comparative Endocrinology. Hong Kong University Press, Hong Kong, pp. 355–357.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (and Other Methods). Version 4. Sinauer, Sunderland, MA.

Swofford, D.L., Olsen, G.J., Waddell, P.J., Hillis, D.M., 1996. Phylogenetic inference, In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics, 2nd edition. Sinauer, Sunderland, Massachusetts, pp. 407–514.

Tan, N.S., Lam, T.J., Ding, J.L., 1996. The first contiguous estrogen receptor gene from a fish, *Oreochromis aureus*: evidence for multiple transcripts. Mol. Cell. Endocrinol. 120, 177–192.

Tchoudakova, A., Pathak, S., Callard, G.V., 1999. Molecular cloning of an estrogen receptor β subtype from the goldfish, *Carassius auratus*. Gen. Comp. Endocrinol. 113, 388–400.

Thomas, P., Pang, Y., Filardo, E.J., Dong, J., 2005. Identity of an estrogen membrane receptor coupled to a G-protein in human breast cancer cells. Endocrinology 146, 624–632.

Thornton, J.W., 2001. Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. PNAS 98, 5671–5676.

Tsai, M.-J., O'Malley, B.W., 1994. Molecular mechanisms of action of steroid/thyroid receptor superfamily members. Annu. Rev. Biochem. 63, 451–486.

Wittbrodt, J., Meyer, A., Schartl, M., 1998. More genes in fish? BioEssays 20, 511–515.