Available online at www.sciencedirect.com
SCIENCE DIRECT®
General and Comparative Endocrinology 140 (2005) 74–83
GENERAL AND COMPARATIVE ENDOCRINOLOGY
www.elsevier.com/locate/ygcen

Characterization of a cetacean aromatase (CYP19) and the phylogeny and functional conservation of vertebrate aromatase

Joanna Y. Wilson[a,*], Andrew G. McArthur[b], John J. Stegeman[a]

[a] *Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA*
[b] *The Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, MA 02543, USA*

Received 11 May 2004; revised 29 September 2004; accepted 12 October 2004
Available online 24 November 2004

Abstract

Aromatase (cytochrome P450 19, CYP19, P450arom) is the enzyme responsible for the production of estrogens, hormones critical for development and reproduction. Aromatase was sequenced from a white-sided dolphin (*Lagenorhynchus acutus*) ovary, transiently transfected into HEK 293 cells, and the expressed protein was characterized for aromatase activity in the presence of androstenedione and testosterone and after exposure to the aromatase inhibitor letrazole. The K_ms for androstenedione and testosterone were 63.5 and 75 nM, respectively, values that are very similar to those reported for other mammalian aromatases. A Bayesian phylogenetic analysis of the vertebrate aromatases was performed on the amino acid sequences of aromatases from fish, amphibians, reptiles, birds, and mammals. Based on known species phylogeny, the cetacean aromatase showed an expected grouping with artiodactyls (cow, sheep, and goat). An analysis of functional divergence showed strong conservation of aromatase across the entire protein, which indicates that the observed sequence divergence is functionally neutral.

Keywords: Cytochrome P450; CYP19; Aromatase; P450arom; Heterologous expression; Phylogeny; Vertebrata; Mammalia; Pisces; Cetacean

1. Introduction

Aromatase (CYP19, P450arom) is the enzyme responsible for the generation of estrogens from their precursor androgens (Corbin et al., 1988). Aromatase is thought to be expressed at least in gonads and brain within vertebrates, with additional expression known in placenta of primates and artiodactyls, and broad expression in humans (Conley and Hinshelwood, 2001; Simpson et al., 1994, 2002). The expression of aromatase is critical for reproductive success. Although loss of function mutations in *CYP19* are rare in humans, they have been linked to infertility (Adashi and Hennebold, 1999; Simpson, 1998), as has been seen in the CYP19 knockout mice (Robertson et al., 2001; Toda et al., 2001b). In humans, where aromatase is expressed in many internal organs, estrogen also appears to be important for vascular function, lipid and carbohydrate metabolism, and bone mineralization (reviewed in Simpson et al., 2002). The extensive involvement of estrogen in both reproductive and non-reproductive functions in mammals is supported by recent studies in aromatase knockout mice (Agarwal et al., 2000; Britt et al., 2001; Fisher et al., 1998; Honda et al., 1998; Jones et al., 2000, 2001; Murata et al., 2002; Robertson et al., 1999; Toda et al., 2001a).

In spite of its importance in reproduction, the aromatase gene has been studied in relatively few mammals and extensively only in humans, where it is thought to play significant roles also in pathophysiologies such as breast cancer and reproductive dysfunction (Adashi and Hennebold, 1999; Chen et al., 1999; Simpson, 1998). Full coding sequences of *CYP19* genes were available from 11 mammalian species (goat, horse, cow, sheep,

* Corresponding author. Fax: +1 508 457 2134.
E-mail address: jwilson@whoi.edu (J.Y. Wilson).

0016-6480/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygcen.2004.10.004

pig, human, marmoset, macaque, mouse, rat, and rabbit). With only single *CYP19* genes evident in human, cow, and horse, *CYP19* is thought to be present as a single gene in most mammals (Conley and Hinshelwood, 2001; Simpson et al., 2002). The pig, with three *CYP19* genes (Choi et al., 1996; Corbin et al., 1999; Graddy et al., 2000), appears to be an exception. Sequencing and characterization of *CYP19* in additional members of the ungulate clade, where multiple aromatase genes occur in at least one species, may enhance our understanding of the conservation and function of this important protein.

Cetaceans (whales and dolphins) are most closely related to ruminant artiodactyls such as the cow (Gatesy and O'Leary, 2001). Many cetacean species were subject to intense hunting and population recovery is not yet certain. The endangered North Atlantic right whale (*Eubalaena glacialis*) appears to have a reproductive shortfall which accounts for at least part of the lack of recovery in this population (Kraus et al., 2001). The reasons for such reproductive deficits are not known but may involve genes important for cetacean reproduction. Experimental studies of marine mammals are precluded; however, in vitro studies could disclose the similarities and differences between reproductive systems of cetaceans and other mammals. It has been suggested that the *CYP19* gene and estradiol function are strongly conserved across the Mammalia, based on the investigations of a few species (Conley and Hinshelwood, 2001). However, whether information from model organisms might apply to understanding reproductive gene function in diverse mammalian wildlife species is not known.

The objective of this research was to establish a molecular foundation for understanding estrogen production in cetaceans and to assess functional differences among aromatases, from both a catalytic and an evolutionary perspective. We cloned the full-length coding sequence of the *CYP19* gene from a white-sided dolphin (*Lagenorhynchus acutus*) ovary and characterized the aromatase activity of this cetacean aromatase expressed in transient transfection experiments. We investigated aromatase diversification by performing a Bayesian phylogenetic analysis of amino acid sequences of 41 aromatases that include mammalian, amphibian, reptile, bird, and fish forms. To examine the evolution of aromatase function, we combined phylogenetic analysis with statistical analysis of sequence divergence.

2. Materials and methods

2.1. Tissue preparation and isolation of dolphin CYP19 cDNA

Total RNA and poly(A)$^+$ mRNA were isolated from the ovary of a white-sided dolphin (*L. acutus*, Accession No. MH-97-534)[1] using Stat60 (Tel-test, Friendswood, TX) and a mini Oligo(dT) poly(A)$^+$ cellulose column spin kit (5′ → 3′ Prime, Boulder, CO), respectively. Two overlapping fragments (702 and 830 bp, covering a total of 1332 bp) of the aromatase gene were amplified by RT-PCR using the GeneAmp RNA PCR kit (Applied Biosystem, Foster City, CA) and degenerate mammalian aromatase primers (Table 1, internal set 1 and 2) with the following protocol: 94 °C for 7 min; 35 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 90 s; and 72 °C for 10 min. Rapid amplification of cDNA ends (RACE) was performed using nested gene-specific primers (Table 1) and poly(A)$^+$ mRNA with the Marathon RACE kit (Clontech, Palo Alto CA). RACE-ready cDNA was generated using poly(A)$^+$ mRNA and the RACE reactions were carried out using a touchdown protocol: 94 °C for 30 s, 5 cycles of 94 °C for 5 s and 72 °C for 3 min, 5 cycles of 94 °C for 5 s and 70 °C for 3 min, and 25 cycles of 94 °C for 5 s and 68 °C for 3 min. Nested RACE reactions were performed as outlined above except that the final cycle was performed for 15 cycles only. The RT-PCR and RACE fragments were gel purified (Geneclean II, Bio 101 Systems, Vista CA), cloned using the pGEM-T Easy vector (Promega, Madison WI), and prepared for sequencing using a QiaPrep miniprep kit (Qiagen, Valencia CA). Clone sequences were generated using a cycle-sequence protocol (Epicentre Technologies, Madison, WI) with IR-labeled vector-primers (M13F and M13R) and resolved on a LI-COR 4200L sequencing apparatus (LI-COR, Lincoln, NE). Multiple clones of each fragment were bi-directionally sequenced (four clones from the 830 bp and three clones of the 702 bp RT-PCR fragments, 14 clones from the 3′RACE, and six clones from the nested 5′RACE). Contig assemblies were generated with Sequencher (GeneCodes, Ann Arbor MI).

2.2. Primers

Degenerate mammalian *CYP19* PCR primers (Table 1) were designed based on a DNA alignment of full-length coding sequences of *CYP19* from cow, horse, pig, mouse, rat, rabbit, and human (see Table 2 for GenBank accession numbers) generated with Clustal (Higgins, 1994). RACE primers (Table 1) were designed based on the consensus dolphin aromatase sequence generated from the initial RT-PCR fragments.

2.3. Phylogenetic and functional analyses

CYP19 amino acid sequences (Table 2) used for phylogenetic analyses were retrieved from GenBank by BLAST searches. Alignment and analysis were

[1] Accession number from stranding response team at the New England Aquarium, Boston MA.

Table 1
Primer sequences and fragment sizes for RT-PCR, RACE, and amplification of the full-length sequence of white-sided dolphin aromatase

Primer name	Primer sequence	Fragment size
Internal set 1		
CYP19_672-691F	5′-GCAAGCTCTCCTYMTCAAAC-3′	~700 bp
CYP19_1353-1373R	5′-AAKCGTCTCARAAGTGTRACC-3′	
Internal set 2		
CYP19_2-23F	5′-TGNTTTTGGAAATGCTGAACCC-3′	~830 bp
CYP19_810-839R	5′-TCCATRBDGTCTTCCAGTTTC-3′	
3′ RACE primers		
CYP19_3′RACE1	5′-ACAAGCTCTTCCTGGGGAGCCCCTTGG-3′	~1.5 kb, ~3 kb
CYP19_3′RACE2	5′-GGATGATGTCATCGATGGCTACCCAGTG-3′	~900 bp, ~1.2 kb
5′ RACE primers		
CYP19_5′RACE	5′-CAACAGGCGCAACTTCGGTCACCATGC-3′	~700 bp[a]
CYP19_5′RACE2	5′-GTCCAAGGGGATCCCCAGGAAGAGCTTG-3′	None
CYP19_5′RACE3	5′-CCTTTCTCATGCATGCCGATGCACTGC-3′	None
Full-length primers		
Fullcds_CYP19F	5′-CGCAAGATGGTTTTGGAAGTGC-3′	1544 bp
Fullcds_CYP19R	5′-AGGGCAGGGACTGACCAAAC-3′	

[a] Fragments were generated with a nested PCR, using the RACE reaction generated with primer CYP19_5′RACE2 and CYP19_5′RACE3.

performed as described elsewhere (McArthur et al., 2003). In brief, sequence alignment was performed using ClustalX 1.81 (Thompson et al., 1997), with manual correction by eye. The stingray sequence was designated as outgroup, based on its basal position in vertebrate divergence. Phylogenetic relationships of the CYP19 sequences were assessed using the computer program MrBayes 3.0b4 (Huelsenbeck and Ronquist, 2001). MC^3 (metropolis coupled, markov chain, monte carlo) estimation of posterior probabilities was performed using non-informative prior probabilities, the JTT + I + Γ (Jones et al., 1992) substitution model with inclusion of unequal amino acid frequencies, and four incrementally heated Markov chains with different random starting trees. The Markov chains were run for 10,000,000 generations, with sampling of topologies every 100 generations. Posterior probabilities of topologies, clades, and parameters were estimated from the sampled topologies after removal of MC^3 burn-in. A functional divergence analysis of the amino acid alignment in the context of the hypothesized phylogenetic tree was performed using the computer program DIVERGE (Gu and Vander Velden, 2002). DIVERGE was used to test the null hypothesis of no changes in site-specific evolutionary rates among CYP19 subclades and to predict sites in the alignment having altered functional constraints. These sites may have been subject to divergent functional evolution. DIVERGE measures change in site-specific evolutionary rates using the coefficient of evolutionary functional divergence (θ), where $\theta = 0$ indicates no change and values approaching $\theta = 1$ reflect increasing functional divergence.

2.4. Transient transfections of dolphin CYP19

The full-length coding region of dolphin *CYP19* was generated by RT-PCR using the primers listed in Table 1. The sequence of the full-length fragment was verified and it was cloned into the pCI-neo mammalian expression vector (Promega, Madison WI). The vector, pCI-neo-dCYP19, and an empty vector, pCI-neo-control, were grown in JM109 *Escherichia coli* cells, and vector DNA was purified using the EndoFree Maxi Prep kit (Qiagen, Valencia CA). Each vector was co-transfected with the pRL-TK *Renilla* luciferase expression vector (Promega, Madison WI) into HEK 293 cells grown in Eagle's MEM (Sigma, St. Louis MO) with 1.5 g L^{-1} sodium bicarbonate, 10 mM sodium pyruvate, 10 mM Hepes, and 10% heat inactivated horse serum (Sigma, St. Louis, MO). Vector DNA was transiently transfected into cells using lipofectamine 2000 reagent (Life Technologies, Carlsbad, CA) diluted with serum free media according to manufacturer's instructions. Transfections were performed 48 h after plating cells and aromatase and luciferase activities were assayed 48 h post-transfection. Transfections were optimized (per 2 cm^2 well) for cell plating density (2×10^5 cells), vector DNA added (1 μg pCI-neo vector and 1 ng pRL-TK vector) and transfection reagent volume (2.5 μl).

Transfected cells were incubated with serum free media with 1 mg ml^{-1} bovine serum albumin (Sigma, St. Louis, MO) and either [1β-^{3}H]androstenedione or [1β,2β-^{3}H]testosterone (NEN Life Sciences, Boston, MA) at 37 °C and 5% CO_2. Testosterone or androstenedione were diluted with cold substrate to a specific activity of 8 or 5.18 μCi $nmol^{-1}$ in ethanol, respectively. The

Table 2
GenBank accession numbers of cytochrome P450 (CYP) sequences used in the present study

Species	Accession No.
Alligator CYP19 (*Alligator mississippiensis*)	AAK31803
Atlantic Halibut CYP19A (*Hippoglossus hippoglossus*)	CAC36394
Bastard Halibut CYP19A (*Paralichthys olivaceus*)	BAA74777
Catfish CYP19A (*Ictalurus punctatus*)	AAB32613
Catfish CYP19B (*Ictalurus punctatus*)	AAL14612
Chicken CYP19 (*Gallus gallus*)	A31916
Cow CYP19 (*Bos ÿabina*)	P46194, Z32741[a]
Dolphin CYP19 (*Lagenorhynchus acutus*)	AY341076
European Seabass CYP19A (*Dicentrarchus labrax*)	CAC21712
Fathead Minnow CYP19B (*Pimephales promelas*)	CAC38767
Frog CYP19 (*Xenopus laevis*)	BAA90529
Gilthead Sea Bream CYP19A (*Sparus aurata*)	AAL27699
Goat CYP19 (*Capra hircus*)	AAN23836
Goldfish CYP19A (*Carassius auratus*)	AAC14013
Goldfish CYP19B (*Carassius auratus*)	BAA23757
Horse CYP19 (*Equus caballus*)	AAC04698, AF031520[a]
Human CYP19 (*Homo sapiens*)	CAA68807, Y07508[a]
Japanese Quail CYP19 (*Coturnix japonica*)	AAN04475
Macaque CYP19 (*Macaca fascicularis*)	BAB64454
Marmoset CYP19 (*Callithrix jacchus*)	AAK58465
Medaka CYP19A (*Oryzias latipes*)	BAA11657
Mouse CYP19 (*Mus musculus*)	NP_031836, G3046857[a]
Mozambique Tilapia CYP19A (*Oreochromis mossambicus*)	AAD31031
Mozambique Tilapia CYP19B (*Oreochromis mossambicus*)	AAD31030
Newt CYP19 (*Pleurodeles waltl*)	AAM95462
Nile Tilapia CYP19A (*Oreochromis niloticus*)	P70091
Nile Tilapia CYP19B (*Oreochromis niloticus*)	AAG49480
Pig CYP19 Type I (*Sus scrofa*)	AAB51388, U92246[a]
Pig CYP19 Type II (*Sus scrofa*)	AAB61697, U52142[a]
Pig CYP19 Type III (*Sus scrofa*)	2210279A
Rabbit CYP19 (*Orycholagus cuniculus*)	CAA92574, Z70301[a]
Rainbow Trout CYP19A (*Oncorhynchus mykiss*)	1806325A
Rainbow Trout CYP19B (*Oncorhynchus mykiss*)	CAC84574
Rat CYP19 (*Rattus norvegicus*)	A36121, M33986[a]
Red Sea Bream CYP19A (*Pagrus major*)	BAB82524
Sheep CYP19 (*Ovis aries*)	CAB40563
Stingray CYP19 (*Dasyatis sabina*)	AAF04617
Turtle CYP19 (*Tarchemys scripta*)	AAG09376
Zebrafinch CYP19 (*Taeniopygia guttata*)	AAB32404
Zebrafish CYP19A (*Danio rerio*)	AAK00643
Zebrafish CYP19B (*Danio rerio*)	AAK00642

[a] Nucleotide sequences used for primer design and nucleotide phylogeny (data not shown).

volume of ethanol was 2% of the total media volume per well. Replicate wells were treated with ethanol alone to account for carrier effects. Additional triplicate wells were incubated with tritiated water to correct for the extraction efficiency and possible water loss during incubation. After incubations were complete, the media were added to a half volume of ice-cold 30% TCA. The cells were washed with PBS, harvested, and lysed for assessment of luciferase activity using either the dual or *Renilla* luciferase assay kit (Promega, Madison WI). Luciferase activity was measured in 20 µl of cell lysate on a TD-20/20 luminometer (Turner Designs, Sunnyvale, CA) after a 2 s delay and a 10 s integration time.

Aromatase activity was measured using the tritiated water release assay (Lephart and Simpson, 1991). Briefly, media/TCA aliquots were centrifuged to precipitate protein and extracted with 2.5 volumes of chloroform. The aqueous phase was mixed with an equal volume of 5% charcoal, 0.5% dextran solution and centrifuged. An aliquot of the aqueous phase was mixed with scintillation fluor (Scintiverse BD, Fisher Scientific, Hampton, NH) and counted on a scintillation counter. Aromatase activity was calculated by converting dpms counted into pmoles water produced per well after correcting for label (testosterone), extraction efficiency, and decay. Aromatase activity was normalized with *Renilla* luciferase activity in each well. Each dose or time point was done in triplicate wells and averaged.

To determine the kinetics of aromatase activity, transfected cells were exposed to 150 nM substrate and media were sampled from each well at 30 min, 1,2, 3, 4, and 6 h. 50% of media was removed from each well by the 6 h time point. For subsequent experiments, transfected cells were exposed for 2 h to either 0–150 nM substrate or 63.5 nM androstenedione and 0–100 nM letrozole (Novartis Pharmaceuticals, East Hanover, NJ).

3. Results

3.1. Dolphin aromatase sequence

A *CYP19* sequence (GenBank Accession No. AY341076) representing approximately 2000 base pairs with an open reading frame of 1509 base pairs was obtained from white-sided dolphin (*L. acutus*) ovarian tissue by RT-PCR and rapid amplification of cDNA ends (RACE, see Table 1 for specific primers). The sequence was confirmed with $\geqslant$6-fold bi-directional coverage within the coding sequence. The full-length cDNA has been cloned and sequenced. The dolphin *CYP19* was most similar to the cow *CYP19*, and ranged from 81 to 94% nucleotide and 78 to 92% amino acid identity to other mammalian full-length coding sequences (rat, rabbit, horse, human, and cow).

3.2. Phylogenetic analysis

After exclusion of poorly aligned regions, the final aromatase amino acid alignment included 436 characters, of which 282 were parsimony informative and 115 constant. The length of the MC^3 burn-in was 8000 of 10,000,000 generations, resulting in a sample of 99,920 trees for estimation of posterior probabilities. Replicate analyses provided very similar estimates of posterior probabilities. Overall resolution of CYP19 phylogeny was high, as indicated by the posterior probabilities (Fig. 1). The dolphin sequence grouped with aromatases from the artiodactyls (Fig. 1) as expected based on phylogenetic relationships within the Mammalia.

3.3. DIVERGE Analysis

Functional divergence over the taxonomic groups represented in Fig. 1 was examined in context of the tree topology, using the DIVERGE software and the rabbit CYP2C5 crystal structure, a mammalian CYP crystal structure suitable for modeling aromatase

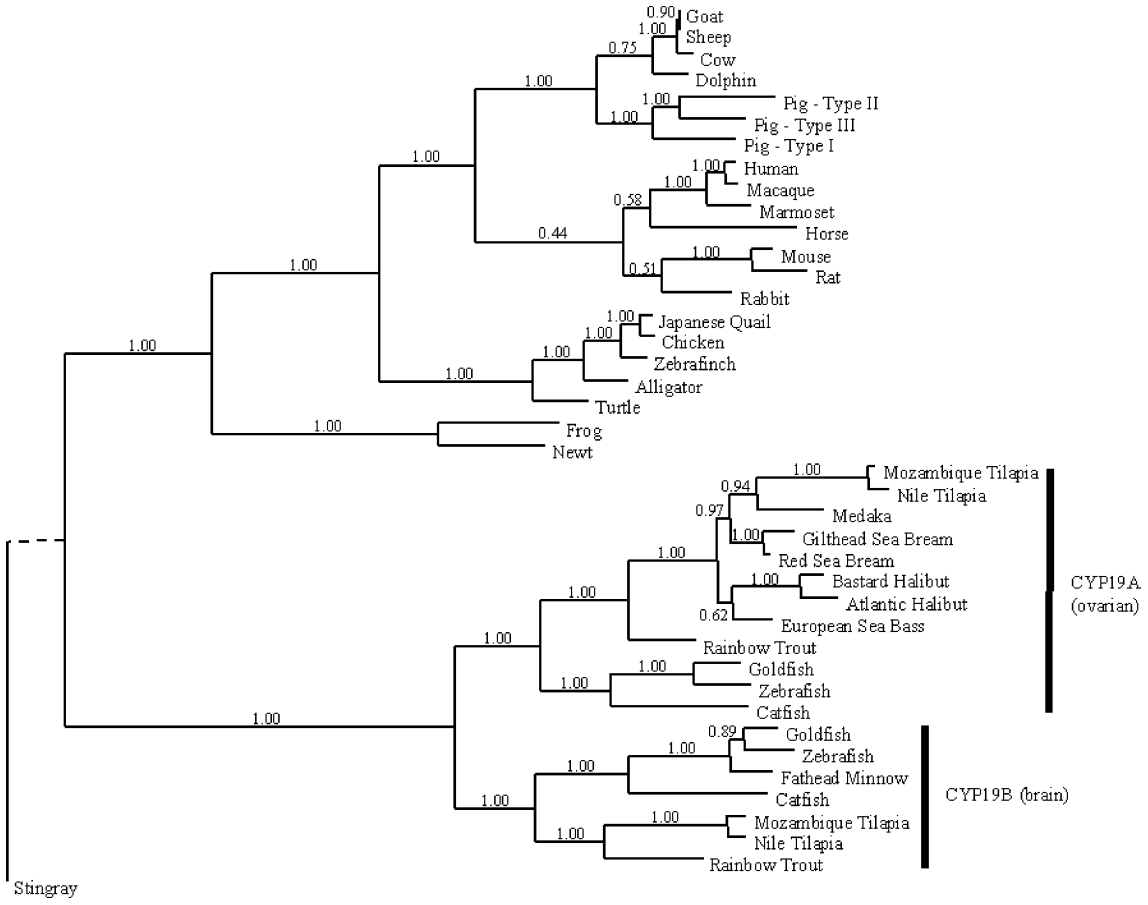

Fig. 1. Phylogeny of vertebrate aromatase. The phylogenetic tree with the highest posterior probability score found by MC^3. A better tree may exist as MC^3 is a stochastic sampling method. Bayesian posterior probabilities for each node are superimposed upon the tree. The tree is rooted using the stingray CYP19 sequence. Horizontal branch lengths are representative of evolutionary change.

Table 3
Tests of the null hypothesis of no change in site-specific evolutionary rates among the four largest CYP19 sub-clades (containing more than four proteins, see Fig. 1), as implemented by the computer program DIVERGE

	Mammal CYP19	Turtle, reptile, bird CYP19	Fish CYP19A	Fish CYP19B
Mammal CYP19	—	0.12 ± 0.11	0.48 ± 0.07	0.41 ± 0.07
Turtle, reptile, bird CYP19	1.25	—	0.20 ± 0.13	0.13 ± 0.16
Fish CYP19A	45.45*	2.47	—	0.00 ± 0.02
Fish CYP19B	30.86*	0.73	0.00	—

Measures of the coefficient of evolutionary functional divergence (θ) and their standard error are presented above the diagonal. Increasing values of θ reflect increasing functional divergence. Likelihood ratio test observed values for the test of the null hypothesis of $\theta = 0$ are presented below the diagonal. Scores with significant rejection ($P < 0.05$) of the null hypothesis are marked with an asterisk.

(Williams et al., 2000). For this analysis, at least four proteins are required for each grouping, thus the functional divergence of the pig sequences could not be specifically tested. Instead, large taxonomic groups were designated: mammals, turtle/bird/reptile, amphibians, and fish (either as one group, or divided into fish CYP19 a and b forms). Between most of these groupings, there was no functional divergence detected (Table 3). The only significant divergence from $\theta = 0$ was found between fish and mammals, although this was marginal (i.e., $\theta < 0.50$, Table 3). This difference was found whether the fish aromatases were grouped as a whole or divided into a and b forms. When divergent residues were plotted across the alignment, they appeared randomly distributed across the entire protein sequence (data not shown). The divergent residues did not cluster relative to the presumed secondary structure of the protein, inferred from the rabbit CYP2C5 crystal structure.

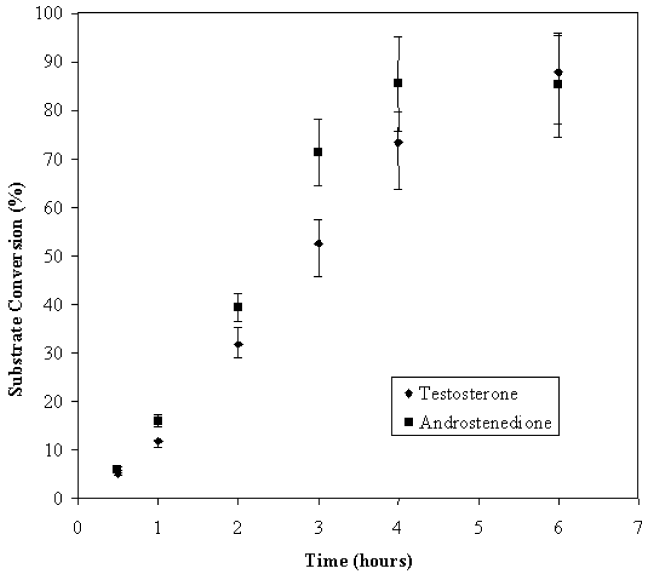

Fig. 2. Tritiated water release from testosterone and androstenedione over time in cells transiently transfected with dolphin aromatase. Assays were performed in triplicate as described in Section 2.

3.4. Transient transfections

The dolphin aromatase was transiently transfected into HEK 293 cells. This cell line does not metabolize androstenedione or testosterone to any detectable degree over 6 h or more of culture (Corbin et al., 1999), which makes HEK 293 cells useful for functional studies of expressed aromatase. HEK 293 cells have been used for both transient and stable transfections of porcine aromatases (Corbin et al., 1999, 2001).

Transfectant-dependent appearance of estradiol was detected in the cell culture media of transfected cells incubated with unlabeled testosterone by radioimmunoassay (data not shown), indicating that the transfection protocol yielded a functional CYP19. The addition of 100 nM letrozole, a specific inhibitor of aromatase in other mammals, inhibited the dolphin aromatase activity by 96% in the presence of 63.5 nM androstenedione (data not shown). The presence of estradiol in the media

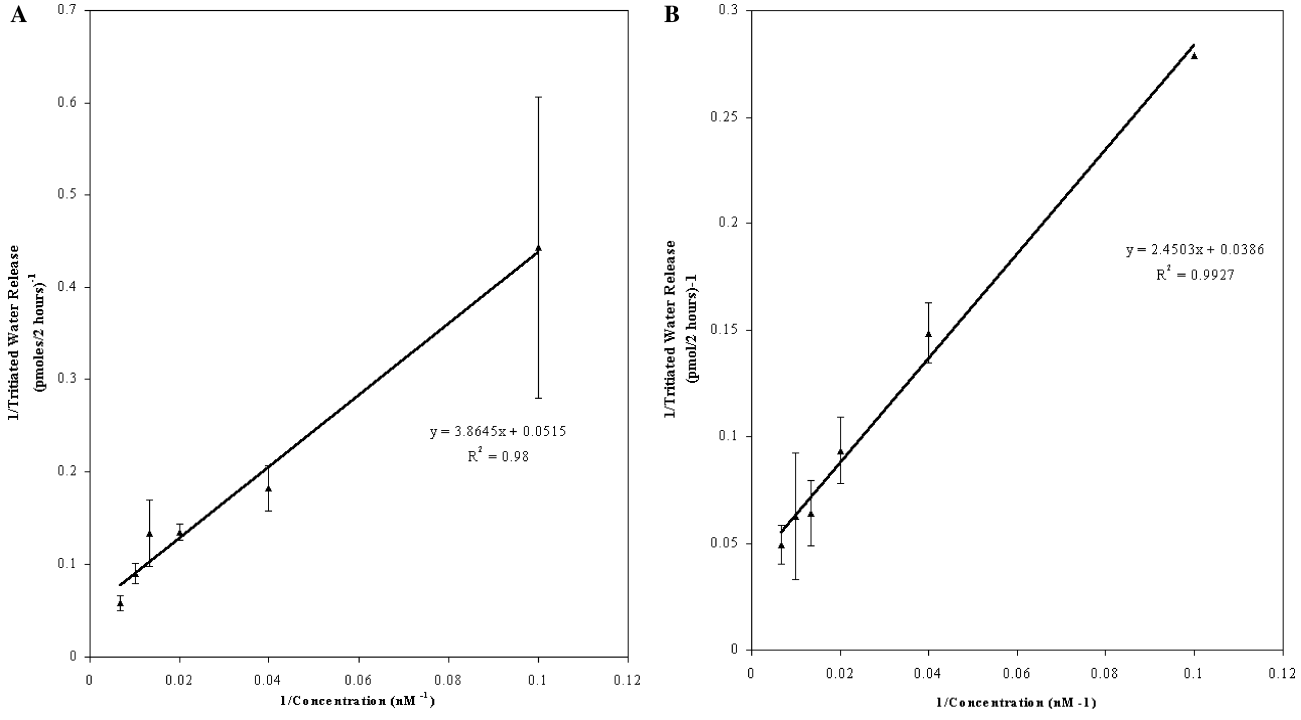

Fig. 3. Aromatase activity at varying concentrations of testosterone (A) and androstenedione (B) in cells transiently transfected with dolphin aromatase. Assays were performed in triplicate as described in Section 2. Aromatase activity was measured with the tritiated water release assay.

Table 4
K_m values for aromatase in a variety of mammalian species

Species, tissue	K_m (nM)	Substrate	System
White-sided dolphin, ovary	75	Testosterone	Transient transfection
	63.5	Androstenedione	
Human, adipose[a]	25	Androstenedione	Stromal cells in culture
Human, placental	55[b]	Testosterone	Stable transfection
	43[c]	Testosterone	Baculovirus transfection
	50[b]	Androstenedione	Stable transfection
	39.9–57.8[d]	Androstenedione	Stable transfection, multiple cell lines
	34[e]	Androstenedione	Yeast expression
Pig, placental[f]	33	Testosterone	Stable transfection
	77	Androstenedione	
Pig, ovary[f]	116	Testosterone	Stable transfection
	104	Androstenedione	

[a] Ackerman et al. (1981).
[b] Corbin et al. (1988).
[c] Stresser et al. (2000).
[d] Zhou et al. (1990).
[e] Pompon et al. (1989).
[f] Corbin et al. (1999).

of initial transfection experiments and the ability of letrozole to nearly completely block aromatase activity in this system confirms that tritiated water production resulted specifically from aromatase activity.

The rates of substrate conversion were linear for up to 4 h with testosterone and 3 h with androstenedione, respectively (Fig. 2). A higher rate of substrate conversion was seen with androstenedione at all time points. The dolphin CYP19 K_m and V_{max} were 75 nM and 19.4 pmol $2\,h^{-1}$ for testosterone and 63.5 nM and 25.9 pmol $2\,h^{-1}$ for androstenedione (Fig. 3 and Table 4).

4. Discussion

4.1. Phylogenetic analysis

The *CYP19* gene is well conserved among mammalian species, with nucleotide identity greater than 80% between dolphin and other mammals, similar to what has been found in studies of mammals and fish (Conley and Hinshelwood, 2001; Tchoudakova and Callard, 1998). Based on both molecular phylogeny and fossil evidence, cetaceans are thought to be most closely related to the ruminant artiodactyls (Gatesy et al., 2001). In agreement with this species phylogeny, the dolphin aromatase is most closely related to the bovine, ovine, and caprine aromatases (Fig. 1). The porcine aromatases are in a sister grouping to the cetartiodactyl clade, which also reflects current molecular evidence of the evolution of the ungulates (Gatesy et al., 2001). Interestingly, the three porcine aromatase sequences are more closely related to each other than to any other aromatase, with a very high posterior probability, indicating that these porcine sequences arose from duplication events only in this lineage, as suggested previously (Conley and Hinshelwood, 2001; Graddy et al., 2000). Further taxon sampling within the ungulates could specifically test this hypothesis. These results were in agreement with phylogenetic reconstructions based on a maximum likelihood analysis of *CYP19* nucleotide sequences from a more limited mammalian alignment (data not shown).

While the aromatase phylogenetic tree followed currently understood species phylogeny, clades with lower posterior probabilities contained species that are traditionally difficult to place in mammalian species phylogenies. This included the rabbit occurring as a potential sister taxon to the rodents (glires), which has been supported by morphological evidence but not always supported by molecular phylogenetics (de Jong, 1999). Perrisodactyls (which include horse), also have been difficult to place in the mammalian tree and are not always sister taxa to the cetartiodactyls (Liu and Miyamoto, 1999; Shoshani and McKenna, 1998).

Areas of well-understood vertebrate phylogeny that had very high support in the aromatase tree (as indicated by posterior probabilities of 1.00) include the grouping of all fish, amphibians, turtles, reptiles, birds, and mammals into appropriate clades. Posterior probabilities are generally higher than non-parametric bootstrap estimates (Huelsenbeck et al., 2002). However, posterior probabilities of 1.00, as seen on several nodes of the aromatase tree, should represent nodes that would also have strong support under bootstrap analysis. Among the fish, the ovarian CYP19a and brain CYP19b formed separate clades, indicating a duplication event early in the fish lines, at least prior to the origin of the bony fishes.

4.2. Functional characterization of dolphin aromatase

Substrate conversion by cells transfected with dolphin *CYP19* was linear for at least 2 h with both androstenedione and testosterone. This is similar to the time course seen with porcine aromatase expressed in the same cell system (Corbin et al., 1999). Additional NADPH-cytochrome P450 reductase was not used in these experiments. Work with pig aromatases transfected into HEK 293 cells showed that additional reductase increased the rate of aromatase activity (Corbin et al., 2001). While reductase may have been limiting in these experiments, the K_m values of cetacean aromatase with both testosterone and androstenedione were very similar to those reported for other mammalian aromatases (see Table 4, Ackerman et al., 1981; Corbin et al., 1988, 1999; Pompon et al., 1989; Stresser et al., 2000; Zhou et al., 1990).

The functional characteristics of pig placental aromatase differ from those of other mammalian aromatases, and from the pig gonadal form, in having slower metabolism of testosterone (i.e., lower V_{max}), lower K_m values for androgens in general, and a higher affinity for testosterone than for androstenedione (Corbin et al., 1999, 2001). Pig gonadal aromatase is sensitive to the typical aromatase inhibitor etomidate while the placental form is not (Corbin et al., 1999). The placental form also has a larger increase in activity with optimal reductase activity (Corbin et al., 2001). Thus, it appears that the pig ovarian form, but not the placental form, is like other mammalian aromatases, including that of the white-sided dolphin.

4.3. DIVERGE analysis

Aromatase shows little functional divergence among all the vertebrates for which sequence is known. The turtle/reptile/bird aromatases did not show significant divergence from either mammals or fish. Likewise, the fish a and b forms were not significantly different from each other. The value for θ was significantly greater than zero only in the comparison between the mammalian and fish aromatases, no matter whether the comparison included all fish aromatases, or separated fish aromatases into brain and ovarian forms. However, the functional differences between mammalian and fish aromatases were marginal ($\theta < 0.5$, Table 3) and the distribution of functionally divergent residues was random across the protein (data not shown). This is in strong contrast to the CYP3 family for example, where the functional differences were pronounced and large clustering of divergent residues enabled the identification of functional regions of the protein that were distinct between taxa (McArthur et al., 2003). Six substrate recognition sites (SRSs) have been identified in the CYP2 family as regions that contain or overlap with residues or fragments thought to determine substrate specificities of CYP2 enzymes, and accumulate more non-synonymous substitution (Gotoh, 1992). Although chimeric constructs and point mutagenesis studies have identified a residue in SRS 1 important for catalytic differences between the gonadal and placental porcine aromatases (Conley et al., 2001, 2002), there is no identifiable variation in SRS regions among divergent aromatases in this study. Since DIVERGE requires at least four sequences per grouping, it was not possible to test for functional divergence specifically within the pig aromatases. It would be interesting to repeat this analysis when sufficient aromatase sequence information is available from other porcine species.

The extensive functional diversity in the CYP3 subfamily is evident in the very large number of genes and the diversity of structures acted on by CYP3s (Maurel, 1996; Waxman, 1999). In contrast, the CYP19 subfamily thus far has a single form in amphibians, reptiles, birds, and most mammals; ovarian, embryonic, and placental forms in pig; and brain and ovarian forms in fish, all active with the same few substrates. As such, it is not surprising that DIVERGE analyses did not see a degree of functional diversity like that in the CYP3 family. Yet, the lack of clustering of divergent residues demonstrates how strongly the entire protein sequence of *CYP19* is conserved. Aromatase proteins may differ at the transcriptional level but the observed sequence divergence within the protein is functionally neutral.

Given the high degree of conservation, one might expect that the catalytic properties and substrate preference would be similar not only among mammals, but also among birds, reptiles, turtles, and amphibians. The substrate preference and K_m's are similar between the white-sided dolphin and other mammalian aromatases (Table 4). Fish aromatases have a preference for testosterone over androstenedione (Zhao et al., 2001), in contrast to all mammalian aromatases studied to date except the placental form in pig (Corbin et al., 1999). This subtlety is apparently below the detection limit of DIVERGE. Studies involving protein modelling (Graham-Lorence et al., 1995), site-directed mutagenesis (Conley et al., 2001, 2002), and naturally occuring mutations (Adashi et al., 1999; Belgorosky et al., 2003; Simpson, 1998; Taylor et al., 1998) will be needed to determine the specific residues of the protein responsible for differences in substrate preference.

The dearth of human subjects with mutations in the *CYP19* gene is consistent with an essential role for aromatase in mammalian physiology. Because of the high degree of conservation in aromatase, once could expect that studies of aromatase in traditional model organisms may be validly extrapolated to other species. Ongoing efforts to study aromatase in non-model species should focus on aromatase copy number, the promoter regions, and expression profiles, to better understand the sites of aromatase expression and similarity in transcriptional control across species. Our findings suggest that experi-

mental studies of aromatase in other mammalian species may be applicable to white-sided dolphins. For wildlife species, such as marine mammals, where experimental work is intractable or unethical, aromatase studies in model organisms should help elucidate the control, function, and importance of this protein in both normal physiology and pathophysiological conditions.

Acknowledgments

The authors thank the Cape Cod Stranding Network, New England Aquarium, and stranding volunteers for help with the sample collection. The authors thank Drs. Heather Handley-Goldstone and Sibel Karchner, WHOI, for help with the RT-PCR and RACE procedures. We gratefully thank Dr. Hilary Morrison, Marine Biological Laboratory, for help with DNA sequencing and Dr. John Trant, Center of Marine Biotechnology, University of Maryland Biotechnology Institute, for advice and discussions on the transfection experiments and HEK 293 cell culture. We thank Dr. Robert Campbell, Ares Advanced Technologies, for initial RIA detection of estradiol in the transfection system. Novartis Pharmaceutical Corporation generously donated the Letrozole used in this study. We thank Dr. Ann Tarrant, Woods Hole Oceanographic Institution, and two anonymous reviewers for useful comments on the manuscript. This work was funded by the Woods Hole Oceanographic Institution (WHOI) Sea Grant Program, under a grant from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce, Grant No. NA86RG0075, Project No. R/B-162 to J.J.S., the Woods Hole Oceanographic Institution Education Department to J.Y.W., and the National Institutes of Health (Grant No. AI51089) and Ellison Medical Foundation to A.G.M. Computational resources were provided by the Josephine Bay Paul Center through funds provided by the W.M. Keck Foundation and G. Unger Vetlesen Foundation. Contribution number 11151 for the Woods Hole Oceanographic Institution.

References

Ackerman, G.E., Smith, M.E., Mendelson, C.R., MacDonald, P.C., Simpson, E.R., 1981. Aromatization of androstenedione by human adipose tissue stromal cells in monolayer culture. J. Clin. Endocrinol. Metab. 53, 412–417.

Adashi, E.Y., Hennebold, J.D., 1999. Single-gene mutations resulting in reproductive dysfunction in women. N. Engl. J. Med. 340 (9), 709–718.

Agarwal, V.R., Sinton, C.M., Liang, C., Fisher, C., German, D.C., Simpson, E.R., 2000. Upregulation of estrogen receptors in the forebrain of aromatase knockout (ArKO) mice. Mol. Cell. Endocrinol. 162 (1–2), 9–16.

Belgorosky, A., Pepe, C., Marino, R., Guercio, G., Saraco, N., Vaiani, E., Rivarola, M., 2003. Hypothalamic–pituitary–ovarian axis during infancy, early and late prepuberty in an aromatase-deficient girl who is a compound heterozygote for two new point mutations of the CYP19 gene. J. Clin. Endocrinol. Metab. 88 (11), 5127–5131.

Britt, K.L., Drummond, A.E., Dyson, M., Wreford, N.G., Jones, M.E., Simpson, E.R., Findlay, J.K., 2001. The ovarian phenotype of the aromatase knockout (ArKO) mouse. J. Steroid Biochem. Mol. Biol. 79 (1–5), 181–185.

Chen, S.Z.D., Okubo, T., Kao, Y.C., Yang, C., 1999. Breast tumor aromatase: functional role and transcriptional regulation. Endocr. Relat. Cancer 6 (2), 149–156.

Choi, I., Simmen, R.C., Simmen, F.A., 1996. Molecular cloning of cytochrome P450 aromatase complementary deoxyribonucleic acid from periimplantation porcine and equine blastocysts identifies multiple novel 5′-untranslated exons expressed in embryos, endometrium, and placenta. Endocrinology 137 (4), 1457–1467.

Conley, A., Hinshelwood, M., 2001. Mammalian aromatases. Reproduction 121 (5), 685–695.

Conley, A., Mapes, S., Corbin, C., Greger, D., Graham, S., 2002. Structural determinants of aromatase cytochrome p450 inhibition in substrate recognition site-1. Mol. Endocrinol. 16 (7), 1456–1468.

Conley, A., Mapes, S., Corbin, C.J., Greger, D., Walters, K., Trant, J., Graham, S., 2001. A comparative approach to structure–function studies of mammalian aromatases. J. Steroid Biochem. Mol. Biol. 79 (1–5), 289–297.

Corbin, C.J., Graham-Lorence, S., McPhaul, M., Mason, J.I., Mendelson, C.R., Simpson, E.R., 1988. Isolation of a full-length cDNA insert encoding human aromatase system cytochrome P-450 and its expression in nonsteroidogenic cells. Proc. Natl. Acad. Sci. USA 85 (23), 8948–8952.

Corbin, C.J., Trant, J.M., Conley, A.J., 2001. Porcine gonadal and placental isozymes of aromatase cytochrome P450: sub-cellular distribution and support by NADPH-cytochrome P450 reductase. Mol. Cell. Endocrinol. 172 (1–2), 115–124.

Corbin, C.J., Trant, J.M., Walters, K.W., Conley, A.J., 1999. Changes in testosterone metabolism associated with the evolution of placental and gonadal isozymes of porcine aromatase cytochrome P450. Endocrinology 140 (11), 5202–5210.

de Jong, W.W., 1999. Molecules remodel the mammalian tree. Trends Ecol. Evol. 13 (7), 270–275.

Fisher, C.R., Graves, K.H., Parlow, A.F., Simpson, E.R., 1998. Characterization of mice deficient in aromatase (ArKO) because of targeted disruption of the cyp19 gene. Proc. Natl. Acad. Sci. USA 95 (12), 6965–6970.

Gatesy, J., O'Leary, M.A., 2001. Deciphering whale origins with molecules and fossils. Trends Ecol. Evol. 16 (10), 562–570.

Gotoh, O., 1992. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. J. Biol. Chem. 267 (1), 83–90.

Graddy, L.G., Kowalski, A.A., Simmen, F.A., Davis, S.L., Baumgartner, W.W., Simmen, R.C., 2000. Multiple isoforms of porcine aromatase are encoded by three distinct genes. J. Steroid Biochem. Mol. Biol. 73 (1–2), 49–57.

Graham-Lorence, S., Amarneh, B., White, R.E., Peterson, J.A., Simpson, E.R., 1995. A three-dimensional model of aromatase cytochrome P450. Protein Sci. 4 (6), 1065–1080.

Gu, X., Vander Velden, K., 2002. DIVERGE: phylogeny-based analysis for functional–structural divergence of a protein family. Bioinformatics 18 (3), 500–501.

Higgins, D.G., 1994. CLUSTAL V: multiple alignment of DNA and protein sequences. Methods Mol. Biol. 25, 307–318.

Honda, S., Harada, N., Ito, S., Takagi, Y., Maeda, S., 1998. Disruption of sexual behavior in male aromatase-deficient mice lacking exons 1 and 2 of the cyp19 gene. Biochem. Biophys. Res. Commun. 252 (2), 445–449.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17 (8), 754–755.

Huelsenbeck, J.P., Larget, B., Miller, R.E., Ronquist, F., 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. Syst. Biol. 51 (5), 673–688.

Jones, D.T., Taylor, W.R., Thornton, J.M., 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8 (3), 275–282.

Jones, M.E., Thorburn, A.W., Britt, K.L., Hewitt, K.N., Misso, M.L., Wreford, N.G., Proietto, J., Oz, O.K., Leury, B.J., Robertson, K.M., Yao, S., Simpson, E.R., 2001. Aromatase-deficient (ArKO) mice accumulate excess adipose tissue. J. Steroid Biochem. Mol. Biol. 79 (1–5), 3–9.

Jones, M.E., Thorburn, A.W., Britt, K.L., Hewitt, K.N., Wreford, N.G., Proietto, J., Oz, O.K., Leury, B.J., Robertson, K.M., Yao, S., Simpson, E.R., 2000. Aromatase-deficient (ArKO) mice have a phenotype of increased adiposity. Proc. Natl. Acad. Sci. USA 97 (23), 12735–12740.

Kraus, S., Hamilton, P., Kenney, R., Knowlton, A., Slay, C., 2001. Reproductive parameters of the North Atlantic right whale. J. Cetacean Res. Manag. Spec. Issue 2, 231–236.

Lephart, E.D., Simpson, E.R., 1991. Assay of aromatase activity. Methods Enzymol. 206, 477–483.

Liu, F.G., Miyamoto, M.M., 1999. Phylogenetic assessment of molecular and morphological data for eutherian mammals. Syst. Biol. 48 (1), 54–64.

Maurel, P., 1996. The CYP3 family. In: Ioannides, C. (Ed.), Cytochromes P450 Metabolic and Toxicological Aspects. CRC Press, Boca Raton, FL, pp. 241–270.

McArthur, A.G., Hegelund, T., Cox, R.L., Stegeman, J.J., Liljenberg, M., Olsson, U., Sundberg, P., Celander, M.C., 2003. Phylogenetic analysis of the cytochrome P450 (CYP3) gene family. J. Mol. Evol. 57 (2), 200–211.

Murata, Y., Robertson, K.M., Jones, M.E., Simpson, E.R., 2002. Effect of estrogen deficiency in the male: the ArKO mouse model. Mol. Cell. Endocrinol. 193 (1–2), 7–12.

Pompon, D., Liu, R.Y., Besman, M.J., Wang, P.L., Shively, J.E., Chen, S., 1989. Expression of human placental aromatase in *Saccharomyces cerevisiae*. Mol. Endocrinol. 3 (9), 1477–1487.

Robertson, K.M., O'Donnell, L., Jones, M.E., Meachem, S.J., Boon, W.C., Fisher, C.R., Graves, K.H., McLachlan, R.I., Simpson, E.R., 1999. Impairment of spermatogenesis in mice lacking a functional aromatase (cyp 19) gene. Proc. Natl. Acad. Sci. USA 96 (14), 7986–7991.

Robertson, K.M., Simpson, E.R., Lacham-Kaplan, O., Jones, M.E., 2001. Characterization of the fertility of male aromatase knockout mice. J. Androl. 22 (5), 825–830.

Shoshani, J., McKenna, M.C., 1998. Higher taxonomic relationships among extant mammals based on morphology, with selected comparisons of results from molecular data. Mol. Physiol. Evol. 9 (3), 572–584.

Simpson, E.R., 1998. Genetic mutations resulting in estrogen insufficiency in the male. Mol. Cell. Endocrinol. 145 (1–2), 55–59.

Simpson, E.R., Clyne, C., Rubin, G., Boon, W.C., Robertson, K., Britt, K., Speed, C., Jones, M., 2002. Aromatase—a brief overview. Annu. Rev. Physiol. 64, 93–127.

Simpson, E.R., Mahendroo, M.S., Means, G.D., Kilgore, M.W., Hinshelwood, M.M., Graham-Lorence, S., Amarneh, B., Ito, Y., Fisher, C.R., Michael, M.D., et al., 1994. Aromatase cytochrome P450, the enzyme responsible for estrogen biosynthesis. Endocr. Rev. 15 (3), 342–355.

Stresser, D.M., Turner, S.D., McNamara, J., Stocker, P., Miller, V.P., Crespi, C.L., Patten, C.J., 2000. A high-throughput screen to identify inhibitors of aromatase (CYP19). Anal. Biochem. 284 (2), 427–430.

Taylor, J.A., Lewis, K.J., Lubahn, D.B., 1998. Estrogen receptor mutations. Mol. Cell. Endocrinol. 145 (1–2), 61–66.

Tchoudakova, A., Callard, G.V., 1998. Identification of multiple CYP19 genes encoding different cytochrome P450 aromatase isozymes in brain and ovary. Endocrinology 139 (4), 2179–2189.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25 (24), 4876–4882.

Toda, K., Takeda, K., Akira, S., Saibara, T., Okada, T., Onishi, S., Shizuta, Y., 2001a. Alternations in hepatic expression of fatty-acid metabolizing enzymes in ArKO mice and their reversal by the treatment with 17β-estradiol or a peroxisome proliferator. J. Steroid Biochem. Mol. Biol. 79 (1–5), 11–17.

Toda, K., Takeda, K., Okada, T., Akira, S., Saibara, T., Kaname, T., Yamamura, K., Onishi, S., Shizuta, Y., 2001b. Targeted disruption of the aromatase P450 gene (Cyp19) in mice and their ovarian and uterine responses to 17β-oestradiol. J. Endocrinol. 170 (1), 99–111.

Waxman, D.J., 1999. P450 gene induction by structurally diverse xenochemicals: central role of nuclear receptors CAR, PXR, and PPAR. Arch. Biochem. Biophys. 369 (1), 11–23.

Williams, P.A., Cosme, J., Sridhar, V., Johnson, E.F., McRee, D.E., 2000. Microsomal cytochrome P450 2C5: comparison to microbial P450s and unique features. J. Inorg. Biochem. 81 (3), 183–190.

Zhao, J., Mak, P., Tchoudakova, A., Callard, G., Chen, S., 2001. Different catalytic properties and inhibitor responses of the goldfish brain and ovary aromatase isozymes. Gen. Comp. Endocrinol. 123 (2), 180–191.

Zhou, D.J., Pompon, D., Chen, S.A., 1990. Stable expression of human aromatase complementary DNA in mammalian cells: a useful system for aromatase inhibitor screening. Cancer Res. 50 (21), 6949–6954.