A Comparative View on Sex Differentiation and Gametogenesis Genes in Lungfish and Coelacanths

Maria Assunta Biscotti[1,†], Mateus Contar Adolfi[2,†], Marco Barucca[1], Mariko Forconi[1], Alberto Pallavicini[3], Marco Gerdol[3], Adriana Canapa[1,†], and Manfred Schartl[2,4,5,*,†]

[1]Dipartimento di Scienze della Vita e dell'Ambiente, Università Politecnica delle Marche, Ancona, Italy

[2]Physiological Chemistry, Biocenter, University of Wuerzburg, Germany

[3]Dipartimento di Scienze della Vita, Università di Trieste, Italy

[4]Comprehensive Cancer Center Mainfranken, University Clinic Wuerzburg, Germany

[5]Hagler Institute of Advanced Study and Department of Biology,Texa A&M University, USA

[†]These authors contributed equally to this work.

*Corresponding author: E-mail: phch1@biozentrum.uni-wuerzburg.de.

Accepted: May 23, 2018

Data deposition: This project has been deposited at GenBank under the accessions MH329951-MH330004.

Abstract

Gonadal sex differentiation and reproduction are the keys to the perpetuation of favorable gene combinations and positively selected traits. In vertebrates, several gonad development features that differentiate tetrapods and fishes are likely to be, at least in part, related to the water-to-land transition. The collection of information from basal sarcopterygians, coelacanths, and lungfishes, is crucial to improve our understanding of the molecular evolution of pathways involved in reproductive functions, since these organisms are generally regarded as "living fossils" and as the direct ancestors of tetrapods. Here, we report for the first time the characterization of >50 genes related to sex differentiation and gametogenesis in *Latimeria menadoensis* and *Protopterus annectens*. Although the expression profiles of most genes is consistent with the intermediate position of basal sarcopterygians between actinopterygian fish and tetrapods, their phylogenetic placement and presence/absence patterns often reveal a closer affinity to the tetrapod orthologs. On the other hand, particular genes, for example, the male gonad factor *gsdf* (*Gonadal Soma-Derived Factor*), provide examples of ancestral traits shared with actinopterygians, which disappeared in the tetrapod lineage.

Key words: sex differentiation, gametogenesis, ovary, testis, evolution, *Protopterus annectens*, *Latimeria menadoensis*.

Introduction

Sexual dimorphism is one of the most pervasive and diverse features of animal morphology, physiology, and behavior. The process of gonad development involves complex and finely tuned mechanisms that regulate gametogenesis, molecular sex determination, and gonad differentiation. Despite the generality of the phenomenon itself, sex determination, and differentiation mechanisms largely vary across metazoan taxa, as the result of repeated, independent lineage-specific evolution, and quick modification of the underlying molecular pathways. Within vertebrates, mammals and birds share a common genetic sex determining mechanism, which also displays a remarkable degree of conservation at the molecular level. Others animal groups are characterized by highly diverse sex determination modes, which either have a genetic basis or depend on other factors, for example, temperature population density, and visual cues. This divergence matches the great variety of the initial triggers at the top of the gene network that direct the development toward a male or female phenotype, as well as the diversity of downstream acting components that implement and translate the initial signal into the development and differentiation of either testis or ovary from the bipotential embryonal gonad anlage (Herpin et al. 2013). This situation is particularly evident in teleost fishes, the largest group of extant chordates, where closely related species often display notably different sex determination mechanisms (Kikuchi and Hamaguchi 2013). Despite this puzzling variability among species, some common principles

are emerging from the analysis of the growing information available about metazoan gonad sex development molecular networks, for example, the recurrent involvement of certain molecules and pathways as primary male sex determinants. Interestingly, some of these commonalities mark teleosts as clearly different from those of other vertebrate classes and, in some instances, of all other tetrapods. These peculiarities might even include the presence or absence of genes that are critical for gonad sex formation and gametogenesis in other branches of the vertebrate phylogenetic tree. The study of sex determination and differentiation mechanisms in basal sarcopterygians represents a key to unveil whether the unique features displayed by teleosts exemplify an ancestral situation or they correspond to teleost-specific specializations. Coelacanths and lungfishes represent the closest extant relatives to the tetrapod lineage, even though the latter have been shown to be phylogenetically closer to the tetrapod stem (Biscotti et al. 2016; Irisarri and Meyer 2016). Moreover, the morphological similarity between the extant sarcopterygian fish and their fossil ancestors make them of outstanding scientific interest, considering that they may also retain, to some extent, components of the molecular machinery of the last common ancestor of tetrapods and fish. Due to the unique phylogenetic position of lungfish, their molecular studies are pivotal in understanding the evolution of vertebrates and the changes that accompanied the water-to-land transition. However, molecular studies in lungfish and coelacanths in particular are severely hampered by the difficult availability and maintenance of specimens in laboratory conditions. Taking into account the impossibility of performing functional studies in these species, the analysis of the gene expression profiles of sex differentiation and gametogenesis genes are the first choice for obtaining insights about gonad determination and differentiation.

Due to the paucity of biological material available from coelacanths, the only study published so far on this topic reported the expression of a few genes in the testis of *Latimeria menadoensis* (Forconi et al. 2013). Whole genome studies in Dipnoi have been long hindered by the enormous size and high repeat content of their genomes. However, the publication of a reference transcriptome for the West African lungfish, *Protopterus annectens* (Biscotti et al. 2016) helped to overcome this major obstacle, finally allowing RNA-Seq-based gene expression studies in this species.

The aim of this study was to expand the knowledge on basal sarcopterygian sex differentiation and gametogenesis by characterizing 51 genes related to this process and assessing their expression levels in gonads. We find that the intermediate phylogenetic placement of lungfish, between acanthopterygian fish and tetrapods, is matched by the patterns of molecular evolution of the sex differentiation and gametogenesis gene network.

Materials and Methods

Male and female *P. annectens* specimens were obtained from a local fish importer. Two testis, three ovaries, and one brain and liver sample each from a male and a female specimen were used for total RNA isolation and sequencing as described previously (Biscotti et al. 2016).

Pieces of the gonads used in the transcriptome analyzes were fixed in 4% formaldehyde solution for 24 h at 4°C. Subsequently, the tissues were dehydrated, embedded in paraffin, and then serially sectioned at 5 μm thickness. The sections were counterstained with hematoxylin and eosin (HE).

The sex and the gonadal maturation stage of the animals used for the generation of transcriptomes were confirmed by histological analysis (fig. 1). The male gonad of *P. annectens* was classified as anastomosing tubular testis type, where tubules do not terminate at the testis periphery (Parenti and Grier 2004). This type is common in the more basal bony fishes, for example, salmonids, cyprinids, and lepisosteids (Uribe et al. 2014). Both male gonadal tissues were at the prereproductive stage (Schulz et al. 2010), with the MG1 containing only few spermatogonias surrounded by Sertoli cells. In MG6, the formation of individual tubules and their lumina are observed, and the germinal epithelium contains higher amounts of spermatogonia. The oogenesis in lungfishes is similar to the teleost (Chaves 1992; Lubzens et al. 2010), and all three female gonads presented previtellogenic and vitellogenic oocytes. No postovulatory follicles were observed.

Sequences corresponding to key sex development-related genes were obtained by tBLASTn sequence homology searches performed against the *P. annectens* and *Latimeria menadoensis* transcriptomes (Canapa et al. 2012; Pallavicini et al. 2013; Biscotti et al. 2016) and the genomes of *L. chalumnae* (Amemiya et al. 2013), *D. rerio* (Howe et al. 2013), and *O. latipes* (Kasahara et al. 2007). The sequence of sex differentiation and gametogenesis-related genes previously characterized in other species were used as queries and similarity thresholds were set on a case-by-case basis. Based on the expression pattern and the so far functional analyzes of those genes, we classified them in male- or female-related genes. The retrieved transcript sequences were translated using the Sequence Translation (https://www.ebi.ac.uk/Tools/st/), permitting to identify UTRs and CDSs. The orthology of each transcript was confirmed using NCBI BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Each transcript was used as query against the database <all nonredundant GenBank CDS translations + PDB + SwissProt + PIR + PRF> excluding environmental samples from WGS projects.

Orthologous sequences used in the phylogenetic analyses were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/) or ENSEMBL (http://www.ensembl.org/index.html) (for accession numbers see supplementary table 1, Supplementary Material online). Multiple alignments of the amino acid

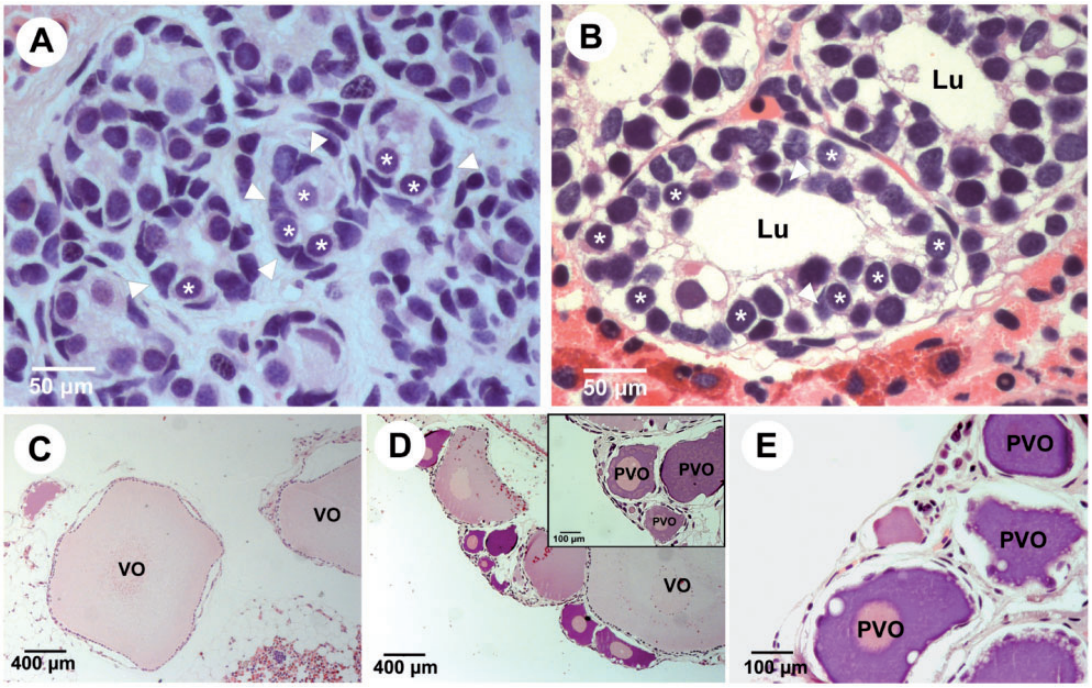

Fig. 1.—Morphology of the gonads of *Protopterus annectens*. (*A*) Testes (MG1) revealing spermatogonias (asterisks) surrounded by Sertoli cells (arrowheads). (*B*) Testes (MG6) with completely formed tubules and lumina (Lu), and germinal epithelium containing spermatogonia (asterisks) and Sertoli cells (arrowheads). (*C*–*E*) Ovarian tissue (FG 2, 3, 4) containing pre–vitellogenic (PVO) and vitellogenic oocytes (VO).

sequences were obtained with ClustalW (http://www.genome.jp/tools/clustalw/) using default parameters.

Phylogenetic analyses were performed by Bayesian inference using MrBayes (version 3.2) (Huelsenbeck et al. 2001). Substitution models as defined by posterior probabilities, stationarity, generations, sampling, burnin, and specific tree building parameters and rooting details are specified in the figure legend description of each tree.

The ω rate (nonsynonymous/synonymous mutations) was obtained with KaKs_calculator (Zhang et al. 2006) according to Goldman and Yang (1994) (supplementary table 2, Supplementary Material online). Poorly aligned positions and divergent regions of alignments were eliminated by Gblocks (Castresana 2000; Talavera and Castresana 2007).

The expression levels of the selected genes in the different tissues and species were determined as described by Biscotti et al. (2016). Briefly, to ensure comparability among species, a set of conserved orthologous genes present in *L. menadoensis*, *P. annectens*, and *D. rerio* was selected as a reference for the calculation of Transcripts Per Million (TPM). RNA-seq data from lungfish (Biscotti et al. 2016) (male and female gonads, brain, and liver), coelacanth (Pallavicini et al. 2013) (male gonad, liver, and skeletal muscle), zebrafish (brain, liver, muscle, male and female gonads), and medaka (male and female gonads) were mapped to the reference nucleotide sequences with the CLC Genomics Workbench v.9 (Qiagen, Hilden, Germany), setting length and similarity fraction parameters to 0.75 and 0.98, respectively. Data from zebrafish and medaka were retrieved from the PhyloFish project database (Pasquier et al. 2016), BioProject PRJNA255848 and PRJNA255889. In medaka, the sequence of the *fgf24* transcript, encoded by a gene present in a poorly assembled genomic region, was retrieved from the de novo assembled transcriptome. Raw reads counts were then converted to TPM values according to the formula suggested by Wagner et al. (2012) (supplementary fig. 1, Supplementary Material online). Considering that absolute TPM expression values, in absence of biological replicates, might be quite variable between individuals and are therefore inappropriate for interspecies comparisons, we assigned expression ranges with the following categories for each gene: absent < 10 TPM, $+$ between 10 and 200, $++$ between 200 and 700, $+++$ between 700 and 2,000, $++++$ over 2,000.

Real-time quantitative PCRs of five genes (*amh*, *ctnnb1*, *dmrt1*, *fgf9*, and *rspo1*, for primers sequences see supplementary table 3, Supplementary Material online) were carried out in order to validate the transcriptome data. Reverse transcription was done using RevertAid First Strand Synthesis kit (Fermentas) and random primers. The reactions were performed with SYBR Green reagent and amplifications were detected with a *Mastercycler ep realplex* (Eppendorf). All results are averages of three independent PCR reactions from cDNA preparations of three females and two males. Transcript levels of target genes were normalized against

Table 1
Identified Genes in Lungfish and Coelacanth

Growth Factors								
amh	*amhr2*	*fgf9*	*fgf16*	*fgf20*	*fgf24*	*gsdf*	*rspo1*	*wnt4*
ptch2	*hhip*	*dhh*	*pdgfa*	*pdgfb*	*pdgfra*	*pdgfrb*	*sdf1*	*cxcr4*
Steroidogenic Enzymes								
srd5a1	*srd5a2*	*srd5a3*	*cyp11b*	*cyp19a1*				
Transcription Factors								
dmrt1	*dmrt3*	*dmrt6*	*foxl2*	*foxl3*	*sf1*	dax1	sox3	sox5
sox8	*sox9*	*sox10*	*wt1*	*gata4*	*lrpprc*	*pax2*		
Meiosis Regulation Genes								
aldh1a1	*aldh1a2*	*aldh1a3*	*cyp26a1*	*cyp26b1*	*cyp26c1*	*stra8*		
Follistatin and β Catenin								
fst	*ctnnb1*							
Steroid Hormone Receptor Genes								
ar	*esra*	*esrb*						

Underlined genes were not found in lungfish.

the lungfish eukaryotic translation initiation factor 4 gamma 1 (*eif4g1*) gene.

Results and Discussion

In this paper, the presence of 51 genes known to be involved in gonad sex differentiation and gametogenesis development in other vertebrates was assessed in the transcriptomes of the lungfish *P. annectens*. This analysis allowed the identification of 44 transcripts (table 1), all showing a complete CDS (supplementary table 4, Supplementary Material online). The five genes used for validation of the transcriptome by qRT-PCR showed mRNA levels consistent with the transcriptome data of the respective genes (supplementary table 5, Supplementary Material online). For comparison's sake we also studied the sarcopterygian fish *Latimeria*, identifying 17 new genes (table 1 and supplementary table 6, Supplementary Material online) in addition to those investigated in our previous work (Forconi et al. 2013). The expression levels (absent, low, intermediate, high) of the identified genes were compared between the two basal sarcopterygians and with the teleosts *Danio rerio* and *Oryzias latipes* to provide the basis for a better understanding of the gene regulatory network governing gametogenesis and sex differentiation and their evolution.

Growth Factors and Their Receptors

Anti-Müllerian hormone (Amh), also known as Müllerian duct-inhibitory substance, is a member of the transforming growth factor-β (TGF-β) superfamily, which causes regression of the Müllerian ducts during testicular differentiation in mammals. The high expression values of *amh* observed in male gonads of several vertebrates have been linked to the proliferation and differentiation of spermatogonia (Skaar et al. 2011; Hu et al. 2015). *amh* is also expressed in females, where it is restricted to the granulosa cells of ovarian follicles (Piprek, Pecio, Laskowska-Kaszub, Kubiak, et al. 2013). Amh mainly acts through the Amhr2 signaling system in the development of gonadal tissue. Orthologous genes have been identified in several teleosts (Klüver et al. 2007; Li et al. 2015; Pfennig et al. 2015), despite the absence of Müllerian ducts in these organisms. In addition, Amh or its type II receptor act as Y-linked male sex determining genes in some teleost species (Hattori et al. 2013; Kikuchi and Hamaguchi 2013; Herpin and Schartl 2015). On the other hand, the Müllerian duct develops in female coelacanths and lungfish (Hildebrand and Goslow 2001) and in a basal group of actinopterygian fishes, the sturgeons (Wrobel 2003). In lungfish and in the teleosts, *amh* is preferentially expressed in male gonads and a robust expression of this gene could be also observed in the testis of *Latimeria* (table 2).

Like coelacanth and most other fish, lungfish possess a bona fide ortholog of *amhr2* (supplementary fig. 2, Supplementary Material online). Unexpectedly, this gene is missing in the zebrafish genome. This appears to be species-specific feature, as *amhr2* orthologs are present in other teleost genomes. In medaka, *amhr2* is highly expressed in testis (supplementary fig. 1, Supplementary Material online), and mutation of this gene leads to overproliferation of

Table 2
Growth Factors and Receptors Generally Related to Male Sex Determination, Differentiation, and Reproductive Functions

	Lungfish		Coelacanth		Zebrafish		Medaka	
	Testis	Ovary	Testis	Ovary	Testis	Ovary	Testis	Ovary
amh	+	−	+	N.R.	++++	−	++	+
amhr2	−	+	+	N.R.	Absent		++	+
gsdf	+	−	+	N.R.	++++	+	++++	++
fgf9	−	+	−	N.R.	Absent		Absent	
fgf16	−	+	−	N.R.	−	−	−	−
fgf20a	Absent		−	N.R.	−	−	−	−
fgf20b					−	−	+	−
fgf24	Absent		−	N.R.	+	+	+	−
sdf1a	++++	+	+	N.R.	+++	−	+	+
sdf1b					+	−	++	+
cxcr4a	++++	+	++	N.R.	+	+	+	+
cxcr4b					++	+	+	−
pdgfaa	+[a]	+[a]	+	N.R.	+	−	−	−
pdgfab							+	−
pdgfra	+	−	+	N.R.	+	+	+	−
pdgfb	+	−	+	N.R.	+	−	+	−
pdgfrb	+	+	−	N.R.	+	−	+	+

[a]Approximately 3-fold higher in ovary than testis.

the germ cells in both testis and ovary, showing an important role of Amh signaling in germ cell proliferation in adult gonad (Morinaga et al. 2007). This gene is expressed in the testis of *Latimeria*, whereas in the lungfish its expression is only found in female gonads (table 2). This finding supports the presence of female expression bias of *amhr2* in lungfish, like in tetrapods (Fagerberg et al. 2014; Jansson et al. 2016), but not in *Latimeria*, where *amhr2* is exclusively expressed in testis, like in teleost fish, where a reverse organ expression pattern has been observed (Lubieniecki et al. 2015).

Gonadal Soma-Derived Factor (GSDF) is another member of the TGF-β superfamily. This gene is particularly interesting because of its crucial function in fish testis development (Zhang et al. 2016). Prior to our identification of a homologous sequence in *Latimeria* (Forconi et al. 2013), *gsdf* was hypothesized to be a teleost innovation. Here, we report a transcript encoding for *gsdf* also in *P. annectens*, whose orthology assessment is strongly supported by phylogenetic inference (fig. 2*A*). Moreover, as a distinctive feature, Gsdfs lack a glycine residue in the CxGxC motif, different from Inha and Amh proteins which have this amino acid (fig. 2*B*). The presence of glycine in this motif is essential for heterodimer formation of glycoprotein hormones (Kinoshita et al. 2006). The functional meaning of this amino acid change for Gsdf is unknown, due to the lack of information about the signaling pathway of this gene. The identification of a *gsdf* sequence in *P. annectens* implies its subsequent loss in the tetrapod lineage. Moreover, the additional identification of a *gsdf* sequence in the cartilaginous fish *C. milii* evidences that this gene was already present in the common ancestor of Chondrichthyes and Osteichthyes.

The high expression of *gsdf* in *P. annectens* testis (table 2) is in line with data previously observed in coelacanth (Forconi et al. 2013) confirming the important role of this gene for male gonad development also in basal sarcopterygians. Furthermore, this pattern is in accordance with the well-described dimorphic expression of this gene in other fish species (Kaneko 2015). *gsdf* can even act as master sex determining gene (Myosho et al. 2012; Zhu et al. 2016) or as a male sex initiator, being activated by Dmrt1/Dmy (Zhang et al. 2016).

In summary, our results point out that a *gsdf* gene with testis-specific function was retained in the sarcopterygian ancestors of tetrapods, only to be lost during the transition from water to land.

R-spondin1 (Rspo1) is a regulator of the WNT/β-catenin signaling pathway. This gene has been characterized as a female sexual regulator in mammals, as it suppresses the male axis and maintains oocyte survival by positively regulating WNT4 signaling. In tetrapods and several teleosts, *rspo1* shows a sexually dimorphic expression pattern with higher values in the female gonad (Zhou et al. 2016), suggesting that Rspo1/Wnt/β-catenin signaling is an ancient pathway of ovarian differentiation in vertebrates. We found a quite high expression for *rspo1* in lungfish testis (table 3), and much lower levels in ovaries. In zebrafish, *rspo1* is equally expressed in ovary and testis, while in medaka no expression was detected. *rspo1* is also expressed in the ovary and testis of

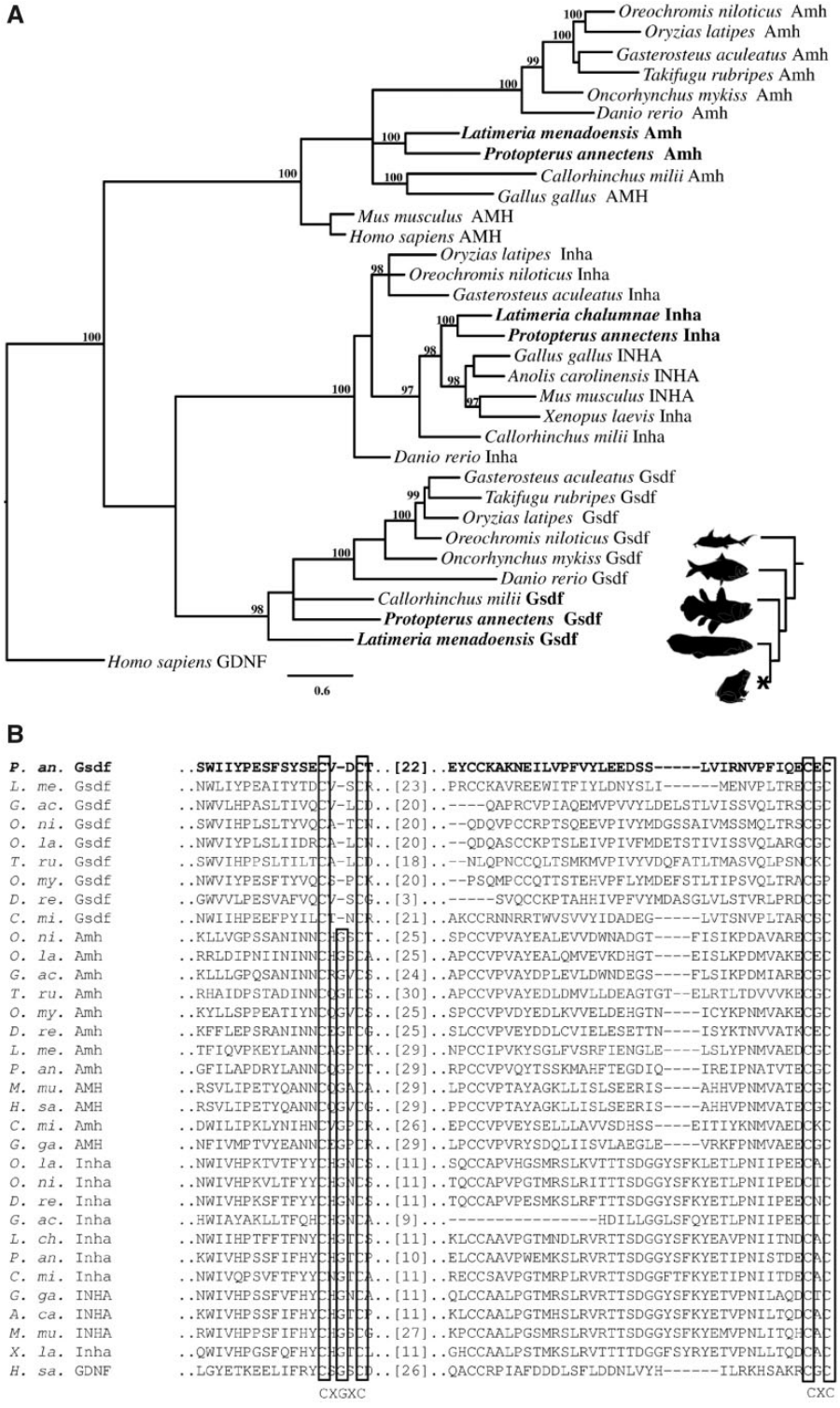

Fig. 2.—Gsdf phylogeny. (*A*) Phylogenetic tree of Gsdf, Amh, and Inha amino acid sequences. Bayesian inference: 2,000,000 generations, sampling every 100, Jones substitution model, stationarity defined as when the average SD of split frequencies approaching 0.007, burn-in set to 5,000, midpoint rooting. Numbers close to nodes indicate posterior probability values. Only values >95 were reported. The sequences in bold were obtained in this work. (*B*) Multiple alignment of Gsdf, Amh, and Inha amino acid sequences showing the CxGxC conserved motif and the lack of the glycine residue, which is diagnostic for Gsdf proteins.

Table 3
Growth Factors and Receptors Generally Related to Female Sex Determination, Differentiation, and Reproductive Functions

	Lungfish		Coelacanth		Zebrafish		Medaka	
	Testis	Ovary	Testis	Ovary	Testis	Ovary	Testis	Ovary
dhh	+	−	+	N.R.	+	−	++	−
ptch2	−	−	−	N.R.	+	−	+	−
hhip	−	+	−	N.R.	++	+	+	+
ctnnb1	+++	++++	+++	N.R.	++++	++++	+++	++
fst	+	−	−	N.R.	+	−	+	−
wnt4a	+	−	−	N.R.	+[a]	+[a]	−	−
wnt4b					+[a]	+[a]	+	−
rspo1	++	+	+	N.R.	+	+	−	−

[a]Approximately 2- to 3-fold higher in testis than ovary.

other teleosts, suggesting that it might be involved in both ovarian and testicular development and spermatogenesis (Wu et al. 2016). In other vertebrates, *RSPO1* is widely expressed (Han et al. 2015) and its transcript has been also detected in adult human testis (Parma et al. 2006). Our findings suggest that the role of *rspo1* in lungfish male gonads is not connected to the action of *wnt4*, as this gene is poorly expressed in lungfish and coelacanth, while both *wnt4* co-orthologs are well expressed in the gonads of zebrafish. Medaka, on the other hand, presents relatively low expression of *wnt4b* in testis.

ß-catenin (Ctnnb1) is the critical molecule that transmits the activity of the female determining pathway signals of Wnt4 and R-spondin to the nucleus to activate effector genes. High expression values were recorded for *ctnnb1* in all four species (table 3). While no strong sex bias in expression is present in zebrafish and medaka, the lungfish ovary shows considerably higher levels than testis, in contrast with the expression pattern of *wnt4* and *rspo1* (table 3).

Follistatin (Fst) encodes a TGFß superfamily binding protein that acts downstream of WNT4 to maintain germ cell survival in the cortical domain of the mammalian ovary (Yao et al. 2004). This gene is present in both basal sarcopterygians. The lack expression of *fst* in lungfish, in line with the absence of detectable expression for *wnt4* (table 3), may be connected to the prereproductive stage of the female gonads.

Fibroblast Growth Factors (Fgf) are a family of small proteins that have multiple functions in regulating cellular development and growth. Numerous studies in mice have shown that FGF9 plays an important role in testis development, creating a positive feedback loop with *Sox9* and antagonizing *Wnt4*, an ovarian factor (Jameson et al. 2012). No orthologs of *fgf9* have been identified in actinopterygian fish so far (Itoh and Konishi 2007). The identification of *fgf9* in *Latimeria* has previously suggested that this gene was present in the common ancestor of Sarcopterygii (Forconi et al. 2013), a hypothesis which is now further supported by the retrieval of this sequence also in *P. annectens* (fig. 3). Moreover, the presence of *fgf9* in *C. milii* dates back the origin of this gene to the common ancestor of Chondrichthyes and Osteichthyes, implying its loss in the lineage of Actinopterygii. Neither *fgf9* nor the other members of the FGF9/16/20 subfamily are expressed in testis of lungfish and coelacanth (table 2). In lungfish, we verified the absence of *fgf20* both at the transcriptome level. Moreover, *fgf9* is strongly expressed in lungfish ovary (table 2). This may indicate that the important function of *fgf9* for testis development is an innovation acquired after the transition from water to land.

To evaluate the possible involvement of another member of the *FGF* gene family as a functional homolog of the tetrapod *Fgf9*, as recently proposed by Leerberg et al. (2017), we took a closer look at *fgf24*, a member of the FGF8/17/18/24 family (fig. 4). The isolation of this gene in the elephant shark supported its presence in the common ancestor of gnathostomes (Jovelin et al. 2010). Within the sarcopterygian lineage, *fgf24* could be only identified in *Latimeria* (Amemiya et al. 2013) but not in tetrapods (Jovelin et al. 2010). In addition, no orthologous sequence to *fgf24* could be retrieved from the *P. annectens* transcriptome, which might either indicate the lack of expression of this gene, or its loss. If confirmed at the genome level, this finding might indicate that *fgf24* was lost in the common ancestor of lungfish and tetrapods. This gene is expressed in the early somatic and in adult gonads of zebrafish Leerberg et al. (2017) (table 2) and was suggested as a plausible candidate to substitute for *fgf9* function in teleosts. In addition, *fgf24* is expressed only in testis of adult medaka.

Stromal-cell-derived factor 1 (Sdf1, also known as Cxcl12) is a chemokine involved in a wide variety of body functions. It binds to CXC receptor 4 (Cxcr4), a G-coupled transmembrane glycoprotein located on the surface membranes of primordial germ cells. The interaction between Sdf1 and its receptor regulates migration and proliferation of various cell types, which importantly comprise also teleosts primordial germ cells. Previous studies assigned a central role to Sdf1 and

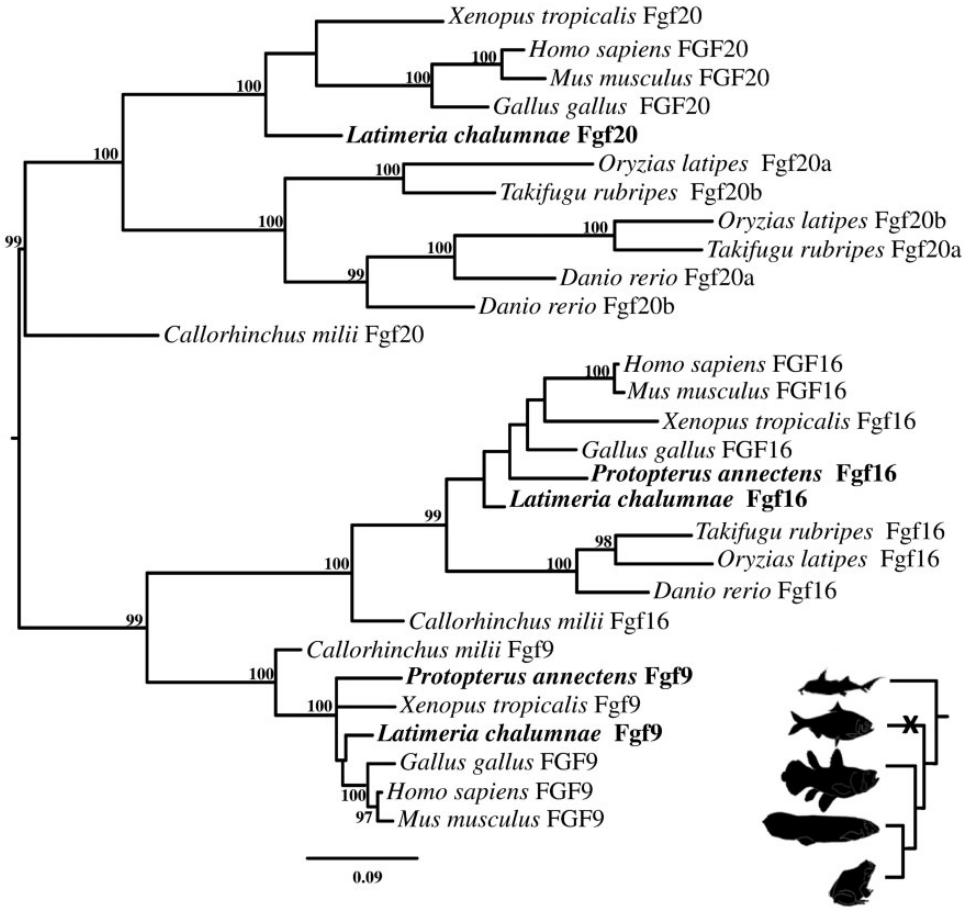

Fig. 3.—Fgf9/16/20 phylogeny. Phylogenetic analysis of Fgf9, Fgf16, Fgf20. Bayesian inference: 1,000,000 generations, sampling every 100, Jones substitution model, stationarity defined as when the average SD of split frequencies approaching 0.0055, burn-in set to 2,500, midpoint rooting. Numbers close to nodes indicate posterior probability values. Only values >95 were reported. The sequences in bold were obtained in this work.

CXCR4 in the development of embryonic and postnatal gonad (Yang et al. 2013; Westernströer et al. 2014; Fernandez et al. 2015). In human adults, the expression of *SDF1* was observed in Sertoli cells. The expression of *Cxcr4* in adult mouse testis (Yang et al. 2013) further suggests a role for the ligand/receptor in spermatogenesis. Both genes are preferentially expressed in lungfish male gonads (table 2), at particularly high levels in the more immature testis sample. In addition, the expression of these genes has been also recently reported in both male and female gonads of the teleost *Oreochromis mossambicus* (Amat-Fernandez et al. 2017).

Platelet-derived growth factors (Pdgf) and their receptors are involved in regulating the development of various tissues (Miura et al. 2002). In particular, they play a fundamental role in the development of gonadal tissue in mammalian fetal and adult testis (Li et al. 2016). In zebrafish and medaka, this gene is preferentially expressed in testis (table 2). However, in both sarcopterygians, the expression of *pdgf*s and their receptors in testis were quite low. The lungfish *pdgfa* and *pdgfrb* orthologs unexpectedly showed the highest expression levels in ovary (table 2).

Steroidogenic Enzymes

Steroid-5-Alpha-Reductase Alpha Polypeptides 1, 2, 3 (Srd5a1, Srd5a2, Srd5a3) are proteins involved in the conversion of testosterone to dihydrotestosterone (DHT), a very potent androgen in teleosts fishes and tetrapods. Unexpectedly, we detected higher expression of *srd5a1* in lungfish ovary, in contrast to its absence in zebrafish and the low expression in female medaka (supplementary table 7, Supplementary Material online). *Srd5a1* is predominantly expressed in the female gonads of frogs but also in rat (Lephart et al. 1992) and human (Haning et al. 1996; Mahendroo and Russell 1999), consistently with our observations in lungfish (supplementary table 7, Supplementary Material online). This finding suggests that Srd5a1 might confer the ability of metabolizing testosterone into DHT in the sarcopterygian lineage, while a different isozyme is likely used in teleost fish.

Cytochrome P450 19A1 (Cyp19a1 or aromatase) catalyzes the final step in the biosynthesis of estrogens, which have a key role in ovarian differentiation and maintenance in fish, amphibians, reptiles, and birds (Bruggeman et al.

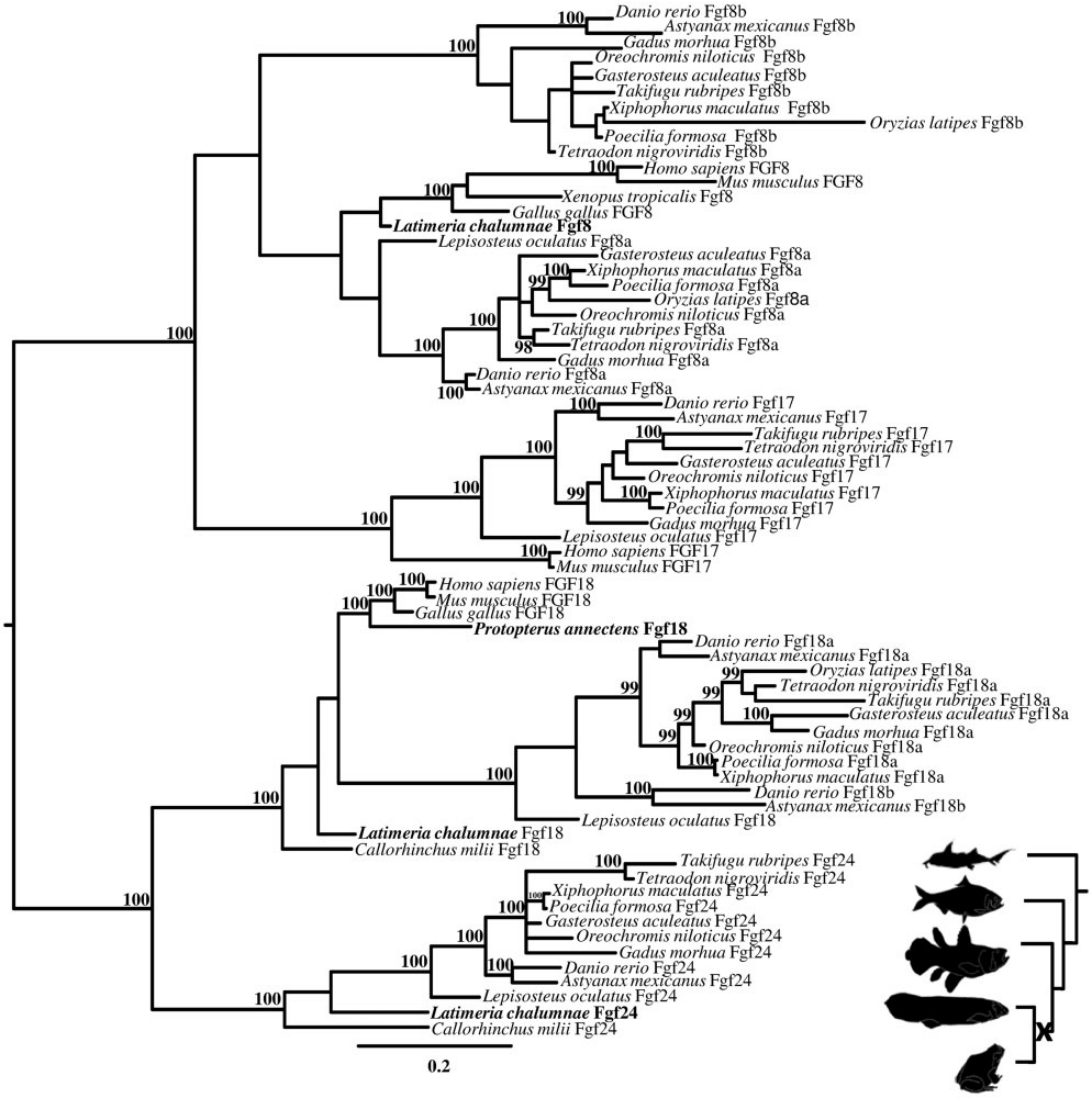

Fig. 4.—Fgf8/17/18/24 phylogeny. Phylogenetic analysis of Fgf8, Fgf17, Fgf18, Fgf24. Bayesian inference: 2,000,000 generations, sampling every 100, Jones substitution model, stationarity defined as when the average SD of split frequencies approaching 0.0062, burn-in set to 5,000, midpoint rooting. Numbers close to nodes indicate posterior probability values. Only values >95 were reported. The sequences in bold were obtained in this work.

2002; Dumond et al. 2008; Lubieniecki et al. 2015). The expression patterns observed in the present study is consistent with a conserved function for this gene also in basal sarcopterygians (supplementary table 8, Supplementary Material online).

Cytochrome P450 11c1 (Cyp11c1). In most teleost fish the predominant androgen is 11-ketotestosterone, which is produced from testosterone by the action of the *cyp11c1* gene. However, tetrapod *cyp11B* homologs function predominantly in adrenal steroid biosynthesis and are not expressed in testis (Schiffer et al. 2015). A gene pertaining to the CYP11 subfamily is present in lungfish as well as in the coelacanth (supplementary fig. 3, Supplementary Material online). The absence of expression in the gonads of these two species is in line with the results of a previous study, which failed to detect 11-ketotestosterone in Australian lungfish (Joss et al. 1996).

Steroid Hormone Receptors

The expression of steroidogenic enzymes is usually connected to the production of biologically active estrogens and androgens. The action of these steroid hormones requires the presence of the cognate receptors, whose transcription is often further up-regulated by the presence of hormones in a positive feedback loop. We found a very unexpected expression pattern for these receptors in lungfish. Indeed, while the *androgen receptor* (*ar*) was expressed in the mature testis of coelacanth and teleost, no expression could be detected in lungfish male gonad nor in any other tissue (supplementary

Table 4
Transcription Factors Generally Related to Male Sex Determination, Differentiation, and Reproductive Functions

	Lungfish		Coelacanth		Zebrafish		Medaka	
	Testis	Ovary	Testis	Ovary	Testis	Ovary	Testis	Ovary
dmrt1	++	+	++	N.R.	++++	+	++	+
dmrt3	–	–	–	N.R.	+	–	+	–
dmrt6	++	+	+++	N.R.	Absent		Absent	
sox3	–	–	–	N.R.	+	–	+	–
sox5	–	–	+	N.R.	+++	–	–	+
sox8a	–	+	–	N.R.	+++	–	–	–
sox8b					++++	+		
sox9a	+[a]	+[a]	++	N.R.	+	–	+	+
sox9b					+[b]	+[b]	+	–
sox10	–	–	+	N.R.	+	–	–	–
gata4	+	–	+	N.R.	+	–	+	+
lrpprc	++	+++	++	N.R.	+++	+++	+	++
sf1	+++	+	+	N.R.	++++	++++	++	+
wt1a	+	–	+	N.R.	++	+	+	+
wt1b							+	+
pax2a	–	+	–	N.R.	+	–	–	–
pax2b					+	–		

[a]In ovary barley above background, but robust expression in testis.
[b]Approximately 3-fold higher in ovary than testis.

table 7, Supplementary Material online). Whether this absence is linked to the nonreproductive stage of the gonads or to the use of another nuclear receptor for sensing androgen levels in lungfish remains an interesting question to be approached in the future. Teleost *esr*s were also preferentially expressed in testis (supplementary table 8, Supplementary Material online). On the other hand, mammalian *ESR*s are produced at very low levels in testis (Cavaco et al. 2009). The exceptionally high levels identified in lungfish male gonads may be related to the premature status of the tissues available for this study (supplementary table 7, Supplementary Material online).

Transcription Factors

Doublesex- and mab-3-related transcriptional factor 1 (Dmrt1), together with Amh and Sox9, is important for vertebrate testicular differentiation. Its expression of *dmrt1* in lungfish (table 4) is in agreement with the dimorphic expression pattern observed from fish to mammals (Webster et al. 2017) and suggest a major role in adult testis function.

In contrast to other vertebrates (Bratuś and Słota 2009), no trace of *dmrt3* expression was detected in the two basal sarcopterygians (table 4). On the other hand, the expression of *dmrt6* in lungfish and coelacanth testis is intriguing (table 4). In mice, this gene is likely involved in directing the transition from spermatogonia to spermatocytes (Zhang, Murphy, et al. 2014). This study allows to extend the taxonomic range distribution of *dmrt6* from tetrapods to basal sarcopterygians. Although no *dmrt6* gene has been annotated yet in any of the fully sequenced teleost genomes, a recent publication reported *dmrt6-like* sequences in spotted gar, cichlids, and siluriforms. This *dmrt6-like* gene was expressed in spermatocytes and its function was shown to be involved in spermatogenesis in tilapia (Zhang, Wang, et al. 2014). The loss of this gene in some teleost lineages could represent another aspect of the extreme variability of the genetic network involved in sex determination and differentiation in fish. Overall, the presence and expression of *dmrt6* in basal sarcopterygians supports an ancient and conserved function for this gene in testis development and function.

The Steroidogenic Factor 1 (Sf1) was mainly expressed in lungfish and coelacanth testis, while in zebrafish it showed a broader tissue distribution (table 4). *Sf1* is expressed in both sexes in mice fetal gonads, but in later developmental stages it continues to be expressed only in testes, where it is activated and maintained by Wilms tumor 1 (WT1). On the other hand, *Sf1* is down-regulated by Forkhead box L2 (FOXL2) in ovary (Takasawa et al. 2014). The low expression of *sf1* observed in lungfish ovary, opposed to the high expression detected in zebrafish female gonad, is in line with the mammalian pattern.

SRY-related HMG box (SOX) family proteins are transcription factors involved in many developmental processes. In particular, *sox3*, *sox5*, *sox8*, *sox9*, and *sox10* have a function in gonad development. *SOX3* belongs to subfamily B and is the ancestor of *SRY* (Marshall-Graves 2008). Although *sox3* is not required for sex determination in mammals

(Weiss et al. 2003), it is the master male sex determining gene in the fish *Oryzias dancena* (Takehana 2014). *sox5*, pertaining to the subfamily D, is a regulator of male germ cells in medaka (Schartl et al. 2018). It is expressed in mice postnatal testis and, together with *Sox9* and *Sox13*, it may play a role in spermatogenesis (Daigle et al. 2015). The other three genes belong to subfamily E: *SOX9* is activated by *SRY* in mammals and is required for sex determination in tetrapods (Jakob and Lovell-Badge 2011); whereas *Sox8* is activated by SOX9 and acts in a positive feedback loop on *Sox9* in testis differentiation, as shown in mice. All five *sox* genes are present in the lungfish transcriptome. In *Latimeria*, in addition to the *sox8/9/10* sequences identified in our previous work (Forconi et al. 2013), we can report the identification of *sox3* (from the genome) and *sox5* (from transcriptome data). The orthology status of the isolated sequences was assessed by phylogenetic analysis (supplementary fig. 4, Supplementary Material online) that assigned each sequence to the corresponding *sox* gene clade.

The tissue distribution of *sox3* transcripts confirms that the almost exclusive role of this gene in brain is also retained in basal sarcopterygians (table 4). Together with SOX1 and SOX2, the other two paralogs of the B1 group of the SOX family, SOX3 plays a key role in central nervous system development in mouse (Rogers et al. 2013). *sox8* was neither expressed in *Latimeria* (Forconi et al. 2013) nor in lungfish testis (table 4). In mammals, the action of SOX8 in the FGF9/SOX9 interaction loop is important for testis maintenance (Barrionuevo et al. 2009). The lack of expression of both *sox8* and *fgf9* hints that this interaction loop, which is crucial in mammalian sex determination, is not yet in place in basal sarcopterygians. Both *sox9* and *wt1* are expressed in lungfish and coelacanth testis, in line with their known pattern of distribution in tetrapods.

Forkhead transcription factors (Fox) are involved in many biological processes, including the regulation of cell proliferation and differentiation, tissue development, and various metabolic processes. Foxl2 plays a key role in the early development of vertebrate female gonads and, in adults, has an ovarian function. In particular FOXL2 suppresses *Dmrt1* for maintaining female fate (Matson et al. 2011) and, together with WNT4, cooperates in regulating *Fst* expression during ovarian development (Garcia-Ortiz et al. 2009). It regulates estrogen synthesis via up-regulation of *Cyp19A1a* (Pannetier et al. 2006). In medaka, the paralogous gene *foxl3* encodes a protein that may play a role in sexual fate decision in germ cells by suppressing spermatogenesis (Nishimura et al. 2015). While both these genes are present in the genome of *L. chalumnae*, only *foxl2* was found in the *P. annectens* transcriptome. The sequences identified were used for phylogenic analyzes (supplementary fig. 5, Supplementary Material online). *foxl2* was expressed in lungfish female gonads (supplementary table 9, Supplementary Material online). No expression of *foxl3* was however detected in coelacanth, lungfish, and zebrafish. Expression of *foxl3* in the gonad was observed only in medaka, indicating a divergent role of this gene between different vertebrates. This gene has not been identified in placental mammals and in amphibians (Bertho et al. 2016) and, except from the study in medaka, its gonadal function remains unclear.

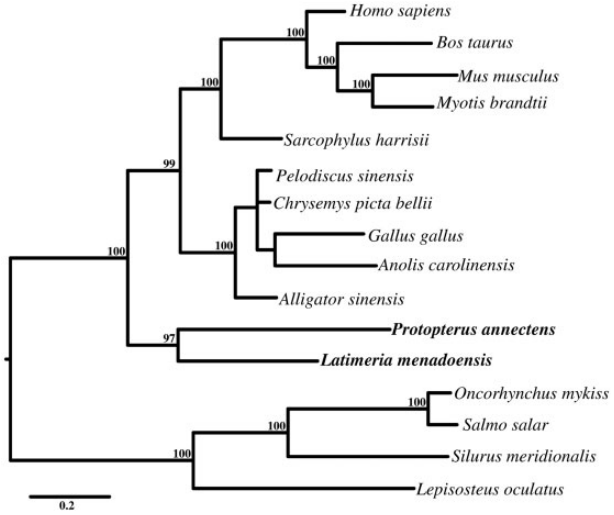

Fig. 5.—Stra8 phylogeny. Phylogenetic analysis of Stra8. Bayesian inference: 1,000,000 generations, sampling every 100, Jones substitution model, stationarity defined as when the average SD of split frequencies approaching 0.0033, burn-in set to 2,500, midpoint rooting. Numbers close to nodes indicate posterior probability values. Only values >95 were reported. The sequences in bold were obtained in this work.

Meiosis Regulation Genes

Meiosis is one of the most important processes in sexual reproduction, as it produces haploid gametes. However, regulation of the meiotic program is still insufficiently understood in most species. In mammals, the morphogen retinoic acid triggers meiotic entry through up-regulation of the *Stimulated by retinoic acid gene 8* (*Stra8*). This gene encodes a retinoic acid-responsive protein that plays a critical role in the regulation of meiotic initiation during spermatogenesis and oogenesis (Feng et al. 2014). Commonly present in tetrapods, this gene is absent in some teleosts species, including zebrafish (Rodríguez-Marí et al. 2013). A more comprehensive analysis of the evolution of *stra8* suggests that this gene was lost in Acanthomorpha and, independently, in the Cypriniform lineage within Actinopterygii (Pasquier et al. 2016). We can however report the presence of this gene in the two basal sarcopterygian species. Phylogenetic analysis positioned the lungfish and coelacanth sequences at a basal position compared with tetrapods (fig. 5).

The level of the retinoic acid as a meiosis signal is determined by the antagonistic action of enzymes involved in its synthesis and degradation: the enzyme ALDH1a2, a member of the retinaldehyde dehydrogenase family, is responsible for

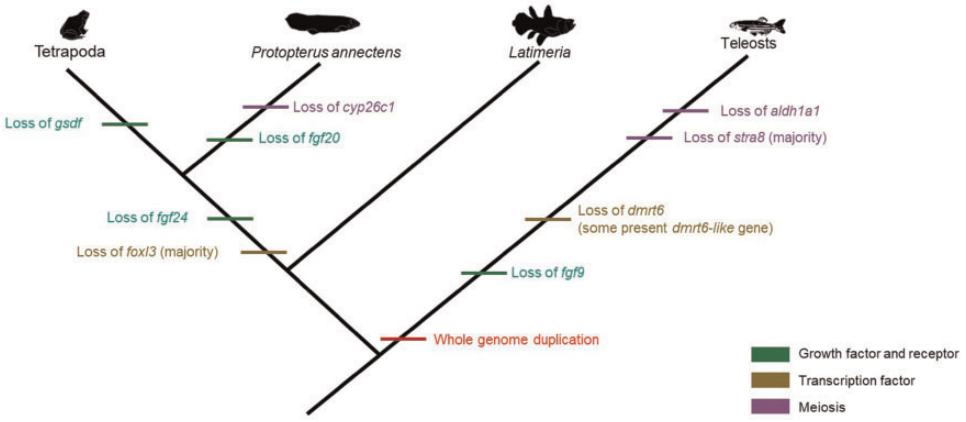

Fig. 6.—Summarizing scheme of sex-related genes lost in the main vertebrate lineages. The scheme shows the sex genes lost in tetrapods, lungfish, coelacanth, and teleosts. The colours are referred to the functional groups, respectively.

its synthesis, while CYP26B1 encodes a major retinoic acid-degrading enzyme. In mice, down-regulation of *Cyp26b1* in females leads to high levels of retinoic acid, which in turn induce the expression of the premeiotic marker *Stra8*, thereby allowing female germ cells to enter meiosis. *Cyp26b1* acts as a meiosis-inhibiting factor in mammals, birds, and amphibians (Bowles and Koopman 2010; Piprek, Pecio, Laskowska-Kaszub, Kloc, et al. 2013; Yu et al. 2013), while in teleosts this role is exerted by *cyp26a1* (Adolfi et al. 2016).

We found high expression of *cyp26b1* and *stra8* in lungfish and coelacanth male gonads suggesting that basal sarcopterygians follow the tetrapod pattern (supplementary table 10, Supplementary Material online). However, significant differences were observed in the expression of *aldh1as*: *aldh1a1* showed higher expression in lungfish male gonads, and is therefore the only gene responsible for retinoic acid synthesis in coelacanth testis (supplementary table 10, Supplementary Material online). Such lineage specific subfunctionalization of retinoic acid metabolizing enzymes has also been reported in a comparative study among different teleost fish (Cañestro et al. 2009). *cyp26c1* was not found in the lungfish transcriptome, marking a possible lineage-specific loss.

Conclusions

Our analysis of sex differentiation and gametogenesis genes in lungfish and coelacanth revealed a complex picture. None of the gene-encoded protein products showed any unusual structural feature, like missing or additional domains or unexpectedly high sequence divergence. We calculated the ω rates for genes that were of special interest due to their appearance/disappearance at the split of sarcopterygians and actinopterygians (*fgf9*, *foxl3*, *gsdf*, and *stra8*), in lungfish, coelacanth and other representative vertebrates (supplementary table 2, Supplementary Material online). All genes appear to evolve under purifying selection (Ka/Ks <1) with the exception of the ω rate obtained between the two *gsdf* co-orthologs in *D. rerio*.

Our analysis of lungfish ovary and testis transcriptomes and the comparison with similar data sets from both zebrafish and medaka, and with the testis transcriptome of coelacanth revealed a situation that reflects the peculiar phylogenetic position of the Dipnoi. We identified some genes whose expression pattern in lungfish and, as far as testis is concerned, also in *Latimeria* is more similar to teleosts than tetrapods. As an example, *gsdf*, which is an important male determining gene in teleosts, but which is missing in the tetrapod lineage, was present and highly expressed in testis of both lungfish and coelacanth.

On the other hand, a number of other genes present in lungfish display a behavior more similar to tetrapods than teleost fishes, for example, *srd5a1*, which was highly expressed in lungfish ovary, like in tetrapods, or *β-catenin*, the transcription factor readout of the female gonad specific tetrapod signaling cascade, which was also highly expressed in lungfish ovary. Intriguingly, the upstream effector R-spondin1 does not follow this pattern. At the same time *wnt4*, which is an ovary specific gene in tetrapods and an upstream regulator of ß-catenin, is highly expressed in teleost testis and did not show significant expression in lungfish and coelacanth gonad samples. *Sf1* is a male specific gene in mammals and lungfish gonad, which was also highly expressed in *Latimeria* testis. *stra8* is absent from many major teleost lineages, but it is present and highly expressed in the testis of basal sarcopterygians, following the tetrapod profile. *Dmrt6*, a gene which has an important function in mammalian spermatogenesis and which is only seldom found in teleost genomes, showed high expression in lungfish and coelacanth testis, indicating functional conservation. The

reasons underpinning the possible loss of such an important gene in most teleosts—similar but independent from *stra8*—are totally unclear.

Finally, some lungfish and coelacanth genes follow neither the teleost nor the tetrapod scheme. *Fgf9* is an important male sex determining gene in tetrapods, which is absent in teleosts. Our analysis demonstrates that *fgf9* has been specifically lost in the teleost lineage and the absence of expression in the lungfish and coelacanth testis may indicate that its role as a male gene was acquired later along the evolution of the tetrapod lineage. Sox8 is a crucial component of the Sox9-driven regulatory loop that determines testis development and Sertoli cell maintenance in mice. Despite the high expression of *sox8* in zebrafish testis, the lack of expression in coelacanth, lungfish, and medaka testis suggests that this regulatory loop is a tetrapod innovation.

Our phylogenetic analyses (fig. 6) pointed out that several genes that are important for sex differentiation and gametogenesis in one branch of the vertebrate tree of life can be totally absent in another branch, further suggesting that other related genes may eventually take over their function, as exemplified by the case of *fgf9* and *fgf24*. In all cases however, the presence/absence of many sex differentiation and gametogenesis genes across different animal groups appears to be mostly due to the lineage-specific loss of ancestral genes rather than to the emergence of novel genes. This implies that the majority of sex differentiation and gametogenesis genes were already present that at the base of vertebrate evolution, after the 1R and 2R whole genome duplication events.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

This work was supported by grants supplied by the Deutsche Forschungsgemeinschaft (SCHA 408/10-1; 12-1) to M.S. This publication was funded by the German Research Foundation (DFG) and the University of Wuerzburg in the funding programme Open Access Publishing. We thank the PhyloSex consortium for making the zebrafish and medaka RNA-seq data available. M.A.B. and A.C. were involved in the identification and characterization of sex genes, phylogenetic analyses, interpreted the data and wrote the paper. M.F. and M.B. did the d*N*/d*S* analyses and contributed to interpretation of the data. A.P. and M.G. subjected in expression analyses. M.C.A. did the histological analyses and contributed to the data interpretation. M.S. conceived the study, interpreted the data and wrote the paper. All authors made contributions to the conception of the study and the experimental design and have given final approval for the version to be published.

Literature Cited

Adolfi MC, Herpin A, Regensburger M, Sacquegno J, Waxman JS. 2016. Retinoic acid and meiosis induction in adult versus embryonic gonads of medaka. Sci Rep. 6:34281.

Amat-Fernandez J, et al. 2017. Molecular characterization of sdf1 and cxcr4 in the Mozambique tilapia, *Oreochromis mossambicus*. Anim Reprod Sci. 176:51–63.

Amemiya CT, et al. 2013. The African coelacanth genome provides insights into tetrapod evolution. Nature 496(7445):311–316.

Barrionuevo F, et al. 2009. Testis cord differentiation after the sex determination stage is independent of Sox9 but fails in the combined absence of Sox9 and Sox8. Dev Biol. 327(2):301–312.

Bertho S, et al. 2016. Foxl2 and its relatives are evolutionary conserved players in gonadal sex differentiation. Sex Dev. 10(3):111–129.

Biscotti MA, et al. 2016. The lungfish transcriptome: a glimpse into molecular evolution events at the transition from water to land. Sci Rep. 6:21571.

Bowles J, Koopman P. 2010. Sex determination in mammalian germ cells: extrinsic versus intrinsic factors. Reproduction 139(6):943–958.

Bratuś A, Słota E. 2009. Comparative cytogenetic and molecular studies of DM domain genes in pig and cattle. Cytogenet Genome Res. 126(1–2):180–185.

Bruggeman V, Van AP, Decuypere E. 2002. Developmental endocrinology of the reproductive axis in the chicken embryo. Comp Biochem Physiol A Mol Integr Physiol. 131(4):839–846.

Canapa A, et al. 2012. Composition and phylogenetic analysis of vitellogenin coding sequences in the Indonesian coelacanth *Latimeria menadoensis*. J Exp Zool B Mol Dev Evol. 318(5):404–416.

Cañestro C, Catchen JM, Rodríguez-Marí A, Yokoi H, Postlethwait JH. 2009. Consequences of lineage-specific gene loss on functional evolution of surviving paralogs: aLDH1A and retinoic acid signaling in vertebrate genomes. PLoS Genet. 5(5):e1000496.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 17(4):540–552.

Cavaco JE, Laurentino SS, Barros A, Sousa M, Socorro S. 2009. Estrogen receptors alpha and beta in human testis: both isoforms are expressed. Syst Biol Reprod Med. 55(4):137–144.

Daigle M, Roumaud P, Martin LJ. 2015. Expressions of Sox9, Sox5, and Sox13 transcription factors in mice testis during postnatal development. Mol Cell Biochem. 407(1–2):209–221.

Dumond H, et al. 2008. Sexual development of the urodele amphibian *Pleurodeles waltl*. Sex Dev. 2(2):104–114.

Fagerberg L, et al. 2014. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteomics 13(2):397–406.

Feng CW, Bowles J, Koopman P. 2014. Control of mammalian germ cell entry into meiosis. Mol Cell Endocrinol. 382(1):488–497.

Fernandez JA, et al. 2015. Primordial germ cell migration in the yellowtail kingfish (*Seriola lalandi*) and identification of stromal cell-derived factor 1. Gen Comp Endocrinol. 213:16–23.

Forconi M, et al. 2013. Characterization of sex determination and sex differentiation genes in *Latimeria*. PLoS One 8(4):e56006.

Garcia-Ortiz JE, et al. 2009. Foxl2 functions in sex determination and histogenesis throughout mouse ovary development. BMC Dev Biol. 9:36.

Goldman N, Yang Z. 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol. 11(5):725–736.

Han YQ, et al. 2015. Molecular cloning and tissue distribution profiles of the chicken R-spondin1 gene. Genet Mol Res. 14(2):3090–3097.

Haning RV Jr, Tantravahi U, Zhao Q, Hackett RJ, Canick JA. 1996. 5alpha-reductase 1 and 2 expression and activity in human ovarian follicles, stroma and corpus luteum as compared to neonatal foreskin. J Steroid Biochem Mol Biol. 59(2):199–204.
Hattori RS, Strüssmann CA, Fernandino JI, Somoza GM. 2013. Genotypic sex determination in teleosts: insights from the testis-determining amhy gene. Gen Comp Endocrinol. 192:55–59.
Herpin A, et al. 2013. Divergent expression regulation of gonad development genes in medaka shows incomplete conservation of the downstream regulatory network of vertebrate sex determination. Mol Biol Evol. 30(10):2328–2346.
Herpin A, Schartl M. 2015. Plasticity of gene-regulatory networks controlling sex determination: of masters, slaves, usual suspects, newcomers, and usurpators. EMBO Rep. 16(10):1260–1274.
Hildebrand M, Goslow GE Jr. 2001. Analysis of vertebrate structure. New York: John Wiley and Sons.
Howe K, et al. 2013. The zebrafish reference genome sequence and its relationship to the human genome. Nature 496(7446):498–503.
Hu Q, Guo W, Gao Y, Tang R, Li D. 2015. Molecular cloning and characterization of amh and dax1 genes and their expression during sex inversion in rice-field eel *Monopterus albus*. Sci Rep. 5:16667.
Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294(5550):2310–2314.
Irisarri I, Meyer A. 2016. The identification of the closest living relative(s) of tetrapods: phylogenomic lessons for resolving short ancient internodes. Syst Biol. 65(6):1057–1075.
Itoh N, Konishi M. 2007. The zebrafish FGF family. Zebrafish 4(3):179–186.
Jakob S, Lovell-Badge R. 2011. Sex determination and the control of Sox9 expression in mammals. FEBS J. 278(7):1002–1009.
Jameson SA, Lin YT, Capel B. 2012. Testis development requires the repression of Wnt4 by Fgf signaling. Dev Biol. 370(1):24–32.
Jansson E, Mattsson A, Goldstone J, Berg C. 2016. Sex-dependent expression of anti-Müllerian hormone (amh) and amh receptor 2 during sex organ differentiation and characterization of the Müllerian duct development in *Xenopus tropicalis*. Gen Comp Endocrinol. 229:132–144.
Joss JM, Edwards A, Kime DE. 1996. In vitro biosynthesis of androgens in the Australian lungfish, *Neoceratodus forsteri*. Gen Comp Endocrinol. 101(3):256–263.
Jovelin R, et al. 2010. Evolution of developmental regulation in the vertebrate FgfD subfamily. J Exp Zool B Mol Dev Evol. 314(1):33–56.
Kaneko H, et al. 2015. Gonadal soma-derived factor (gsdf), a TGF-beta superfamily gene, induces testis differentiation in the teleost fish *Oreochromis niloticus*. Mol Cell Endocrinol. 415:87–99.,
Kasahara M, et al. 2007. The medaka draft genome and insights into vertebrate genome evolution. Nature 447(7145):714–719.
Kikuchi K, Hamaguchi S. 2013. Novel sex-determining genes in fish and sex chromosome evolution. Dev Dyn. 242(4):339–353.
Kinoshita K, Kusunoki M, Miyai K. 2006. Analysis of the three dimensional structure of the CXGXC motif in the CMGCC and CAGYC regions of alpha-and beta-subunits of human chorionic gonadotropin: importance of glycine residue (G) in the motif. Endocr J. 53(1):51–58.
Klüver N, et al. 2007. Differential expression of anti-müllerian hormone (amh) and anti-müllerian hormone receptor type II (amhrII) in the teleost medaka. Dev Dyn. 236(1):271–281.
Leerberg DM, Sano K, Draper BW. 2017. Fibroblast growth factor signaling is required for early somatic gonad development in zebrafish. PLoS Genet. 13(9):e1006993.
Lephart ED, Doody KJ, McPhaul MJ, Simpson ER. 1992. Inverse relationship between ovarian aromatase cytochrome P450 and 5 alpha-reductase enzyme activities and mRNA levels during the estrous cycle in the rat. J Steroid Biochem Mol Biol. 42(5):439–447.
Li M, et al. 2015. A tandem duplicate of anti-Mullerian hormone with a missense SNP on the Y chromosome is essential for male sex determination in nile tilapia, *Oreochromis niloticus*. PLoS Genet. 11(11):e1005678.
Li X, et al. 2016. Regulation of seminiferous tubule-associated stem Leydig cells in adult rat testes. Proc Natl Acad Sci U S A. 113(10):2666–2671.
Lubieniecki KP, et al. 2015. Expression analysis of sex-determining pathway genes during development in male and female Atlantic salmon (*Salmo salar*). Physiol Genomics 47(12):581–587.
Lubzens E, Young G, Bobe J, Cerdà J. 2010. Oogenesis in teleosts: how eggs are formed. Gen Comp Endocrinol. 165(3):367–189.
Mahendroo MS, Russell DW. 1999. Male and female isoenzymes of steroid 5alpha-reductase. Rev Reprod. 4(3):179–183.
Marshall-Graves JA. 2008. Weird animal genomes and the evolution of vertebrate sex and sex chromosomes. Annu Rev Genet. 42(1):565–586.
Matson CK, et al. 2011. DMRT1 prevents female reprogramming in the postnatal mammalian testis. Nature 476(7358):101–104.
Miura T, Miura C, Konda Y, Yamauchi K. 2002. Spermatogenesis-preventing substance in Japanese eel. Development 129(11):2689–2697.
Morinaga C, et al. 2007. The hotei mutation of medaka in the anti-Mullerian hormone receptor causes the dysregulation of germ cell and sexual development. Proc Natl Acad Sci U S A. 104(23):9691–9696.
Myosho T, et al. 2012. Tracing the emergence of a novel sex-determining gene in medaka, *Oryzias luzonensis*. Genetics 191(1):163–170.
Nishimura T, et al. 2015. Sex determination. Foxl3 is a germ cell-intrinsic factor involved in sperm-egg fate decision in medaka. Science 349(6245):328–331.
Pallavicini A, et al. 2013. Analysis of the transcriptome of the Indonesian coelacanth *Latimeria menadoensis*. BMC Genomics 14:538.
Pannetier M, et al. 2006. FOXL2 activates P450 aromatase gene transcription: towards a better characterization of the early steps of mammalian ovarian development. J Mol Endocrinol. 36(3):399–413.
Parenti LR, Grier HJ. 2004. Evolution and phylogeny of gonad morphology in bony fishes. Integr Comp Biol. 44(5):333–348.
Parma P, et al. 2006. R-spondin1 is essential in sex determination, skin differentiation and malignancy. Nat Genet. 38(11):1304–1309.
Pasquier J, et al. 2016. Gene evolution and gene expression after whole genome duplication in fish: the PhyloFish database. BMC Genomics 17:368.
Pfennig F, Standke A, Gutzeit HO. 2015. The role of Amh signaling in teleost fish–Multiple functions not restricted to the gonads. Gen Comp Endocrinol. 223:87–107.
Piprek RP, Pecio A, Laskowska-Kaszub K, Kubiak JZ, Szymura JM. 2013. Sexual dimorphism of AMH, DMRT1 and RSPO1 localization in the developing gonads of six anuran species. Int J Dev Biol. 57(11–12):891–895.
Piprek RP, Pecio A, Laskowska-Kaszub K, Kloc M, et al. 2013. Retinoic acid homeostasis regulates meiotic entry in developing anuran gonads and in Bidder's organ through Raldh2 and Cyp26b1 proteins. Mech Dev. 130(11–12):613–627.
Rodríguez-Marí A, et al. 2013. Retinoic acid metabolic genes, meiosis, and gonadal sex differentiation in zebrafish. PLoS One 8(9):e73951.
Rogers N, et al. 2013. Expression of the murine transcription factor SOX3 during embryonic and adult neurogenesis. Gene Expr Patterns 13(7):240–248.
Schartl M, et al. 2018. Sox5 is involved in germ-cell regulation and sex determination in medaka following co-option of nested transposable elements. BMC Biol. 16(1):16.
Schiffer L, Anderko S, Hannemann F, Eiden-Plach A, Bernhardt R. 2015. The CYP11B subfamily. J Steroid Biochem Mol Biol. 151:38–51.
Schulz RW, et al. 2010. Spermatogenesis in fish. Gen Comp Endocrinol. 165(3):390–411.

Skaar KS, et al. 2011. Proteolytically activated, recombinant anti-mullerian hormone inhibits androgen secretion, proliferation, and differentiation of spermatogonia in adult zebrafish testis organ cultures. Endocrinology 152(9):3527–3540.
Takasawa K, et al. 2014. FOXL2 transcriptionally represses Sf1 expression by antagonizing WT1 during ovarian development in mice. FASEB J. 28(5):2020–2028.
Takehana Y, et al. 2014. Co-option of Sox3 as the male-determining factor on the Y chromosome in the fish *Oryzias dancena*. Nat Commun. 5:4157.
Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. System Biol. 56(4):564–577.
Uribe MC, Grier HJ, Mejía-Roa V. 2014. Comparative testicular structure and spermatogenesis in bony fishes. Spermatogenesis 4(3):e983400.
Wagner GP, Kin K, Lynch VJ. 2012. Measurement of mRNA abundance using RNA-seq data: rPKM measure is inconsistent among samples. Theory Biosci. 131(4):281–285.
Webster KA, et al. 2017. Dmrt1 is necessary for male sexual development in zebrafish. Dev Biol. 422(1):33–46.
Weiss J, et al. 2003. Sox3 is required for gonadal function, but not sex determination, in males and females. Mol Cell Biol. 23(22):8084–8091.
Westernströer B, et al. 2014. Profiling of Cxcl12 receptors, Cxcr4 and Cxcr7 in murine testis development and a spermatogenic depletion model indicates a role for Cxcr7 in controlling Cxcl12 activity. PLoS One 9(12):e112598.
Wrobel KH. 2003. The genus *Acipenser* as a model for vertebrate urogenital development: the muellerian duct. Anat Embryol. 206(4):255–271.
Wu L, Yang P, Luo F, Wang D, Zhou L. 2016. R-spondin1 signaling pathway is required for both the ovarian and testicular development in a teleosts, Nile tilapia (*Oreochromis niloticus*). Gen Comp Endocrinol. 230-231:177–185.
Yang QE, Kim D, Kaucher A, Oatley MJ, Oatley JM. 2013. CXCL12-CXCR4 signaling is required for the maintenance of mouse spermatogonial stem cells. J Cell Sci. 126(Pt 4):1009–1020.
Yao HH, et al. 2004. Follistatin operates downstream of Wnt4 in mammalian ovary organogenesis. Dev Dyn. 230(2):210–215.
Yu M, et al. 2013. RALDH2, the enzyme for retinoic acid synthesis, mediates meiotic initiation in germ cells of the female embryonic chickens. Amino Acids 44(2):405–412.
Zhang T, Murphy MW, Gearhart MD, Bardwell VJ, Zarkower D. 2014. The mammalian Doublesex homolog DMRT6 coordinates the transition between mitotic and meiotic developmental programs during spermatogenesis. Development 141(19):3662–3671.
Zhang X, Wang H, et al. 2014. Isolation of doublesex- and mab-3-related transcription factor 6 and its involvement in spermatogenesis in tilapia. Biol Reprod. 91(6):136.
Zhang X, et al. 2016. Autosomal gsdf acts as a male sex initiator in the fish medaka. Sci Rep. 6:19738.
Zhang Z, et al. 2006. KaKs calculator: calculating Ka and Ks through model selection and model averaging. Genomics Proteomics Bioinformatics 4(4):259–263.
Zhou L, et al. 2016. Rspo1-activated signalling molecules are sufficient to induce ovarian differentiation in XY medaka (*Oryzias latipes*). Sci Rep. 6:19543.
Zhu Y, Wang C, Chen X, Guan G. 2016. Identification of gonadal soma-derived factor involvement in *Monopterus albus* (protogynous rice field eel) sex change. Mol Biol Rep. 43(7):629–637.

Associate editor: Maria Costantini