The Genetics of Sexual Development in East African Cichlid Fishes

Inauguraldissertation

Zur Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Corina Heule

aus Altstätten, Schweiz

Basel, 2015

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Walter Salzburger
(Universität Basel)

und

Prof. Judith Mank
(University College London)

Basel, den 12.11.2013

Prof. Dr. Jörg Schibler
(Dekan, Universität Basel)

Contents

Introduction

The cichlid fishes of East Africa

Cichlid fishes occur in tropical freshwater lakes and river systems around the world (Salzburger and Meyer 2004; Salzburger 2009). With ~ 3000 species (Turner *et al.* 2001; Turner 2007) they represent the largest family of all ray-finned fishes (Actinopterygii) (~ 31,000 species http://www.catalogueoflife.org). Remarkably, almost half of the cichlid species inhabit the Great Lakes and surrounding rivers of East Africa. Lake Malawi harbors ~ 600 species, Lake Victoria ~ 500 and Lake Tanganyika ~ 180 (Turner 2007). This richness in species likely evolved within a few million years only (Kocher 2004), and the species flocks of cichlids represent an outstanding example of adaptive radiations (Seehausen 2006). In addition to the exceptionally high species number, cichlid fishes in the Great Lakes are extremely diverse in terms of ecology, morphology and breeding behavior and therefore represent a major model system in evolutionary biology (Kornfield and Smith 2000; Kocher 2004; Seehausen 2006; Salzburger 2009; Santos and Salzburger 2012).

The phenotypic diversity is reflected in e.g., the trophic morphology, which correlates with preferred food items (Fryer and Iles 1972; Liem 1973; 1980; Muschick *et al.* 2012). Besides cichlids that feed on algae, mollusks or fish, there exist species with a highly specialized diet such as scales, eyes or fins of other fish (Fryer and Iles 1972; Muschick *et al.* 2012). Cichlids can also be colorful and many species show a distinct sexual color dimorphism with rather dull females and brightly colored males. One color trait particularly studied are so-called egg-spots: conspicuous yellow, orange or red ovoid markings with an outer transparent circle, generally on the anal fin of males, thought to be important for courtship behavior and intra-sexual communication (Wickler 1962a; b; Fryer and Iles 1972; Hert 1989; Goldschmidt 1991; Theis *et al.* 2012). Egg-spots vary in size, number, intensity, and position on the fin (Theis *et al*. unpublished). In addition to these diverse phenotypic traits, cichlids show distinct types of breeding behaviors: substrate-breeding, shell-breeding and mouth-breeding, which can be accompanied by different strategies such as cooperative or parasitic reproductive behavior and paternal/maternal/biparental brood care (as in (Goodwin *et al.* 1998; Taborsky 2001; Barlow 2008; Schütz *et al.* 2010)).

The oldest of the cichlid radiations, the assemblage in Lake Tanganyika, is genetically, morphologically and behaviorally more diverse than the cichlid communities of the other lakes (Salzburger *et al.* 2002; Kocher 2004). In contrast to the other lakes, Lake Tanganyika harbors more than one lineage, so-called tribe (Salzburger *et al.* 2002; Salzburger and Meyer 2004). All the cichlid species in Lake Malawi and Victoria belong to the lineage of the haplochromines and are derived from Lake Tanganyika tribes (Salzburger *et al.* 2005; Sturmbauer *et al.* 2011). Haplochromines are sexually dimorphic maternal mouthbrooders and include the 'modern haplochromines', which additionally all carry anal fin egg-spots. (Fryer and Iles 1972; Salzburger *et al.* 2005; Salzburger *et al.* 2007).

Astatotilapia burtoni (Günther 1894) is a basal member of this rapidly radiating and especially species-rich lineage and is a model species in different fields of biology including evolution, ecology and behavior (Salzburger *et al.* 2007; Siegel *et al.* 2007; Theis *et al.* 2012; Colombo *et al.* 2013). It occurs in affluent rivers and close to estuaries in the shore area of Lake Tanganyika (Fernald and Hirata 1977). With a sequenced genome, *A. burtoni* offers new possibilities to address scientific approaches, especially in the study of the genetic basis underlying specific traits. In my doctoral work, *A. burtoni* was subject in several studies that aimed to answer questions regarding egg-spot development (Chapter 1) and expression patterns of sex-specific genes in adults (Chapter 3) and in developing males (Chapter 4).

Sex determination mechanisms

In general, different strategies for sexual reproduction exist: unisexuality (all-female species), hermaphroditism (sequential, serial and simultaneous) and gonochorism (two separate sexes at all life stages). These strategies can be controlled genetically, environmentally or by a combination thereof (for more details and references see Chapter 2). Mammals, birds, amphibians and reptiles are gonochorists with the following sex-determining systems: mammals possess an XX-XY sex-determining system with a male sex-determiner on the Y-chromosome whereas birds have a ZZ-ZW sex-determining system with a yet not fully resolved underlying genetic mechanism (for reviews see (Kashimada and Koopman 2010; Chue and Smith 2011)). Amphibians display different genetic systems (including a variety of XX-XY and ZZ-ZW systems) and reptiles additionally exhibit environmental (temperature dependent) sex determination (for a review see (Sarre *et al.* 2011)).

In gonochoristic fish, all of the above-mentioned types and even more complex sex determination mechanisms exist (Chapter 2). For example the platyfish has three types of

sex chromosomes (X, Y and W) and depending on the combination, it is either a male (XY or YY) or a female (XX, XW, YW) (Schultheis *et al.* 2006). The zebrafish is likely to have a polygenic sex determination system, which is governed by strain-specific sex-associated loci and influenced by differential sensitivity to environmental factors (Shang *et al.* 2006; Bradley *et al.* 2011; Anderson *et al.* 2012). Various salmonid species share the same male sex-determiner (Yano *et al.* 2012; Yano *et al.* 2013). Surprisingly, this locus maps on different linkage groups between and sometimes even within species, supposedly due to transposition events (Phillips 2013). In addition to gonochorism, the other strategies for sexual reproduction, unisexuality and hermaphroditism, are also seen in fish (Chapter 2). A famous example for a sex-changing fish is the clownfish (*Amphiprion percula*), a protandrous species, living in social groups with one dominant female, one dominant male and up to four subordinate males. Once the female dies or disappears, the dominant male undergoes sex change and becomes a female; and one of the subordinate males then takes over the role of the dominant male (Buston 2004).

Sex determination in cichlids

Most of the studies on sex determination and differentiation in cichlids have been conducted in one of the more basal species, the Nile tilapia (*Oreochromis niloticus*). This species is widely distributed in African lakes and rivers and has a huge commercial impact as a food fish. Males grow faster and are therefore more marketable (Baroiller and Jalabert 1989). The Nile tilapia possesses an XX-XY sex determination system, which can substantially be influenced by environmental factors, i.e., temperature changes during specific time windows. At the same time the Nile tilapia, and other fish, can be sex-reversed by exogenous hormones (Baroiller *et al.* 2009). This sensitivity enables the establishment of all-male broods either by treatment with androgens directly or by making use of YY-males that only produce male offspring. These super-males are descendants of estrogen-induced sex-reversed males (morphological females with an XY genotype) and normal males (XY) (Baroiller and Jalabert 1989; Scott *et al.* 1989). Other cichlid species investigated with regard to sex determination are the *mbuna* cichlids from Lake Malawi. A new female sex-determiner on a potential W-chromosome invaded the genomes of theses cichlids with an ancestral XX-XY system, likely resolving a sexual conflict over coloration (Roberts *et al.* 2009; Ser *et al.* 2010; Parnell and Streelman 2012). Another studied species is a Lake Victoria cichlid, in which sex determination is substantially influenced by B-chromosomes (Yoshida *et al.* 2011). These are small supernumerary chromosomes that exist in more than 2000 eukaryotic species additional to the standard A-chromosomes (autosomes and sex

chromosomes) (Camacho 2005).

These examples illustrate that sex determination mechanisms are diverse in cichlids and at the same time they highlight the complexity thereof. For a better understanding I would like to define the following terms: In agreement with (Uller and Helanterä 2011) we do not make a clear distinction between sex determination and sex differentiation, but rather see these processes as a continuum of sexual development. The classical division between the two terms might be too strict for most of the fish species because of the observed plasticity in sexual development of fish, i.e., the sex of an individual can be determined genetically (sex determination) but formation of gonads (sex differentiation) does not follow its fate due to influences of environmental factors. In addition, sex in fish can be functionally reversed even in adult stages (Paul-Prasanth *et al.* 2013). As master sex-determining gene (master regulator, sex determination gene) we refer to a gene that is sufficient to initiate sexual development of a male or a female with no other genetic or environmental factors changing the phenotypic pathway (in accordance with (Yano *et al.* 2012)). Note, that fish with polygenic sex determination (Bradley *et al.* 2011; Anderson *et al.* 2012; Parnell and Streelman 2012) do not have a master sex determination gene in this sense.

Outline of my PhD thesis

Chapter 1

In this paper, we studied the development of anal fin egg-spots in *Astatotilapia burtoni*. Differential group and individual set-ups were used to investigate the morphological development of this putative evolutionary key-innovation with the help of picture series. Based on the photographs of one male, we defined four stages of egg-spot formation. Variable raising conditions (different densities in the group and different ages in the individual set-up) led to differential growth rates and thereby slightly altered timing of the four stages. Standard body length was found to be a better predictor for the development of egg-spots than age under the given circumstances of differential raising conditions.

The staging of egg-spot formation was only possible because we kept fish individually, which allowed us to sex fish retrospectively. Six out of the 13 individual fish were males. A possibility to identify males before the experiment i.e., with the help of sex-specific markers, would have allowed us to increase sample size. In addition, a molecular investigation of the formation of egg-spots using gene expression analysis always contains female individuals

as long as fish cannot be sexed morphologically.
These limitations awakened my interest to identify genetic differences between male and female *A. burtoni* and therefore led me to investigate sexual development in cichlids, the main topic of my PhD thesis.

Chapter 2

In this review we summarize and discuss the different kinds of sex determination mechanisms reported in gonochoristic teleost fishes. We first provide a brief overview on environmental sex determination systems with a focus on temperature dependent sex determination. We then review the diversity of genetic sex determination mechanisms. In fish, these involve sex chromosomes, the action of monogenic or polygenic regulators and the influence of B-chromosomes. Sex chromosomal systems include simple male and female heterogamety (XX-XY and ZZ-ZW) and more complex mechanisms (with more than one pair or more than two types of sex chromosomes). In particular, we discuss the five master sex determination genes discovered in fish and how sexual development cascades might have evolved.

Chapter 3

In this paper we investigated a set of sexual development candidate genes for patterns of sequence evolution and gene expression in a representative subset of East African cichlids. Expression of all candidate genes was measured in brain and gonad tissue of adult male and female cichlid fishes. The most astonishing result was the shift in expression of the two copies of the aromatase *cyp19a1*. Instead of the presumably conserved ovary- and brain-specific expression of the A- and B-copy, respectively, we found tribe-specific expression shifts towards testis. The A-copy was expressed in testis in the lineage of the Ectodini and the B-copy was highly expressed in the testis of the Haplochromini (Haplochromines, see above).

Chapter 4

This chapter describes a main project of my PhD thesis. The search for a sex-specific marker in *Astatotilapia burtoni* led from the candidate gene approach in adult cichlids (described in Chapter 3) to an extension of these results to a series of gene expression profiling during development. The aim was to characterize expression patterns of the same set of candidate genes during early male and female development with the help of all-male and all-female broods. In a first step, we treated young fish with hormones in order to sex-reverse them. This was successful for the estrogen treatment only and resulted in all fish growing to morphological females in this experimental part. Subsequent crossing experiments provided evidence for an XX-XY sex-determining system in *A. burtoni*

and revealed a putative YY-male that only produced male offspring. This super-male was crossed to untreated XX-females in order to obtain all-males broods, which were then used to profile gene expression of the candidate genes in heads and trunks (as proxies for brain and gonad tissue) during male development. We described genes presumably involved in early and late testis, as well as in brain formation. Two of the genes showing a peak early in development and in the trunks only were investigated in more detail because their expression profile suggested a position at, or at least close to, the top of the sexual development cascade. Genomic sequence analysis excluded these genes as the master sex-determiner in *A. burtoni*. Nonetheless, with a putative transcription factor binding site analysis, we identified possible regulators acting upstream of those early testis genes and thus representing promising candidates for the sex determination gene in this species.

References

Anderson, J. L., A. Rodríguez-Marí, I. Braasch, A. Amores, P. Hohenlohe *et al.*, 2012 Multiple sex-associated regions and a putative sex chromosome in zebrafish revealed by RAD mapping and population genomics. PLoS One 7: e40701.

Barlow, G., 2008 *The cichlid fishes: nature's grand experiment in evolution*. Basic Books.

Baroiller, J.-F., and B. Jalabert, 1989 Contribution of research in reproductive physiology to the culture of tilapias. Aquat. Living Resour. 2: 105-116.

Baroiller, J. F., H. D'Cotta, E. Bezault, S. Wessels and G. Hoerstgen-Schwark, 2009 Tilapia sex determination: where temperature and genetics meet. Comp. Biochem. Physiol. Part A Mol. Integr. Physiol. 153: 30-38.

Bradley, K. M., J. P. Breyer, D. B. Melville, K. W. Broman, E. W. Knapik *et al.*, 2011 An SNP-based linkage map for zebrafish reveals sex determination loci. G3 (Bethesda) 1: 3-9.

Buston, P. M., 2004 Territory inheritance in clownfish. Proc Biol Sci 271 Suppl 4: S252-254.

Camacho, J. P. M., 2005 B chromosomes, pp. 223-286 in *The evolution of the genome*, edited by T. R. Gregory. Academic Press, Burlington.

Chue, J., and C. A. Smith, 2011 Sex determination and sexual differentiation in the avian model. FEBS J. 278: 1027-1034.

Colombo, M., E. T. Diepeveen, M. Muschick, M. E. Santos, A. Indermaur *et al.*, 2013 The ecological and genetic basis of convergent thick-lipped phenotypes in cichlid fishes. Mol. Ecol. 22: 670-684.

Fernald, R. D., and N. R. Hirata, 1977 Field study of Haplochromis burtoni: habitats and co-habitant. Environ. Biol. Fishes 2: 299-308.

Fryer, G., and T. D. Iles, 1972 *The cichlid fishes of the Great Lakes of Africa : their biology and evolution*. [s.n.], Edinburgh.

Goldschmidt, T., 1991 Egg mimics in haplochromine cichlids (Pisces, Perciformes) from Lake Victoria. Ethology 88: 177-190.

Goodwin, N. B., S. Balshine-Earn and J. D. Reynolds, 1998 Evolutionary transitions in parental care in cichlid fish. Proc Biol Sci 265: 2265-2272.

Hert, E., 1989 The function of egg-spots in an African mouth-brooding cichlid fish. Anim. Behav. 37, Part 5: 726-732.

Kashimada, K., and P. Koopman, 2010 *Sry*: the master switch in mammalian sex determination. Development 137: 3921-3930.

Kocher, T. D., 2004 Adaptive evolution and explosive speciation: the cichlid fish model. Nat. Rev. Genet. 5: 288-298.

Kornfield, I., and P. F. Smith, 2000 Africa cichlid fishes: model systems for evolutionary biology. Annu. Rev. Ecol. Syst.

Liem, K. F., 1973 Evolutionary strategies and morphological innovations: cichlid pharyngeal jaws. Syst. Biol. 22: 425-441.

Liem, K. F., 1980 Adaptive significance of intra- and interspecific differences in the feeding

repertoires of cichlid fishes. Am. Zool. 20: 295-314.

Muschick, M., A. Indermaur and W. Salzburger, 2012 Convergent evolution within an adaptive radiation of cichlid fishes. Curr. Biol. 22: 2362-2368.

Parnell, N. F., and J. T. Streelman, 2012 Genetic interactions controlling sex and color establish the potential for sexual conflict in Lake Malawi cichlid fishes. Heredity 110: 239-246.

Paul-Prasanth, B., R. K. Bhandari, T. Kobayashi, R. Horiguchi, Y. Kobayashi *et al.*, 2013 Estrogen oversees the maintenance of the female genetic program in terminally differentiated gonochorists. Sci. Rep. 3: 2862.

Phillips, R. B., 2013 Evolution of the sex chromosomes in salmonid fishes. Cytogenet. Genome Res.: 177-185.

Roberts, R. B., J. R. Ser and T. D. Kocher, 2009 Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. Science 326: 998-1001.

Salzburger, W., 2009 The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. Mol. Ecol. 18: 169-185.

Salzburger, W., I. Braasch and A. Meyer, 2007 Adaptive sequence evolution in a color gene involved in the formation of the characteristic egg-dummies of male haplochromine cichlid fishes. BMC Biol. 5: 51.

Salzburger, W., T. Mack, E. Verheyen and A. Meyer, 2005 Out of Tanganyika: Genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. BMC Evol. Biol. 5: 17.

Salzburger, W., and A. Meyer, 2004 The species flocks of East African cichlid fishes: recent advances in molecular phylogenetics and population genetics. Naturwissenschaften 91: 277-290.

Salzburger, W., A. Meyer, S. Baric, E. Verheyen and C. Sturmbauer, 2002 Phylogeny of the Lake Tanganyika cichlid species flock and its relationship to the Central and East African haplochromine cichlid fish faunas. Syst. Biol. 51: 113-135.

Santos, M. E., and W. Salzburger, 2012 Evolution. How cichlids diversify. Science 338: 619-621.

Sarre, S. D., T. Ezaz and A. Georges, 2011 Transitions between sex-determining systems in reptiles and amphibians. Annu. Rev. Genomics Hum. Genet. 12: 391-406.

Schultheis, C., Q. Zhou, A. Froschauer, I. Nanda, Y. Selz *et al.*, 2006 Molecular analysis of the sex-determining region of the platyfish *Xiphophorus maculatus*. Zebrafish 3: 299-309.

Schütz, D., G. Pachler, E. Ripmeester, O. Goffinet and M. Taborsky, 2010 Reproductive investment of giants and dwarfs: specialized tactics in a cichlid fish with alternative male morphs. Funct. Ecol. 24: 131-140.

Scott, A. G., D. J. Penman, J. A. Beardmore and D. O. F. Skibinski, 1989 The 'YY' supermale in *Oreochromis niloticus* (L.) and its potential in aquaculture. Aquaculture 78: 237-251.

Seehausen, O., 2006 African cichlid fish: a model system in adaptive radiation research.

Proc. R. Soc. Biol. Sci. Ser. B 273: 1987-1998.

Ser, J. R., R. B. Roberts and T. D. Kocher, 2010 Multiple interacting loci control sex determination in Lake Malawi cichlid fish. Evolution 64: 486-501.

Shang, E. H. H., R. M. K. Yu and R. S. S. Wu, 2006 Hypoxia affects sex differentiation and development, leading to a male-dominated population in zebrafish (*Danio rerio*). Environ. Sci. Technol. 40: 3118-3122.

Siegel, N., S. Hoegg, W. Salzburger, I. Braasch and A. Meyer, 2007 Comparative genomics of ParaHox clusters of teleost fishes: gene cluster breakup and the retention of gene sets following whole genome duplications. BMC Genomics 8: 312.

Sturmbauer, C., M. Husemann and P. D. Danley, 2011 Explosive speciation and adaptive radiation of East African cichlid fishes, pp. 333-362 in *Biodiversity Hotspots*, edited by F. Zachos and J. Habel. Springer Berlin Heidelberg, Berlin, Heidelberg.

Taborsky, M., 2001 The evolution of bourgeois, parasitic, and cooperative reproductive behaviors in fishes. J. Hered. 92: 100-110.

Theis, A., W. Salzburger and B. Egger, 2012 The function of anal fin egg-spots in the cichlid fish *Astatotilapia burtoni*. PLoS One 7: e29878.

Turner, G. F., 2007 Adaptive radiation of cichlid fish. Curr. Biol. 17: R827-R831.

Turner, G. F., O. Seehausen, M. E. Knight, C. J. Allender and R. L. Robinson, 2001 How many species of cichlid fishes are there in African lakes? Mol. Ecol. 10: 793-806.

Uller, T., and H. Helanterä, 2011 From the origin of sex-determining factors to the evolution of sex-determining systems. Q. Rev. Biol. 86: 163-180.

Wickler, W., 1962a 'Egg-dummies' as natural releasers in mouth-breeding cichlids. Nature 194: 1092-1093.

Wickler, W., 1962b Zur Stammesgeschichte funktionell korrelierter Organ- und Verhaltensmerkmale: Ei-Attrappen und Maulbrüten bei afrikanischen Cichliden. Zeitschrift für Tierpsychologie 19: 129-164.

Yano, A., R. Guyomard, B. Nicol, E. Jouanno, E. Quillet *et al.*, 2012 An immune-related gene evolved into the master sex-determining gene in rainbow trout, *Oncorhynchus mykiss*. Curr. Biol. 22: 1423-1428.

Yano, A., B. Nicol, E. Jouanno, E. Quillet, A. Fostier *et al.*, 2013 The sexually dimorphic on the Y-chromosome gene (*sdY*) is a conserved male-specific Y-chromosome sequence in many salmonids. Evol. Appl. 6: 486-496.

Yoshida, K., Y. Terai, S. Mizoiri, M. Aibara, H. Nishihara *et al.*, 2011 B chromosomes have a functional effect on female sex determination in Lake Victoria cichlid fishes. PLoS Genet. 7: e1002203.

Chapter 1

The ontogenetic development of egg-spots in the haplochromine cichlid fish *Astatotilapia burtoni*

Corina Heule and Walter Salzburger

CH performed the experiments, drafted the manuscript and prepared the figures. WS helped with the experimental design of the study and the preparation of the manuscript.

Published in *Journal of Fish Biology*

Journal of FISH BIOLOGY
Published for the Fisheries Society of the British Isles

Journal of Fish Biology (2011) **78**, 1588–1593
doi:10.1111/j.1095-8649.2011.02944.x, available online at wileyonlinelibrary.com

The ontogenetic development of egg-spots in the haplochromine cichlid fish *Astatotilapia burtoni*

C. Heule and W. Salzburger*

Zoological Institute, University of Basel, Vesalgasse 1, 4051 Basel, Switzerland

(Received 17 November 2010, Accepted 9 February 2011)

A series of *Astatotilapia burtoni* photographs were used to investigate the ontogenetic development of male egg-spots, a putative evolutionary key innovation of haplochromine cichlids. Four stages of egg-spot development were defined and all males had developed true egg-spots (stage 4) before reaching a standard length of 25 mm. Raising condition only slightly influenced the timing of the first appearance of true egg-spots.

Key words: adaptive radiation; key innovation; ontogeny; egg-dummies; photographic series.

The species flocks of cichlid fishes in the East African Great Lakes Victoria, Malawi and Tanganyika represent the most species-rich and eco-morphologically diverse animal adaptive radiations (Seehausen, 2006; Salzburger, 2009). At least 1500 cichlid species have evolved in East Africa in only the last few thousands to several millions of years (Salzburger, 2009). The species differ greatly in ecologically relevant, *i.e.* naturally selected, structures such as body morphology, jaw and tooth shape, and also in sexually selected traits such as colouration (Salzburger, 2009). The vast majority of East Africa's lacustrine and riverine cichlids and all species of the species flocks of Lakes Victoria and Malawi belong to a single 'tribe' of cichlids, the Haplochromini (Salzburger *et al.,* 2005). Haplochromine cichlids are characterized by a slight modification of the pharyngeal jaw apparatus, pronounced sexual colour dimorphism, derived maternal mouth-brooding behaviour and, in most species, anal fin egg-spots (Fryer & Iles, 1972; Greenwood, 1979; Salzburger *et al*., 2005). Their reproductive system involving polychromatism, female mouth-brooding and egg-spots has been suggested to be one of the evolutionary key innovations responsible for the haplochromines' propensity for explosive speciation (Salzburger *et al.,* 2005, 2007).

Egg-spots or 'egg-dummies' are conspicuous orange, yellow or red spots with an outer transparent circle, which appear on the anal fin of most males of the derived 'modern haplochromines' (Wickler, 1962*a*; Fryer & Iles, 1972; Salzburger *et al*.,

*Author to whom correspondence should be addressed. Tel.: +41 61 267 03 03; email: walter.salzburger@unibas.ch

2005). Note that egg-spots may also occur on females' anal fins, but in a less perceptible manner. Egg-spots are involved in the modern haplochromines' courtship behaviour. Usually, a male initiates courtship by displaying his nuptial colouration (and egg-spots) while quivering in order to attract the female. The female lays a batch of eggs in the pit, which was created by the male in his territory. She then swims a circle and picks the eggs up in her mouth. Seemingly attracted by the colourful egg-spots on the male anal fin, she approaches the male's genital opening where she attempts to take up these putative 'egg-dummies'. In this moment the male releases the sperm, which fertilizes the eggs in the female's buccal cavity (Wickler, 1962*a*, *b*). The female then carries the larvae until fry are fully developed (which may last for several weeks). This leads to fewer but larger eggs that are well protected in the buccal cavity of the mouth-brooding female. Despite the importance of this trait in cichlid evolution (Salzburger *et al.,* 2007), little is known about the developmental basis of egg-spot formation.

In this study, the ontogenetic development of egg-spots was investigated in Burton's haplo *Astatotilapia burtoni* (Günther 1894), a cichlid species living in Lake Tanganyika and its surrounding affluent rivers. It belongs to the modern haplochromines and is, hence, a maternal mouth-brooder (Salzburger *et al.,* 2005). *Astatotilapia burtoni* has been established as a model system to study questions from various fields of biology such as behaviour, ecology, genetics, genomics and development (Salzburger *et al*., 2008). Most importantly, its genome is one of the five cichlid genomes that are currently being sequenced.

Almost 50 years ago, Wickler observed that egg-spots in *A. burtoni* occur only later in ontogeny (Wickler, 1962*b*) and that, in the beginning, the anal fin is transparent. Egg-spot formation starts with a yellowish pattern covering the fin, which then forms blotches that finally become surrounded by an outer transparent circle. The present study focused on the ontogenetic development of egg-spots in *A. burtoni* raised under different conditions. Of main interest were (1) the timing of egg-spot formation with respect to age, body size and mass and (2) the influence of raising condition on egg-spot development.

In order to answer these questions, egg-spot formation was investigated in two main set-ups: a group set-up and an individual series. For the group set-up, *c*. 30 fry of an inbred laboratory strain were taken out of a female's mouth and kept in a net measuring 16·5 × 13 × 13 cm that was placed in a small tank (group 1). In addition, *c*. 30 fish were kept in a small tank measuring 40 × 25 × 25 cm (group 2). They were already 2 months old when starting the experiment. The fish in both tanks were fed with a similar amount of regular flake food. The following procedure was performed every 2–3 days during a period of 12 and 8 weeks, respectively: Six random individuals of each group were anaesthetized with 130 mg l^{-1} MS222 (Sigma-Aldrich; www.sigmaaldrich.com) for 1 min, weighed with a scale (KERN 440-3A, Kern & Sohn Gmbh; www.Kern-Sohn.com) and photographed with a Nikon D40 (www.nikon.com), using a macro lens (Tokina AT-X MACRO 90 mm 1:2.5; www.tokinalens.com) in a specially designed photo-cuvette. This cuvette made it possible to take photographs of the fish in water, so the fish were less stressed. In addition, the fins were well spread in these cuvettes. Usually the photographs were taken in the morning and then used to measure the standard length (L_S) (using the TPSDIG software from Rohlf, 2006) ($n = 158$) and to define the timing of egg-spot development ($n = 1500$). Data were displayed with R (R Developmental Core

Team; www.r-project.org). For the individual series, 14 fish of age 18 dpf (days post fertilization) were kept isolated in nets measuring 16·5 × 13 × 13 cm that were all placed in the same tank. All individuals were photographed and weighed twice a week for a period of 14 weeks (series A) following the above-mentioned procedure. As seven individuals escaped, six fish of about the same size but of a different age (90 dpf) were added in week 7 (series B). These individual fish were also fed with about the same amount of regular flake food. A total of 2854 photographs were taken from the fish of the individual series to define the timing of egg-spot formation; 121 of these were used to measure L_S in males. Additionally, a movie file was created from individual images showing the ontogeny of one male (18–113 dpf) (Movie S1, Supporting information).

The analysis of the group set-up revealed differences in growth rate between group 1 and group 2 [Fig. 1(a)]. The same was true for series A and series B of the individual series [Fig. 1(a)]. The individual series was used to define stages of egg-spot development in *A. burtoni* males, because here fish could be sexed retrospectively. Therefore, only males were used for further analyses. Of the 13 fish raised individually, six were males. The ontogenetic development of egg-spots in one male over 3 months (first 24 photographs) and one additional photograph after a year are shown in Fig. 2. The photographs were inspected by eye and four stages were defined, which are highlighted in different colours. The first stage (green border strip in Fig. 2) was categorized by a transparent anal fin with some yellow pigments only, particularly at the first three fin rays. The second stage (orange) was characterized by the occurrence of orange pigments (arrowhead in Fig. 2). In the third stage (blue) the orange pigments accumulated into amorphic spots. Only in the fourth stage (red), fully developed egg-spots with an outer transparent circle became visible. In all cases, also in the group treatment, there was an initial number

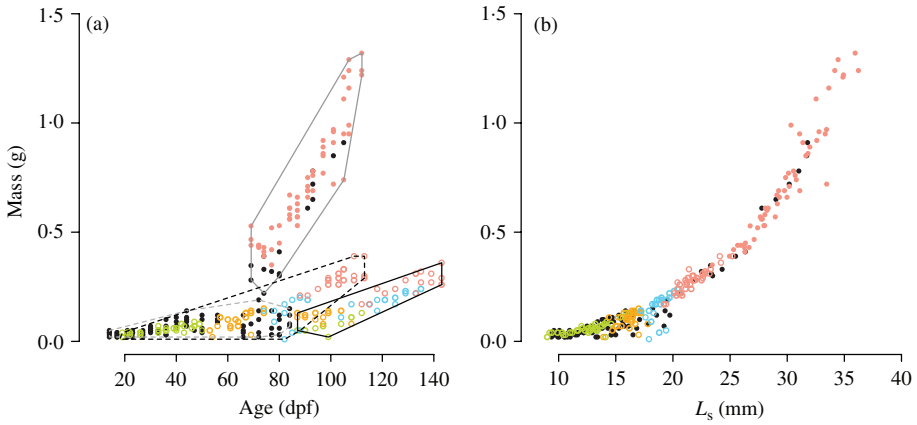

FIG. 1. Ontogenetic development of egg-spots in *Astatotilapia burtoni*, plotted as mass (a) at age in days post fertilization (dpf) and (b) at standard length (L_S). •, group set-up stages 1–3; •, group set-up stage 4; ○, individual series stage 1; ○, individual series stage 2; ○, individual series stage 3; ○, individual series stage 4. The outlines indicate the different set-ups. - - - -, group 1; ——, group 2; - - - -, individual series A; ——, individual series B.

of three egg-spots. Only later in development did this number increase. Females developed egg-spot-like structures that were less conspicuous and occurred much later.

The remaining photographs of the individual series were then inspected by eye and classified into one of the four stages defined above. The same procedure was performed with photographs of the group set-up, although here the main focus was on the identification of the age, size and mass at which true egg-spots in males appeared (*i.e.* stage 4). Note that the classification into stages 1–3 was more problematic in the group set-up, as unlike in the individual series, fish could not be sexed retrospectively. As mentioned above, the distinct egg-spot morphology of females could bias the classification. The data from the group set-up and the individual series were then jointly plotted (Fig. 1). The correlation between age and mass [Fig. 1(a)] was confounded by treatment effects in the different set-ups of groups 1 and 2 and series A and B, whereas treatment did not have an apparent effect on the correlation between L_S and mass [Fig. 1(b)]. As a consequence, L_S was used to describe egg-spot appearance.

This study showed that, in *A. burtoni*, growth rate and the appearance of egg-spots do not depend on age, but rather on the condition in which the fish had been raised. For example, in group 1, where density was higher compared to group 2, fish were smaller at the same age [Fig. 1(a)]. Also, in the individual series, fish of series B were smaller at the same age compared to series A. Note that fish of series B were added 3 months after starting the experiment, but had been raised before in a small tank at higher density than the isolated fish from series A. In the group set-up, growth differences between the groups had an effect on pigmentation, as the orange pigmentation (stage 2) was detectable at a smaller size in group 1 compared to group 2. In contrast, in the individual series, the first detectable orange pigments appeared at the same L_S, but at a much younger age in the fish from series A compared to those of series B. This variation in the timing of the occurrence of the orange pigmentation might be explained by differences in food uptake. Egg-spots are a costly trait, as the orange colour comes from carotenoids imbedded in chromatophores. It is these carotenoids that cannot be synthesized by fishes, so they have to be taken up with nutrition (Endler, 1983; Kodric-Brown, 1985). As the isolated fish did not have to compete for food, they probably had more nutrition available. This would explain why the fish that had been isolated for their whole life (series A) were able to develop their egg-spots at a younger age than those of series B. Alternatively, isolation as such and light condition (it was darker in the small nets than in glass aquaria) might have had an effect on pigmentation. For example, Goldschmidt (1991) has shown that light intensity influenced egg-spot size in Lake Victoria cichlids.

Taken together, four stages of egg-spot formation were defined in *A. burtoni* (Fig. 2). True egg-spots (stage 4) were first observed in an individual male at $L_S = 19{\cdot}2$ mm in the individual series. Although egg-spots appeared at slightly smaller L_S in the individual series, all males showed egg-spots at 25 mm L_S [Fig. 1(b)]. Any fish >25 mm and without egg-spots [black dots in Fig. 1(b)] was most certainly a female. On the basis of the observed differences in pigmentation between group set-up and individual series, it is suggested that individually marked fish are raised in group tanks to rule out condition effects (*e.g.* isolation and light). Future studies on egg-spot development should include other haplochromine species.

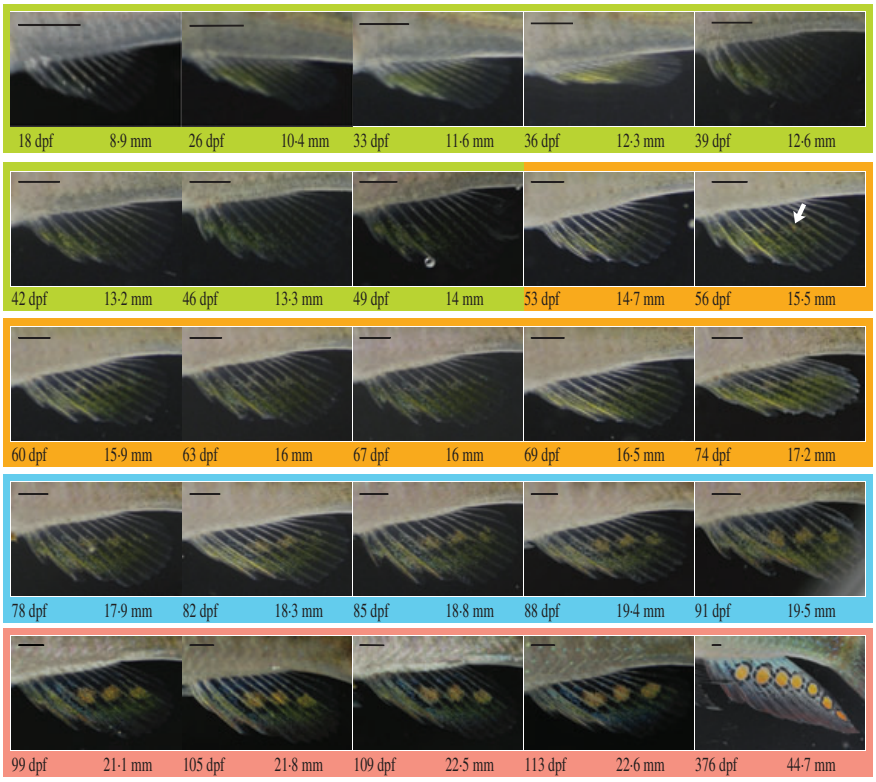

FIG. 2. Pictures series showing the anal fin of an individual male of *Astatotilapia burtoni*. The age (in days post fertilization, dpf) and the standard length, L_S, at every measurement are given below each photograph; scale bars represent 1 mm. Colours indicate the four stages of egg-spot development. ⇨, orange pigments.

We would like to thank E. Santos, M. Matschiner, the Assistant Editor C. Brown, the Associated Editor I. Ahnesjö and two anonymous referees for valuable comments, M. Muschick for discussion and help with taking photographs, A. Indermaur for discussion and fish maintenance and V. Mislin for the construction of photo-cuvettes. Experiments have been performed under permission from the cantonal veterinary office Basel (licence no. 2317). Financial support was provided from the European Research Council (ERC), Starting Grant 'INTERGENADAPT' and the University of Basel.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Movie S1. Egg-spot development in *Astatotilapia burtoni*.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

Endler, J. A. (1983). Natural and sexual selection on color patterns in poeciliid fishes. *Environmental Biology of Fishes* **9,** 173–190.

Fryer, G. & Iles, T. (1972). *The cichlid fishes of the Great Lakes of Africa: Their Biology and Evolution*. Edinburgh: Oliver & Boyd.

Goldschmidt, T. (1991). Egg mimics in haplochromine cichlids (Pisces, Perciformes) from Lake Victoria. *Ethology* **88,** 177–190. doi: 10.1111/j.1439-0310.1991.tb00273.x

Greenwood, P. (1979). Towards a phyletic classification of the 'genus' Haplochromis (Pisces, Cichlidae) and related taxa. *Bulletin of the British Museum (Natural History). Zoology* **35,** 265–322.

Günther, A. C. L. G. (1894). Descriptions of the reptiles and fishes collected by Mr. E. Coode-Hore on Lake Tanganyika. *Proceedings of the Zoological Society of London* **1893** (Part 4), 628–632.

Kodric-Brown, A. (1985). Female preference and sexual selection for male coloration in the guppy (*Poecilia reticulata*). *Behavioral Ecology and Sociobiology* **17,** 199–205.

Rohlf, F. J. (2006). *tpsDIG*, Version 2.10. Stony Brook, NY: Department of Ecology and Evolution, State University of New York.

Salzburger, W. (2009). The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. *Molecular Ecology* **18,** 169–185.

Salzburger, W., Mack, T., Verheyen, E. & Meyer, A. (2005). Out of Tanganyika: genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. *BMC Evolutionary Biology* **5,** 17.

Salzburger, W., Braasch, I. & Meyer, A. (2007). Adaptive sequence evolution in a color gene involved in the formation of the characteristic egg-dummies of male haplochromine cichlid fishes. *BMC Biology* **5,** 51.

Salzburger, W., Renn, S. C., Steinke, D., Braasch, I., Hofmann, H. A. & Meyer, A. (2008). Annotation of expressed sequence tags for the East African cichlid fish *Astatotilapia burtoni* and evolutionary analyses of cichlid ORFs. *BMC Genomics* **9,** 96.

Seehausen, O. (2006). African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society B* **273,** 1987–1998.

Wickler, W. (1962*a*). Egg-dummies as natural releasers in mouth-breeding cichlids. *Nature* **194,** 1092–1093.

Wickler, W. (1962*b*). Zur Stammesgeschichte funktionell korrelierter Organ- und Verhaltensmerkmale: Ei-Attrappen und Maulbrüten bei afrikanischen Cichliden. *Zeitschrift für Tierpsychologie* **19,** 129–164.

Chapter 2

Genetics of sexual development – an evolutionary playground for fish

Corina Heule, Walter Salzburger and Astrid Böhne

CH reviewed the literature and drafted the chapters 1 – 3 in the manuscript. Chapter 4 was written by CH and AB together. AB reviewed the literature, drafted chapter 5 and prepared the figures. WS received the invitation of this review from *Genetics* editorial board and helped drafting the manuscript.

Published in *Genetics* as an 'Invited Review'

Genetics of Sexual Development: An Evolutionary Playground for Fish

Corina Heule, Walter Salzburger, and Astrid Böhne[1]
Zoological Institute, University of Basel, 4051 Basel, Switzerland

ABSTRACT Teleost fishes are the most species-rich clade of vertebrates and feature an overwhelming diversity of sex-determining mechanisms, classically grouped into environmental and genetic systems. Here, we review the recent findings in the field of sex determination in fish. In the past few years, several new master regulators of sex determination and other factors involved in sexual development have been discovered in teleosts. These data point toward a greater genetic plasticity in generating the male and female sex than previously appreciated and implicate novel gene pathways in the initial regulation of the sexual fate. Overall, it seems that sex determination in fish does not resort to a single genetic cascade but is rather regulated along a continuum of environmental and heritable factors.

IN contrast to mammals and birds, cold-blooded vertebrates, and among them teleost fishes in particular, show a variety of strategies for sexual reproduction (Figure 1), ranging from unisexuality (all-female species) to hermaphroditism (sequential, serial, and simultaneous, including outcrossing and selfing species) to gonochorism (two separate sexes at all life stages). The underlying phenotypes are regulated by a variety of sex determination (SD) mechanisms that have classically been divided into two main categories: genetic sex determination (GSD) and environmental sex determination (ESD) (Figure 2).

Environmental factors impacting sex determination in fish are water pH, oxygen concentration, growth rate, density, social state, and, most commonly, temperature (for a detailed review on ESD see, *e.g.*, Baroiller *et al.* 2009b and Stelkens and Wedekind 2010). As indicated in Figure 2, GSD systems in fish compose a variety of different mechanisms and have been reviewed in detail elsewhere (*e.g.*, Devlin and Nagahama 2002; Volff *et al.* 2007).

The GSD systems that have received the most scientific attention so far are those involving sex chromosomes, which either may be distinguishable cytologically (heteromorphic) or appear identical (homomorphic). In both cases, one sex is heterogametic (possessing two different sex chromosomes and hence producing two types of gametes) and the other one homogametic (a genotype with two copies of the same sex chromosome, producing only one type of gamete). A male-heterogametic system is called an XX-XY system, and female-heterogametic systems are denoted as ZZ-ZW. Both types of heterogamety exist in teleosts and are even found side by side in closely related species [*e.g.*, tilapias (Cnaani *et al.* 2008), ricefishes (Takehana *et al.* 2008), or sticklebacks (Ross *et al.* 2009)]; for more details on the phylogenetic distribution of GSD mechanisms in teleost fish, see Mank *et al.* (2006). Note that sex chromosomes in fish are mostly homomorphic and not differentiated (Ohno 1974), which is in contrast to the degenerated Y and W chromosomes in mammals (Graves 2006) and birds (Takagi and Sasaki 1974), respectively. This is one possible explanation for the viable combination of different sex chromosomal systems within a single species or population of fish (Parnell and Streelman 2013) and could be a mechanistic reason why sex chromosome turnovers occur easily and frequently in this group (Mank and Avise 2009). Additionally, fish can have more complex sex chromosomal systems involving more than one chromosome pair (see Figure 2). Even within a single fish species, more than two sex chromosomes may occur at the same time, or more than two types of sex chromosomes may co-exist in the same species (Schultheis *et al.* 2006; Cioffi *et al.* 2013), which can sometimes be due to chromosome fusions (Kitano and Peichel 2012).

Detailed insights on the gene level for GSD/sex chromosomal systems are currently available for only a limited number of fish species, and all but one of these cases involve

doi: 10.1534/genetics.114.161158
[1]Corresponding author: Zoological Institute, University of Basel, Vesalgasse 1, 4051 Basel, Switzerland. E-mail: astrid.boehne@unibas.ch

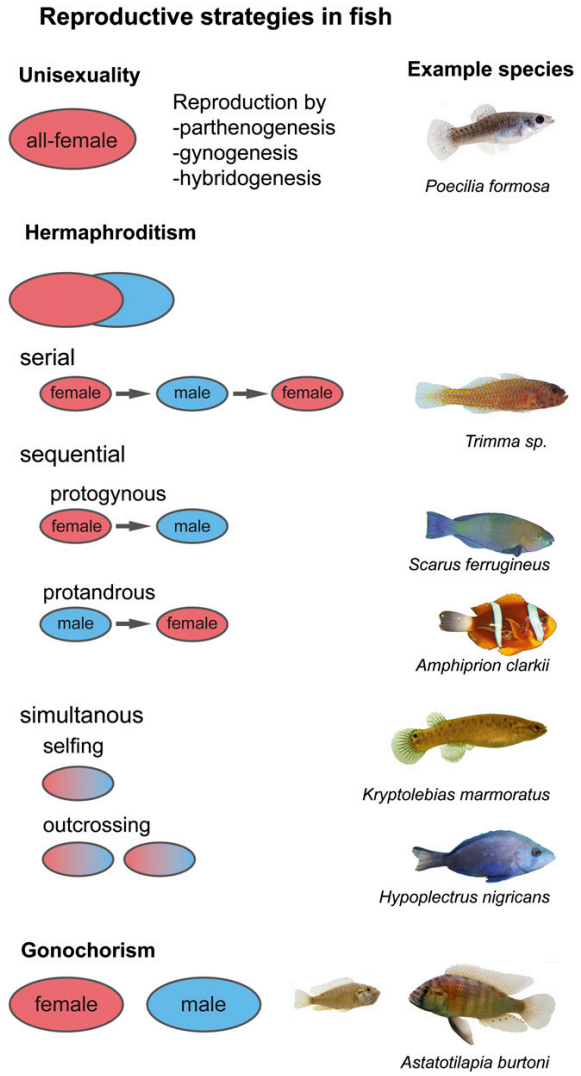

Figure 1 Reproductive strategies in fish. Fish can be grouped according to their reproductive strategy into unisexuals, hermaphrodites, and gonochorists. Further subdivisions of these three categories are shown with pictures of species exemplifying the strategies. Fish images: *Amphiprion clarkii* courtesy of Sara Mae Stieb; *Hypoplectrus nigricans* courtesy of Oscar Puebla; *Scarus ferrugineus* courtesy of Moritz Muschick; *Astatotilapia burtoni* courtesy of Anya Theis; *Poecilia formosa* and *Kryptolebias marmoratus* courtesy of Manfred Schartl; *Trimma* sp. courtesy of Rick Winterbottom [serial hermaphroditism has been described in several species of the genus *Trimma* (Kuwamura and Nakashima 1998; Sakurai *et al.* 2009; and references therein)].

a rather simple genetic system with male heterogamety and one major sex determiner (see below). The only exception is the widely used model species zebrafish (*Danio rerio*), which has a polyfactorial SD system implicating four different chromosomes (chromosomes 3, 4, 5, and 16) (Bradley *et al.* 2011; Anderson *et al.* 2012) and also environmental cues (Shang *et al.* 2006).

In this review, we focus on newly described genetic sex-determining systems and possible mechanisms allowing their emergence in fishes, which are the most successful group of vertebrates with ~30,000 species.

Genetic Basis of Sex Determination

Classic view on sexual development

The conventional definition of GSD states that the sex of an individual is fixed upon fertilization by inherited genetic factors. The establishment of the corresponding sexual phenotype is subsequently achieved through sexual development. The process of sexual development has classically been divided into sex determination and sex differentiation. Determination is understood as the "master" switch (initial inherited factor) that causes the first steps of the SD cascade, which then activate further downstream genes of sexual differentiation, which themselves regulate steroid hormone production, eventually leading to one functional gonad type with the corresponding sexual phenotype (Figure 3, top). In this scenario, the master switch functions as a presence–absence signal, in which the presence initiates the cascade to actively produce one sex, whereas the absence causes the other sexual differentiation program, which has long been considered as the default sex. However, this view has recently been challenged in, for example, mammals, where female development also needs active gene regulation and feedback loops (Munger and Capel 2012).

In the classic view, the SD cascade is considered to have evolved stepwise, in a retrograde fashion, with the downstream genetic network evolving first (and remaining conserved over time and between species), and new genes subsequently and independently added to the top of the hierarchy in different species/lineages (bottom-up theory) (Wilkins 1995, 2005) (Figure 3, top). That major changes occur frequently at the top of the cascade are substantiated by the observation that even closely related fish species (*e.g.*, of the same genus/family) do not share the same master SD gene (*e.g.*, the ricefishes, genus *Oryzias*) (Kondo *et al.* 2004).

However, the hypothesis of the existence of a conserved genetic cascade of SD common to all vertebrates governed by different lineage or species-specific genes ["masters change, slaves remain (Graham *et al.* 2003)] is largely based on sequence conservation in the downstream genes and conserved expression patterns for a handful of genes such as the testis-specific *dmrt1* (Smith *et al.* 1999).

Sex as a threshold phenotype based on a sex-determining network

More and more data reveal that, especially in teleost fishes, sex seems to be a rather plastic phenotype with natural and inducible sex reversal even after the genetic fixation of one sex at fertilization (Paul-Prasanth *et al.* 2013). These results thus question the scenario of a stepwise, hierarchical cascade with a fixed outcome. Instead, it has been suggested that sex should be interpreted as a threshold trait with

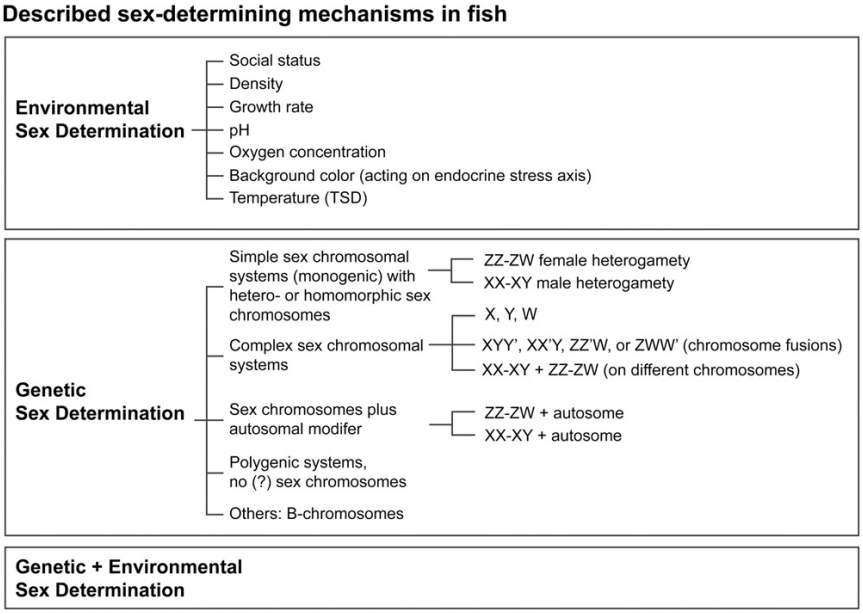

Figure 2 Sex-determining mechanisms in fish. Sex-determining systems in fish have been broadly classified into environmental and genetic sex determination. For both classes, the currently described subsystems are shown.

triggers (genetic, parental, or environmental) acting on a "parliament of interacting gene networks" (Crews and Bull 2009) and hormones (Uller and Helanterä 2011; Schwanz *et al.* 2013), resulting in a rather complex interplay between different signals (Figure 3, bottom). Importantly, gene networks are not necessarily hierarchical but rather modular (for review and further extension of this idea, see Salazar-Ciudad 2009 and Barabasi and Oltvai 2004). Assuming an interconnected network of genes and other factors impacting sex, the emergence of new sex determiners should be viewed from a developmental perspective (Figure 3, bottom), which allows for the emergence of major effect loci at different levels/nodes of the network.

In the next section, we present the current models for the evolution of such major-effect loci and our extension of these to a network concept of sex determination.

Suggested Mechanisms for the Emergence of New Master Sex Determination Genes

Based on the view of sex determination as a hierarchical cascade, the model of Schartl (2004) suggests that a new master SD gene can arise via a combination of gene duplication, sequence change (coding or regulatory), and (up-) recruitment [Figure 4, scenario 1 after (Schartl 2004)]. However, similar mechanisms of genetic changes, based mainly on duplication and subsequent mutation and especially without the loss of one of the duplicates, could also work with sex determination relying on a network of genes and must not necessarily occur only at the top of a cascade (Figure 4, scenario 2). These two ideas are rather mechanistic. More generally, the observation that similar/related genes, which are implicated in sexual development in several species, are recruited as master regulators (mainly dm-domain and HMG-domain transcription factors) led Marshall Graves and Peichel (2010) to put forward the "limited options" hypothesis. They suggested that a pool of genes/chromosomes "good at doing the job" might exist and be preferentially reused repeatedly in the vertebrate lineage (although in different combinations) rather than that "any" gene could be recruited as a master SD gene. As indicated in Figure 2, teleost fish are a particularly attractive system to study these models for the evolution of SD networks and their master genes.

Currently, detailed genetic information on male SD genes is available for six teleost species that we will discuss in detail in the next paragraph.

New master sex-determining genes in teleosts: dmY and gsdf in the Oryzias genus:

The best-studied SD gene in fish is the second vertebrate master SD gene that was described after the discovery of the mammalian *Sry* gene, namely *dmrt1by/dmy, DM-domain gene on the Y chromosome*, which acts as the master sex determiner in two species of the genus *Oryzias* [the medaka, *Oryzias latipes*, and the Malabar ricefish, *O. curvinotus* (Matsuda *et al.* 2002; Nanda *et al.* 2002)]. *Dmy* is not present in any other fish species studied so far, not even in the closely related *Oryzias* species (Kondo *et al.* 2003). *Dmy* arose in the common ancestor of *O. latipes* and its sister species *O. curvinotus* and *O. luzonensis* between 10 ± 2 and 18 ± 2 million years ago through a segmental duplication of a small autosomal region containing its precursor gene *dmrt1* [a gene with a well-described function in testis in vertebrates (Matson and Zarkower 2012) and some neighboring genes (Kondo *et al.* 2004) (see Figure 5A for illustration)]. This was

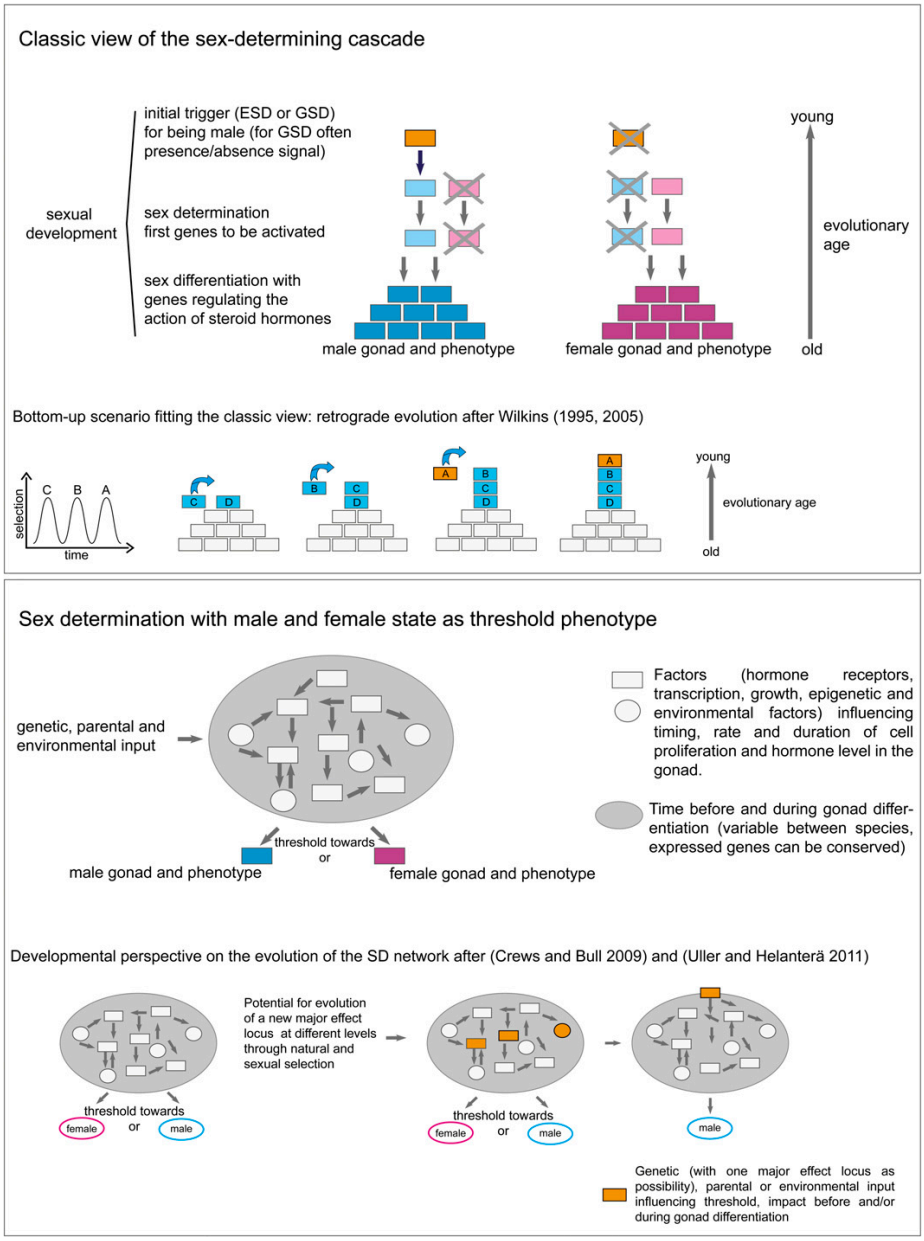

Figure 3 Two views on the sex-determining cascade. Classic view on the sex-determining cascade: The prevailing view on sexual development is the one of an initial trigger (environmental or genetic, mostly a presence/absence signal) initiating the sex determination cascade that activates sex differentiation, finally causing the establishment of one gonad type and the corresponding sexual phenotype. In this scenario, the presence of the initial trigger (here exemplified for a male master determiner) activates one cascade, whereas its absence leads to the other sex. Based on this cascade assumption, genes are added stepwise to the exisiting cascade. The cascade thus evolves in a retrograde fashion. The last (*i.e.*, the most downstream) step is the first one to be selected for. This evolution leads to the genetic network of sexual development divided into two steps: sex determination and differentiation (Wilkins 1995, 2005). Sex determination with male and female state as threshold phenotype: Based on a developmental perspective (Crews and Bull 2009; Uller and Helanterä 2011), sexual development is not split into determination and differentiation but rather controlled by a combination of different heritable and external factors influencing cell proliferation and hormone levels with a male and female threshold. Determining regulators evolve via canalization toward major-effect loci influencing the male/female threshold. Note that under this model without a strict hierarchical cascade, major-effect loci could emerge at all levels and are not imposed at the very top of the cascade.

followed by the insertion of the duplicated region on the proto-Y chromosome, directly creating a region of suppressed recombination. Outside of the relatively small sex-determining region (258 kb), in which *dmy* remained the only functional gene, recombination is not suppressed between the medaka X and Y chromosomes (Brunner *et al.* 2001; Kondo *et al.* 2006). *Dmy* is often cited as a typical example of an up-recruited SD gene, as is *DM-W* in the frog *Xenopus laevis* and possibly *dmrt1* in birds. It is derived from a gene that has already been implicated in testis formation and male development, *dmrt1*, agreeing with the idea of limited options for controlling sex determination. The mechanism of its evolution as a master SD gene likely followed the one described in Figure 4B.

In a closely related *Oryzias* species, *O. luzonensis*, which has lost *dmY*, Myosho *et al.* (2012) identified a derived version of the *gonadal soma derived growth factor* (*gsdf*) as the male master sex determiner. The gene has a Y- and X-chromosome-specific allele, called $gsdf^Y$ and $gsdf^X$. Gene expression experiments revealed an overexpression of $gsdf^Y$ during the time period of sex determination, whereas expression levels of both alleles were similar later in development.

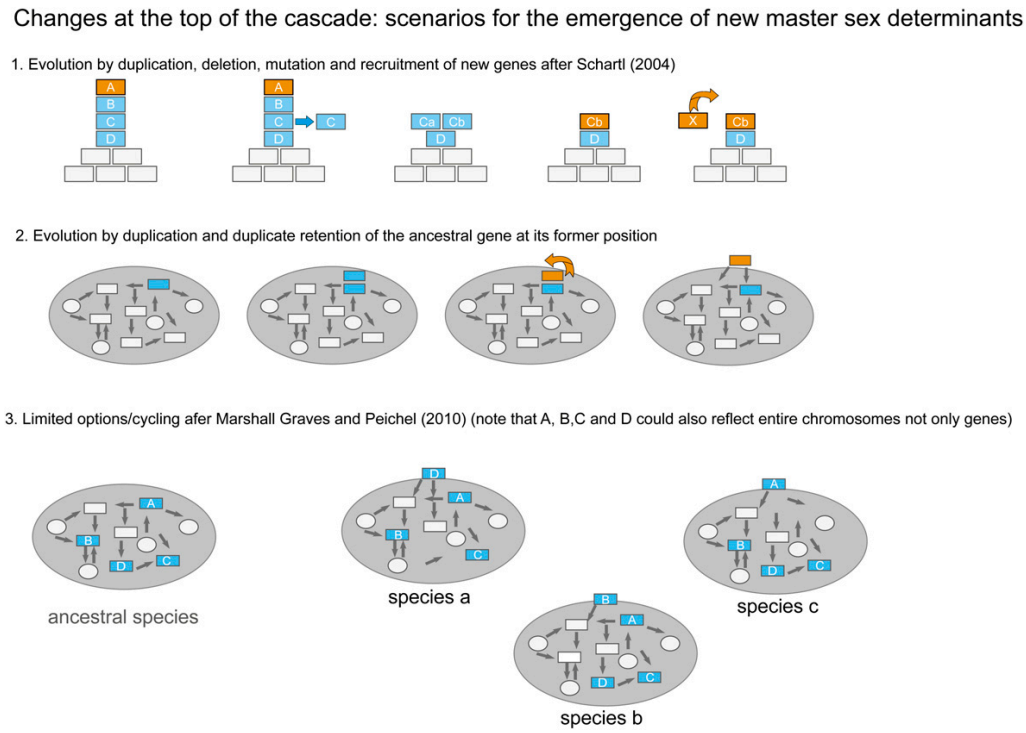

Figure 4 Models for the emergence of new genetic master sex determiners. Three mechanisms mainly influencing the evolution of master SD genes are shown. (1) Based on the classic cascade view, new genes formed by gene duplication or mutation in existing genes can be up-recruited to the top of the cascade (Schartl 2004). (2) Based on the view of SD as a genetic network, gene duplication and/or mutation of a member of the network could create potential material for the evolution of a new master gene without the loss of the ancestral gene. (3) The limited options theory after Marshall Graves and Peichel (2010) proposes that a pool of genes or entire chromosomes are reused in different species to become master determiners. Note that, as shown here, this model does not rely on a hierarchical cascade view of SD.

gsdf^{Y} and *gsdf*^{X} vary by 12 nucleotide substitutions, but all are silent. The male expression is thus likely caused by a mutation in a *cis*-regulatory region of *gsdf*^{Y}. Myosho *et al.* (2012) speculate that this mutation could involve a change in a putative binding site for steroidogenic factor 1 (Sf1), causing a loss of the binding site in *gsdf*^{Y} but not in *gsdf*^{X}. This change could lead to the observed higher expression during an earlier stage of sexual development compared to its ancestral counterpart. Transgenic transformation with *gsdf*^{Y} constructs led to XX males in *O. luzonenis* (as well as in its sister species *O. latipes* with a different SD gene), confirming a master sex-determining role for this gene. In *O. luzonensis*, *gsdf*^{Y} thus seems to have functionally replaced *dmY*. Note that the *O. luzonensis* genome harbors a pseudogenized copy of *dmrt1* called *Oludmrt1p*, in addition to the autosomal *dmrt1* gene (*Oludmrt1*). However, this pseudogene is located on a chromosome that is not syntenic to the Y chromosome of *O. latipes*. It is therefore likely that *O. luzonensis* has lost *dmY* from its genome and that *Oludmrt1p* is an independent pseudogene copy of *dmrt1* (Kondo *et al.* 2004; Tanaka *et al.* 2007).

In contrast to *dmY*, *gsdf* is not a transcription factor but a member of the TGF-β superfamily and specific to fish [it is present in all available teleost genomes and in the sarcopterygian *Latimeria* but not in tetrapods (Forconi *et al.* 2013)]. In *O. latipes* males, *gsdf* is transcribed at 6 days post fertilization in the primodial gonad and colocalized with *dmy* in the somatic cells (Sertoli cells) of the adult male gonad (Shibata *et al.* 2010). This suggests that *gsdf* acts rather early [indeed, it is expressed earlier than *dmrt1*] (Kobayashi *et al.* 2004)], which possibly facilitated its recruitment as a master SD gene. Functional data for *gsdf* are still limited, but it has been implicated with the proliferation of primordial germ cells (PGCs) and spermatogonia in the rainbow trout (Sawatari *et al.* 2007) and it possibly acts during female-to-male sex change, promoting proliferation of spermatogonia and spermatogenesis in the wrasse *Halichoeres trimaculatus* (Horiguchi *et al.* 2013). Furthermore, *gsdf* turns out to be a good candidate for the sex-determining locus in a member of another fish genus, the sablefish *Anoplopoma fimbria* (Rondeau *et al.* 2013), suggesting a central role for this gene in SD in fish and calling for further studies. The model of Uller and Helanterä (2011) (see Figure 3) suggests that factors important in the network of sex determination (and possibly driving the threshold toward one sex) are those influencing timing, rate, and duration of cell proliferation. With the information available for *gsdf*, this gene might be one such factor, and this could explain

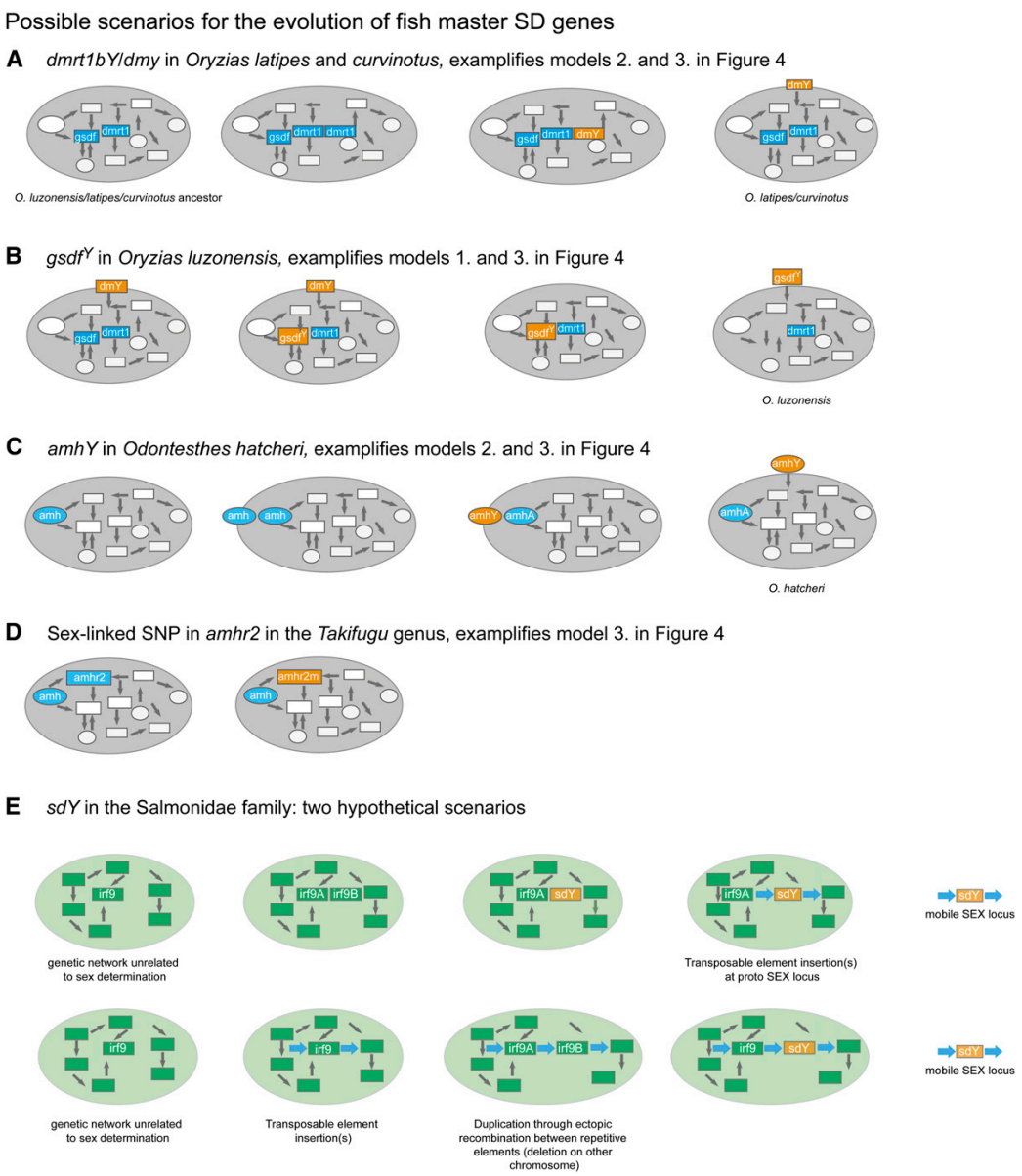

Figure 5 Possible scenarios for the evolution of fish SD genes. The figure illustrates possible evolutionary paths for the currently known master SD genes in teleost fish: (A) *dmrt1bY/dmy*, (B) *gsdf*^{Y}, (C) *amhY*, (D) *amhr2*, and (E) *sdY*. Horizontal arrows indicate gene duplications; color change of boxes to orange indicates acquisition of master-determining function via mutation (coding or regulatory). Green boxes in E indicate that the genetic network that *irf9* belongs to was not related to sex determination before the emergence of *sdY*.

why a gene that is not a transcription factor has been recruited as a master SD gene and could, at least in fish, represent support for the limited options hypothesis.

amhY and amhr2: Two other newly described sex determiners, *amhy* and *amhr2*, also belong to the TGF-β superfamily. *amhy* is a gene duplicate of *amh* (*anti-Müllerian hormone*) found on the Y chromosome of the Patagonian pejerrey, *Odontesthes hatcheri*. This gene has a 557-bp long insertion in the third intron and 92.2% sequence identity on the amino acid level compared to the autosomal copy (*amha* for *amh autosomal*) (Hattori *et al.* 2012). Hattori *et al.* (2012) showed that *amhy* is expressed during gonadal formation starting at 6 days after fertilization whereas *amha* is expressed later (12 weeks after fertilization). *amhy* transcripts are located in somatic cells (likely Sertoli cells) surrounding germ cells in the developing male gonad, a typical expression pattern for SD genes. Knockdown of *amhy* in XY

embryos led to an up-regulation of female factors and, consequently, to the development of ovaries (Hattori *et al.* 2012).

The *amhr2* (*amh receptor 2*) gene has been identified as possible third nontranscription factor master SD gene in the tiger pufferfish, *Takifugu rubripes* (Kamiya *et al.* 2012). Here, a single nucleotide polymorphism (SNP) in the kinase domain of *amhr2* has been associated with sex, with males being heterozygous and females homozygous. This sex-specific SNP, which changes an amino acid (His384 to Asp384), is also associated with sex in two other *Takifugu* species (*T. pardalis* and *T. poecilonotus*), which diverged around 5 million years ago from *T. rubripes*. The authors suggest that sex in the tiger pufferfish is determined by a combination of the two *amhr2* alleles, with the Y allele being dominant.

AMH is a secreted glycoprotein of the TGF-β superfamily best studied in tetrapods, where it is produced by the Sertoli cells and causes regression of Müllerian ducts during male sexual development (Josso *et al.* 2001; Rey *et al.* 2003; Johnson *et al.* 2008). AMH signaling functions through its primary receptor, AMHR2, which, upon AMH-binding, induces the assembly of a receptor complex eventually, activating expression of target genes (Josso *et al.* 2001) (see Figure 6 for details).

Although fish possess *amh* and *amhr2*, they lack Müllerian ducts. In fish, the AMH-signaling pathway has been implicated in PGC proliferation and spermatogenesis, similar to the function proposed for Gsdf (Morinaga *et al.* 2007). In general, AMH signaling in fish could decrease the number of germ cells (Morinaga *et al.* 2007; Lee *et al.* 2009), which in some species, such as the medaka or zebrafish, are crucial for gonadal fate determination (Kurokawa *et al.* 2007; Siegfried and Nüsslein-Volhard 2008) with more germ cells leading to a female gonad. However, this seems not to be a general mechanism in fish (Fujimoto *et al.* 2010; Goto *et al.* 2012). Another proposed mechanism of action for AMH/AMHR2 in fish is through the activation of the aromatase enzyme that synthesizes estrogens from androgens as has been described in mammals (Di Clemente *et al.* 1992). Higher estrogen levels result in ovary differentiation in fish (Devlin and Nagahama 2002). A suppression of the estrogen production by AMH/AMHR2 signaling on the aromatase could thus decrease estrogen levels and so promote testis formation.

amhy, like *dmY*, exemplifies the mechanism of gene duplication for the formation of new master SD genes. Based on the limited data available, members of the TGF-β superfamily (*gsdf*, *amh*, and *amhr2*) could be part of the limited options for master SD genes, at least in fish. However, the options for the evolution of new SD genes in fish could be less limited as the next example suggests.

sdY in the salmonid family: The fourth newly discovered sex regulator in fish is *sdY* (*sexually dimorphic on the Y chromosome*) in the rainbow trout, *Oncorhynchus myskiss* (Yano *et al.* 2012). This gene is the first example of a SD gene not related to a gene that was already implicated with sexual development and, hence, could be an example for functional *de novo* evolution (Yano *et al.* 2012). *sdY* shows sequence similarity to *irf9* (*interferon regulatory factor 9*), a transcription factor that acts in the immune system (Yano *et al.* 2012). The presence of *sdY* was confirmed in 15 salmonid species, and male sex linkage was found for 13 of these (Yano *et al.* 2013). Interestingly, the Y chromosomes in the investigated salmonid species are not syntenic, and hence the SD locus is located on different chromosomes. Before the discovery of *sdY*, the idea of a (likely small) conserved but mobile SD locus had been put forward for the salmonid family as a possible explanation for the fact that different chromosomes showed sex linkage in different species (Davidson *et al.* 2009) and even within species (Moghadam *et al.* 2007; Küttner *et al.* 2011; Eisbrenner *et al.* 2013).

To conclude this section, support for the limited-options theory comes certainly from *dmy* [duplication of *dmrt1*, a gene involved in development of testis in vertebrates and retention of *dmrt1* at its ancestral position (Kondo *et al.* 2006)], and possibly also *amhy* [duplication of *amh*, a gene that is responsible for the degeneration of the Müllerian ducts in male mammals (Klattig and Englert 2007) and retention of its ancestor *amha*], $gsdf^{Y}$ [derived from *gsdf*, a gene involved in the proliferation and differentiation of germ cells in fish (Gautier *et al.* 2011a)], and *amhr2* [a receptor involved in the AMH pathway in vertebrates (Klattig and Englert 2007)]. By contrast, *irf9*, the next most closely related gene to the salmonid SD gene *sdY*, has not been implicated in SD before and so could have been added on top of a conserved genetic network at a later stage.

Further Flexibility in the Network: Variation in Downstream Gene Usage

Until recently, the steps following the initiation of sex determination were thought to follow a rather well-orchestrated and conserved cascade of genes (some of them candidates for the limited options) controlling sexual development mainly via the action of steroid hormones (Lange *et al.* 2002; Nakamura 2010; Angelopoulou *et al.* 2012; Morohashi *et al.* 2013). Among the most prominent examples of apparently conserved genes of sexual development are *dmrt1* (Matson and Zarkower 2012) and *sox9* (Morrish and Sinclair 2002; Kobayashi *et al.* 2005) for testis formation and *wnt4* (Smith *et al.* 2008) and the aromatase *cyp19a1* (Valenzuela *et al.* 2003; Diotel *et al.* 2010) for ovary development. However, recent data from various fish species suggest that, in addition to the flexibility at the top of the cascade, even these key elements are not as conserved as previously assumed. This is evidenced by varying and species-specific expression patterns between sexes and throughout development (see, *e.g.*, Vizziano *et al.* 2007; Ijiri *et al.* 2008; Hale *et al.* 2011; Herpin *et al.* 2013; and Table 1). For example, *sox9* and the aromatase *cyp19a1* show expression patterns in East African cichlid fishes that are not consistent with conserved testis and ovary functions, respectively (Böhne *et al.* 2013).

The case of the aromatase *cyp19a1* is particularly striking. As already mentioned, this enzyme converts androgens

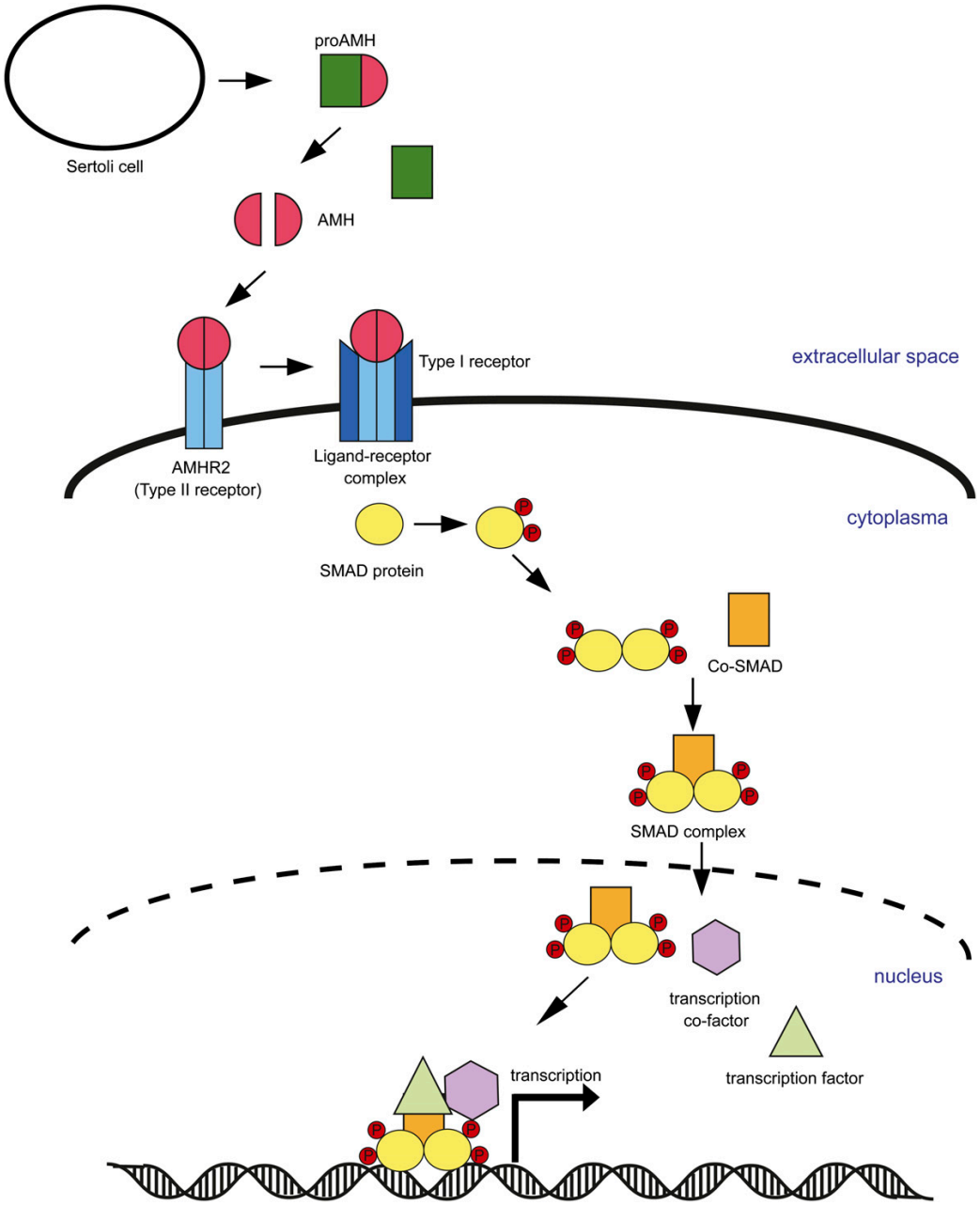

Figure 6 AMH/AMHR2-signaling pathway. Sertoli cells secrete a precursor form of AMH. After cleavage, an AMH dimer binds to AMHR2, which activates a type I receptor (currently not characterized; mechanism derived from comparisons with other receptors of the TGF-β superfamily). The ligand-receptor complex phosphorylates SMAD proteins, which then form a complex by incorporating Co-SMADs. This complex translocates into the nucleus and regulates gene expression together with transcription factors and cofactors. Figure modified after Fan *et al.* (2012) and Kikuchi and Hamaguchi (2013) and references therein.

into estrogens and was thought to control the female and male pathway by its presence or absence in expression, respectively (Guiguen *et al.* 2010). After the teleost-specific genome duplication, teleost fish retained two copies of this gene (*cyp19a1A* and *cyp19a1B*), which are active in ovary and brain tissue, respectively. Some derived East African cichlid species express the otherwise ovary-specific gene *cyp19a1A* in testis (Böhne *et al.* 2013). This finding asks for a critical revision of the hormonal control of sexual differentiation.

Table 1 Variation in gene expression profiles of sexual development genes in major teleost model species

Gene	Nile tilapia (*O. niloticus*)	Medaka (*O. latipes*)	Patagonian pejerrey (*O. hatcheri*)	Rainbow trout (*O. mykiss*)	Zebrafish (*D. rerio*)	Burton's haplo (*A. burtoni*)
amh	♂	♀/♂	♂ (first *amhy*, then *amha*)	♂	♂	♂
cyp19a1A	♀	♀	♀	♀	♀	♀
cyp19a1B	♀	♀	NA	♂	♂/♀	♂
dmrt1	♂	♂	♀/♂	♂	♂	♂
gsdf	♂	♂	NA	♂	♂/♀	♀
sf-1	first ♀, then ♂	♀	♂/♀	♂	♂	♂
sox9A	♂	♀	NA	♂	♂	♂
sox9B	♂	♂	NA	♂	♀	♀

Data: *Oreochromis niloticus* (Kobayashi *et al.* 2008; Ijiri *et al.* 2008; Baroiller *et al.* 2009a; Poonlaphdecha *et al.* 2011, 2013; Tao *et al.* 2013; http://www.uniprot.org/uniprot/E9RGV7), *Oryzias latipes* (Suzuki *et al.* 2004; Klüver *et al.* 2005; Nakamoto *et al.* 2007; Shibata *et al.* 2010; Siegfried 2010; Okubo *et al.* 2011; Nakamura *et al.* 2012), *Odontesthes hatcheri* (Hattori *et al.* 2013), *Oncorhynchus mykiss* (Baron *et al.* 2005; Sawatari *et al.* 2007; Gohin *et al.* 2011; Vizziano-Cantonnet *et al.* 2011), *D. rerio* (Chiang *et al.* 2001; Guo *et al.* 2005; Von Hofsten and Olsson 2005; Kallivretaki *et al.* 2007; Siegfried 2010; Gautier *et al.* 2011b; Nakamura *et al.* 2012), *Astatotilapia burtoni* (Böhne *et al.* 2013; A. Böhne and W. Salzburger, unpublished results). ♂, male specific or overexpressed; ♀, female specific or overexpressed; NA, not tested.

Interlinking GSD and ESD

Interestingly, steroid hormones, or the regulation of their production, are a putative link between ESD and GSD (Nakamura 2010; Angelopoulou *et al.* 2012), with *cyp19a1* being a prime candidate (Valenzuela *et al.* 2013). Navarro-Martín *et al.* (2011) have recently investigated the methylation levels of the aromatase (*cyp19a1A*) promoter in the European sea bass, *Dicentrarchus labrax*, a species with a polygenic SD system influenced by temperature (Vandeputte *et al.* 2007). They found that the *cyp19a1A* promoter in males features twice as many methylated sites compared to females, leading to an under-expression of this gene in males. The methylation rate is positively correlated with water temperature, and temperature-masculinized fish (*i.e.*, sex-reversed genetic females) have a higher methylation rate and lower levels of *cyp19a1A* expression.

In some species thought to have GSD, temperature has been shown to impact sex ratios (*e.g.*, the Nile tilapia) (Baroiller *et al.* 2009a). Furthermore, some species previously thought to have ESD were later found to have genetic mechanisms of sex determination (Ospina-Álvarez and Piferrer 2008). This has led to a controversial discussion about the occurrence of true or obligate temperature-dependent sex determination (TSD) in fish. A critical evaluation of the available data on the prevalence of TSD in fish (Ospina-Álvarez and Piferrer 2008) suggested two prerequisites to a true TSD system: (1) there must be the absence of sex chromosomes and (2) the temperature shifts influencing sex ratios must occur during decisive developmental steps. Based on these assumptions, TSD in fish appears to be much less widespread than previously thought with many observed cases involving (extreme) temperatures that lie outside the naturally occurring range (Ospina-Álvarez and Piferrer 2008). However, the two assumptions seem rather strict, given the fact that sex in fish is often not stable but can change even after juvenile stages. Baroiller *et al.* (2009b) suggested, in a less strict way, the classification of TSD fish species into three categories: (1) species with strong GSD and an absence of, or weak sensitivity to, temperature; (2) species that are highly sensitive to temperature or other environmental factors; and (3) species that are sensitive to several environmental cues. This definition seems to better reflect the empirical data and offers another part of the framework toward a new understanding of the mechanisms driving sex determination systems and their evolution.

Conclusion

Considering the recent findings of a substantial flexibility in expression and timing in the downstream network of sexual differentiation, the diversity of master SD genes found in fish and the co-influence of genetic and environmental factors, we think that it is time to redefine our understanding of the SD cascade. An attempt into this direction has recently been made by proposing that one should not strictly distinguish between sex determination and sex differentiation (Uller and Helanterä 2011). Instead, the entire process should be seen as a continuum in which all factors (genetic, maternal, and environmental) work together in a nonhierarchical network to ultimately form a male or female phenotype. This process would be canalized toward the male or female fate above a certain sex-specific threshold of genetic and environmental interactions influencing cell proliferation and hormonal levels (Figure 3, top). An extreme canalization of this process can result in the evolution of one or more major-effect locus/loci, such as the master SD genes described earlier. However, this does not have to happen, as reflected by the polygenic SD system in the zebrafish.

Interestingly, this view also offers a developmental explanation for how factors that, at the first glance, seemed unlikely master regulators, such as the members of the TGF-β family *gsdf*, *amh*, and *amhr2*, could become important in sex determination. The evolution of the TGF-β family members as master SD genes, although not acting as transcription factors, might simply reflect their previous involvement in the regulation of timing, rate, and duration of cell proliferation (in the developing gonad) and interaction with steroidogenesis, as described for their ancestral genes (Teixeira *et al.* 1999; Yamamoto *et al.* 2002; Sawatari *et al.* 2007; Shiraishi *et al.* 2008). For example, *amha* and *amhy* are spatially colocalized but not temporally (Hattori *et al.* 2012), providing

one example of how a change in timing of expression could lead to the up-recruitment of a seemingly downstream factor.

To understand the evolution of sex-determining networks—and possibly there is not only one cascade of SD but rather several network modules that are slightly rewired to ever-new combinations—we have to understand the different steps of sexual development and their interconnections (see Table 1 for a first approach to understanding shifts in gene expression in fish). In this context, the network modules could also be considered as limited options; however, we are far away from understanding where the limits really are, especially in fish.

We suggest a critical rethinking of the current models for the evolution of sex-determining genes with a stronger focus on the interactions further downstream in the network. All currently proposed models aim to explain switches in the master control genes without considering the dynamics of the subsequent networks. Hence, much effort has been put toward the identification of new "masters," especially in fish. Why SD systems are so extremely divergent and why turnovers of sex-determining systems are so seemingly frequent in fish remains to be answered. We propose regarding the SD pathway as a whole and not considering only its master switches as an isolated phenomenon when investigating these systems.

Acknowledgments

The authors thank Catherine Peichel for discussion and valuable comments on the manuscript. WS was supported by the University of Basel, the Swiss National Science Foundation (SNF) and the European Research Council (ERC), AB was supported by the Forschungsfonds of the University of Basel and the VolkswagenStiftung.

Literature Cited

Anderson, J. L., A. Rodríguez Marí, I. Braasch, A. Amores, P. Hohenlohe *et al.*, 2012 Multiple sex-associated regions and a putative sex chromosome in zebrafish revealed by RAD mapping and population genomics. PLoS ONE 7: e40701.

Angelopoulou, R., G. Lavranos, and P. Manolakou, 2012 Sex determination strategies in 2012: Towards a common regulatory model? Reprod. Biol. Endocrinol. 10: 13.

Barabasi, A.-L., and Z. N. Oltvai, 2004 Network biology: understanding the cell's functional organization. Nat. Rev. Genet. 5: 101–113.

Baroiller, J. F., H. D'Cotta, E. Bezault, S. Wessels, and G. Hoerstgen-Schwark, 2009a Tilapia sex determination: where temperature and genetics meet. Comp. Biochem. Physiol. A Mol. Integr. Physiol. 153: 30–38.

Baroiller, J. F., H. D'Cotta, and E. Saillant, 2009b Environmental effects on fish sex determination and differentiation. Sex Dev. 3: 118–135.

Baron, D., R. Houlgatte, A. Fostier, and Y. Guiguen, 2005 Large-scale temporal gene expression profiling during gonadal differentiation and early gametogenesis in rainbow trout. Biol. Reprod. 73: 959–966.

Böhne, A., C. Heule, N. Boileau, and W. Salzburger, 2013 Expression and sequence evolution of aromatase *cyp19a1* and other sexual development genes in East African cichlid fishes. Mol. Biol. Evol. 30: 2268–2285.

Bradley, K. M., J. P. Breyer, D. B. Melville, K. W. Broman, E. W. Knapik *et al.*, 2011 An SNP-based linkage map for zebrafish reveals sex determination loci. G3 (Bethesda) 1: 3–9.

Brunner, B., U. Hornung, Z. Shan, I. Nanda, M. Kondo *et al.*, 2001 Genomic organization and expression of the doublesex-related gene cluster in vertebrates and detection of putative regulatory regions for DMRT1. Genomics 77: 8–17.

Chiang, E. F.-L., Y.-L. Yan, Y. Guiguen, J. Postlethwait, and B.-c. Chung, 2001 Two Cyp19 (P450 aromatase) genes on duplicated zebrafish chromosomes are expressed in ovary or brain. Mol. Biol. Evol. 18: 542–550.

Cioffi, M. B., T. Liehr, V. Trifonov, W. F. Molina, and L. A. Bertollo, 2013 Independent sex chromosome evolution in lower vertebrates: a molecular cytogenetic overview in the Erythrinidae fish family. Cytogenet. Genome Res. 141: 186–194.

Cnaani, A., B. Y. Lee, N. Zilberman, C. Ozouf-Costaz, G. Hulata *et al.*, 2008 Genetics of sex determination in tilapiine species. Sex Dev. 2: 43–54.

Crews, D., and J. J. Bull, 2009 Mode and tempo in environmental sex determination in vertebrates. Semin. Cell Dev. Biol. 20: 251–255.

Davidson, W. S., T. K. Huang, K. Fujiki, K. R. von Schalburg, and B. F. Koop, 2009 The sex determining loci and sex chromosomes in the family Salmonidae. Sex Dev. 3: 78–87.

Devlin, R. H., and Y. Nagahama, 2002 Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. Aquaculture 208: 191–364.

di Clemente, N., S. Ghaffari, R. B. Pepinsky, C. Pieau, N. Josso *et al.*, 1992 A quantitative and interspecific test for biological activity of anti-Mullerian hormone: the fetal ovary aromatase assay. Development 114: 721–727.

Diotel, N., Y. L. Page, K. Mouriec, S.-K. Tong, E. Pellegrini *et al.*, 2010 Aromatase in the brain of teleost fish: expression, regulation and putative functions. Front. Neuroendocrinol. 31: 172–192.

Eisbrenner, W. D., N. Botwright, M. Cook, E. A. Davidson, S. Dominik *et al.*, 2013 Evidence for multiple sex-determining loci in Tasmanian Atlantic salmon (*Salmo salar*) Heredity (Edinb.). DOI: 10.1038/hdy.2013.55.

Fan, Y.-S., Y.-J. Hu, and W.-X. Yang, 2012 TGF-β superfamily: how does it regulate testis development. Mol. Biol. Rep. 39: 4727–4741.

Forconi, M., A. Canapa, M. Barucca, M. A. Biscotti, T. Capriglione *et al.*, 2013 Characterization of sex determination and sex differentiation genes in *Latimeria*. PLoS ONE 8: e56006.

Fujimoto, T., T. Nishimura, R. Goto-Kazeto, Y. Kawakami, E. Yamaha *et al.*, 2010 Sexual dimorphism of gonadal structure and gene expression in germ cell-deficient loach, a teleost fish. Proc. Natl. Acad. Sci. USA 107: 17211–17216.

Gautier, A., F. Le Gac, and J.-J. Lareyre, 2011a The *gsdf* gene locus harbors evolutionary conserved and clustered genes preferentially expressed in fish previtellogenic oocytes. Gene 472: 7–17.

Gautier, A., F. Sohm, J.-S. Joly, F. Le Gac, and J.-J. Lareyre, 2011b The proximal promoter region of the zebrafish *gsdf* gene is sufficient to mimic the spatio-temporal expression pattern of the endogenous gene in Sertoli and granulosa cells. Biol. Reprod. 85: 1240–1251.

Gohin, M., P. Bodinier, A. Fostier, F. Chesnel, and J. Bobe, 2011 Aromatase is expressed and active in the rainbow trout oocyte during final oocyte maturation. Mol. Reprod. Dev. 78: 510–518.

Goto, R., T. Saito, T. Takeda, T. Fujimoto, M. Takagi *et al.*, 2012 Germ cells are not the primary factor for sexual fate determination in goldfish. Dev. Biol. 370: 98–109.

Graham, P., J. K. M. Penn, and P. Schedl, 2003 Masters change, slaves remain. Bioessays 25: 1–4.

Graves, J. A. M., 2006 Sex chromosome specialization and degeneration in mammals. Cell 124: 901–914.
Guiguen, Y., A. Fostier, F. Piferrer, and C.-F. Chang, 2010 Ovarian aromatase and estrogens: a pivotal role for gonadal sex differentiation and sex change in fish. Gen. Comp. Endocrinol. 165: 352–366.
Guo, Y., H. Cheng, X. Huang, S. Gao, H. Yu *et al.*, 2005 Gene structure, multiple alternative splicing, and expression in gonads of zebrafish *Dmrt1*. Biochem. Biophys. Res. Commun. 330: 950–957.
Hale, M., P. Xu, J. Scardina, P. Wheeler, G. Thorgaard *et al.*, 2011 Differential gene expression in male and female rainbow trout embryos prior to the onset of gross morphological differentiation of the gonads. BMC Genomics 12: 404.
Hattori, R. S., Y. Murai, M. Oura, S. Masuda, S. K. Majhi *et al.*, 2012 A Y-linked anti-Müllerian hormone duplication takes over a critical role in sex determination. Proc. Natl. Acad. Sci. USA 109: 2955–2959.
Hattori, R. S., C. A. Strussmann, J. I. Fernandino, and G. M. Somoza, 2013 Genotypic sex determination in teleosts: insights from the testis-determining *amhy* gene. Gen. Comp. Endocrinol. 192: 55–59.
Herpin, A., M. C. Adolfi, B. Nicol, M. Hinzmann, C. Schmidt *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: 2328–2346.
Horiguchi, R., R. Nozu, T. Hirai, Y. Kobayashi, Y. Nagahama *et al.*, 2013 Characterization of gonadal soma-derived factor expression during sex change in the protogynous wrasse, *Halichoeres trimaculatus*. Dev. Dyn. 242: 388–399.
Ijiri, S., H. Kaneko, T. Kobayashi, D.-S. Wang, F. Sakai *et al.*, 2008 Sexual dimorphic expression of genes in gonads during early differentiation of a teleost fish, the Nile tilapia *Oreochromis niloticus*. Biol. Reprod. 78: 333–341.
Johnson, P. A., T. R. Kent, M. E. Urick, and J. R. Giles, 2008 Expression and regulation of anti-Mullerian hormone in an oviparous species, the hen. Biol. Reprod. 78: 13–19.
Josso, N., N. di Clemente, and L. Gouédard, 2001 Anti-Müllerian hormone and its receptors. Mol. Cell. Endocrinol. 179: 25–32.
Kallivretaki, E., R. I. L. Eggen, S. C. F. Neuhauss, O. Kah, and H. Segner, 2007 The zebrafish, brain-specific, aromatase *cyp19a2* is neither expressed nor distributed in a sexually dimorphic manner during sexual differentiation. Dev. Dyn. 236: 3155–3166.
Kamiya, T., W. Kai, S. Tasumi, A. Oka, T. Matsunaga *et al.*, 2012 A trans-species missense SNP in *amhr2* is associated with sex determination in the tiger pufferfish, *Takifugu rubripes* (Fugu). PLoS Genet. 8: e1002798.
Kikuchi, K., and S. Hamaguchi, 2013 Novel sex-determining genes in fish and sex chromosome evolution. Dev. Dyn. 242: 339–353.
Kitano, J., and C. Peichel, 2012 Turnover of sex chromosomes and speciation in fishes. Environ. Biol. Fishes 94: 549–558.
Klattig, J., and C. Englert, 2007 The Müllerian duct: recent insights into its development and regression. Sex Dev. 1: 271–278.
Klüver, N., M. Kondo, A. Herpin, H. Mitani, and M. Schartl, 2005 Divergent expression patterns of *Sox9* duplicates in teleosts indicate a lineage specific subfunctionalization. Dev. Genes Evol. 215: 297–305.
Kobayashi, A., H. A. O. Chang, M.-C. Chaboissier, A. Schedl, and R. R. Behringer, 2005 Sox9 in testis determination. Ann. N. Y. Acad. Sci. 1061: 9–17.
Kobayashi, T., M. Matsuda, H. Kajiura-Kobayashi, A. Suzuki, N. Saito *et al.*, 2004 Two DM domain genes, *DMY* and *DMRT1*, involved in testicular differentiation and development in the medaka, *Oryzias latipes*. Dev. Dyn. 231: 518–526.
Kobayashi, T., H. Kajiura-Kobayashi, G. Guan, and Y. Nagahama, 2008 Sexual dimorphic expression of *DMRT1* and *Sox9a* during gonadal differentiation and hormone-induced sex reversal in the teleost fish Nile tilapia (*Oreochromis niloticus*). Dev. Dyn. 237: 297–306.
Kondo, M., I. Nanda, U. Hornung, S. Asakawa, N. Shimizu *et al.*, 2003 Absence of the candidate male sex-determining gene *dmrt1b(Y)* of medaka from other fish species. Curr. Biol. 13: 416–420.
Kondo, M., I. Nanda, U. Hornung, M. Schmid, and M. Schartl, 2004 Evolutionary origin of the medaka Y chromosome. Curr. Biol. 14: 1664–1669.
Kondo, M., U. Hornung, I. Nanda, S. Imai, T. Sasaki *et al.*, 2006 Genomic organization of the sex-determining and adjacent regions of the sex chromosomes of medaka. Genome Res. 16: 815–826.
Kurokawa, H., D. Saito, S. Nakamura, Y. Katoh-Fukui, K. Ohta *et al.*, 2007 Germ cells are essential for sexual dimorphism in the medaka gonad. Proc. Natl. Acad. Sci. USA 104: 16958–16963.
Küttner, E., J. Nilsson, S. Skúlason, S. Gunnarsson, M. M. Ferguson *et al.*, 2011 Sex chromosome polymorphisms in Arctic charr and their evolutionary origins. Genome 54: 852–861.
Kuwamura, T., and Y. Nakashima, 1998 New aspects of sex change among reef fishes: recent studies in Japan. Environ. Biol. Fishes 52: 125–135.
Lange, I. G., A. Hartel, and H. H. D. Meyer, 2002 Evolution of oestrogen functions in vertebrates. J. Steroid Biochem. Mol. Biol. 83: 219–226.
Lee, K. H., A. Yamaguchi, H. Rashid, K. Kadomura, S. Yasumoto *et al.*, 2009 Germ cell degeneration in high-temperature treated pufferfish, *Takifugu rubripes*. Sex Dev. 3: 225–232.
Mank, J. E., and J. C. Avise, 2009 Evolutionary diversity and turnover of sex determination in teleost fishes. Sex Dev. 3: 60–67.
Mank, J. E., D. E. L. Promislow, and J. C. Avise, 2006 Evolution of alternative sex-determining mechanisms in teleost fishes. Biol. J. Linn. Soc. Lond. 87: 83–93.
Marshall Graves, J. A., and C. L. Peichel, 2010 Are homologies in vertebrate sex determination due to shared ancestry or to limited options? Genome Biol. 11: 205.
Matson, C. K., and D. Zarkower, 2012 Sex and the singular DM domain: insights into sexual regulation, evolution and plasticity. Nat. Rev. Genet. 13: 163–174.
Matsuda, M., Y. Nagahama, A. Shinomiya, T. Sato, C. Matsuda *et al.*, 2002 *DMY* is a Y-specific DM-domain gene required for male development in the medaka fish. Nature 417: 559–563.
Moghadam, H. K., M. M. Ferguson, and R. G. Danzmann, 2007 Linkage variation at the sex-determining locus within Fraser strain Arctic charr *Salvelinus alpinus*. J. Fish Biol. 71: 294–301.
Morinaga, C., D. Saito, S. Nakamura, T. Sasaki, S. Asakawa *et al.*, 2007 The hotei mutation of medaka in the anti-Müllerian hormone receptor causes the dysregulation of germ cell and sexual development. Proc. Natl. Acad. Sci. USA 104: 9691–9696.
Morohashi, K., T. Baba, and M. Tanaka, 2013 Steroid hormones and the development of reproductive organs. Sex Dev. 7: 61–79.
Morrish, B., and A. Sinclair, 2002 Vertebrate sex determination: many means to an end. Reproduction 124: 447–457.
Munger, S. C., and B. Capel, 2012 Sex and the circuitry: progress toward a systems-level understanding of vertebrate sex determination. Wiley Interdiscip. Rev. Syst. Biol. Med. 4: 401–412.
Myosho, T., H. Otake, H. Masuyama, M. Matsuda, Y. Kuroki *et al.*, 2012 Tracing the emergence of a novel sex-determining gene in medaka, *Oryzias luzonensis*. Genetics 191: 163–170.
Nakamoto, M., D.-S. Wang, A. Suzuki, M. Matsuda, Y. Nagahama *et al.*, 2007 Dax1 suppresses P450arom expression in medaka ovarian follicles. Mol. Reprod. Dev. 74: 1239–1246.

Nakamura, M., 2010 The mechanism of sex determination in vertebrates: Are sex steroids the key-factor? J. Exp. Zool. A Ecol. Genet. Physiol. 313: 381–398.
Nakamura, S., I. Watakabe, T. Nishimura, A. Toyoda, Y. Taniguchi *et al.*, 2012 Analysis of medaka *sox9* orthologue reveals a conserved role in germ cell maintenance. PLoS ONE 7: e29982.
Nanda, I., M. Kondo, U. Hornung, S. Asakawa, C. Winkler *et al.*, 2002 A duplicated copy of *DMRT1* in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*. Proc. Natl. Acad. Sci. USA 99: 11778–11783.
Navarro-Martín, L., J. Viñas, L. Ribas, N. Díaz, A. Gutiérrez *et al.*, 2011 DNA methylation of the gonadal aromatase c*yp19a* promoter is involved in temperature-dependent sex ratio shifts in the European sea bass. PLoS Genet. 7: e1002447.
Ohno, S., 1974 *Animal Cytogenetics*. Borntraeger, Berlin.
Okubo, K., A. Takeuchi, R. Chaube, B. Paul-Prasanth, S. Kanda *et al.*, 2011 Sex differences in aromatase gene expression in the medaka brain. J. Neuroendocrinol. 23: 412–423.
Ospina-Álvarez, N., and F. Piferrer, 2008 Temperature-dependent sex determination in fish revisited: prevalence, a single sex ratio response pattern, and possible effects of climate change. PLoS ONE 3: e2837.
Parnell, N. F., and J. T. Streelman, 2013 Genetic interactions controlling sex and color establish the potential for sexual conflict in Lake Malawi cichlid fishes. Heredity (Edinb.) 110: 239–246.
Paul-Prasanth, B., R. K. Bhandari, T. Kobayashi, R. Horiguchi, Y. Kobayashi *et al.*, 2013 Estrogen oversees the maintenance of the female genetic program in terminally differentiated gonochorists. Sci. Rep. **3:** 2862.
Poonlaphdecha, S., E. Pepey, S. H. Huang, M. Canonne, L. Soler *et al.*, 2011 Elevated *amh* gene expression in the brain of male tilapia (*Oreochromis niloticus*) during testis differentiation. Sex Dev. 5: 33–47.
Poonlaphdecha, S., E. Pepey, M. Canonne, H. de Verdal, J. F. Baroiller *et al.*, 2013 Temperature induced-masculinisation in the Nile tilapia causes rapid up-regulation of both *dmrt1* and *amh* expressions. Gen. Comp. Endocrinol.
Rey, R., C. Lukas-Croisier, C. Lasala, and P. Bedecarrás, 2003 AMH/MIS: what we know already about the gene, the protein and its regulation. Mol. Cell. Endocrinol. 211: 21–31.
Rondeau, E., A. Messmer, D. Sanderson, S. Jantzen, K. von Schalburg *et al.*, 2013 Genomics of sablefish (*Anoplopoma fimbria*): expressed genes, mitochondrial phylogeny, linkage map and identification of a putative sex gene. BMC Genomics 14: 452.
Ross, J. A., J. R. Urton, J. Boland, M. D. Shapiro, and C. L. Peichel, 2009 Turnover of sex chromosomes in the stickleback fishes (Gasterosteidae). PLoS Genet. 5: e1000391.
Sakurai, M., S. Nakakoji, H. Manabe, S.-i. Dewa, A. Shinomiya *et al.*, 2009 Bi-directional sex change and gonad structure in the gobiid fish *Trimma yanagitai*. Ichthyol. Res. 56: 82–86.
Salazar-Ciudad, I., 2009 Looking at the origin of phenotypic variation from pattern formation gene networks. J. Biosci. 34: 573–587.
Sawatari, E., S. Shikina, T. Takeuchi, and G. Yoshizaki, 2007 A novel transforming growth factor-β superfamily member expressed in gonadal somatic cells enhances primordial germ cell and spermatogonial proliferation in rainbow trout (*Oncorhynchus mykiss*). Dev. Biol. 301: 266–275.
Schartl, M., 2004 Sex chromosome evolution in non-mammalian vertebrates. Curr. Opin. Genet. Dev. 14: 634–641.
Schultheis, C., Q. Zhou, A. Froschauer, I. Nanda, Y. Selz *et al.*, 2006 Molecular analysis of the sex-determining region of the platyfish *Xiphophorus maculatus*. Zebrafish 3: 299–309.
Schwanz, L. E., T. Ezaz, B. Gruber, and A. Georges, 2013 Novel evolutionary pathways of sex-determining mechanisms. J. Evol. Biol. 26: 2544–2557.
Shang, E. H. H., R. M. K. Yu, and R. S. S. Wu, 2006 Hypoxia affects sex differentiation and development, leading to a male-dominated population in zebrafish (*Danio rerio*). Environ. Sci. Technol. 40: 3118–3122.
Shibata, Y., B. Paul-Prasanth, A. Suzuki, T. Usami, M. Nakamoto *et al.*, 2010 Expression of gonadal soma derived factor (GSDF) is spatially and temporally correlated with early testicular differentiation in medaka. Gene Expr. Patterns 10: 283–289.
Shiraishi, E., N. Yoshinaga, T. Miura, H. Yokoi, Y. Wakamatsu *et al.*, 2008 Müllerian inhibiting substance is required for germ cell proliferation during early gonadal differentiation in medaka (*Oryzias latipes*). Endocrinology 149: 1813–1819.
Siegfried, K. R., 2010 In search of determinants: gene expression during gonadal sex differentiation. J. Fish Biol. 76: 1879–1902.
Siegfried, K. R., and C. Nüsslein-Volhard, 2008 Germ line control of female sex determination in zebrafish. Dev. Biol. 324: 277–287.
Smith, C. A., P. J. McClive, P. S. Western, K. J. Reed, and A. H. Sinclair, 1999 Evolution: conservation of a sex-determining gene. Nature 402: 601–602.
Smith, C., C. Shoemaker, K. Roeszler, J. Queen, D. Crews *et al.*, 2008 Cloning and expression of *R-Spondin1* in different vertebrates suggests a conserved role in ovarian development. BMC Dev. Biol. 8: 72.
Stelkens, R. B., and C. Wedekind, 2010 Environmental sex reversal, Trojan sex genes, and sex ratio adjustment: conditions and population consequences. Mol. Ecol. 19: 627–646.
Suzuki, A., M. Tanaka, and N. Shibata, 2004 Expression of aromatase mRNA and effects of aromatase inhibitor during ovarian development in the medaka, *Oryzias latipes*. J. Exp. Zool. A Comp. Exp. Biol. 301A: 266–273.
Takagi, N., and M. Sasaki, 1974 A phylogenetic study of bird karyotypes. Chromosoma 46: 91–120.
Takehana, Y., S. Hamaguchi, and M. Sakaizumi, 2008 Different origins of ZZ/ZW sex chromosomes in closely related medaka fishes, *Oryzias javanicus* and *O. hubbsi*. Chromosome Res. 16: 801–811.
Tanaka, K., Y. Takehana, K. Naruse, S. Hamaguchi, and M. Sakaizumi, 2007 Evidence for different origins of sex chromosomes in closely related *Oryzias* fishes: substitution of the master sex-determining gene. Genetics 177: 2075–2081.
Tao, W., J. Yuan, L. Zhou, L. Sun, Y. Sun *et al.*, 2013 Characterization of gonadal transcriptomes from Nile tilapia (*Oreochromis niloticus*) reveals differentially expressed genes. PLoS ONE 8: e63604.
Teixeira, J., E. Fynn-Thompson, A. H. Payne, and P. K. Donahoe, 1999 Müllerian-inhibiting substance regulates androgen synthesis at the transcriptional level. Endocrinology 140: 4732–4738.
Uller, T., and H. Helanterä, 2011 From the origin of sex-determining factors to the evolution of sex-determining systems. Q. Rev. Biol. 86: 163–180.
Valenzuela, N., D. C. Adams, and F. J. Janzen, 2003 Pattern does not equal process: Exactly when is sex environmentally determined? Am. Nat. 161: 676–683.
Valenzuela, N., J. Neuwald, and R. Literman, 2013 Transcriptional evolution underlying vertebrate sexual development. Dev. Dyn. 242: 307–319.
Vandeputte, M., M. Dupont-Nivet, H. Chavanne, and B. Chatain, 2007 A polygenic hypothesis for sex determination in the European sea bass *Dicentrarchus labrax*. Genetics 176: 1049–1057.
Vizziano, D., G. Randuineau, D. Baron, C. Cauty, and Y. Guiguen, 2007 Characterization of early molecular sex differentiation in rainbow trout, *Oncorhynchus mykiss*. Dev. Dyn. 236: 2198–2206.
Vizziano-Cantonnet, D., I. Anglade, E. Pellegrini, M.-M. Gueguen, A. Fostier *et al.*, 2011 Sexual dimorphism in the brain aromatase expression and activity, and in the central expression of other steroidogenic enzymes during the period of sex differentiation

in monosex rainbow trout populations. Gen. Comp. Endocrinol. 170: 346–355.
Volff, J. N., I. Nanda, and M. Schartl, 2007 Governing sex determination in fish: regulatory putsches and ephemeral dictators. Sex Dev. 1: 85–99.
von Hofsten, J., and P.-E. Olsson, 2005 Zebrafish sex determination and differentiation: involvement of FTZ-F1 genes. Reprod. Biol. Endocrinol. 3: 63.
Wilkins, A., 1995 Moving up the hierarchy: A hypothesis on the evolution of a genetic sex determination pathway. Bioessays 17: 71–77.
Wilkins, A. S., 2005 Recasting developmental evolution in terms of genetic pathway and network evolution . . . and the implications for comparative biology. Brain Res. Bull. 66: 495–509.
Yamamoto, N., L. K. Christenson, J. M. McAllister, and J. F. Strauss, 2002 Growth differentiation factor-9 inhibits 3′5′-adenosine monophosphate-stimulated steroidogenesis in human granulosa and theca cells. J. Clin. Endocrinol. Metab. 87: 2849–2856.
Yano, A., R. Guyomard, B. Nicol, E. Jouanno, E. Quillet *et al.*, 2012 An immune-related gene evolved into the master sex-determining gene in rainbow trout, *Oncorhynchus mykiss*. Curr. Biol. 22: 1423–1428.
Yano, A., B. Nicol, E. Jouanno, E. Quillet, A. Fostier *et al.*, 2013 The sexually dimorphic on the Y-chromosome gene (*sdY*) is a conserved male-specific Y-chromosome sequence in many salmonids. Evol. Appl. 6: 486–496.

Communicating editor: A. S. Wilkins

Chapter 3

Expression and sequence evolution of aromatase *cyp19a1* and other sexual development genes in East African cichlid fishes

Astrid Böhne, **Corina Heule**, Nicolas Boileau and Walter Salzburger

AB, CH and WS designed the study and wrote the manuscript. AB performed sequence analysis. AB and CH performed qRT-PCR expression analysis. AB and NB sequenced the *cyp19a1A* locus.

Published in *Molecular Evolution and Biology*

Expression and Sequence Evolution of Aromatase *cyp19a1* and Other Sexual Development Genes in East African Cichlid Fishes

Astrid Böhne,*[1] Corina Heule,[1] Nicolas Boileau,[1] and Walter Salzburger[1]

[1]Zoological Institute, University of Basel, Basel, Switzerland

***Corresponding author:** E-mail: astrid.boehne@unibas.ch.

Associate editor: John Parsch

Sequences newly reported in this article have been deposited in GenBank (accession nos. KC684559–KC684586).

Abstract

Sex determination mechanisms are highly variable across teleost fishes and sexual development is often plastic. Nevertheless, downstream factors establishing the two sexes are presumably conserved. Here, we study sequence evolution and gene expression of core genes of sexual development in a prime model system in evolutionary biology, the East African cichlid fishes. Using the available five cichlid genomes, we test for signs of positive selection in 28 genes including duplicates from the teleost whole-genome duplication, and examine the expression of these candidate genes in three cichlid species. We then focus on a particularly striking case, the A- and B-copies of the aromatase *cyp19a1*, and detect different evolutionary trajectories: *cyp19a1A* evolved under strong positive selection, whereas *cyp19a1B* remained conserved at the protein level, yet is subject to regulatory changes at its transcription start sites. Importantly, we find shifts in gene expression in both copies. *Cyp19a1* is considered the most conserved ovary-factor in vertebrates, and in all teleosts investigated so far, *cyp19a1A* and *cyp19a1B* are expressed in ovaries and the brain, respectively. This is not the case in cichlids, where we find new expression patterns in two derived lineages: the A-copy gained a novel testis-function in the Ectodine lineage, whereas the B-copy is overexpressed in the testis of the speciest-richest cichlid group, the Haplochromini. This suggests that even key factors of sexual development, including the sex steroid pathway, are not conserved in fish, supporting the idea that flexibility in sexual determination and differentiation may be a driving force of speciation.

Key words: **sex determination, adaptive evolution, cichlid fishes, aromatase, developmental system drift, gene duplication.**

Article

Introduction

Sexual development in animals involves the steps of sex determination, the initial decision to develop either ovaries or testes, and sex differentiation, the subsequent developmental program establishing the male and female phenotype through the action of steroid hormones. Interestingly, the master regulatory triggers of sex determination are not conserved throughout the animal kingdom, and sex differentiation can be rather plastic (Cutting et al. 2013). This is particularly exemplified by teleost fishes, in which all sorts of natural hermaphroditism occur and sex changes can be induced through a variety of external factors (Godwin 2010, and references therein).

Contrary to the observed variation in the initial triggers of sex determination, the downstream genetic factors of sexual development are thought to be conserved between species because they are part of complex genetic networks with often pleiotropic effects. Their modifications presumably have deleterious effects, whereas changes "at the top of the hierarchies" are apparently better tolerated (Marín and Baker 1998). Previous studies have thus mainly focused on sex differentiation and the ubiquitous downstream genes (Valenzuela et al. 2013). Of particular interest here are the presumably conserved genetic and enzymatic network regulating the production of sex steroid hormones in vertebrates, where estrogens mediate ovarian development and androgens testis differentiation. For example, the action of *cyp19a1*, which aromatizes androgens into estrogens, is considered a general feature of ovarian development in fish, amphibians, reptiles, and mammals (for a review see Nakamura 2010), whereas *sox9* appears to be its major counterpart in testis development (Valenzuela et al. 2013).

Sex-determining systems in teleosts, the taxonomically largest group of vertebrates, changed frequently during evolution and vary even between closely related species (Mank et al. 2006). So far, five male master sex-determining genes (i.e., equivalents to the mammalian *sry*) have been identified: *DMY/dmrt1bY* in medaka (*Oryzias latipes*) and its sister species *O. curvinotus* (Matsuda et al. 2002, 2003; Nanda et al. 2002); $Gsdf^{Y}$, a member of the transforming growth factor-beta superfamily, in *O. luzonensis* (Myosho et al. 2012); *sdY* in the rainbow trout (Yano et al. 2012); *amhy*, (Y-chromosomal copy of the *anti-Müllerian hormone*) in the Patagonian pejerrey (*Odontesthes hatcheri*) (Hattori et al. 2012); and in three *Takifugu* species (*T. rubripes*, *T. pardalis*, and *T. poecilonotus*) an association between male sex and a (proto-)Y-chromosomal single nucleotide polymorphism in the *amh receptor type II* (*amhr2*) was found (Kamiya et al. 2012). Four of these genes, *DMY/dmrt1bY*, $Gsdf^{Y}$, *amhy* and *amhr2*, exemplify the scenario that genes already implicated downstream in a network (such as the genetic cascade of sex determination) are more readily uprecruited to the top of the hierarchy than

nonimplicated genes—a hypothesis originally proposed by Wilkins (1995). Sex determination and differentiation have also been studied in various other model fish species including zebrafish (Anderson et al. 2012), the poeciliid fishes guppy (Tripathi et al. 2009) and platyfish (Böhne et al. 2009), with a main focus on understanding and characterizing the genetic control over the establishment of the two sexes. Also, the downstream factors or particular genetic pathways of, for example, gonad development have been studied in several species (for review see Nakamura 2010; Angelopoulou et al. 2012). Another field of research that has received much attention is sex change in fish and, particularly, the identification of hormones and chemical components (including pollutants) that influence it (for review see Scholz and Klüver 2009). Thus, a considerable body of literature on sex determination and differentiation in fish is available, including data on many candidate genes.

Cichlid fishes of the East African Great Lakes Malawi (LM), Victoria (LV), and Tanganyika (LT) are a prime model system in evolutionary biology and provide an exceptional opportunity to study organismal diversification (Kocher 2004). Each of the Great Lakes harbors cichlid species assemblages counting hundreds of endemic species, which evolved within a few millions of years (LT; Salzburger et al. 2002), to less than 150,000 years (LV; Verheyen et al. 2003) only. This opens up a unique possibility for comparative studies at different evolutionary time scales, involving repeated patterns of diversification. As for the majority of fish species, the triggers of sex determination in cichlids are largely unknown. Yet, it becomes clear that also in this group various mechanisms exist, including genetic systems and environmental triggers such as water pH and temperature (Baroiller 2009; Reddon and Hurd 2013). The known genetic factors in cichlids include, for example, sex determination via B-chromosomes (small supernumerary/accessory chromosomes, Yoshida et al. 2011) and male and female heterogametic sex chromosome systems, with the possibility of both systems co-existing within a single species (Roberts et al. 2009). The best-studied cichlid species with respect to sexual development is the Nile tilapia (*Oreochromis niloticus*), a member of a more basal lineage, widely distributed in rivers and lakes of Africa. The Nile tilapia has an XX-XY sex-determining system, which can substantially be influenced by temperature (Baroiller 2009). For this species, expression profiles of key genes of sexual development are available (Ijiri et al. 2008), which is not the case for other cichlid species such as the radiating lineages in East Africa. The only exception to some extent is the LT cichlid *Astatotilapia burtoni*, in which the relation between hormones and behavior has been examined (Renn et al. 2012). To the best of our knowledge, sex determination mechanisms and sex differentiation have not been studied in a phylogenetically representative set of East African lake cichlids. To this end, it is particularly important to focus on the cichlid assemblage of LT, which is the oldest of the three lakes and, hence, home to the genetically, morphologically and ecologically most diverse cichlid species flock (Salzburger et al. 2002). LT's 250 mostly endemic species (Snoeks 2001) are currently classified into 12 to 16 lineages, so-called tribes (Muschick et al. 2012). The taxonomic situation is quite different in LM and LV, where the adaptive radiations consist of only one of these tribes, the Haplochromini (Salzburger et al. 2002).

Genes involved in sex determination (including its master regulators), differentiation, and reproduction often evolve under a positive selection regime (Sorhannus and Kosakovsky Pond 2006; Hasselmann et al. 2008; Morgan et al. 2010; Sobrinho and de Brito 2010). This has been explained by selection for reproductive isolation (which may ultimately lead to speciation, Vacquier and Swanson 2011), intersexual (genomic) conflict (Rice and Holland 1997), sperm competition (Pizzari and Birkhead 2002), and/or male-female coevolution (Swanson et al. 2003). This general trend seems to hold true also for cichlids, in which previous scans for positive selection uncovered three genes with putative roles in sexual development and reproduction: an aromatase gene (*cyp19a1A*, Gerrard and Meyer 2007), a gene with a gonad-specific expression profile (*SPP120*, Gerrard and Meyer 2007), and a meiotic maturation factor (*20-beta-hydroxysteroid dehydrogenase*, Baldo et al. 2011).

Here, we investigate 28 candidate genes in detail in cichlids. The genes were recovered from a thorough literature survey focusing on genes involved in sex determination and differentiation in vertebrates, and, especially in teleosts (see supplementary table S1 for details, Supplementary Material online). These candidates include gene copies that have emerged in the course of the fish-specific whole-genome duplication (FSGD; Meyer and Schartl 1999). To our knowledge, this has not been done before for genes involved in sexual development, although it seems crucial as gene duplications open the routes to partitioning gene functions and to neofunctionalization. We first investigate the candidate genes for sequence evolution using the available cichlid genomes (including species from all three lakes) and study their expression in ovary, testis and brain tissue in three species from LT (*A. burtoni*, *Ophthalmotilapia ventralis*, and *Neolamprologus pulcher*), which were chosen with respect to available genetic and genomic resources (Baldo et al. 2011 and BROAD; www.broadinstitute.org/models/tilapia, last accessed July 23, 2013). Sequence data analysis of key genes are then extended to a phylogenetically representative set of 32 species belonging to 13 different tribes and the expression analysis is extended to nine East African cichlid species, representing four tribes, including species from all three lakes.

Results

Sequence Evolution of Candidate Genes: Signs of Positive Selection in East African Cichlids

To test for signs of positive selection in our set of 28 candidate genes, we used the annotated sequences of the Nile tilapia genome as reference to recover full-length coding sequences from all available cichlid genomes (*A. burtoni* and *N. brichardi* from LT, *Metriaclima zebra* from LM, and *Pundamilia nyererei* from LV; BROAD institute, see supplementary table S2 for details, Supplementary Material online). Using tilapia as outgroup, we tested for signs of positive selection specific to the

East African cichlid radiations. We calculated mean gene wide Ka/Ks (nonsynonymous to synonymous substitution rates; see supplementary table S3 [Supplementary Material online] and Material and Methods for details) for all pairwise comparisons. The two genes with the highest gene-wide Ka/Ks values were *cyp19a1A*, as also detected by Gerrard and Meyer in an expressed sequence tag (EST) wide screen (Gerrard and Meyer 2007; discussed earlier), and *nanos1A*, which has not been studied in this context in cichlids so far. Note that a role for a *nanos* family member in germ cell differentiation has been proposed in tilapia (Kobayashi 2010).

One drawback with gene-wide Ka/Ks analyses is that they might miss signals of adaptive sequence evolution that are restricted to certain parts of a gene only; it is generally unlikely that all positions of a given gene experience positive selection at the same time, and many cases are known where positive selection is restricted to particular domains of a gene (Salzburger et al. 2007). To account for this, we analyzed our gene set using Selecton (Doron-Faigenboim et al. 2005; Stern et al. 2007) to test for site-wise selection, again using Nile tilapia as reference. We detected positively selected sites in the following genes: *cyp19a1A* (in 60 amino acid positions out of 511), *foxl2B* (9 of 268), *sox9A* (1 of 500), *sox9B* (7 of 484), *tdrd1* (96 of 1,164), and *wnt4A* (2 of 352). The results for *cyp19a1A* and *tdrd1* remained significant after comparison with the alternative hypotheses (i.e., no positive selection, see supplementary fig. S1 [Supplementary Material online] for details).

Sequence Evolution of *Cyp19a1* in Cichlids

As the initial analyses based on the five cichlid genomes pointed out new patterns of sequence evolution for *cyp19a1A*, whereas we could not detect a single site under positive selection in the B-copy (*cyp19a1B*; supplementary fig. S1, Supplementary Material online), we decided to further study these two genes in East African cichlids. Overall, *cyp19a1B* is more conserved among the five cichlid genomes: Cyp19a1A differs in 12.2% (average over all pairwise comparisons to tilapia, done with KaKs_calculator, Zhang et al. 2006) of all amino acid sites (3% synonymous changes and 9.2% nonsynonymous changes), whereas Cyp19a1B varies in 9.8% of all amino acid sites (4.8% synonymous changes and 5% nonsynonymous changes). Especially, nonsynonymous substitutions occur more often in Cyp19a1A compared to Cyp19a1B.

We then extended our analyses to a more representative set of East African cichlids. To this end, we used Ion Torrent next-generation amplicon sequencing (LifeTechnologies) to re-sequence a 3.6 kb genomic fragment of *cyp19a1A* in 28 cichlid species belonging to 12 different tribes (see Materials and Methods and supplementary table S4 [Supplementary Material online], GENBANK accession numbers KC684559–KC684586). The *cyp19a1A* coding sequences of all 28 newly sequenced species plus the sequences obtained from the cichlid genomes (*O. niloticus* as reference, *N. pulcher*, *A. burtoni*, and *P. nyererei*) were subjected to tests for positive selection as described earlier. The overall gene-wide Ka/Ks of 0.91 resulting from this analysis was similar to the value obtained with the limited data set. However, when checking in the comprehensive data set for site-wise selection with Selecton, we detected an additional 50 amino acid sites (adding up to a total of 108 out of 511 sites) under positive selection. All but three of the sites suggested to have evolved under positive selection in the initial analysis were confirmed by this second, more extensive analysis.

Placing the sites under a positive selection regime into the structural context of the well-studied aromatase protein (fig. 1) revealed that they do not cluster in a specific region of the coding sequence but are rather distributed over the entire protein and also occur in major functional domains. Additionally, the observed amino acid changes (marked with arrows above the strongest selected sites, fig. 1) are mostly not conservative. This suggests that *cyp19a1A* has experienced important changes in its functionality in East African cichlids.

To obtain a more detailed understanding of *cyp19a1A* sequence evolution in East African cichlids, we investigated positive selection in the phylogenetic context using HyPhy (Pond et al. 2005) and Codeml (Yang 2007) (fig. 2 and table 1).

These analyses uncovered several branches in the phylogeny with particularly strong adaptive sequence evolution. One of these branches represents the ancestor of the adaptive radiation of cichlids in the East African Great lakes (branch after split from the outgroup *O. niloticus*; light gray box, fig. 2). The signal of positive selection gets even stronger when the basal Tanganyikan tribes Bathybatini, Trematocarini, and Boulengerochromini are excluded (dark gray box, fig. 2). A strong signal of positive selection was also found in the lineage leading to the taxonomically largest cichlid tribe, the Haplochromini (dark green box, fig. 2, remember that all cichlids of LV and LM belong to this group). The so detected amino acid positions are consistent with the ones shown in figure 1 and fall into the two first categories of strongest positive selection of Selecton (fig. 2 and table 1).

Comparative Sequence Analysis of *Cyp19a1A* and *B*: Transcription Start Site Variation and Promoter Evolution

Using cichlid transcriptome databases (access over BROAD cichlid genome consortium, assembled RNA-seq data for the species with sequenced genomes), we investigated the transcription start sites (TSSs) of both *cyp19a1* gene copies. *Cyp19a1* transcripts were found in transcriptome databases derived from brain, gonad and mixed embryonic tissue. Interestingly, there are substantial differences between the two genes: In *cyp19a1A*, the TSSs of the three presumably functional transcripts (i.e., transcripts comprising the start codon marked with an orange box in supplementary fig. S2, Supplementary Material online) found in *O. niloticus*, *A. burtoni*, and *M. zebra* are 20 nucleotides apart of each other. Quite to the contrary, the TSSs of *cyp19a1B* are located more than 1,200 bp apart from each other in the transcripts found for *A. burtoni*, *P. nyererei*, and *M. zebra*. A closer inspection revealed two alternative splice forms of *cyp19a1B*:

Fig. 1. Placement of site-specific selection on the protein sequence of Cyp19a1A. Selection values shown were calculated on a multispecies alignment covering the LT cichlid assemblage using Selecton. Structural information on the Cyp19a1A protein was taken from (Graham-Lorence et al. 1995; Chiang et al. 2001; Castro et al. 2005). Occurring amino acid changes are indicated above the sites under strongest positive selection as detected with Selecton (category 1).

transcription of *cyp19a1B* in cichlids can start inside intron 1 or mark the beginning of the untranslated exon 1, in which case intron 1 is correctly spliced out (see supplementary fig. S2 [Supplementary Material online] for the intron location). For the two basal species *O. niloticus* and *N. brichardi*, we only detected the splice variant with the TSS inside exon 1.

Regulation of *Cyp19a1A* and *B* Expression in (Cichlid) Fish

The *cyp19a1A* and *B* promoter regions have previously been studied in several teleost species (Callard et al. 2001; Diotel et al. 2010) just as the promoter of the single copy *cyp19a1* gene in tetrapods (Hinshelwood et al. 2000; Nakagawa and Iwabuchi 2012). In a first step, to characterize general patterns in the molecular evolution of the *cyp19a1* upstream regions in teleosts, we compared the existing cichlid sequences with all other available teleost aromatase promoters (fig. 3 and table 2). Note that *cyp19a1B* is absent from the available *M. zebra* (LM) genome assembly version 1, so that we excluded this species in further analysis. Using Vista plots of nucleotide similarity (Mayor et al. 2000; Frazer et al. 2004), we identified three conserved regions in the upstream sequences of both gene copies marked green, yellow, and brown in figure 3. Here, we define the upstream region as 5′-sequence between the start codon of the *cyp19a1A* and *B* genes and their respective adjacent gene. For *cyp19a1A*, we analyzed the 3,394 bp until its neighboring gene *gliomedinA* (*gldnA*); for *cyp19a1B*, we analyzed 2,571 bp to *gldnB*. We inspected the three conserved regions with regard to common transcription factor binding sites focusing on factors with a known function/response role in brain/central nervous system (CNS) and gonads/germ cells, as these are the tissues in which the aromatase genes are described to be expressed. Although remarkable differences in the expression patterns of *cyp19a1A* and *B* exist between teleosts, and especially between cichlids (discussed later), the *cyp19a1* promoters still show conserved regions between all teleosts, underlining their general functional importance.

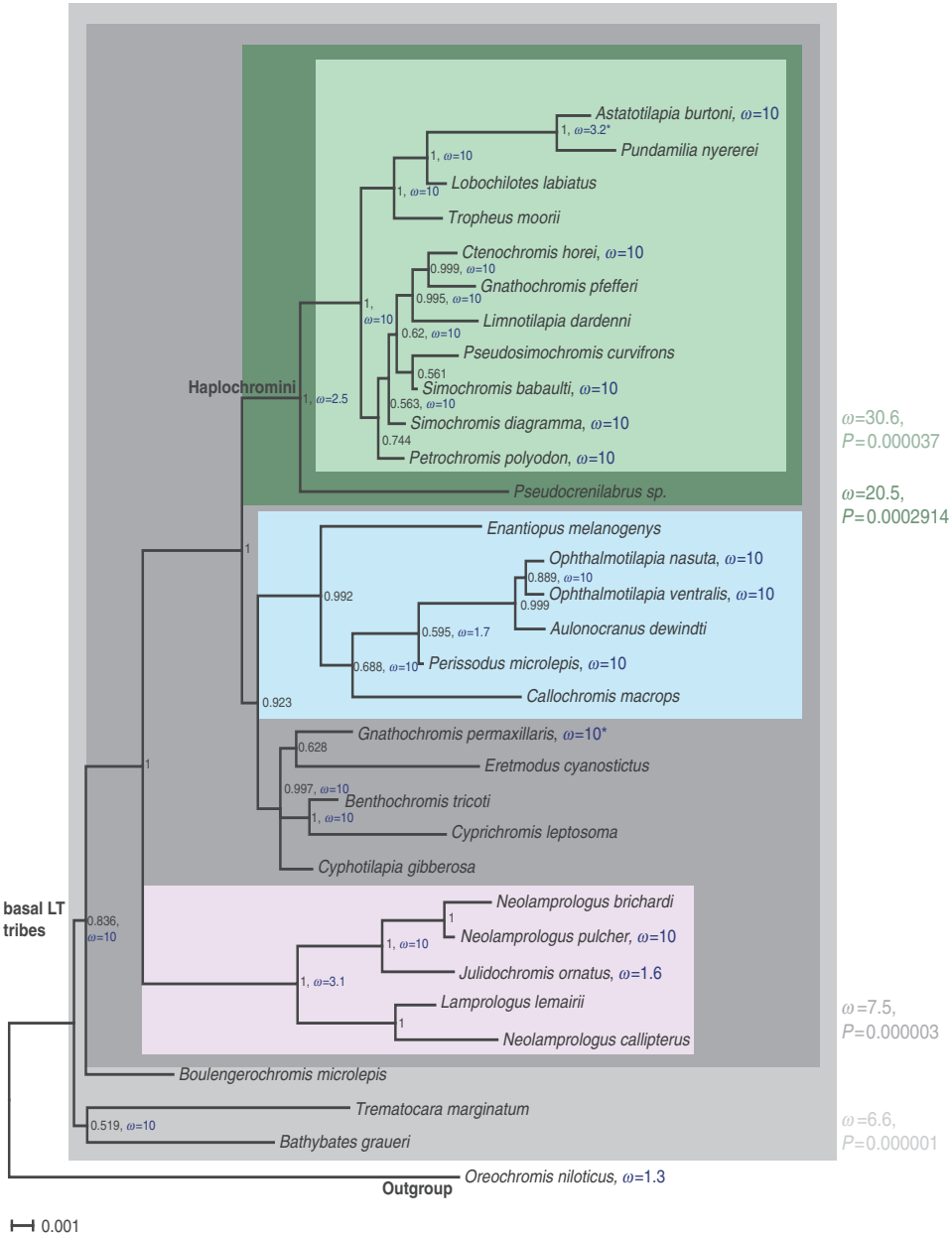

Fig. 2. Ka/Ks rates on a phylogenetic tree of aromatase *cyp19a1A* in cichlids. Phylogenetic tree of aromatase *cyp19a1A* of representatives of the major East African cichlid lineages showing the Ka/Ks rates reconstructed with HyPhy and branch-site selection as calculated with codeml. Nodes are labeled with Bayesian posterior probabilities as obtained with MrBayes. Colored boxes denote clades that were tested against all other branches with codeml (see Materials and Methods for details). Significant obtained Ka/Ks values (ω) are shown next to the boxes in the respective colors. ω values denoted in blue were calculated with Branch-site REL in HyPhy. Table 1 presents the sites that are under positive selection as detected by codeml in the different clades that show a significant ω with a probability (BEB) of more than 50%. The values obtained with Selecton are given in addition.

Putative transcription factor binding sites inside the conserved regions and shared between teleosts (underlined text in table 2) comprise factors belonging to the two investigated categories (brain/CNS and gonads/germ cells). None of these categories seems to dominate, however, neither in *cyp19a1A* nor in *cyp19a1B*.

Putative Transcription Factor Binding Sites in Teleost Conserved Upstream Regions of *Cyp19a1A* and *B*

Using MatInspector (Genomatix Software GmbH), we identified putative estrogen response elements (EREs) in the upstream regions of both *cyp19a1* duplicates, which is consistent with previous studies in fish (Diotel et al. 2010). In the

Table 1. Cyp19a1A Amino Acid Sites under Positive Selection in Different Phylogenetic Clades of East African Cichlids.

Cyp19a1A Amino Acid Position in *O. niloticus* Reference Protein	Category of Positive Selection on Selecton Scale[a]	BEB of Positive Selection in Clades with Significant ω[b] as Tested with Codeml			
		1[c] (Light Gray)	2 (Dark Gray)	3 (Dark Green)	4 (Light Green)
2	1	0.701	0.712	0.941	0.955
22	2		0.5	0.81	0.852
123	1				0.624
157	1	0.563	0.569		
169	1	0.646			
191	1	0.998	0.998	0.885	
205	1	0.978	0.98	0.535	0.628
257	1	0.858			
314	1			0.772	0.831
362	1	0.816	0.827	0.621	0.696
370	1	0.777	0.813		
461	1	0.859		0.659	0.721
471	1	0.979	0.98	0.944	0.956
507	1	0.628			

NOTE.—BEB, Bayes Empirical Bayes.

[a]1 and 2 represent the two categories of strongest positive selection as detected with Selecton corresponding to Ka/Ks values over 2.7 for category 1 and 1–2.7 for category 2 (fig. 1).

[b]ω = Ka/Ks.

[c]The phylogenetic categories (1, light gray; 2, dark gray; 3, dark green; 4, light green) correspond to the ones shown in figure 2.

first teleost-conserved block upstream of *cyp19a1A* (green box, fig. 3), we also detected putative Nr5a2 (Pezzi et al. 2004) and Fox-family binding sites (Wang et al. 2010) shared between teleosts. Here, we also detected partially conserved sites belonging to the category "brain/CNS transcription factors," such as neuronal factors of the Oct family including Oct-6 and the transcriptional repressor DREAM. Oct-6 is known to be expressed in brain and testis, to play a role in myelination and to act as an RNA polymerase II distal enhancer (Hofmann et al. 2010). DREAM plays a role in nociception in the brain and is broadly expressed in testis (Rivas et al. 2004).

Tilapia, platyfish, and medaka contain, in the first block, a potential site for Nr1f1, which is important for the development of the cerebellum, regulation of (lipid) metabolism, lymphocyte development, inflammatory responses, and possibly myogenesis or muscle function (Jetten 2009). Finally, we detected putative CREB (cAMP response element-binding protein)-binding sites in a subset of species. CREB can act as transcriptional activator or repressor and acts in neurogenesis, neuronal survival and plasticity (Barco and Marie 2011; Gruart et al. 2012). Note that CREB binding sites have been described in the *cyp19a1A* promoter before (Chang et al. 2005). Finally, we detected shared putative Sox-family binding sites in the second conserved block (Callard et al. 2001).

Interestingly, all three conserved blocks of the *cyp19a1B* teleost promoter contain putative homeobox domain transcription factor binding sites, including sites for members of the HOX family. Additionally, we detected possible binding sites for the Fox-gene-family in block one, which is also the case in the *cyp19a1A* promoter (blocks 1 and 2). In three teleost species (platyfish, medaka, and stickleback), we found conserved putative sites for Hmx3 (involved in specification of neuronal cell types, required in the hypothalamic-pituitary axis), RFX1 (transcriptional activator of MHC classII genes), and PCE-1 (an element usually found in promoter regions of photoreceptor genes). However, these sites were not detected in tilapia or pufferfishes.

A closer inspection of the cichlid sequences only (supplementary fig. S2, Supplementary Material online) revealed that these contain, in the conserved upstream regions of both gene copies, putative binding sites for SF-1, several Sox-family members and several androgen response elements in addition to the ERE and estrogen receptor (ER)-related sites. The promoter region of *cyp19a1A* in cichlids contains also putative binding sites for Nr5a2 and Dmrt1. Furthermore, we confirmed the existence of putative binding sites for Nkx6-1 in both *cyp19a1* promoters in cichlids, just as previously described for *cyp19a1B* in tilapia, zebrafish, and medaka (Chang et al. 2005; Diotel et al. 2010).

Expression Profiling of *Cyp19a1A* and *B* and Other Sexual Development Genes in Cichlids

In a first step, we tested by quantitative real-time polymerase chain reaction (qRT-PCR), the expression of our set of candidate genes in three different species from LT (*A. burtoni*, *O. ventralis*, and *N. pulcher*, for species choice see Materials and Methods) belonging to three different cichlid lineages. We obtained expression profiles in both gonad and brain tissue for 23 of these candidate genes (see Materials and Methods for details on our working procedure; the remaining five genes, *amh*, *arA*, *arB*, *dax1b*, and *tdrd1* could not successfully be amplified on cDNA from gonadal tissue). Brain tissue was included in these experiments, because several of the

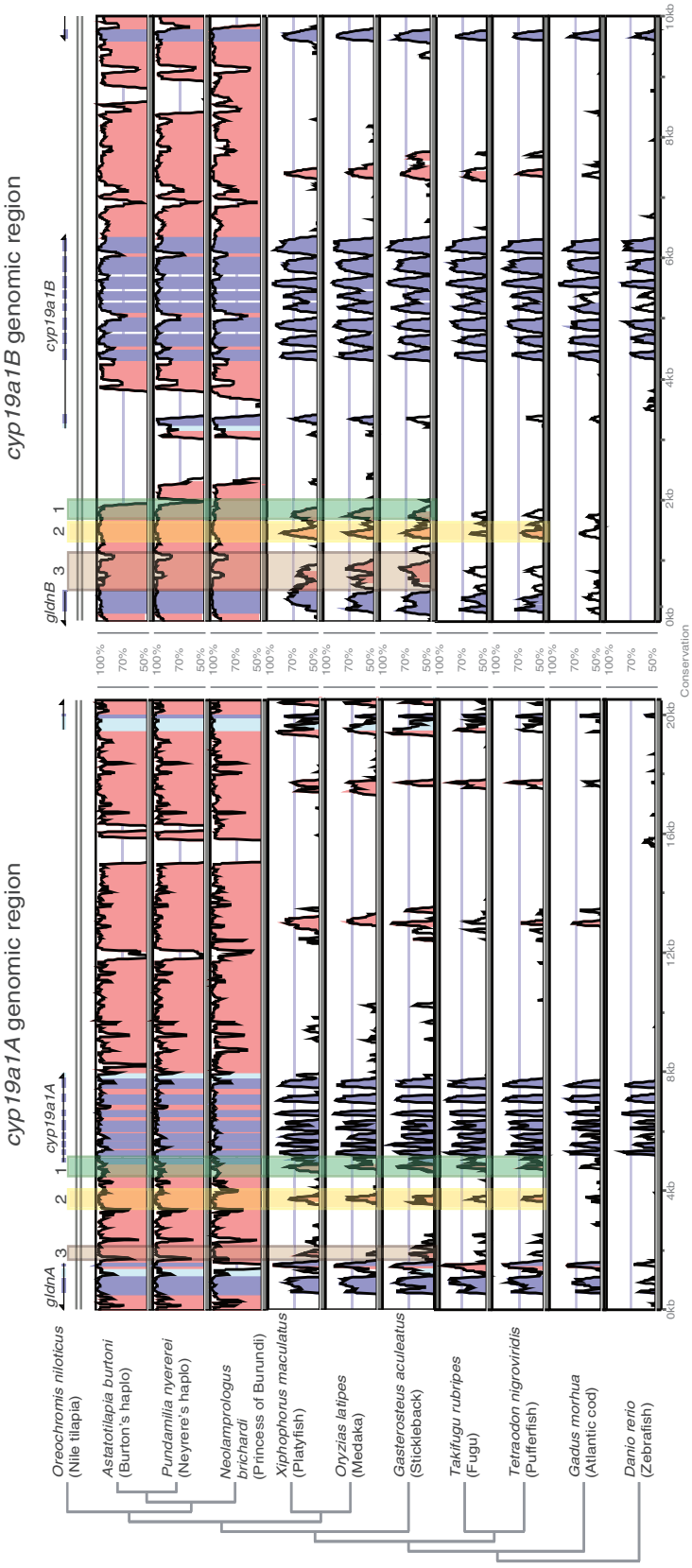

FIG. 3. Comparison of the homologous regions surrounding *cyp19a1A* and *cyp19a1B* in teleosts. Shuffle-LAGAN Vista plots for *cyp19a1A* and *cyp19a1B*. Blue peaks indicate conserved coding regions, pink peaks represent conserved noncoding regions. Light blue regions correspond to UTRs. Green, yellow, and brown boxes surround conserved possible promoter regions for both genes (in the text referred to as conserved boxes). Putative transcription factor binding sites for factors known to be active in gonads/germ cells and brain/CNS identified in these regions are indicated in table 2.

Table 2. Putative Transcription Factor Binding Sites in the Teleost Promoter Regions of *cyp19a1A* and *B*.

	cyp19a1A
	Block 1[a]
Nile tilapia	NF1 Nr5a2 ERR Sox5 FoxF2 FAC1 Oct-F Nr1f1 DREAM NGN/NeuroD
Platyfish	MEL1 Hmx3 Nr5a2 ERR Nr5a2 ERR Sox5 FoxA1/2 Sry Oct-6 Nr1f1 DREAM CREB Myt1l
Medaka	SF1 ERR SPZ1 Nr5a2 ERRalpha Sox5 FoxA1/2 FAC1 Oct-6 Nr1f1 DREAM
Stickleback	HMGI(Y)
Fugu	SPI1 Nr5a2 Sox5 XBP1 Dec-2 ARNT FoxA1/2 ATF2 CREB
Pufferfish	Sox5 HRE Dec-2 ARNT FoxF2 ATF2 CREB
	Block 2
Nile tilapia	Pou5f1 Sox5 AP1
Platyfish	Sox5 AP1 Hbox-F[b] FoxJ1
Medaka	Sox5 Hbox-F HRE
Stickleback	RFX1 FoxK2
Pufferfish	CREB
	Block 3
Nile tilapia	MEL1 Nanog MyF4 Ascl1
Platyfish	PLAG1 Meis1 Evi-1 HBP-1
Stickleback	SPI-1 GABP WT1 MEL1 Nr1d1
	cyp19a1B
	Block 1
Nile tilapia	Nkx2-5 RFX1/3 MyT1l HoxC9 FoxF1 Nkx2-5 HoxC9 FoxA1/2 Sox9 SPI-1
Platyfish	Hmx3 MyT1 RFX1 HoxC9 FoxJ1a HoxC9 SPI-1
Medaka	Hmx3 RFX1 Six4 HoxC9 Hmx2 HMGA1/2 Hox-F[c]
Stickleback	Hmx3 RFX1 FoxF1 MyF6
	Block 2
Nile tilapia	Hhex Msx-1/2 Brn-3 Nkx6-1 Nmp4 AP1 Sox2 ERE-F Hbox-F HMGA1/2 PLAG1
Platyfish	Sox30 SPZ1 TGIF1 ERE-F Hbox-F PCE-1 SPI-1
Medaka	MEL1 Etv4 TGIF1 ERE-F S8 Hbox-F PCE-1 ERR Nr5a2
Stickleback	ERE-F Hbox-F HMGA1/2 PCE-1
Pufferfish	ERE-F AP1 Hox9A
	Block 3
Nile tilapia	E2a Ascl1 Meis1 GSH-2 Phox2a/b SPI-1 Gsh-1 Elf5 HoxC9 Sox-F MyT1l MyT1l
Platyfish	TGIF DREAM MEL1
Medaka	RFX1 ZIC2 NGN/NeuroD HRE Hbox-F
Stickleback	Elf5 MyT1l

[a]Blocks correspond to the regions shown in figure 3.

[b]Binding sites that are shared between at least two species are underlined.

[c]F indicates that the binding site is predicted to bind several members of the same family. For the sake of space, not all family members are listed but the family is indicated by –F.

candidate genes are putative members of the hypothalamic–pituitary–gonadal axis, especially those involved in sex steroid synthesis, and hence, are expected to be expressed in the brain. This expectation can also be made for the adult brain because the fish brain produces sex steroid hormones necessary for sexual differentiation, plasticity, and reproduction throughout the entire lifespan of a fish unlike the mammalian brain, which is irreversibly sexualized during early development.

Table 3 shows a summary of the expression data for the three tested cichlid species (for box plots and detailed values of all expression tests and statistics see supplementary material S2, Supplementary Material online). Most of the tested gene candidates show—at least to some extent—expression patterns as described for other species (e.g., *dmrt1* is also a testis gene in cichlids). The tissue of highest expression level is most often conserved between the examined species. However, we also found deviations from described expression patterns.

We cannot confirm a role for *rspo1* as an ovary factor in adult cichlids, where it is mainly expressed in the brain. Rspo1 has been suggested to act as an activator of the wnt/beta-catenin pathway in the developing ovary (Smith et al. 2008) and is overexpressed in zebrafish ovaries (Zhang et al. 2011). Our data do, however, support a role for other members of this pathway, *ctnnb1A* and *B* and *wnt4a*, in the adult ovary, whereas *wnt4B*, just like *rspo1*, is predominantly expressed in the brain and testis; furthermore, *ctnnb1A* seems to also have experienced a shift in expression toward the brain in adult cichlids. Therefore, the role of the wnt-pathway in ovary development seems to be less conserved than previously thought, and, in addition, to differ between paralogous gene copies.

Table 3. Expression Patterns of Candidate Genes in Cichlids (Based on qRT-PCR, Calculated after Simon [2003], See Supplementary Material S2, Supplementary Material online).

Gene	Expression in *Astatotilapia burtoni*	Expression in *Ophthalmotilapia ventralis*	Expression in *Neolamprologus pulcher*
ctnnb1A[a], *catenin beta A*, intracellular signal transducer, part of the wnt-pathway in ovaries	Brain > gonads	Ovary and brain > testis	Ovary and brain > testis
ctnnb1B, *catenin beta B*	Ovary > brain > testis	Ovary > brain > testis	Ovary > female brain > male brain > testis
cyp11b2, *cytochrome P450 family 11 subfamily B polypeptide 2*, androgen synthesis, testis factor	Testis > female brain > male brain > ovary	Testis > brain and ovary	Testis > brain > ovary
cyp19a1A, *aromatase cytochrome P450 family 19 subfamily A polypeptide 1 A*, estrogen synthesis, ovarian aromatase	Ovary > brain and testis	Gonads > brain	Ovary > testis > brain
cyp19a1B, *aromatase cytochrome P450 family 19 subfamily A polypeptide 1*, brain aromatase	Testis > brain and ovary	Male brain > female brain > gonads	Brain > gonads
dax1A, *dosage-sensitive sex reversal adrenal hypoplasia critical region on chromosome X gene 1 A*, orphan receptor, regulates steroidogenesis, interacts with *sf-1*	Ovary > testis > brain	Ovary and brain > testis	Ovary and brain > testis
dmrt1, *doublesex and mab-3 related transcription factor 1*, male sex/testis differentiation	Testis > ovary	Testis > ovary	Testis > ovary > female brain > male brain
figla, *factor in the germline alpha*, germline transcription factor, acts in folliculogenesis	Ovary > testis	Ovary > testis	Ovary > testis
foxl2A, *forkhead transcription factor L2 A*, transcription factor, regulates *cyp19a1A*, ovary development and function	Ovary > brain > testis	Ovary > brain > testis	Ovary > brain > testis
foxl2B, *forkhead transcription factor L2 B*, discussed earlier	Gonads	N.A.	N.A.
gata4, GATA *binding protein 4*, transcription factor, gonad development, and function	Ovary > testis	Gonads > brain	Gonads > brain
nanos1A, brain and germ cell development	Brain > ovary > testis	Brain > ovary > testis	Brain > ovary > testis
nanos1B, discussed earlier	Brain > ovary > testis	Brain > ovary > testis	Brain > gonads
nr5a2 (*lrh-1*), *nuclear receptor subfamily 5*, *group A*, *member 2*, acts in steroidogenesis	Brain > gonads	Brain > gonads	Brain > ovary > testis
nr5a5, *nuclear receptor subfamily 5*, *group A*, *member 5*, related to *sf-1*, gene of the *nr5a* family, regulation of sterol/steroid metabolism	Not expressed	Brain and ovary	Brain > ovary > testis
rspo1, *R-spondin-1*, activates wnt/beta-catenin pathway in ovary differentiation	Brain > gonads	Brain > testis > ovary	Brain > gonads
sf-1 (*nr5a1*) *steroidogenic factor 1*, orphan nuclear receptor, gonadal differentiation, regulates steroidogenesis	Gonads > brain	N.A.[b]	N.A.[b]
sox9A, *SRY* (*sex determining region Y*)-*box 9 A*, transcription factor, regulates *amh* expression together with sf-1, gonad development, Sox9 is a conserved vertebrate testis factor	Brain > gonads	Brain > ovary > testis	Brain > gonads
sox9B, *SRY* (*sex determining region Y*)-*box 9 B*	Ovary > brain > testis	Ovary and brain > testis	Brain > ovary > testis
wnt4A, wingless-*type MMTV integration site family member 4A*, female reproductive development	Ovary > brain > testis	Ovary and brain > testis	N.A.[b]
wnt4B, *wingless-type MMTV integration site family member 4B*	Brain and testis	Brain and testis > ovary	Brain > testis > ovary
wt1A, *Wilms tumor 1 A*, gonad formation, testis differentiation, interacts with *sf-1*	Gonads > brain	Gonads > brain	Ovary > testis > brain
wt1B, *Wilms tumor 1 B*	N.A.[b]	Gonads > brain	Ovary > testis > male brain > female brain

NOTE.—N.A., not applicable.

[a]*A* and *B* denote gene copies derived from the FSGD.

[b]Primer do not amplify successfully.

In mammals, Sox9 is indispensable for testis differentiation (Kanai et al. 2005) and *sox9* is the first gene to show male sex-specific expression in many species (for a review see Morais da Silva et al. 1996). In fish, both *sox9* gene copies were found to be (highly) expressed in developing and adult testis (Zhou et al. 2003; Johnsen et al. 2013). In our set of cichlids, *sox9A* seems to be important in brain tissue and *sox9B* in ovaries and to some extent in the brain.

The most intriguing expression patterns observed in our data set involved once more the aromatase gene *cyp19a1*. Although the brain expression patterns of *cyp19a1B* in two of our study species (*N. pulcher* and *O. ventralis*) are consistent with the general trend in teleosts, that is, expression of the A-copy in ovaries and the B-copy in the brain (Callard et al. 2001; Cao et al. 2012), the member of the derived and exceptionally species-rich haplochromine cichlids (*A. burtoni*) shows a different pattern. In *A. burtoni*, *cyp19a1B* is overexpressed in testis and not in the brain, a pattern so far not described. Also *cyp19a1A* shows a deviation from the general trend in teleosts, in this case in the Ectodini *O. ventralis*, where we detected a similar expression level in testis and in ovaries (the other two cichlids show overexpression in ovaries, like other teleosts).

Because of the observed new expression patterns and the strong signal for positive selected sites in *cyp19a1A*, we decided to study both gene copies in more detail and more species focusing on a phylogenetically representative set of the East African cichlids. To this end, we performed qRT-PCR expression analysis in adults of six additional cichlid species: a representative of an extra cichlid tribe (*Perissodus microlepis*, tribe Perissodini) as well as more representatives of the Ectodini (*Enantiopus melanogenys*), the Lamprologini (*Altolamprologus fasciatus*), and, particularly, the Haplochromini (*Ctenochromis horei*, *Pseudotropheus pulpican*, and *Haplochromis obliquidens*). Thus, we now also included members of the two major haplochromine adaptive radiations, that of lakes Victoria (*H. obliquidens*) and Malawi (*P. pulpican*). The results of these experiments are shown in figure 4 (for further details on values and statistics see supplementary material S3, Supplementary Material online).

First, we confirmed that the newly detected overexpression of *cyp19a1B* in testis of *A. burtoni* is representative for haplochromine cichlids in general, indicating that this shift in gene expression has been acquired in the ancestor of the most species-rich group of cichlids.

Second, the expression of *cyp19a1A* in testis in *O. ventralis* also seems to be a more general pattern in Ectodini, as evidenced by the analysis of a second representative of this tribe, *E. melanogenys*. The expression patterns observed in Lamprologini and Perissodini, however, resemble the presumably ancestral patterns known from other teleosts.

Discussion

In this study, we investigated supposedly conserved candidate genes implicated in sex determination, sex differentiation, and reproduction in East African cichlid fishes. We found substantial variation in adult expression patterns for several of these genes, suggesting that not only the master initial regulators but also the downstream genetic factors of sex determination are less conserved than previously thought. We could show that this variation depends on the species investigated, and that expression can vary even between closely related species. Furthermore, variation in sequence evolution and expression can act on the level of additional gene copies such as the ones that emerged from the FSGD (e.g., see our data on *sox9A* and *B* and *wnt4A* and *B*), opening the path for sub- and neofunctionalization. This suggests that not only the master regulatory triggers of sex determination are labile in the animal kingdom but also that the downstream genetic factors can experience substantial changes in coding sequence and function. This adds a further level of plasticity to the processes of sexual development, at least in teleost fish.

Aromatase Gene Expression in Cichlids

The aromatase *cyp19a1* is considered one of the most important, if not the key gene in vertebrate sexual development (Diotel et al. 2010; Valenzuela et al. 2013). *Cyp19a1* is the sensitive gene that reacts to temperature changes and thus induces sexual development in reptiles (Merchant-Larios and Díaz-Hernández 2013), it has been implicated with sex change in fish (Liu et al. 2009; Nozu et al. 2009) and it is one of the first genes to show sex differences in expression in developing tilapia (Ijiri et al. 2008). *Cyp19a1* is a fundamental component of the estrogen pathway and it has been suggested that hormonal actions, such as the ones of estrogen, are so important that sex determination relies on them exclusively, regardless of their initial activation (Angelopoulou et al. 2012; Merchant-Larios and Díaz-Hernández 2013).

In adult vertebrates, *cyp19a1* typically comes in two flavors, the brain and the ovary aromatase. In tetrapods, this target specificity is achieved through the generation of alternative 5′-UTR transcripts derived from a single gene, mediated by tissue-specific promoters (Toda and Shizuta 1993; Golovine et al. 2003); an exception is the pig, which has three copies of *cyp19a1* (Graddy et al. 2000). Teleosts, on the other hand, have two copies of the gene generated by the FSGD and generally express the A-copy in ovaries and the B-copy in the brain (an exception are the European and the Japanese eels that only retained one copy after the FSGD, which is expressed in both, the brain and gonads, Tzchori et al. 2004; Jeng et al. 2012). The high expression levels in the fish brain, especially of *cyp19a1B*, are explained by its regulation through a feedback loop of estrogen on EREs in its promoter (Callard et al. 2001; Diotel et al. 2010, and references therein). In general, the functions of *cyp19a1A* and *cyp19a1B* have been suggested to be highly conserved among teleosts (Callard et al. 2001).

Here, we show that East African cichlid fishes deviate from this trend. *Cyp19a1A*, which shows strong signs of positive selection in cichlids, has experienced a shift in expression from ovary-specific to testis and ovaries in the Ectodini of LT. This pattern clearly contradicts the general assumption that an upregulation of *cyp19a1A* is needed to initiate and maintain ovary function in fish, whereas its downregulation induces testis development (Guiguen et al. 2010). Furthermore, the

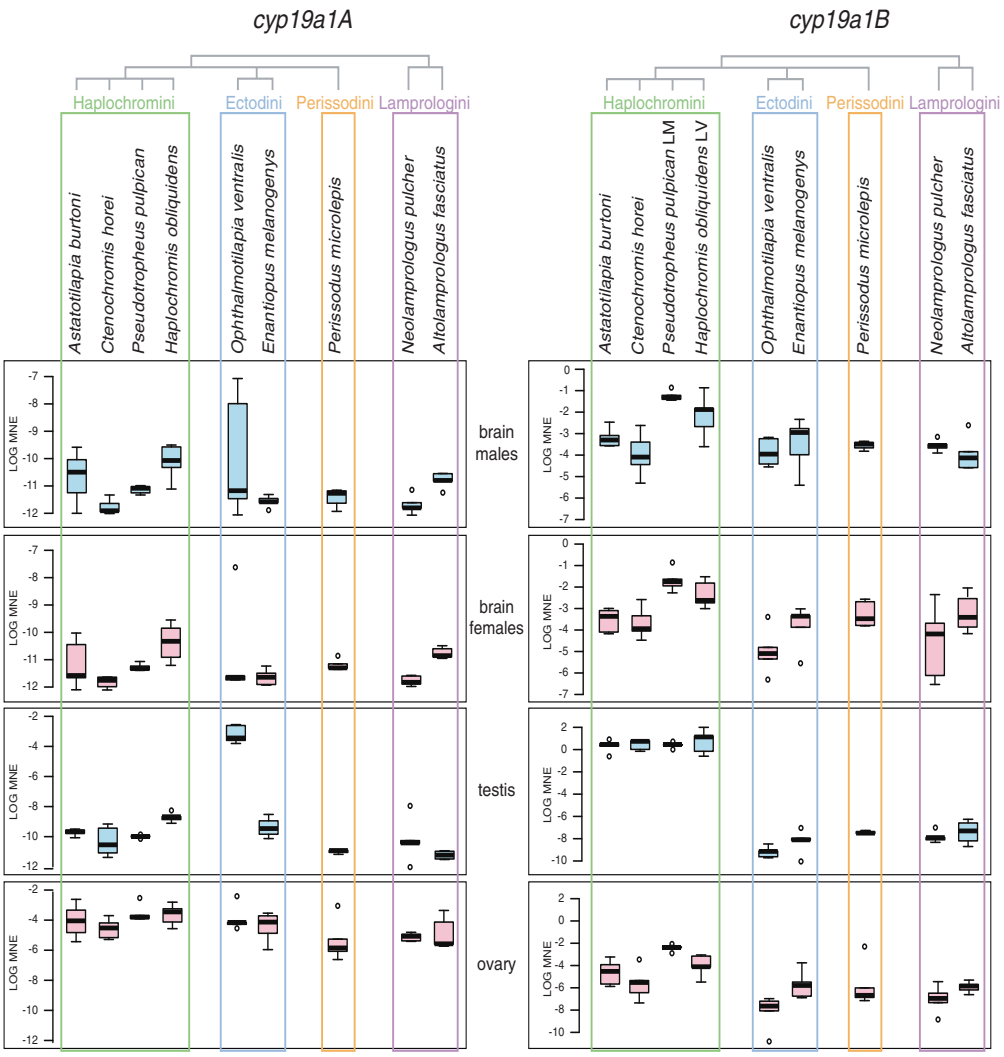

Fig. 4. qRT-PCR expression data for aromatase *cyp19a1A* and *B* in different East African cichlid species. qRT-PCR expression data for *cyp19a1A* and *cyp19a1B* are shown as mean normalized expression (MNE, calculated as described in Simon 2003) on a logarithmic scale tested in male and female brain tissue, testis, and ovary in species representing four major cichlid lineages (Haplochromini, Ectodini, Perissodini, and Lamprologini). *Pseudotropheus pulpican* is a species of Lake Malawi and *Haplochromis obliquidens* of Lake Victoria. All other species are from LT. Sample size is five for each tissue and sex except for *Perissodus microlepis* (only four males available and six females) and *Altolamprologus fasciatus* (only four males available). For details on values, raw data, and statistics see supplementary material S3, Supplementary Material online.

protein domains of Cyp19a1a could be substantially influenced by the changes occurring in its coding sequence in East African cichlids and hence its functionality in the estrogen pathway remains to be determined in these fish.

Cyp19a1B, on the other hand, shows high sequence conservation in the investigated cichlids suggesting that it preserved its function in testosterone aromatization to estrogen. Still, this gene has experienced a shift in expression from brain to testis in the Haplochromini. It thus appears that in this most species-rich group of cichlids the main site of adult estrogen production is the testis and not the brain. The haplochromines are the cichlid lineage displaying the most pronounced sexual color dimorphism, which has often been implicated with sexual selection via female choice, and they show the most derived female mouthbrooding behavior (Salzburger et al. 2005). Whether the unprecedented recruitment of the otherwise brain-specific gene *cyp19a1B* in testis is causally linked to some of these features of haplochromine cichlids would need to be investigated in the future.

Adult ectodine cichlids mainly express *cyp19a1A* in the gonads and haplochromine cichlids overexpress *cyp19a1B* in the testis. This should lead, at least in the haplochromines, to high estrogen levels in the testis, whereas the functionality of Cyp19a1A in estrogen aromatization to testosterone can be questioned.

High estrogen levels in the haplochromine cichlid *A. burtoni* and in many other vertebrates are linked to aggressive behavior (Trainor et al. 2006); however, the steroids effecting

aggressive (but not reproductive) behavior are apparently not of gonadal origin in *A. burtoni* (Soma et al. 1996; Huffman et al. 2013), leaving the function of high aromatase activity in the testis of this species unexplained. Yet, none of the studies on estrogen levels in *A. burtoni* is specific for one of the two *cyp19a1* copies (Huffman et al. 2013). Pharmacologically blocking aromatase (presumably both forms), decreases aggressive behavior, whereas reproductive behavior is not influenced (Huffman et al. 2013, and references therein). We suggest further critical investigation of the usage of such inhibiting substances with respect to the sequence changes that we observed in *cyp19a1A*.

Estrogen production via aromatase in the testis of fish has been described to be involved in spermatogenesis, where estrogen is necessary for the renewal of spermatagonial stem cells and thought to be necessary for germ cell proliferation and possibly differentiation (Schulz et al. 2010). Yet, overexpression of aromatase in testis compared with ovary or brain has not been described in fish before (see table 2 in Piferrer and Blázquez 2005). Furthermore, an increase of estrogen levels has been shown to impair testis function in fish and substantially influence reproduction (Kobayashi et al. 2011), leaving the high expression levels of aromatase, and especially *cyp19a1B* in haplochromine cichlids, hard to explain with the currently available data on estrogen actions on testis/spermiogenesis.

Considering our analysis of the promoter regions of *cyp19a1A* and *B* in cichlids, we think that the presence of SF-1 and Dmrt-1 binding sites in the *cyp19a1A* promoter and the expression of both factors in the testis of cichlids (fig. 3 and table 3) is consistent with a repressive action of these two transcription factors on *cyp19a1A* expression in testis. Conversely, *cyp19a1B* is overexpressed in testis in haplochromines, suggesting that it may not be suppressed by SF-1 in testis, though the cichlid *cyp19a1B* promoter contains a putative binding site for this factor and *sf-1* is overexpressed in testis. The main expression of *nr5a2* in the brain of *A. burtoni* contradicts its activating action on *cyp19a1A* in gonads, although there is a conserved potential binding site for it in the *A* promoter. Still, Nr5a2 could drive *cyp19a1A* expression in the brain or earlier during development in the gonads.

Further studies should unravel the functionality of the two aromatase copies in cichlids, their function throughout brain and gonad development and their action on adult steroid production involved in aggression and reproduction. Also the regulation of their expression by neuronal/gonadal transcription factors as well as through estrogens and androgens (as suggested by the presence of response elements for both steroid types in the cichlid promoter regions of *cyp19a1A* and *B*) need further investigation.

Contrasting Sequence Evolution Trajectories of the Two Aromatase Gene Copies in Cichlids

The relative contribution to phenotypic divergence of changes in the *cis*-regulatory elements of a gene versus changes in the coding sequence is still nebulous and a matter of debate (Hoekstra and Coyne 2007), although it has been argued that rapid evolution in closely related taxa such as cichlids is more likely to be mediated by regulatory changes (Baldo et al. 2011). Here, we show that even two copies of the same ancestral gene emerging from a duplication event may pursue these distinct evolutionary trajectories, both likely leading to functional divergence. Cyp19a1A has undergone various, presumably adaptive changes in its amino acid sequence during cichlid evolution that might substantially influence its function. *Cyp19a1B*, on the other hand, seems to be under strong purifying selection in cichlids, indicating that functional constraints act on at least one of the duplicate gene copies, though the effects of aromatase action remain illusive. Still, also the B-copy shows new evolutionary paths in cichlids: in addition to expression shift in Haplochromini (overexpression in testis) we could show the usage of different TSSs in Ectodini and Haplochromini, a feature already described for the single copy *cyp19a1* gene in tetrapods (mentioned earlier). This shift seems to be caused by *cis* or *trans* regulatory changes. The other, more basal cichlids investigated here show conserved expression patterns in both genes resembling other teleosts. This illustrates that paralogs can be subjects to differential evolutionary fates also at late, that is, derived stages of an evolutionary radiation.

More Flexibility in the Sex-Determining Cascade: *Sox9* Expression in Cichlids—More than a Testis Factor

In mammals, Sox9 drives testis differentiation (Kanai et al. 2005) and a similar role has been suggested for this gene in other vertebrate groups, after the observation that it is often the first gene to show male sex-specific expression (Morais da Silva et al. 1996). In teleost fish, both *sox9* gene copies were described to be (highly) expressed in developing and adult testis (Zhou et al. 2003; Johnsen et al. 2013). Apparently, however, the expression pattern is more flexible in adult fish, it varies by species and depends on the gene copy investigated. In line with this, we show that *sox9A* seems to rather play a role in the brain and *sox9B* in ovaries and also to some extent in the brain. This illustrates that substantial differences exist in the expression pattern of *sox9* between teleost fishes and underlines once more the importance of studying both duplicates of a single copy tetrapod gene, when present in fish, and not to conclude a function from one copy only. This is also illustrated by our data on *wnt4A* and *B*. *Sox9A* and *B* are obviously another example of a sub- or neofunctionalization of gene copies after duplication (Ohno 1970).

Sexual Development Is Plastic in Teleost Fishes

It has been proposed that a set of genes forms a common regulatory network of sexual differentiation in vertebrates (Angelopoulou et al. 2012; Valenzuela et al. 2013). This study shows that some of the genes implicated in sexual development and reproduction are indeed conserved between lineages. The function of these genes, however, as reflected by their varying expression patterns in different

species (this study and see Cutting et al. 2013 for a review) and changing coding sequence as we show here for the aromatase gene copies, seems to be less conserved than previously thought.

Plasticity in sexual development has been studied in detail in sex-changing fish and species that possess temperature dependent sex-determining mechanisms, where gene expression changes in response to temperature have been assessed. Most other available studies on sexual development have focused on one gene or a particular pathway of a few interacting genes. Our work was based on a broad screen of many genes and with respect to both, gene expression and sequence evolution, in a prime model system in evolutionary biology, the East African cichlids. We detected substantial differences in expression within different cichlid species and between cichlids, or certain lineages thereof, and other teleosts. The most intriguing outcome of our study is the lineage-specific shifts in gene expression in the two copies of the aromatase *cyp19a1*, a major core gene of sexual development, behavior, and reproduction in chordates. That the changes in gene expression in the two *cyp19a1* copies occurred in derived cichlid lineages, but not in the ancestral tribes, suggests that functional divergence between two gene copies derived from the FSGD adds another level of flexibility to the sex-determining cascade of teleosts.

Taken together, most of the studied genes are part of the cascade of sex determination, differentiation, and reproduction across teleosts but their usage may differ substantially at different taxonomic levels. This can be interpreted as a developmental system drift (DSD), a change in the molecular pathway with preservation of the morphological/phenotypic outcome (True and Haag 2001), in this case the establishment, maintenance, and functioning of separate sexes. Indeed, changes of master control genes in the sex-determining cascade between closely related taxa have been called "the most striking examples of diversity in a conserved regulatory process" (True and Haag 2001), exemplifying DSD. Our data and also those of others (Wang and Sommer 2011; Valenzuela et al. 2013) indicate that this diversity not only affects the top of the sex-determining cascade but also includes sexual differentiation genes, hormone pathways and the molecular networks necessary for the development of reproductive organs and behavior. The occurring modifications can thus affect all levels of a developmental process and may depend on *cis* or *trans* changes, novel wiring, new patterns of coding sequence evolution, different gene usage and *de novo* recruitment and evolution of genes.

Especially the observed sequence and expression changes of the aromatase gene copies could have implications on the evolution of sex chromosomes, which might be driven by the action of hormones (Howard 2002). Our study once more demonstrates that understanding evolutionary shifts in sex-determining systems requires the characterization of downstream gene variability (Uller and Helantera 2011). Cichlid fishes are, in this respect, an ideal model system since many closely related, hence genetically very similar species, with different sex-determining mechanisms and all sorts of reproductive strategies exist.

Materials and Methods

Tissue Sampling and Nucleic Acid Extraction

The presented data set comprises DNA sequence data for 33 species of East African cichlid fishes (see supplementary table S4, Supplementary Material online) and expression data for nine species (*A. burtoni*, *C. horei*, *O. ventralis*, *E. melanogenys*, *N. pulcher*, *Alt. fasciatus*, *P. microlepis*, *H. obliquidens*, and *Pse. pulpican*). The investigated species belong to 13 different tribes and thus cover a great portion of the taxonomic and phylogenetic diversity of East African cichlids (Salzburger et al. 2002, 2005). The species that were tested for candidate gene expression were chosen with regards to available genetic resources and their respective phylogenetic positions. All candidate genes were tested in *A. burtoni* (a member of the Haplochromini), *O. ventralis* (Ectodini), and *N. pulcher* (Lamprologini). These three cichlid tribes differ remarkably in many life history aspects and with regards to sexual development and, especially, their reproductive strategies. *A. burtoni* and *O. ventralis* are maternal mouthbreeders, whereas *N. pulcher* is a bi-parental substrate spawner within a cooperative helper group. Additionally, *A. burtoni* shows a rather strong sexual color dimorphism.

Samples of the species from LT were collected during field expeditions to the lake and surrounding rivers between 2007 and 2012 under a research permit issued by the Lake Tanganyika Research Unit, Department of Fisheries, Mpulungu, Zambia. Tissue samples of *N. callipterus* were kindly provided by M. Taborsky. Adult and mature specimens for two additional species (*H. obliquidens* from Lake Victoria and *P. pulpican* from Lake Malawi) were derived from the aquarium trade. Life fish were kept under standard conditions (12 h light, 12 h dark, 25 °C) at the animal facility of the Salzburger group, University of Basel, Switzerland. Research was performed under the cantonal veterinary permit no. 2317.

All samples for the expression studies were collected during the above mentioned field expeditions in 2011/2012. Tissues (brain, gonad, gill, anal fin, and liver) were collected from wild caught adult fish within one hour to reduce gene expression shifts caused by stress. Tissues were directly stored in RNAlater (LifeTechnologies). DNA samples (muscle and fin tissue) were taken at the same time and stored in 100% ethanol. Fish were sexed based on external morphological traits and visual gonad inspection. Samples were only included in the study when ovary or testis structure could be determined unambiguously. For the mouth-breeding species only non-breeding females were dissected. Given the sampling conditions, reliable determination of GSI (gonadosomatic index) or plasma steroid levels to infer reproductive status of the invested specimens was not feasible. Note however, that previous studies on stable laboratory populations of male *A. burtoni* have shown that changes in gonad mass are not necessarily reflecting reproductive maturity but are rather signs of interstitial cell division (Huffman et al. 2012, and references therein). Additionally, testis of both subordinate and dominant males contains all spermatogenic stages (Maruska and Fernald 2011) and sperm proliferation does not differ

between them (Kustan et al. 2011). A reliable assessment of reproductive status and dominance over gene expression analysis is, however, possible (Huffman et al. 2012) and sex-specific expression in our data set is consistent between samples of one sex.

Genomic DNA was extracted from ethanol preserved muscle or fin tissue applying manually a Proteinase K digestion followed by sodium chloride extraction and ethanol precipitation (Laird et al. 1991) or using the robotic workstation BioSprint 96 (Qiagen) following the manufacturer's instructions.

Samples stored in RNAlater were incubated overnight in Trizol (LifeTechnologies) at 4 °C (this initial incubation step increases RNA yield after long-term storage in RNAlater) prior to extraction and subsequently homogenized in Trizol using a FastPrep24 (MP Biomedicals Europe). Total RNA was extracted following the Trizol protocol. RNA concentration and quality was measured using a NanoDrop1000 Spectrophotometer (ThermoScientific). RNA samples were treated with DNA-free Kit (LifeTechnologies) as recommended by the manufacturer.

Reverse Transcription and Quantitative Real-Time PCR

DNase-treated RNA was reverse transcribed using the High capacity RNA-to-cDNA kit (LifeTechnologies) following the manufacturer's protocol. Real-time PCR experiments were run on a StepOnePlus Real-Time system (LifeTechnologies) on a final cDNA concentration of 0.5 ng/µl for the experiment shown in table 3 and 1 ng/µl for the experiment shown in figure 4 with 200 nM final primer concentration and the SYBR Green master (Rox) dye (Roche) in 20 µl final volume with the following cycling conditions: 95 °C 10 min, 40 cycles 95 °C 10 s, and 58 °C 30 s for experiments shown in column 3 of table 3 and 95 °C 10 min, 40 cycles 95 °C 15 s, and 58 °C 60 s for all other experiments shown). All qRT-PCR amplifications included a melt curve step after cycling. Primers were constructed in conserved regions chosen based on the *A. burtoni* and *N. brichardi* (*N. brichardi* forms together with the expression investigated for *N. pulcher*, the *N. brichardi*/*pulcher* species complex, Duftner et al. 2007) genome sequences and sequenced amplicons of the corresponding regions for *O. ventralis* to guarantee equal binding efficiencies between species.

All used primers (see supplementary table S5 [Supplementary Material online] for primer sequences) were initially validated on serial dilutions of factor 5 of *A. burtoni* juvenile whole body mixed cDNA (34 individuals of 7.8–24 mm standard length, 3 weeks to 2 months age, laboratory strain, equimolar pooled RNA) and for *N. brichardi*/*pulcher* specific primers on serial dilutions of factor 5 of *N. brichardi* juvenile whole body mixed cDNA (six individuals pooled, 18.4–24.6 mm standard length, 3 months age, kindly provided by H. Gante, University of Basel) starting with 10 ng/µl DNA final concentration. qRT-PCR efficiencies (E) were calculated for each reaction from the slope of the standard curve using the equation $E = 10^{(-1/\text{slope})}$ as implemented in the StepOnePlus software (LifeTechnologies) with an efficiency of 2 being equal to 100% ($E\% = [10^{(-1/\text{slope})} - 1] \times 100$) and an indicator of a robust assay.

Expression was measured as comparative C_T experiments using *rpl7* (*ribosomal protein L7*) as control gene and gills as reference tissue. *Rpl7* was chosen after an initial comparison with two other presumably house keeping genes (*ef1a* and *beta-actin*). *rpl7* showed the most consistent expression between samples (i.e., smallest C_T range) and has been used previously in fish qRT-PCR experiments (Böhne et al. 2010). Gill tissue was chosen as reference for the delta-delta-C_T method, because it did not show differences between sexes and expressed the genes of interest. Furthermore, expression in gills was most consistent between species and in comparison to liver and anal fin tissue. For quantification, data were first analyzed using the delta-delta-C_T method (Livak and Schmittgen 2001) and custom R-scripts. As lowest threshold for expression we applied 37 cycles. As a second, and this time reference tissue independent, analysis method, mean normalized gene expression was calculated as described in (Simon 2003) again using *rpl7* as control gene.

All experiments were carried out with three technical replicates. The details of all experiments including used samples, raw data, graphs, and statistics are given in supplementary materials S2 and S3, Supplementary Material online.

Cyp19a1A Gene Sequencing

A fragment including the entire coding region was amplified by PCR on genomic DNA of 28 cichlid species from LT. The species cover 12 tribes of the cichlid assemblage of this lake; the number of species sequenced per tribe was chosen according to species-richness of the tribe. When available, a male and a female specimen were sequenced (supplementary table S4, Supplementary Material online).

Long-range PCR reactions were carried out using Phusion Master Mix (New England Biolabs) in 40 µl reaction volume on 50 ng genomic DNA with the primers for 5′-TGAACTAGG TCCTGTAAACCCAAGG-3′ and rev 5′-AAGACTTTTGCACAG AACAGTAGG-3′ (cycling conditions: 98 °C 30 s followed by 35 cycles of 98 °C for 10 s, 64 °C for 30 s, 72 °C for 4 min and with a final elongation of 10 min at 72 °C, primers ordered from Microsynth).

PCR products were visualized using SYBRSafe DNA Gel Stain (LifeTechnologies) on a 1.5% agarose gel. Thirty microliters of PCR product was gel purified using QIAquick Gel Extraction Kit (Qiagen). 100 ng of purified PCR product were used for barcoded library preparations using NEBNextFast DNA fragmentation kit (fragmentation time 28 min) and Ion Xpress Plus Fragment Library kit and Ion Xpress Barcode Adapters (LifeTechnologies) according to manufacturer's recommendation. Libraries were pooled after barcoding for subsequent analysis. Quantification prior to sequencing was done with the Ion Library Quantitation Kit (LifeTechnologies) on a StepOnePlus Real-Time system (LifeTechnologies).

Libraries were sequenced on three 314 chips on an Ion PGM Sequencer (LifeTechnologies). Demultiplexed sequences

were loaded into Geneious (Biomatters Ltd.) and mapped to the reference amplicon reconstructed from the *A. burtoni* genome (BROAD) as closest available relative for all self-sequenced species. Alignments were corrected manually. The subsequent sequence analyses were performed in Geneious using MUSCLE (Edgar 2004) as alignment algorithm. Alignments per species were combined to construct one consensus sequence per species with 75% identical threshold for ambiguous sites. Sequences for *O. niloticus*, *P. nyererei*, and *N. brichardi* were retrieved from the cichlid genomes (BROAD) and included in the alignment.

The final alignment based on the grouped nucleotide sequences was trimmed and used for the reconstruction of a phylogenetic tree with MrBayes (5 million generations, default settings, version 3.2.1., Ronquist et al. 2012) under the GTR + I + gamma model as determined with jModelTest2 (Darriba et al. 2012). The obtained tree topology was in accordance with a maximum likelihood tree (1,000 bootstraps) obtained with PhyML (Guindon et al. 2010) (data not shown). An alignment covering only the coding region from start to stop codon according to the Nile tilapia sequence available in Ensembl (ENSONIT00000000198) was used for selection analysis.

In Silico Sequence Analysis

Sequences of the candidate genes were identified using the reciprocal best hit BLAST method with annotated sequences from the medaka (*O. latipes*) genome assembly (http://www.ensembl.org/Oryzias_latipes, last accessed July 23, 2013) as queries against the Nile tilapia (*Oreochromis niloticus*) EST data available in Genbank (www.ncbi.nlm.nih.gov, last accessed July 23, 2013) and, upon availability, against the Nile tilapia genome (http://www.ensembl.org/Oreochromis_niloticus, last accessed July 23, 2013). The retrieved tilapia sequences were used as queries in a local BLASTn search against the cichlid fish genomes (BROAD) as references. If the reciprocal best hit method was not conclusive, especially to distinguish cichlid/teleost copies of the same original tetrapod gene, phylogenetic analysis were carried out (data not shown) including more teleost sequences to characterize the A- and B-copies according to the annotation used for medaka in Ensembl (www.ensembl.org, last accessed July 23, 2013).

In-frame alignments of the coding sequences from start codon (included) to stop codon (excluded for Selecton analysis, see below) were performed with CodonCodeAligner (CodonCode Corporation). All alignments were checked and corrected manually with respect to exon–intron boundaries and in-frame indels.

Overall Ka/Ks estimates were calculated using KaKs_calculator (Zhang et al. 2006) with the MA method for each pairwise comparison to the Nile tilapia sequence over the entire length (supplementary tables S2 and S3, Supplementary Material online).

Site-wise Ka/Ks estimates were performed with Selecton (Doron-Faigenboim et al. 2005; Stern et al. 2007) using the tilapia sequence as query. The M8 (M8, beta + $w \geq 1$) model enabled for positive selection was used (Yang et al. 2000). When positively selected sites were detected under this model, it was tested against the null model (no positive selection, M8a, beta + w = 1, Swanson et al. 2003).

The distribution of positively selected sites in *cyp19a1A* along the phylogeny was tested with HyPhy (Pond et al. 2005) and codeml, which is part of the PAML package (Yang 2007). To test for branch-specific signs of positive selection we run the Branch-site REL model (Kosakovsky Pond et al. 2011) as implemented in HyPhy on www.Datamonkey.org (last accessed July 23, 2013) (Pond and Frost 2005; Delport et al. 2010). To test for site- and branch-specific signs of selection, we used codeml with the branch-site model and a test for positive selection (Yang and Nielsen 2002; Yang et al. 2005; Zhang et al. 2005) to test different clades against all other clades of the phylogeny (fig. 2).

Promoter Analysis

Promoter analyses were carried out on the flanking regions of the *cyp19a1A* and *cyp19a1B* genes exported from all available teleost genomes on www.ensembl.org (last accessed July 23, 2013) (Release 69, October 2012) and on the sequences available for *A. burtoni*, *P. nyererei*, and *N. brichardi* (BROAD). Annotation files were exported from Ensembl or, whenever necessary, created manually for the nonannotated genomes. The annotation for *gldnB* on the Nile tilapia sequence was added manually, as this gene is not annotated in the current genome release. Annotation was done using the coding sequence present in Genbank (Accession number XM_003443944). Alignments for all teleost promoters were performed using mVISTA (Mayor et al. 2000; Frazer et al. 2004) with Shuffle-LAGAN as alignment algorithm. Alignments including only cichlid sequences were performed with MUSCLE (Edgar 2004) as implemented in Geneious (Biomatters Ltd). Putative transcription factor binding sites were annotated according to (Tong and Chung 2003; Yoshiura et al. 2003; Chang et al. 2005; Ohmuro-Matsuyama et al. 2007; Wang et al. 2007; Diotel et al. 2010) and with the help of MatInspector as implemented in the Genomatix software suite V2.6 (Genomatix Software GmbH) with a matrix similarity cutoff of >0.9. We selected only transcription factors that fell in the categories brain and/or CNSs and ovary, testis, and germ cells. Note that putative transcription factor binding sites that are described as "transcription factor –F" (e.g., Hbox-F) include more than three members of the same family predicted to bind at the same site and are thus abbreviated with –F for family.

Supplementary Material

Supplementary material S1–S3, tables S1–S5, and figures S1 and S2 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

A.B., C.H., and W.S. designed the study and wrote the manuscript. A.B. performed sequence analysis. A.B. and C.H. performed qRT-PCR expression analysis. A.B. and N.B. sequenced the *cyp19a1A* locus. The authors thank A. Indermaur, F. Kim,

B. Meyer, B. Egger, and A. Theis for help with sampling; B. Meyer and L. Baldo for helpful discussion on data analysis; F. Piferrer for helpful comments on *cyp19a1*, the Broad Institute as well as the Cichlid Genome Consortium for providing access to cichlid genomes and transcriptomes, and two anonymous referees and the associate editor for valuable comments on a previous version of this manuscript. This work was supported by the Forschungsfond Universität Basel, the Volkswagenstiftung (Postdoctoral Fellowship Evolutionary Biology, grant number 86 031), and the DAAD (German academic exchange service, grant number D/10/52114) to A.B.; travel grant of the Swiss Academy of Sciences (SZG-Reisestipendium) to C.H.; and the European Research Council (ERC, Starting Grant "INTERGENADAPT") and the Swiss National Science Foundation to W.S. This research was done at the Zoological Institute, University of Basel, Basel, Switzerland.

References

Anderson JL, Rodríguez Marí A, Braasch I, Amores A, Hohenlohe P, Batzel P, Postlethwait JH. 2012. Multiple sex-associated regions and a putative sex chromosome in zebrafish revealed by RAD mapping and population genomics. *PLoS One* 7:e40701.

Angelopoulou R, Lavranos G, Manolakou P. 2012. Sex determination strategies in 2012: towards a common regulatory model? *Reprod Biol Endocrinol.* 10:13.

Baldo L, Santos ME, Salzburger W. 2011. Comparative transcriptomics of Eastern African cichlid fishes shows signs of positive selection and a large contribution of untranslated regions to genetic diversity. *Genome Biol Evol.* 3:443–455.

Barco A, Marie H. 2011. Genetic approaches to investigate the role of CREB in neuronal plasticity and memory. *Mol Neurobiol.* 44:330–349.

Baroiller JF. 2009. Tilapia sex determination: where temperature and genetics meet. *Comp Biochem Physiol A Mol Integr Physiol.* 153:30–38.

Böhne A, Darras A, D'Cotta H, Baroiller J-F, Galiana-Arnoux D, Volff J-N. 2010. The vertebrate makorin ubiquitin ligase gene family has been shaped by large-scale duplication and retroposition from an ancestral gonad-specific, maternal-effect gene. *BMC Genomics* 11:721.

Böhne A, Schultheis C, Galiana-Arnoux D, et al. (20 co-authors). 2009. Molecular analysis of the sex chromosomes of the platyfish *Xiphophorus maculatus*: towards the identification of a new type of master sexual regulator in vertebrates. *Integr Zool.* 4:277–284.

Callard GV, Tchoudakova AV, Kishida M, Wood E. 2001. Differential tissue distribution, developmental programming, estrogen regulation and promoter characteristics of *cyp19* genes in teleost fish. *J Steroid Biochem Mol Biol.* 79:305–314.

Cao M, Duan J, Cheng N, Zhong X, Wang Z, Hu W, Zhao H. 2012. Sexually dimorphic and ontogenetic expression of *dmrt1*, *cyp19a1a* and *cyp19a1b* in *Gobiocypris rarus*. *Comp Biochem Physiol A Mol Integr Physiol.* 162:303–309.

Castro LF, Santos M, Reis-Henriques M. 2005. The genomic environment around the Aromatase gene: evolutionary insights. *BMC Evol Biol.* 5:43.

Chang X, Kobayashi T, Senthilkumaran B, Kobayashi-Kajura H, Sudhakumari CC, Nagahama Y. 2005. Two types of aromatase with different encoding genes, tissue distribution and developmental expression in Nile tilapia (*Oreochromis niloticus*). *Gen Comp Endocrinol.* 141:101–115.

Chiang EF-L, Yan Y-L, Guiguen Y, Postlethwait J, Chung B-C. 2001. Two Cyp19 (P450 aromatase) genes on duplicated zebrafish chromosomes are expressed in ovary or brain. *Mol Biol Evol.* 18:542–550.

Cutting A, Chue J, Smith CA. 2013. Just how conserved is vertebrate sex determination? *Dev Dyn.* 242:380–387.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods.* 9:772–772.

Delport W, Poon AFY, Frost SDW, Kosakovsky Pond SL. 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* 26:2455–2457.

Diotel N, Page YL, Mouriec K, et al. (11 co-authors). 2010. Aromatase in the brain of teleost fish: expression, regulation and putative functions. *Front Neuroendocrinol.* 31:172–192.

Doron-Faigenboim A, Stern A, Mayrose I, Bacharach E, Pupko T. 2005. Selecton: a server for detecting evolutionary forces at a single amino-acid site. *Bioinformatics* 21:2101–2103.

Duftner N, Sefc KM, Koblmüller S, Salzburger W, Taborsky M, Sturmbauer C. 2007. Parallel evolution of facial stripe patterns in the *Neolamprologus brichardi/pulcher* species complex endemic to Lake Tanganyika. *Mol Phylogenet Evol.* 45:706–715.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–1797.

Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004. VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* 1:W273–W279.

Gerrard DT, Meyer A. 2007. Positive selection and gene conversion in SPP120, a fertilization-related gene, during the East African cichlid fish radiation. *Mol Biol Evol.* 24:2286–2297.

Godwin J. 2010. Neuroendocrinology of sexual plasticity in teleost fishes. *Front Neuroendocrinol.* 31:203–216.

Golovine K, Schwerin M, Vanselow J. 2003. Three different promoters control expression of the aromatase cytochrome P450 gene (*cyp19*) in mouse gonads and brain. *Biol Reprod.* 68:978–984.

Graddy LG, Kowalski AA, Simmen FA, Davis SLF, Baumgartner WW, Simmen RCM. 2000. Multiple isoforms of porcine aromatase are encoded by three distinct genes. *J Steroid Biochem Mol Biol.* 73:49–57.

Graham-Lorence S, Peterson JA, Amarneh B, Simpson ER, White RE. 1995. A three-dimensional model of aromatase cytochrome P450. *Protein Sci.* 4:1065–1080.

Gruart A, Benito E, Delgado-García JM, Barco A. 2012. Enhanced cAMP response element-binding protein activity increases neuronal excitability, hippocampal long-term potentiation, and classical eyeblink conditioning in alert behaving mice. *J Neurosci.* 32:17431–17441.

Guiguen Y, Fostier A, Piferrer F, Chang C-F. 2010. Ovarian aromatase and estrogens: a pivotal role for gonadal sex differentiation and sex change in fish. *Gen Comp Endocrinol.* 165:352–366.

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 59:307–321.

Hasselmann M, Gempe T, Schiøtt M, Nunes-Silva CG, Otte M, Beye M. 2008. Evidence for the evolutionary nascence of a novel sex determination pathway in honeybees. *Nature* 454:519–522.

Hattori RS, Murai Y, Oura M, Masuda S, Majhi SK, Sakamoto T, Fernandino JI, Somoza GM, Yokota M, Strüssmann CA. 2012. A Y-linked anti-Müllerian hormone duplication takes over a critical role in sex determination. *Proc Natl Acad Sci U S A.* 109:2955–2959.

Hinshelwood MM, Smith ME, Murry BA, Mendelson CR. 2000. A 278 bp region just upstream of the human *CYP19* (aromatase) gene mediates ovary-specific expression in transgenic mice. *Endocrinology* 141:2050–2053.

Hoekstra HE, Coyne JA. 2007. The locus of evolution: evo devo and the genetics of adaptation. *Evolution* 61:995–1016.

Hofmann E, Reichart U, Gausterer C, Guelly C, Meijer D, Muller M, Strobl B. 2010. Octamer-binding factor 6 (*Oct-6/Pou3f1*) is induced by interferon and contributes to dsRNA-mediated transcriptional responses. *BMC Cell Biol.* 11:61.

Howard JM. 2002. "Mitochondrial Eve", "Y Chromosome Adam", testosterone, and human evolution. *Rivista di Biologia* 95:319–325.

Huffman LS, Mitchell MM, O'Connell LA, Hofmann HA. 2012. Rising StARs: behavioral, hormonal, and molecular responses to social challenge and opportunity. *Horm Behav.* 61:631–641.

Huffman LS, O'Connell LA, Hofmann HA. 2013. Aromatase regulates aggression in the African cichlid fish *Astatotilapia burtoni*. *Physiol Behav.* 112–113:77–83.

Ijiri S, Kaneko H, Kobayashi T, Wang D-S, Sakai F, Paul-Prasanth B, Nakamura M, Nagahama Y. 2008. Sexual dimorphic expression of genes in gonads during early differentiation of a teleost fish, the Nile tilapia *Oreochromis niloticus*. *Biol Reprod.* 78:333–341.

Jeng S-R, Yueh W-S, Pen Y-T, Gueguen M-M, Pasquier J, Dufour S, Chang C-F, Kah O. 2012. Expression of aromatase in radial glial cells in the brain of the Japanese eel provides insight into the evolution of the *cyp191a* gene in actinopterygians. *PLoS One* 7:e44750.

Jetten AM. 2009. Retinoid-related orphan receptors (RORs): critical roles in development, immunity, circadian rhythm, and cellular metabolism. *Nucl Recept Signal.* 7:e003.

Johnsen H, Tveiten H, Torgersen JS, Andersen Ø. 2013. Divergent and sex-dimorphic expression of the paralogs of the Sox9-Amh-Cyp19a1 regulatory cascade in developing and adult Atlantic cod (*Gadus morhua L.*). *Mol Reprod Dev.* 80:358–370.

Kamiya T, Kai W, Tasumi S, et al. (16 co-authors). 2012. A trans-species missense SNP in *amhr2* is associated with sex determination in the tiger pufferfish, *Takifugu rubripes* (Fugu). *PLoS Genet.* 8:e1002798.

Kanai Y, Hiramatsu R, Matoba S, Kidokoro T. 2005. From SRY to SOX9: mammalian testis differentiation. *J Biochem.* 138:13–19.

Kobayashi T. 2010. *In vitro* germ cell differentiation during sex differentiation in a teleost fish. *Int J Dev Biol.* 54:105–111.

Kobayashi Y, Nozu R, Nakamura M. 2011. Role of estrogen in spermatogenesis in initial phase males of the three-spot wrasse (*Halichoeres trimaculatus*): effect of aromatase inhibitor on the testis. *Dev Dyn.* 240:116–121.

Kocher TD. 2004. Adaptive evolution and explosive speciation: the cichlid fish model. *Nat Rev Genet.* 5:288–298.

Kosakovsky Pond SL, Murrell B, Fourment M, Frost SDW, Delport W, Scheffler K. 2011. A random effects branch-site model for detecting episodic diversifying selection. *Mol Biol Evol.* 28:3033–3043.

Kustan JM, Maruska KP, Fernald RD. 2011. Subordinate male cichlids retain reproductive competence during social suppression. *Proc Biol Sci.* 279:434–443.

Laird PW, Zijderveld A, Linders K, Rudnicki MA, Jaenisch R, Berns A. 1991. Simplified mammalian DNA isolation procedure. *Nucleic Acids Res.* 19:4293.

Liu J-F, Guiguen Y, Liu S-J. 2009. Aromatase (P450arom) and 11β-hydroxylase (P45011β) genes are differentially expressed during the sex change process of the protogynous rice field eel, *Monopterus albus*. *Fish Physiol Biochem.* 35:511–518.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25:402–408.

Mank JE, Promislow DEL, Avise JC. 2006. Evolution of alternative sex-determining mechanisms in teleost fishes. *Biol J Linn Soc Lond.* 87: 83–93.

Marín I, Baker SB. 1998. The evolutionary dynamics of sex determination. *Science* 281:1990–1994.

Maruska KP, Fernald RD. 2011. Plasticity of the reproductive axis caused by social status change in an African cichlid fish: II. Testicular gene expression and spermatogenesis. *Endocrinology* 152:291–302.

Matsuda M, Nagahama Y, Shinomiya A, et al. (13 co-authors). 2002. *DMY* is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature* 417:559–563.

Matsuda M, Sato T, Toyazaki Y, Nagahama Y, Hamaguchi S, Sakaizumi M. 2003. *Oryzias curvinotus* has *DMY*, a gene that is required for male development in the medaka, *O. latipes*. *Zool Sci.* 20:159–161.

Mayor C, Brudno M, Schwartz JR, Poliakov A, Rubin EM, Frazer KA, Pachter LS, Dubchak I. 2000. VISTA: visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* 16:1046.

Merchant-Larios H, Díaz-Hernández V. 2013. Environmental sex determination mechanisms in reptiles. *Sex Dev.* 7:95–103.

Meyer A, Schartl M. 1999. Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol.* 11:699–704.

Morais da Silva S, Hacker A, Harley V, Goodfellow P, Swain A, Lovell-Badge R. 1996. *Sox9* expression during gonadal development implies a conserved role for the gene in testis differentiation in mammals and birds. *Nat Genet.* 14:62–68.

Morgan C, Loughran N, Walsh T, Harrison A, O'Connell M. 2010. Positive selection neighboring functionally essential sites and disease-implicated regions of mammalian reproductive proteins. *BMC Evol Biol.* 10:39.

Muschick M, Indermaur A, Salzburger W. 2012. Convergent evolution within an adaptive radiation of cichlid fishes. *Curr Biol.* 22: 2362–2368.

Myosho T, Otake H, Masuyama H, Matsuda M, Kuroki Y, Fujiyama A, Naruse K, Hamaguchi S, Sakaizumi M. 2012. Tracing the emergence of a novel sex-determining gene in medaka, *Oryzias luzonensis*. *Genetics* 191:163–170.

Nakagawa T, Iwabuchi J. 2012. Brain-specific promoter/exon I.f of the *cyp19a1* (aromatase) gene in *Xenopus laevis*. *J Steroid Biochem Mol Biol.* 132:247–255.

Nakamura M. 2010. The mechanism of sex determination in vertebrates—are sex steroids the key-factor? *J Exp Zool A Ecol Genet Physiol.* 313:381–398.

Nanda I, Kondo M, Hornung U, et al. (12 co-authors). 2002. A duplicated copy of *DMRT1* in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*. *Proc Natl Acad Sci U S A.* 99: 11778–11783.

Nozu R, Kojima Y, Nakamura M. 2009. Short term treatment with aromatase inhibitor induces sex change in the protogynous wrasse, *Halichoeres trimaculatus*. *Gen Comp Endocrinol.* 161:360–364.

Ohmuro-Matsuyama Y, Okubo K, Matsuda M, Ijiri S, Wang D, Guan G, Suzuki T, Matsuyama M, Morohashi K-i, Nagahama Y. 2007. Liver receptor homologue-1 (LRH-1) activates the promoter of brain aromatase (*cyp19a2*) in a teleost fish, the medaka, *Oryzias latipes*. *Mol Reprod Dev.* 74:1065–1071.

Ohno S. 1970. Evolution by gene duplication. New York: Springer Verlag.

Pezzi V, Sirianni R, Chimento A, Maggiolini M, Bourguiba S, Delalande C, Carreau S, Andò S, Simpson ER, Clyne CD. 2004. Differential expression of steroidogenic factor-1/adrenal 4 binding protein and liver receptor homolog-1 (LRH-1)/fetoprotein transcription factor in the rat testis: LRH-1 as a potential regulator of testicular aromatase expression. *Endocrinology* 145:2186–2196.

Piferrer F, Blázquez M. 2005. Aromatase distribution and regulation in fish. *Fish Physiol Biochem.* 31:215–226.

Pizzari T, Birkhead TR. 2002. The sexually-selected sperm hypothesis: sexbiased inheritance and sexual antagonism. *Biol Rev.* 77:183–209.

Pond SLK, Frost SDW. 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21: 2531–2533.

Pond SLK, Frost SDW, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21:676–679.

Reddon AR, Hurd PL. 2013. Water pH during early development influences sex ratio and male morph in a West African cichlid fish, *Pelvicachromis pulcher*. *Zoology (Jena)* 116:139–143.

Renn SCP, Fraser EJ, Aubin-Horth N, Trainor BC, Hofmann HA. 2012. Females of an African cichlid fish display male-typical social dominance behavior and elevated androgens in the absence of males. *Horm Behav.* 61:496–503.

Rice WR, Holland B. 1997. The enemies within: intergenomic conflict, interlocus contest evolution (ICE), and the intraspecific Red Queen. *Behav Ecol Sociobiol.* 41:1–10.

Rivas M, Mellström B, Naranjo JR, Santisteban P. 2004. Transcriptional repressor DREAM interacts with thyroid transcription factor-1 and regulates thyroglobulin gene expression. *J Biol Chem.* 279: 33114–33122.

Roberts RB, Ser JR, Kocher TD. 2009. Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. *Science* 326: 998–1001.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol.* 61:539–542.

Salzburger W, Braasch I, Meyer A. 2007. Adaptive sequence evolution in a color gene involved in the formation of the characteristic egg-dummies of male haplochromine cichlid fishes. *BMC Biol.* 5:51.

Salzburger W, Mack T, Verheyen E, Meyer A. 2005. Out of Tanganyika: genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. *BMC Evol Biol.* 5:17.

Salzburger W, Meyer A, Baric S, Verheyen E, Sturmbauer C. 2002. Phylogeny of the Lake Tanganyika cichlid species flock and its relationship to the Central and East African Haplochromine cichlid fish faunas. *Syst Biol.* 51:113–135.

Scholz S, Klüver N. 2009. Effects of endocrine disrupters on sexual, gonadal development in fish. *Sex Dev.* 3:136–151.

Schulz RW, de França LR, Lareyre J-J, LeGac F, Chiarini-Garcia H, Nobrega RH, Miura T. 2010. Spermatogenesis in fish. *Gen Comp Endocrinol.* 165:390–411.

Simon P. 2003. Q-Gene: processing quantitative real-time RT–PCR data. *Bioinformatics* 19:1439–1440.

Smith C, Shoemaker C, Roeszler K, Queen J, Crews D, Sinclair A. 2008. Cloning and expression of R-Spondin1 in different vertebrates suggests a conserved role in ovarian development. *BMC Dev Biol.* 8:72.

Snoeks J. 2001. Cichlid diversity, speciation and systematics: examples from the Great African Lakes. *J Aquaric Aquat Sci.* 9:150–166.

Sobrinho I, de Brito R. 2010. Evidence for positive selection in the gene *fruitless* in *Anastrepha* fruit flies. *BMC Evol Biol.* 10:293.

Soma KK, Francis RC, Wingfield JC, Fernald RD. 1996. Androgen regulation of hypothalamic neurons containing gonadotropin-releasing hormone in a cichlid fish: integration with social cues. *Horm Behav.* 30:216–226.

Sorhannus U, Kosakovsky Pond S. 2006. Evidence for positive selection on a sexual reproduction gene in the diatom genus *Thalassiosira* (Bacillariophyta). *J Mol Evol.* 63:231–239.

Stern A, Doron-Faigenboim A, Erez E, Martz E, Bacharach E, Pupko T. 2007. Selecton 2007: advanced models for detecting positive and purifying selection using a Bayesian inference approach. *Nucleic Acids Res.* 35:W506–W511.

Swanson WJ, Nielsen R, Yang Q. 2003. Pervasive adaptive evolution in mammalian fertilization proteins. *Mol Biol Evol.* 20:18–20.

Toda K, Shizuta Y. 1993. Molecular cloning of a cDNA showing alternative splicing of the 5′-untranslated sequence of mRNA for human aromatase P-450. *Eur J Biochem.* 213:383–389.

Tong S-K, Chung BC. 2003. Analysis of zebrafish *cyp19* promoters. *J Steroid Biochem Mol Biol.* 86:381–386.

Trainor BC, Kyomen HH, Marler CA. 2006. Estrogenic encounters: how interactions between aromatase and the environment modulate aggression. *Front Neuroendocrinol.* 27:170–179.

Tripathi N, Hoffmann M, Weigel D, Dreyer C. 2009. Linkage analysis reveals the independent origin of Poeciliid sex chromosomes and a case of atypical sex inheritance in the guppy (*Poecilia reticulata*). *Genetics* 182:365–374.

True JR, Haag ES. 2001. Developmental system drift and flexibility in evolutionary trajectories. *Evol Dev.* 3:109–119.

Tzchori I, Degani G, Hurvitz A, Moav B. 2004. Cloning and developmental expression of the cytochrome P450 aromatase gene (*CYP19*) in the European eel (*Anguilla anguilla*). *Gen Comp Endocrinol.* 138: 271–280.

Uller T, Helantera H. 2011. From the origin of sex-determining factors to the evolution of sex-determining systems. *Q Rev Biol.* 86:163–180.

Vacquier VD, Swanson WJ. 2011. Selection in the rapid evolution of gamete recognition proteins in marine invertebrates. *Cold Spring Harb Perspect Biol.* 3:a002931.

Valenzuela N, Neuwald JL, Literman R. 2013. Transcriptional evolution underlying vertebrate sexual development. *Dev Dyn.* 242:307–319.

Verheyen E, Salzburger W, Snoeks J, Meyer A. 2003. Origin of the superflock of cichlid fishes from Lake Victoria, East Africa. *Science* 300: 325–329.

Wang D-S, Kobayashi T, Zhou L-Y, Paul-Prasanth B, Ijiri S, Sakai F, Okubo K, Morohashi K-i, Nagahama Y. 2007. Foxl2 up-regulates aromatase gene transcription in a female-specific manner by binding to the promoter as well as interacting with Ad4 binding protein/steroidogenic factor 1. *Mol Endocrinol.* 21:712–725.

Wang D-S, Zhou L-Y, Kobayashi T, Matsuda M, Shibata Y, Sakai F, Nagahama Y. 2010. Doublesex- and mab-3-related transcription factor-1 repression of aromatase transcription, a possible mechanism favoring the male pathway in tilapia. *Endocrinology* 151: 1331–1340.

Wang X, Sommer RJ. 2011. Antagonism of LIN-17/frizzled and LIN-18/Ryk in nematode vulva induction reveals evolutionary alterations in core developmental pathways. *PLoS Biol.* 9:e1001110.

Wilkins A. 1995. Moving up the hierarchy: a hypothesis on the evolution of a genetic sex determination pathway. *BioEssays* 17: 71–77.

Yang Z. 2007. PAML 4: a program package for phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 24:1586–1591.

Yang Z, Nielsen R. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol Biol Evol.* 19:908–917.

Yang Z, Nielsen R, Goldman N, Pedersen AM. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155:431–449.

Yang Z, Wong WSW, Nielsen R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol Biol Evol.* 22: 1107–1118.

Yano A, Guyomard R, Nicol B, Jouanno E, Quillet E, Klopp C, Cabau C, Bouchez O, Fostier A, Guiguen Y. 2012. An immune-related gene evolved into the master sex-determining gene in rainbow trout, *Oncorhynchus mykiss*. *Curr Biol.* 22:1423–1428.

Yoshida K, Terai Y, Mizoiri S, et al. (11 co-authors). 2011. B chromosomes have a functional effect on female sex determination in Lake Victoria cichlid fishes. *PLoS Genet.* 7:e1002203.

Yoshiura Y, Senthilkumaran B, Watanabe M, Oba Y, Kobayashi T, Nagahama Y. 2003. Synergistic expression of Ad4BP/SF-1 and cytochrome P-450 Aromatase (ovarian type) in the ovary of Nile tilapia, *Oreochromis niloticus*, during vitellogenesis suggests transcriptional interaction. *Biol Reprod.* 68:1545–1553.

Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol.* 22:2472–2479.

Zhang Y, Li F, Sun D, Liu J, Liu N, Yu Q. 2011. Molecular analysis shows differential expression of R-spondin1 in zebrafish (*Danio rerio*) gonads. *Mol Biol Rep.* 38:275–282.

Zhang Z, Li J, Zhao XQ, Wang J, Wong GK, Yu J. 2006. KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. *Genomics Proteomics Bioinformatics* 4:259–263.

Zhou R, Liu L, Guo Y, Yu H, Cheng H, Huang X, Tiersch TR, Berta P. 2003. Similar gene structure of two *Sox9a* genes and their expression patterns during gonadal differentiation in a teleost fish, rice field eel (*Monopterus albus*). *Mol Reprod Dev.* 66:211–217.

Chapter 4

Genetics and timing of sex determination in the East African cichlid fish *Astatotilapia burtoni*

Corina Heule, Carolin Göppert, Walter Salzburger and Astrid Böhne

CH, WS and AB designed the study and wrote the manuscript. CH and AB performed hormone treatments and crossings. CH performed the qRT-PCR expression and *wnt4B* coding region analysis. CH and AB conducted the promoter and transcription factor binding site analysis. CG sequenced *wt1A* coding and promoter regions. All authors read and approved the final manuscript.

Published in *BMC Genetics*

Heule et al. BMC Genetics _####################_
DOI 10.1186/s12863-014-0140-5

BMC Genetics

RESEARCH ARTICLE Open Access

Genetics and timing of sex determination in the East African cichlid fish *Astatotilapia burtoni*

Corina Heule, Carolin Göppert, Walter Salzburger and Astrid Böhne*

Abstract

Background: The factors determining sex are diverse in vertebrates and especially so in teleost fishes. Only a handful of master sex-determining genes have been identified, however great efforts have been undertaken to characterize the subsequent genetic network of sex differentiation in various organisms. East African cichlids offer an ideal model system to study the complexity of sexual development, since many different sex-determining mechanisms occur in closely related species of this fish family. Here, we investigated the sex-determining system and gene expression profiles during male development of *Astatotilapia burtoni*, a member of the rapidly radiating and exceptionally species-rich haplochromine lineage.

Results: Crossing experiments with hormonally sex-reversed fish provided evidence for an XX-XY sex determination system in *A. burtoni*. Resultant all-male broods were used to assess gene expression patterns throughout development of a set of candidate genes, previously characterized in adult cichlids only.

Conclusions: We could identify the onset of gonad sexual differentiation at 11–12 dpf. The expression profiles identified *wnt4B* and *wt1A* as the earliest gonad markers in *A. burtoni*. Furthermore we identified late testis genes (*cyp19a1A*, *gsdf*, *dmrt1* and *gata4*), and brain markers (*ctnnb1A*, *ctnnb1B*, dax1A, *foxl2*, *foxl3*, *nanos1A*, *nanos1B*, *rspo1*, *sf-1*, *sox9A* and *sox9B*).

Keywords: Sexual development, Cichlidae, Adaptive radiation, Speciation, Gene expression profiles

Background

Sexual development encompasses sex determination and sex differentiation and can be viewed as a complex genetic network that is initiated by a sex-determining trigger mediating the expression of sex differentiation genes, which ultimately establish the male or female phenotype [1]. In teleost fishes, with over 25,000 species the largest vertebrate group, sex determination mechanisms are much more variable compared to other vertebrates [2]. So far, six master sex-determining genes have been identified in teleosts, namely *dmy/dmrt1bY* in *Oryzias latipes* and *O. curvinotus* [3,4], $gsdf^{Y}$ in *O. luzonensis* [5], *sox3* in *O. dancena* [6], *amhy* in *Odontesthes hatcheri* [7], *amhr2* in *Takifugu rubripes* [8] and *sdY* in *Oncorhynchus mykiss* and several other salmonids [9,10]. In addition to this variation in the initial regulators, we and others could show recently that also the subsequent genetic steps of sex differentiation are not conserved in fishes, asking for further investigation of the mechanisms of sexual development in this group of animals [11,12].

Master sex-determining genes are thought to be expressed early in development, thus marking the initial time point of the sexual development cascade. Their expression then either decreases directly after (comparable to the expression pattern shown in Figure 1A and in particular described for *dmy/dmrt1bY* in *O. latipes* [13]) or is maintained during the juvenile stage (as suggested for *amhy* [7] and *sdY* [9]). To the best of our knowledge, there is no example of a sex determination gene that is still highly expressed in adult fish. However, expression studies on several fish sex determination genes covering the development from embryo to adults are lacking, and in mammals, the sex-determining gene *sry* is expressed in adult testis of mouse and rat [14,15].

Sex differentiation genes, on the other hand, can act at different time points after their initiation until sexual maturity (i.e., until gonads are fully developed) or even afterwards, e.g., by being involved in gonad maintenance and function (Figure 1B and exemplified by *dmrt1* [16-18]).

* Correspondence: astrid.boehne@unibas.ch
Zoological Institute, University of Basel, Vesalgasse 1, 4051 Basel, Switzerland

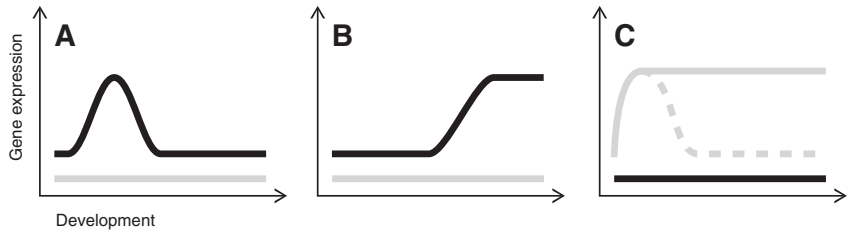

Figure 1 Schematic expression patterns of sexual development genes. The graphs show possible expression profiles post fertilization in the developing brain (grey line) and testis (black line). **(A)** Early testis genes (including sex-determining genes) are highly expressed before and/or at the onset of gonadal formation and subsequently down regulated. **(B)** Late testis genes are expressed later in development, mainly during the formation and maintenance of gonads. **(C)** Brain genes are higher expressed in the brains than in the testis, forming the brain and/or influencing the sexual development gene network via the action of hormones. Their expression can be maintained (continuous line) or decreased (dashed line) after the first increase in expression.

Similarly to gene expression patterns in the gonads, sex differentiation genes can be expressed in the brain as part of the hypothalamus-pituitary-gonadal axis, and hence can -like gonad genes- follow one of the two patterns shown in Figure 1C.

In general, gonads are formed by the interplay of sexual development genes and the action of hormones [19-22]. This can be a rather plastic process, especially in fish, making it more difficult to classify sex differentiation genes according to their expression profiles and also questioning a separation between sex determination and differentiation [23].

Cichlid fishes, and the species flocks of cichlids in the East African Great lakes in particular, are an excellent model system in evolutionary biology, with hundreds of closely related species showing a high degree of diversity in morphology, behavior and ecology [24-27]. This diversity also seems to apply to sex determination systems, as evidenced by data suggesting that different mechanisms occur in cichlids including sex determination via environmental (temperature and pH) and genetic factors (single gene or polygenic actions), or a combination thereof [28-33]. The best-studied cichlid in terms of sexual development is the widely distributed and farmed Nile tilapia (*Oreochromis niloticus*), which has an XX-XY sex-determining system that can strongly be influenced by temperature [34]. There are two time windows (2–3 days post fertilization, dpf, and 10–20 dpf), in which temperature and steroid hormones can override genetic sex determination in the Nile tilapia, with the actual critical time period of gonad differentiation at 9 to 15 dpf [34 and references therein]. Studies of sexual development in the Nile tilapia encompass both, genetic and morphological data, and therefore make this species a good reference system.

Here, we focused on another cichlid species, *Astatotilapia burtoni,* which inhabits Lake Tanganyika, and its affluent rivers, and is a model system especially in behavioral but also genetic research (e.g., [35]). This sexually dimorphic species, in which males are larger and brightly colored whereas females are rather dull, belongs to the most derived and species-rich lineage of East African cichlids, the haplochromines. Like the Nile tilapia, *A. burtoni* is a maternal mouthbrooder; the female incubates the fertilized eggs in her buccal cavity at least until hatching. Because of different developmental pace, the sexual development of *A. burtoni* cannot be compared in exact (day to day) time steps to the Nile tilapia. Although Nile tilapia and *A. burtoni* embryos hatch approximately at the same age (5–6 dpf [36] and 4–7 dpf, [37], respectively), Nile tilapia embryos start free swimming earlier than *A. burtoni* embryos (12 and 14 dpf, respectively [36,37]) but become sexually mature later (at the age of 22–24 weeks [38] compared to 13–14 weeks in the here used *A. burtoni* strain, personal observation). Until now, the embryonic and juvenile development of *A. burtoni* has not been studied in detail. Even though *A. burtoni* is one of the five cichlid species with a sequenced genome [39], neither the sex-determining system nor the time window of sex determination have been characterized.

Based on the assumption that sex is determined genetically, we used a common approach to infer male or female heterogamety. We generated mono-sex fish groups over steroid hormone treatments via food and conducted crossing experiments. The resultant sex ratios point to an XX-XY sex-determining system in *A. burtoni*. Subsequent crossings were carried out to generate a YY-supermale to sire male-only offspring. Making use of candidate genes expressed in brain and gonad tissue of adult *A. burtoni* [11], we studied changes in gene expression throughout male sexual development. Without prior knowledge on the time window of actual sex determination in this species, we decided to investigate gene expression as early as possible starting at 7 dpf. We profiled expression of sexual development genes from 7–48 dpf using high throughput quantitative real-time polymerase chain reaction on single individuals. Most of the gene expression

profiles corresponded to one of the following patterns: early testis genes, late testis genes and brain/head genes (Figure 1).

Results

Generating all-male broods in *A. burtoni*

Sexual development in fish is plastic and sex reversal can be induced in a variety of species even after reaching sexual maturity [40]. For these purposes, steroid hormones or hormone synthesis inhibitors can be administered over the surrounding water or via food supply. Here, we fed four *A. burtoni* broods with estrogen treated flake food during four weeks of development in order to obtain all-female broods. We started treatment at the earliest feeding point of this species, at around 14 dpf. This procedure has been carried out successfully in another cichlid species, the Nile tilapia (personal communication H D'Cotta), which starts feeding at around 12 dpf [36]. After treatment, we obtained 100% morphological females in all broods. These natural female and feminized fish were used for crossings with untreated, normal males. Among the offspring of these individual crossings, four broods showed a ~ 1 : 3 (female : male) sex ratio, whereas other crosses, likely derived from normal females, which can morphologically not be distinguished from sex-reversed individuals, had a sex ratio of approximately 1 : 1. This is a strong indication for an XY-XX system in *A. burtoni* (Figure 2). Note that a ZZ-ZW female heterogametic sex determination system can be ruled out for *A. burtoni*, because sex-reversed ZZ females would have produced only males in the first generation of crossings, all of our crosses however contained at least 1/3 female offspring.

Crossings of sex-reversed XY fish (phenotypic females) with normal, XY-males should lead to the following types and proportion of offspring: one quarter of XX-females, two quarters of XY-males and one quarter of YY-males (super-males) (Figure 2). Note that, morphologically, the two types of males should be undistinguishable.

Subsequent crossings of all males of one of the broods with a 1:3 sex ratio to normal females revealed one male that only produced male offspring, suggesting that it is indeed a YY-male, lending further support to an XX/XY sex determination system in this species.

Expression profiles of sexual development genes

We crossed the YY-super-male to XX-females to produce all-male broods, which we used to investigate

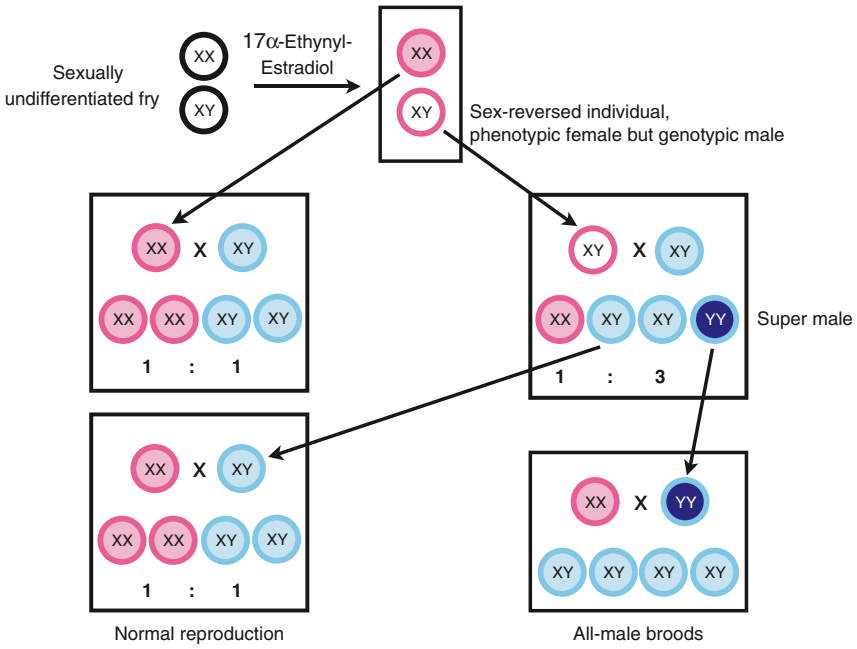

Figure 2 Crossing scheme to obtain all-male broods from estrogen sex-reversed fish. Sexually undifferentiated fry including both, XX- and XY-genotypes, were treated with estrogen resulting in only phenotypical females. These females – genetic females and sex-reversed genotypic males – were then crossed to untreated genotypic males. Crossing of XX-females to untreated males (left site) reflect the normal reproduction with a sex ratio of 1:1, with the corresponding female XX and male XY genotypes in the offspring. Crossing of sex-reversed XY-females to untreated XY-males led to a sex ratio of 1:3 with the genotypes XX (phenotype female), XY (phenotype male), YY (super male, phenotype male). These two types of phenotypically undistinguishable males were back-crossed to normal XX-females resulting again in either a 1:1 sex ratio (for XY-males) or in all-male broods (for the YY-male). Pink and blue outer circles denote phenotypic females and males, respectively.

expression patterns of sex differentiation genes during early male development. In similar experiments in the Nile tilapia, the spurious occurrence of females in the offspring of super-males has been reported [41]. To allow a potential detection of such spontaneously occurring phenotypic females in these broods, gene expression was measured in individual samples rather than pooling samples. To our knowledge, this is the first study that used a large number of individual samples in a dense sampling scheme for establishing the gene expression profiles of a set of candidate genes for sexual development (24 genes tested in 88 individuals sampled at 22 time points during a period of 40 days). Fish were dissected from the yolk and separated in head and trunk, as proxies for developing brain and gonad. Single organ dissection is not possible at these early stages of development, especially if gene expression is to be accessed on an individual basis. The chosen approach has already successfully been applied in other species [5,7,42-47].

The relative expression of a set of candidate genes, previously tested in brain and gonad tissue of adult cichlid fishes [11], plus one additional gene, *gsdf*, was profiled during male development. These genes are candidates for sex determination and differentiation as suggested by their described function in fish and tetrapods. This gene list includes, wherever existing, the two paralogous gene copies emerging from the fish-specific whole-genome duplication [48].

The brain and the gonads are the main tissues acting in sexual development. In addition, sexually dimorphic expression can be observed in the brain even earlier than in the gonad, a pattern already described in cichlids [43,49]. Samples were taken between 7 and 48 dpf, with a daily sampling at the beginning of the experiment (7 – 20 dpf) and then every third (during 20 – 38 dpf) and afterwards every fifth day (38 – 48 dpf,) as day-to-day changes are more prominent early in development [36]. We then used the Fluidigm system to test the expression of the 24 candidate genes. Gene expression was calculated as fold change in gene expression using the delta-delta-CT method [50], compared to expression in a juvenile tissue pool (Figure 3 and Additional file 1) or relative to the mean of the four biological replicates at the first sampling point at 7 dpf (Additional file 2). For each sampling point the fold change in gene expression in heads and trunks of four individuals was calculated. For details on sample sizes for each gene see Additional file 3.

The expression profile of a known testis-specific gene (*dmrt1*) in all tested trunks strongly suggests that all individuals were indeed males and that none of the offspring was a female. In addition, we raised fish that were not used for the gene expression experiment to adulthood/maturity and confirmed that all of them were males. We hence did not detect any occurrence of spurious females.

We investigated gene expression patterns according to the expression profiles explained in Figure 1 and compared expression between heads and trunks. Figure 3 shows the most prominent examples for the expression profiles early testis genes, late testis genes and (early) brain genes (for all expression profiles see Additional files 1 and 2). In the following, we describe the results in more detail.

Testis and brain markers

From all 24 candidate genes, only *wnt4B* and *wt1A* are likely to represent early testis genes, i.e., showing a peak in expression early in development and in trunks only (Figure 3A, corresponding to the profile shown in Figure 1A). *Cyp19a1A*, *gsdf* and *dmrt1* appeared as late testis genes with an increase in trunk expression over time (Figure 3B, corresponding to the profile shown in Figure 1B). *Gata4* showed a similar increase in expression in trunks starting earlier as the other genes, around 15 dpf (see Additional file 1). In total, we detected 12 'brain' genes (*ctnnb1A*, *ctnnb1B*, *cyp19a1B*, *dax1A*, *foxl2A/foxl2*, *foxl2B*, *nanos1A*, *nanos1B*, *rspo1*, *sf-1*, *sox9A* and *sox9B*). For illustration purposes, we show the results for both gene copies of *wnt4*, *wt1* and *cyp19a1* in Figure 3.

Wnt4A and *wnt4B* – different fates for gene copies

Wnt4A showed higher expression levels in heads than in trunks, whereas *wnt4B* showed the opposite signature with a higher expression in trunks than in heads. Also in adult males, *wnt4A* is significantly higher expressed in brain compared to testis tissue [11]. In adult cichlids, there is a detectable difference in gene expression between the two paralogs of *wnt4*, with the A-copy being ovary- and the B-copy being testis-specific [11]. W*nt4B* was one of only two genes with the earliest peak of expression in trunks (7 – 15 dpf), resembling the pattern of a sex-determining gene.

Wt1A and *wt1B* – testis genes with different temporal patterns

Wt1A and *wt1B* are both higher expressed in trunks than in heads throughout the experimental time period, which is congruent with the pattern observed in adult males of *A. burtoni* [11]. *Wt1A* is the second gene that showed an expression peak in trunks at the beginning of development (between 7 and 15 dpf) but in contrast to *wnt4B* at the same time point also an increase of expression in heads (Figure 3A).

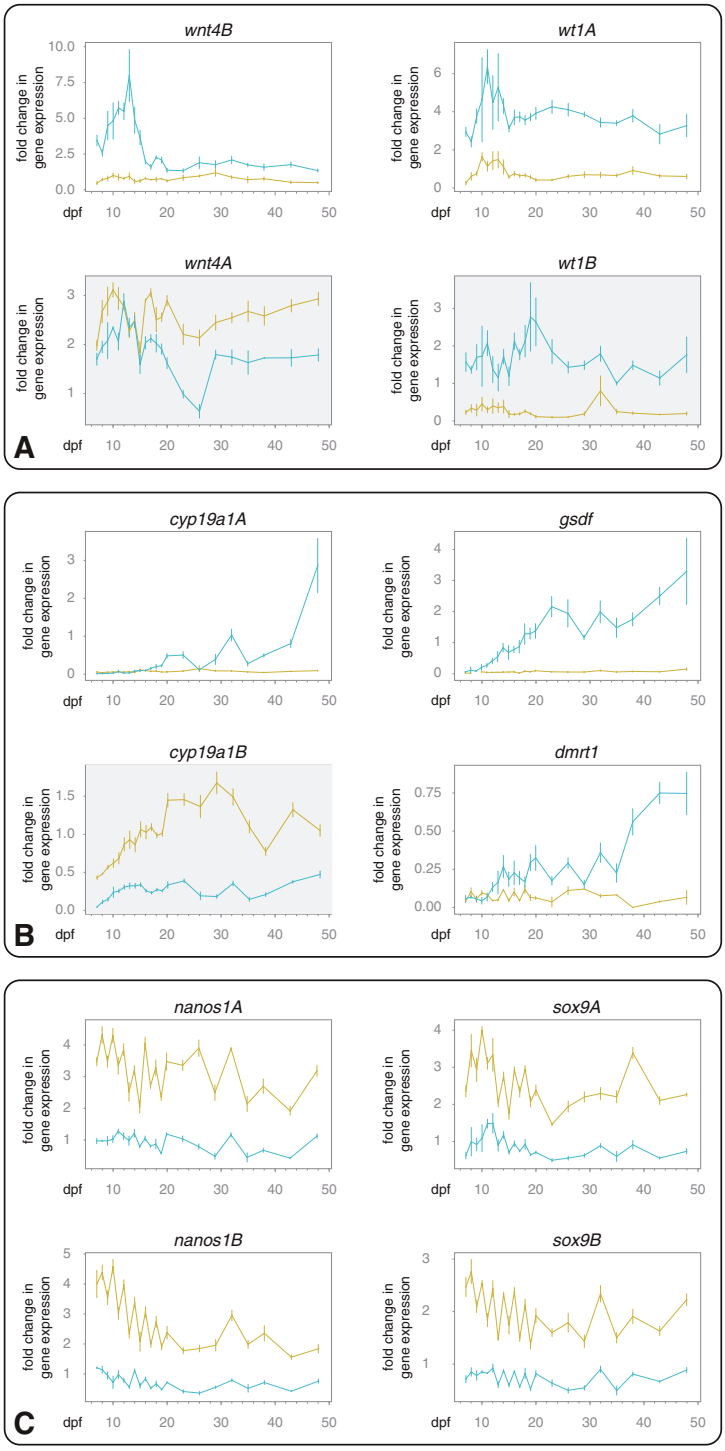

Figure 3 Gene expression of sexual development genes in heads and trunks of developing male *A. burtoni*. **(A)** *Wnt4B* and *wt1A* were the only detected early testis genes, here shown with their paralogous gene copies *wnt4A* and *wt1B* (grey background). **(B)** *Cyp19a1A*, *gsdf* and *dmrt1* are examples of late testis genes, *cyp19a1B* is the teleost specific paralog of *cyp19a1A* (grey background). **(C)** *Nanos1A*, *nanos1B*, *sox9A* and *sox9B* are examples for brain genes. Gene expression is shown as fold change (Livak) ± SE in heads (green) and trunks (blue) from 7 – 48 dpf using *rpl7* as reference gene and a juvenile tissue mix as reference tissue (see Additional file 3 for further details).

Dmrt1 and *gsdf* - late testis genes possibly important for gonad maintenance

Dmrt1 is known as *the* conserved vertebrate testis gene [51] and also shows testis-specificity in adult *A. burtoni* [11]. We found similar levels of gene expression in heads and trunks early in development (7 – 11 dpf) followed by an increase (12 – 48 dpf) in expression in trunks only, pointing to a later function in testis development (Figure 3B). In many of the head samples *dmrt1* expression could not be detected (see Additional file 3 for details), which is consistent with previous results in adult brains [11].

Gsdf (*gonadal soma-derived factor*) is a sexual development gene only existing in fish [52], which has received considerable attention recently. In the above-mentioned *O. luzonensis*, Y- and X-chromosome specific alleles have been identified for this gene ($gsdf^Y$and $gsdf^X$, respectively), with the former turning out to be the master sex determiner in this species [5]. In another species, the sablefish *Anoplopoma fimbria*, *gsdf* seems to be a strong candidate for the sex-determining locus, too [53]. Furthermore in medaka, *gsdf* expression has been implicated with early testicular differentiation [54].

In *A. burtoni* the expression profile of *gsdf* resembled that of *dmrt1*, with a constant increase of expression in trunks after a short time of low expression (7–10 dpf), and constant low expression in heads (Figure 3B). Just as for *dmrt1*, in some of the head samples, *gsdf* expression could not be detected (see Additional file 3 for details).

The aromatases *cyp19a1A* and *cyp19a1B*

The expression pattern of the aromatase *cyp19a1A* in the heads remained similar over time whereas its expression in trunks increased constantly. The expression of *cyp19a1B* was always higher in heads than in trunks, with an increase in expression in both tissues during 7 – 11 dpf, followed by a stable period (12 – 43 dpf), and then the expression in trunks increased again (48 dpf). The expression pattern of *cyp19a1A* in adults of *A. burtoni* in brain and gonad tissue shows no difference, and the expression pattern of *cyp19a1B* shows a significant testis-specific over-expression [11]. In developing *A. burtoni* males, *cyp19a1A* seems to play a role in the gonads. The testis-specific expression of *cyp19a1B* seen in adults only becomes established after 48 dpf, with a start of rising expression detected in our experiments after 40 dpf.

Markers of the developing brain

As mentioned above, we detected 12 'brain' genes. The strongest differences in expression between heads and trunks, and hence likely representing brain up-regulated genes, were found for *nanos1A*, *nanos1B*, *sox9A* and *sox9B* (Figure 3C). This is consistent with the expression patterns seen in adult males of *A. burtoni*, where a significantly higher expression in brain tissue than in the testis has been found [11]. The expression level of *nanos1B* in heads was highest at 7 dpf and then decreased (comparable to Figure 1C, dashed line). *Sox9*, similar to *dmrt1*, is considered a prominent example for a gene generally involved in testis formation and function [55,56]. However, this does not seem to be the case in developing and adult *A. burtoni*.

Investigation of the early testis markers: Sequence and promoter analysis of *wnt4B* and *wt1A*

As the *wnt4B* and *wt1A* expression showed a peak early in development (7 – 15 dpf) and then decreased to a constantly low level, thus mimicking the expression of a potential sex determination gene, we decided to investigate these genes' sequences in detail in *A. burtoni*. For *wnt4B*, we sequenced the entire genic region, whereas for *wt1A* we focused on the coding region only, due to the large size of the region (~ 20 kb). A sequence comparison of the coding region of males and females did not show any allelic differences between the sexes for both genes. Also the intronic sequences of *wnt4B* did not show any sex-specific differences. However, gene expression could still be differently regulated due to sex-specific changes in the promoter region of the genes. To identify the potential promoter regions of *wnt4B* and *wt1A* we compared the upstream sequences of the two genes in the accessible teleost fish genomes using Vista plots of nucleotide similarity [57,58] (Figures 4 and 5). The 5' neighboring gene to *wnt4B* is *chd4b*, which is located ~13 kb upstream. We created VistaPlots comprising this entire region. The next annotated gene 5' of *wt1A* is more than 50 kb upstream. We thus decided to focus our analysis on the region 20 kb upstream to *wt1A*.

In an additional step, after *in silico* definition of a core conserved upstream region of *wnt4B* (see colored blocks in Figure 4), we sequenced ~ 7 kb of this promoter in *A. burtoni* males and females of our lab strain. We also obtained ~ 4 kb upstream sequence for *wt1A*. Again, no differences between the sexes were found in the upstream regions of *wnt4B* and *wt1A*. For *wt1A* we detected two alleles with one of them having a 223 bp deletion compared to the reference genome. However, neither the deletion nor any other detected heterozygous site segregated with sex.

Transcription factor binding-sites in *wnt4B* and *wt1A* potential promoters

To identify genes regulating *wnt4B* and *wt1A* expression and, thereby, possibly being more upstream in the sex-determining cascade, we performed a transcription factor binding-site analysis of the two conserved regions in *wnt4B* (blocks 1 and 2 in Figure 4) and the one conserved region

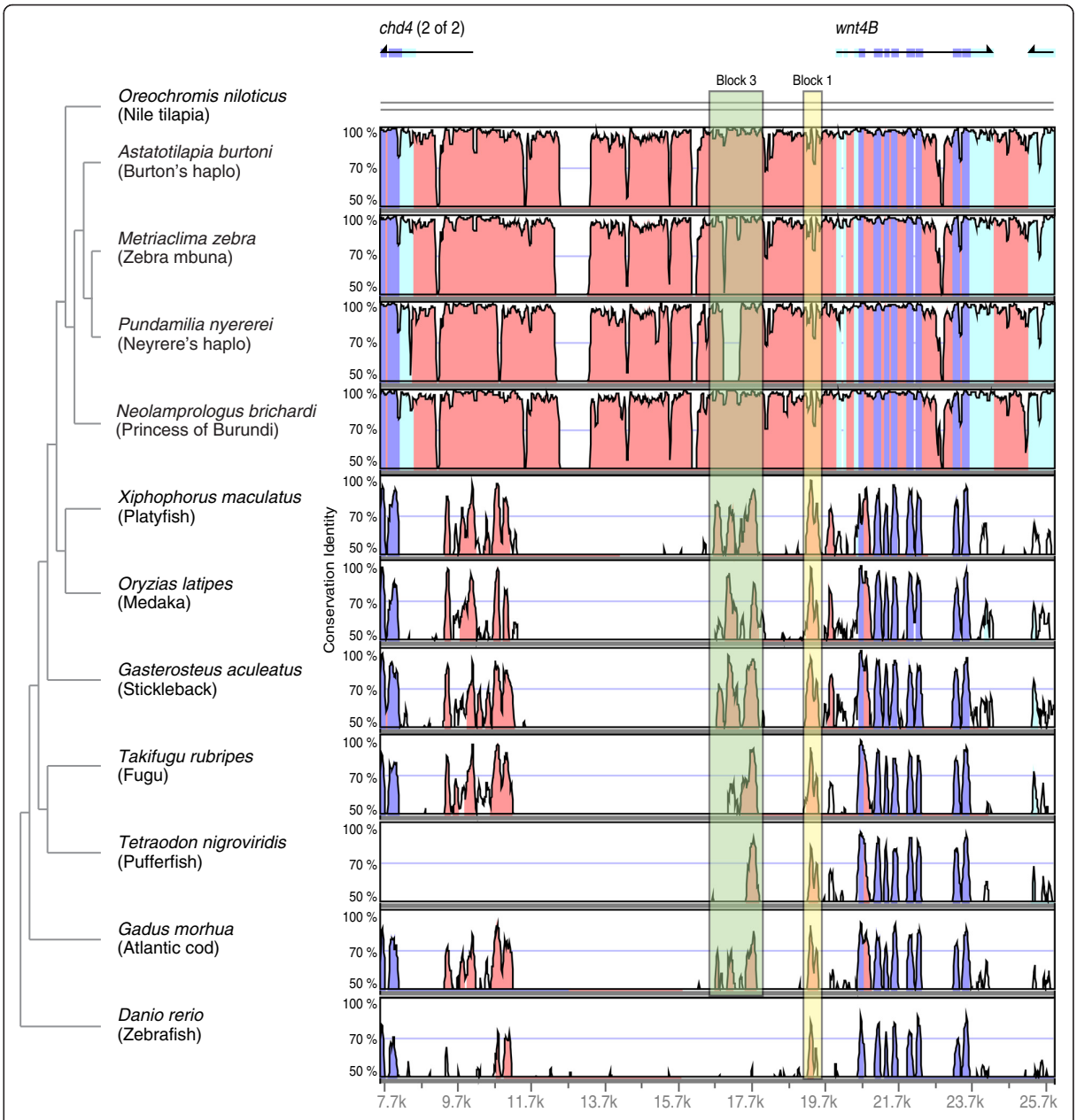

Figure 4 Comparison of the *wnt4B* upstream region. Shuffle-LAGAN Vista plots [57,58] for *wnt4B* and its 5' adjacent gene *chd4*. Peaks indicate conservation identity of sequences above 50% across the tested species. Blue stands for coding and pink for noncoding regions, respectively. Light blue regions represent UTRs. Yellow block 1 and green block 2 were investigated in the process of transcription factor binding site analysis (Table 1).

in *wt1A* (yellow block in Figure 5) using MatInspector. We focused on transcription factors with a described function in gonads, germ cells, brain and/or central nervous system and compared the putative binding sites of *A. burtoni* with the ones present in all other available fish genomes. Tables 1 and 2 show all putative binding-sites detected in the *A. burtoni* sequence and indicate, in which other species these sites have been detected (for a complete table with all putative transcription factor binding-sites including non-conserved sites in all tested species, see Additional files 4 and 5).

Interestingly, we identified several conserved binding sites for transcription factors that have been implicated with sexual development before. For *wnt4B* we found

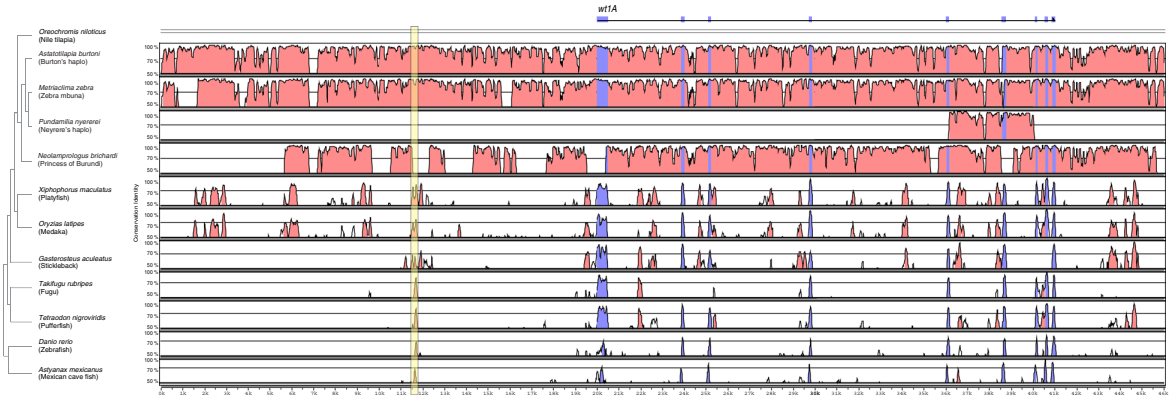

Figure 5 Comparison of the *wt1A* upstream region. Shuffle-LAGAN Vista plots [57,58] for *wt1A*. Peaks indicate conservation identity of sequences above 50% across the tested species. Blue stands for coding and pink for noncoding regions, respectively. The yellow block was investigated in the process of transcription factor binding site analysis (Table 2).

that six out of seven species show a conserved putative binding site for Wt1 in block 2 (Table 1). This fits well with our own expression data (Figure 3) as well as other studies in fish [59,60], which support an involvement of *wt1A* in early testis formation. Other promising upstream candidates of *wnt4B* are Sox30 and the androgen receptor (AR). Sox30 is expressed specifically in gonads of the Nile tilapia, with one isoform being even limited to the developing testis [61]. The androgen receptor can bind testosterone and dihydrotestosterone and thereby plays an important role in controlling male development [62]. Interestingly, *ar* is higher expressed in brains of dominant *A. burtoni* males than in subordinate males [63]. In the developing gonads of the Nile tilapia the expression levels of *ar* in males and females are similar [17].

Remarkably, we found putative transcription factor binding sites for two of our candidate genes: *wt1* (discussed above and Figure 3A) and *sf-1* (Additional file 1). However, the expression pattern of *sf-1* in developing testis (expression in trunks) does not support its putative role as a direct regulator of *wnt4B*, as it was expressed at low levels during the experimental time period (Additional file 1). The expression profiles in heads, on the other hand, showed high expression at the beginning (7 – 12 dpf), with a constant decrease afterwards (as in Figure 1C, dashed line; and Additional file 1). *Sf-1* might thus be an example of an early brain gene influencing sexual development via other factors than *wnt4B*.

In contrast to *wnt4B*, we could identify only one small conserved block upstream of *wt1A*. We did not find a binding-site for any of our candidate genes or an obvious transcription factor already known to play a role in sexual development or any binding site only present in *A. burtoni* in that block. However, we found a broad range of neuronal transcription factors and binding sites for members of the dm-domain family, here *dmrt2*, which might have a female sex-specific role in adult cichlids [64]. As for *wnt4B*, we also found a binding site for a Sox-family member, here Sox6.

Interestingly, we found binding sites for several members of the forkhead transcription factor family (Foxa1, Foxp1, Fkhrl1 alias Foxo3 and Foxp1), which are known as regulators of development and reproduction. Together with *foxl2* and *foxl3*, they were also among the candidate genes in our expression assay.

Discussion

Here we provide first experimental proof for a male sex-determining (XX-XY) system in the haplochromine cichlid *Astatotilapia burtoni,* making use of hormonal sex-reversal and the subsequent generation of mono-sex broods. Offspring from male-only broods were investigated for gene expression patterns to define the window of sex determination in *A. burtoni*, which seems to take place at 11–12 dpf.

Throughout larval development, we decided to investigate gene expression in whole heads and trunks, including also other tissues than brains and gonads. Similar studies have been conducted in the Nile tilapia, which revealed that expression of sexual development genes in brains and testis is comparable to the one in heads and trunks, respectively [42,43].

We chose this approach in order to assess the individual gene expression level rather than pooling samples. Furthermore, the timing of morphological development, especially of gonads but also brain structures, is unknown in *A. burtoni* and no marker of gonad differentiation is available for this species, making an early single tissue dissection physiologically and technically impossible. By using whole trunks we made sure that we did

Table 1 Predicted transcription factor binding sites in the *wnt4B* promoter region of *A. burtoni*

Block 1								Block 2						
A. burtoni	*X. maculatus*	*O. latipes*	*G. aculeatus*	*T. rubripes*	*T. nigroviridis*	*G. morhua*	*D. rerio*	*A. burtoni*	*X. maculatus*	*O. latipes*	*G. aculeatus*	*T. rubripes*	*T. nigroviridis*	*G. morhua*
AR								**AR**			x			
E4BP4	x		x	x	x	x	x	**Creb**	x	x	x	x	x	x
Ets1							x	**Creb1**	x	x				
Foxa-1 and 2	x	x	x			x	x	Dbp						
Foxk2	x			x			x	**Dec2**	x	x	x			x
Gsh2								**Egr1**		x	x			
Helt								**ESRRA**	x		x	x		x
Isl1								**Evi1**				x	x	x
Meis1	x	x	x	x	x	x	x	**FAC1**	x	x	x	x	x	x
Myt1	x	x	x	x		x	x	FoxC1						
Pax6								**Hmx2**					x	x
Plag1	x	x	x		x	x		**Hmx3**	x			x		
Satb1								**Hre**	x		x		x	x
Sox6	x	x	x			x		Ir1						
Spi1		x	x	x	x	x	x	**Irx5**	x	x				
Tef	x		x	x	x	x	x	**Mef3**	x					
								Meis1			x	x	x	x
								MEL1	x		x		x	x
								Myf5						x
								Myf6		x		x	x	x
								Nanog	x	x		x	x	x
								NF1	x	x	x	x		x
								Nkx2-5	x		x			x
								Nkx6-1						x
								Nur-Family			x	x		
								Pax4	x		x			x
								Plag1	x	x	x	x	x	x
								Pou5f1	x	x	x			x
								Rfx4	x		x			
								Rfx7	x					
								SF1			x			x
								Sox3	x	x	x	x	x	x
								Sox30		x			x	

Table 1 Predicted transcription factor binding sites in the *wnt4B* promoter region of *A. burtoni* *(Continued)*

	Sox6	x	x	x	x	x	x
	Ttf1						
	Wt1	x	x	x	x		x
	Ybx1						
	Zfp67	x	x				x
	Zic2					x	x

Blocks correspond to the green and yellow regions in Figure 4. Bold binding sites are shared with at least one other species. "x" denotes the detection of the binding site in the respective species.

Table 2 Predicted transcription factor binding-sites in the *wt1A* promoter region of *A. burtoni*

A. burtoni	*M. zebra*	*N. brichardi*	*O. niloticus*	*X. maculatus*	*O. latipes*	*G. aculeatus*	*T. rubripes*	*T. nigroviridis*	*D. rerio*	*A. mexicanus*
AP1	x		x					x		
ATF1	x			x						
ATF6	x	x	x							
Atoh1	x		x	x	x				x	x
Barx1	x		x	x						
Bcl6b	x	x	x							
Creb	x	x	x		x	x				
Creb1	x		x	x						
Dlx2	x		x						x	
Dlx3	x		x	x					x	
Dmrt2	x		x						x	
dre	x		x							
E2a	x		x	x	x		x	x	x	x
Elf3	x		x	x	x		x			
eng1a	x		x							
eng2a	x		x	x						
Evi1	x	x	x	x					x	
FAC1	x		x							
Fkhrl1	x	x	x							
Foxa1	x		x							
Foxp1	x		x						x	
foxp2	x		x							
gli3	x		x	x						
gr	x		x	x		x				
Gsh2	x		x	x					x	
Hif1	x	x	x							
hlf	x									
HOX/PBX binding sites	x		x				x			
hoxb9	x		x	x						
ISL LIM homeobox 2	x		x	x					x	
Isx	x		x	x						
lhx2b	x		x							
Meis1	x	x	x	x	x			x		x
Meis1b and Hoxa9 heterodimeric complexes	x	x	x						x	
MEL1	x	x	x		x			x	x	x
Myf5	x		x				x			
MyoD	x		x							
Nk2-3	x		x			x	x		x	
Nkx2-5	x	x	x	x	x				x	
Nkx2-9	x		x			x	x		x	
Nkx5-1	x		x	x					x	x
Nobox	x		x	x					x	
nr2c1	x		x	x	x		x		x	
nrf2	x		x							

Table 2 Predicted transcription factor binding-sites in the *wt1A* promoter region of *A. burtoni* (Continued)

nrsf	x			x	x						x	
Pax6	x	x		x	x	x			x	x	x	x
pce1	x			x	x						x	
Plag1	x			x		x			x			
Pou3f2	x										x	
S8	x			x	x						x	
six1b	x			x								
Sox6	x	x		x	x		x		x	x		
Tax/CREB complex	x			x							x	
Tgif	x			x	x						x	
Zfp67	x			x	x	x			x			

Shown are binding sites in the conserved region marked in yellow in Figure 5. Bold binding sites are shared with at least one other species. "x" denotes the detection of the binding site in the respective species.

have testis tissue in our samples starting from the onset of gonad formation. Another reason why it was important to test gene expression on an individual level is the possible occurrence of spurious XY-females in offspring derived from super-males, which has been described for other cichlids [41,65]. Furthermore, the sex determining gene(s) are not yet identified in *A. burtoni* and additional minor factors influencing sexual fate (environmental or genetic) cannot be ruled out.

After careful inspection of all raw and analyzed data, we did not find any evidence of females in the broods sired by the super-male, i.e., there was no individual with opposing expression patterns at a given sampling point. Especially the expression of the conserved testis factor *dmrt1* in the trunks is a good indicator for male gonad functioning, which is also evidenced by similar profiles in the Nile tilapia (increase in expression of *dmrt1* in testis [17]). A developing ovary would likely have contradicted this trend in gene expression.

Concerning the heads, we cannot rule out the possibility that the expression in other tissues than the brain is picked up by our experiment. For example, if the expression level of a gene is higher in eyes than in testis and higher in testis than in brains (corresponding to: “eyes > testis > brains”), then the overall head expression would be higher compared to trunks (and hence lead to the wrong classification into a brain gene). Having a closer look at the 12 "brain genes" identified by our approach, they either still show a higher expression in the brain than the testis in adult *A. burtoni* (eight genes) or have the reversed expression pattern in adults (four genes) [11]. Thus, the expression levels that we measured in the heads for these four genes (*dax1A*, *foxl2B*, *sf1* and *cyp19a1B*) might not be truly brain specific. Alternatively, the expression pattern may change later in development with an up-regulation in testis and/or down regulation in brain. We think that the latter is more likely, since *foxl2B*, *sf1* and *cyp19a1B* indeed showed a late increase in trunk/testis expression in our experiment, which might further increase beyond the period tested here.

Comparing between gene expression patterns within our experiment, we can show, once more, that paralogous gene copies derived from the fish specific whole genome duplication can evolve different functions, reflected by differences in tissue specificity. In our dataset this is true for *wnt4A* and *B* and *cyp19a1A* and *B*, with each of them having one copy being over-expressed in the heads and one in the trunks. However, we also observed a retention of the same (and hence probably ancestral) expression pattern in both gene copies, for example with very similar expression patterns for *nanos1A* and *B* and *sox9A* and *B*, which is also true in the adult stage [11].

Our main goal was to identify genetic markers for the time window of sex determination in *A. burtoni*. This critical time period, in which the decision if the bipotential embryonic gonad develops towards ovaries or testes is made, has so far been characterized in only one cichlid, the Nile tilapia, where it takes place at 9 to 15 dpf [17,34]. The trunk expression peaks of *wt1A* and *wnt4B* at 11 and 12 dpf suggested that also in *A. burtoni* the time window of sex determination takes place early in development, before any major signs of differentiated gonads become visible. In addition, the narrowness of the expression peak indicated that this time window is rather short. Note that our initial hormone treatment roughly started at the same time point likely accounting for the successful 100% sex reversal.

From the two genes with this early expression peak, especially *wnt4* received some attention in the research of sex determination. Female up-regulation or male down-regulation of *wnt4* expression have been described to be important for promoting ovarian development and function in mammals [66-68]. Also in the developing male gonad *wnt4* is needed for Sertoli cell differentiation, a

crucial step for testis determination [69]. Still, data from teleost fish are largely lacking for *wnt4* and especially for the two teleost paralogs.

Wt1 plays a role in testis differentiation and sex determination in mammals [70,71]. In the medaka, both genes, *wt1a* and *b*, are important for primordial germ cell maintenance, a crucial regulatory mechanism in gonad differentiation in fish [72]. In the Nile tilapia, *wt1a* is up-regulated in the developing male gonad [59]. Hence, *wt1* might act early in gonad differentiation also in other species.

Our sequence analysis of coding and promoter sequence of *wnt4B* and *wt1A* did not reveal any nucleotide difference associated with sex and thus ruled out the two genes as initial genetic regulator of sex determination in *A. burtoni*. However, it is very likely that they represent one of the first members of the sex determination network to be activated during the critical time point of sex determination. Interestingly, the promoter sequence of *wnt4B* contains a potential binding site for *wt1*, meaning that the two genes might functionally interact. Our promoter analysis further suggested that the *androgen receptor* (*ar*), *steroidogenic factor 1* (*sf1*) and *sox3*, three genes with a well-described function in male specific processes [70,73], might regulate *wnt4B* expression. Note that *ar* has two predicted binding-sites in the *wnt4B* promoter, with one being species-specific to *A. burtoni,* and that *sox3* has been co-opted as a master sex-determining gene in another fish species [6]. We did not detect any such obvious candidate among the possible transcriptional regulators of *wt1A*.

Conclusion

In this study, we investigated the expression profiles of sexual development genes in the East African cichlid fish *Astatotilapia burtoni* during early male development. Based on hormonal treatment and subsequent crossing experiments we provided evidence that a male master determiner defines sex in *A. burtoni*. We identified early testis genes, late testis genes and male brain genes (Figures 1 and 3). The earliest testis markers *wnt4B* and *wt1A* were investigated in more detail, as they are strong candidates for the role of the sex-determining gene in *A. burtoni*, due to their expression pattern. Genomic sequences of males and females showed no differences, neither in the coding nor in their promoter region, ruling them out as an initial genetic male determiner. Nonetheless, we suggest that both have an important function early in the sexual development cascade and might even be one of the first targets of the still unknown sex determination factor. A transcription factor binding site analysis revealed possible candidates for master regulators of sexual development in *A. burtoni* such as *sox30*, *ar* and *sf-1*. Future investigations of these candidates, including sequence and expression analyses, together with similar gene expression experiments in female *A. burtoni* should shed more light on the complex cascade of sexual development to finally uncover the master sex-determining gene in this model cichlid species.

Methods

All experiments involving animals were performed in accordance with public regulations under the permits no. 2317 and no. 2620 issued by the cantonal veterinary office of the canton Basel-Stadt (Switzerland).

Estrogen treatment

Animals used in this study were derived from a lab strain of the species *A. burtoni*, an East African cichlid fish from Lake Tanganyika and its surrounding affluent rivers, reared at the fish facility of the Zoological Institute of the University of Basel at 24°C with a 12 hours dark–light cycle.

We treated four clutches of *A. burtoni* with 17α-EthynylEstradiol (E-4876, Sigma) for feminization (protocol kindly provided by H. D'Cotta; see also [1]). 15 mg 17α-EthynylEstradiol were dissolved in 100 ml of 100% ethanol, poured onto 100 g flake food (sera vipan®) and dried at 37°C. From 14 dpf (which is the date when the first fish in the clutches started feeding after the yolk had been absorbed) fish were fed three times a day during four weeks with the hormone treated food. Feeding with 17α-EthynylEstradiol treated food resulted in 100% morphological females in all broods. Amongst these morphological females, we expected (assuming an XX-XY sex determination system) that roughly half of the individuals would have an XX (female) and the other half an XY (male) genotype. Treated fish were subsequently crossed with untreated, normal males. Among the offspring of these individual crossings, several broods showed a 1: 3 (female : male) sex ratio indicative of an XY genotype of the mother (feminized genetic male). Of these crosses, all male offspring was further crossed to normal females. One of these crosses resulted in all male offspring, suggesting that the father was a YY-supermale. For an overview of the crossing design see Figure 2.

Tissue sampling

The potential YY-male resulting from the above mentioned experiment was crossed to untreated females of the lab strain to produce all-male broods. The resulting eggs were collected within an hour after fertilization from the female's mouth and incubated in an Erlenmeyer at 24°C with constant airflow in a 12 hours dark–light cycle. Four individuals were sampled at each of the sampling points at the following days post fertilization: 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 23, 26, 29, 32, 35,

38, 43 and 48. This sampling scheme, with a denser sampling early in development, was chosen because development progresses faster in early stages compared to later stages [36]. Eight clutches were needed to obtain a total of 88 fish. Individuals were photographed for length measurements with a Leica DFC 310 FX (Leica Microsystems). At these early developmental time points, the sampled fish are too small (~ 5 mm standard length) to dissect single organs. To guarantee sufficient RNA material, we thus separated embryos into heads and trunks as proxies for developing brain and gonad tissue, an approach widely used in other fish species [5,7,42-47]. Dissected tissues were stored in Trizol at −80°C until further proceeding.

RNA extraction, DNase treatment and cDNA synthesis

Thawed samples were homogenized using a FastPrep®24 beat beater (MP Biomedicals Europe). Total RNA was extracted following the Trizol protocol. RNA quality and concentration were measured using a NanoDrop 1000 spectrophotometer (ThermoScientific). The RNA was stored at −80°C until further use. RNA samples were treated with DNA-*free*™ Kit (LifeTechnologies) as recommended by the manufacturer. DNase-treated RNA was reverse transcribed using the High Capacity RNA-to-cDNA™ Kit (LifeTechnologies) according to the manufacturer's protocol and diluted to a concentration of 5 ng/μl of cDNA for further procedure.

qRT-PCR expression experiments

In addition to 24 primers (23 candidate genes and *rpl7* as a reference gene) described in [11], and to the primers for *ef1a* and *rpsA3* (used as further reference genes, described in [74]) a primer pair for *gsdf* (as a candidate gene, Forward 5'- CCACCATGGCCTTTGCATTC -3' and Reverse 5'- TCACAGGTGCCAAGGTGAGT -3') was designed and validated for *A. burtoni* following the procedure described in [11] *Rpl7*, *rpsA3* and *ef1a* were tested as possible reference genes. *RpsA3* and *rpl7* showed high stability over all samples (whereas *ef1a* showed slightly more variation). Subsequently, *rpl7* was chosen as a reference gene in the analysis of the qRT-PCR experiments.

Prior to the qRT-PCR experiment, a specific target amplification (multiplex-amplification to increase the amount of targets of interest) was carried out as follows: 2.6 μl TaqMan PreAmp Master Mix (LifeTechnologies), 1.3 μl of a 200 nM mix of all primer pairs and 1.3 μl cDNA were pre-amplified in a thermo cycler (LifeTechnologies) (cycling conditions: 1 × 95°C for 10 minutes, 14 × 95°C for 15 seconds and 58°C for 4 minutes) and diluted 1 : 5 with Low EDTA buffer. The sample premix [2.5 μl TaqMan Gene Expression Mastermix (LifeTechnologies), 0.25 μl DNA Binding Dye Loading Reagent (Fluidigm), 0.25 μl Eva Green (Biotium), 0.75 μl Low EDTA buffer, 1.25 μl of cDNA] and the Assay mix [2.5 μl Assay Loading Reagent (Fluidigm), 0.25 μl Low EDTA buffer, 2.25 μl of 20 μM primer pair] were pipetted on a primed 96 × 96 chip and the plate was loaded in the IFC controller both according to Fluidigm protocols. Expression profiles of the candidate genes in heads and trunks of *A. burtoni* were measured using a Fluidigm BioMark™ assay (HD Systems) at the Genetic Diversity Centre (GDC) of the ETH Zurich with the following cycling conditions: 95°C for 10 minutes, 40 cycles of 95°C for 15 seconds and 58°C for 1 minute. All reactions were followed by a melt curve step to ensure primer specificity and detect possible erroneous amplification. The experiment included three technical replicates of all samples and four biological replicates of all the juvenile samples. Expression data was first analyzed using the Fluidigm Real-Time PCR analysis software to detect technical outliers and for the inspection of melt curves. As outliers we identified samples that showed a deviation from the other samples over all genes, what could easily be seen in the heat map generated by the software. This can happen if an integrated fluidic circuit on the Fuidigm system is blocked by an air bubble. The fold change in expression of the candidate genes in the samples was then calculated with the delta-delta-CT method [50] using custom R scripts. For normalization, the CT values of the reference gene *rpl7* and the mean CT value of a juvenile tissue mix were used. In an additional analysis the fold change was calculated and plotted relative to the mean of the four technical replicates at the first sampling point at 7 dpf (Additional file 2).

Wnt4B and *wt1A* sequencing

DNA from adult males and females of *A. burtoni* lab-strain individuals was extracted from fin clip samples by applying a Proteinase K digestion followed by sodium chloride extraction and ethanol precipitation as described in [75]. To sequence the coding and promoter region of *wnt4B*, nine primer pairs (one of them with two different reverse primers) were designed based on the *A. burtoni* genome [39] using GenScript. The genomic region of *wt1a* spans more than 20 kb in the Nile tilapia genome (over www.ensembl.org) here used as reference for annotation of the *wt1a* coding sequence in the non-annotated *A. burtoni* genome. We thus decided to focus on the coding region for sequencing and constructed primer pairs to amplify each of the nine exons. To sequence the potential promoter region of *wt1A*, six additional primer pairs were constructed covering ~ 4 kb upstream of *wt1a*. The adjacent annotated gene, *depdc7*, is located ~55 kb upstream of *wt1A* in the Nile tilapia genome. PCR reactions were carried out on nine individuals per sex for *wnt4A* and eight individuals per sex for *wt1A* using REDTaq DNA Polymerase (Sigma-Aldrich)

and Phusion Master Mix (New England Biolabs) (for primer sequences and cycling conditions see Additional file 6). PCR products were visualized with GELRed (Biotium) on 1.5% agarose gels. Fragments were sequenced on a 3130xl capillary sequencer (Applied Biosystems) and alignments were performed with CodonCodeAligner (CodonCode Corporation), manually inspected and compared to the corresponding region in the *A. burtoni* genome.

Wnt4B and *wt1A* promoter analysis

Promoter analysis was carried out on the upstream regions of *wnt4B* and the *wt1A* sequences of all the available teleost genomes over www.ensembl.org (release 62) and on the cichlid genome sequences of *A. burtoni*, *Neolamprologus brichardi*, *Orechromis niloticus*, *Pundamilia nyererei* and *Metriaclima zebra* [39]. For *wnt4B* we extracted ~13 kb upstream region until it's next neighboring gene, *chd4*. For *wt1A* we analyzed ~20 kb upstream sequence. Alignments were done with mVISTA [57,58] using Shuffle-LAGAN as alignment algorithm. The Nile tilapia sequence was used as a reference. Putative transcription factor binding sites for *A. burtoni* and the sequenced teleost genomes were identified using MatInspecor (Genomatix Software GmbH). We selected transcription factors that showed a matrix similarity > 0.9 and that belonged to one of the following categories: testis, ovary, germ cell, brain and/or central nervous system. Abbreviated names of transcription factors were taken from Genbank. Tables 1 and 2 show all factors detected in *A. burtoni* and their conservation in the other investigated teleost genomes (indicated by an "x" in Table 1). The complete list with all detected binding sites in all species is shown in Additional files 4 and 5.

Additional files

Additional file 1: Expression data of additional sexual development genes during development of *A. burtoni*. Gene expression as fold change (Livak) ± SE in heads (light green) and trunks (dark blue) from 7 – 48 dpf using *rpl7* as reference gene and a juvenile tissue mix as reference tissue. For details on sample size see Additional file 3.

Additional file 2: Expression data of all candidate genes during development of *A. burtoni*. Gene expression as fold change (Livak) ± SE in heads and trunks relative to the first sampling point at 7 dpf. For details on sample size see Additional file 3.

Additional file 3: Sample sizes for qRT-PCR experiment if other than four. Sample size of trunks at 8, 10 and 11 dpf is three for all genes (and two for *wnt4B*) and therefore not depicted here. Besides *sf-1* (trunk tissue of three individuals at 12 and 13 dpf) and *gata4* (head tissue of three individuals at 38 dpf) all the missing data can be accounted to not detectable expression of *dmrt1* and *gsdf* in heads.

Additional file 4: Putative transcription factor binding sites in the conserved promoter regions (block 1 and block 2 as in Figure 4) of *wnt4B* in teleost genomes. We chose transcription factors with a Matrix similarity > 0.9 and described in the tissues testis, ovary, germ cells, brain and/or central nervous system. Included species are *A. burtoni*, *Xiphophorus maculatus*, *Oryzias latipes*, *Gasterosteus aculeatus*, *Takifugu rubripes*, *Tetraodon nigroviridis*, *Gadus morhua* and *Danio rerio* for block 1 and without *Danio rerio* for block 2. Abbreviated names of transcription factors were taken from Genbank.

Additional file 5: Putative transcription factor binding sites in the conserved promoter region (yellow block in Figure 5) within 20 kb upstream of *wt1A* in teleost genomes. We chose transcription factors with a Matrix similarity > 0.9 and described in the tissues testis, ovary, germ cells, brain and/or central nervous system. Included species are *A. burtoni*, *Metriaclima zebra*, *Neolamprologus brichardi*, *Oreochromis niloticus*, *Xiphophorus maculatus*, *Oryzias latipes*, *Gasterosteus aculeatus*, *Takifugu rubripes*, *Tetraodon nigroviridis*, *Danio rerio* and *Astyanax mexicanus*. Abbreviated names of transcription factors were taken from Genbank.

Additional file 6: Primer sequences and cycling conditions used for sequencing of *wnt4B* and *wt1A* coding and promoter sequence. For amplicon six of *wnt4B* a second reverse primer (reverse 6_2) was designed closer towards forward 6 to ensure complete sequencing of this DNA stretch.

Abbreviations

CT: Threshold cycle; dpf: days post fertilization; qRT-PCR: quantitative real-time polymerase chain reaction; RT: Room temperature; UTR: Untranslated region.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

AB and WS designed the study, AB, WS and CH wrote the manuscript. CH and AB performed hormone treatments and crossings. CH performed the qRT-PCR expression and *wnt4B* sequence analysis. CH and AB conducted the promoter and transcription factor binding site analysis. CG sequenced *wt1A* coding and promoter regions. All authors read and approved the final manuscript.

Authors' information

CH was a PhD student, CG is a master student and AB is a postdoctoral researcher in the group of WS. CH, CG and AB investigate sex determination and differentiation and their evolution in teleost fish using cichlids as a model system. WS is a Professor of Zoology and Evolutionary Biology at the University of Basel. He and his team focus on the genetic basis of adaptation, evolutionary novelties and diversification mainly in cichlid fishes.

Acknowledgements

We thank N. Boileau for support in the lab and A. Minder, K. Eschbach and the GDC at the ETH Zurich for access to the Fluidigm and help during the experiments. The authors thank A. Indermaur, A. Rüegg and Y. Kläfiger for fish keeping. This work was funded by the Forschungsfond Universität Basel, the Volkswagenstiftung (Postdoctoral Fellowship Evolutionary Biology, grant number 86 031), and the DAAD (German academic exchange service, grant number D/10/54114) to AB; and the European Research Council (ERC, Starting Grant "INTERGENADAPT" and Consolidator Grant "CICHLID ~ X") and the Swiss National Science Foundation to WS.

Received: 20 November 2014 Accepted: 1 December 2014

References

1. Devlin RH, Nagahama Y: **Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences.** *Aquaculture* 2002, **208**(3):191–364.
2. Cutting A, Chue J, Smith CA: **Just how conserved is vertebrate sex determination?** *Dev Dyn* 2013, **242**(4):380–387.
3. Matsuda M, Nagahama Y, Shinomiya A, Sato T, Matsuda C, Kobayashi T, Morrey CE, Shibata N, Asakawa S, Shimizu N, Hori H, Hamaguchi S, Sakaizuma M: ***DMY* is a Y-specific DM-domain gene required for male development in the medaka fish.** *Nature* 2002, **417:**559–563.
4. Nanda I, Kondo M, Hornung U, Asakawa S, Winkler C, Shimizu A, Shan Z, Haaf T, Shimizu N, Shima A, Schmid M, Schartl M: **A duplicated copy of *DMRT1* in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*.** *Proc Natl Acad Sci U S A* 2002, **99**(18):11778–11783.

5. Myosho T, Otake H, Masuyama H, Matsuda M, Kuroki Y, Fujiyama A, Naruse K, Hamaguchi S, Sakaizumi M: **Tracing the emergence of a novel sex-determining gene in medaka *Oryzias luzonensis*.** *Genetics* 2012, **191**(1):163–170.
6. Takehana Y, Matsuda M, Myosho T, Suster ML, Kawakami K, Shin-I T, Kohara Y, Kuroki Y, Toyoda A, Fujiyama A, Hamaguchi S, Sakaizumi M, Naruse K: **Co-option of *Sox3* as the male-determining factor on the Y chromosome in the fish *Oryzias dancena*.** *Nat Commun* 2014, **5:**4157.
7. Hattori RS, Murai Y, Oura M, Masuda S, Majhi SK, Sakamoto T, Fernandino JI, Somoza GM, Yokota M, Strüssmann CA: **A Y-linked anti-Müllerian hormone duplication takes over a critical role in sex determination.** *Proc Natl Acad Sci U S A* 2012, **109**(8):2955–2959.
8. Kamiya T, Kai W, Tasumi S, Oka A, Matsunaga T, Mizuno N, Fujita M, Suetake H, Suzuki S, Hosoya S, Tohari S, Brenner S, Miyadi T, Venkatesh B, Suzuki Y, Kikuchi K: **A trans-species missense SNP in *amhr2* is associated with sex determination in the tiger pufferfish, *Takifugu rubripes* (Fugu).** *PLoS Genet* 2012, **8**(7):e1002798.
9. Yano A, Guyomard R, Nicol B, Jouanno E, Quillet E, Klopp C, Cabau C, Bouchez O, Fostier A, Guiguen Y: **An immune-related gene evolved into the master sex-determining gene in rainbow trout, *Oncorhynchus mykiss*.** *Curr Biol* 2012, **22**(15):1423–1428.
10. Yano A, Nicol B, Jouanno E, Quillet E, Fostier A, Guyomard R, Guiguen Y: **The sexually dimorphic on the Y-chromosome gene (*sdY*) is a conserved male-specific Y-chromosome sequence in many salmonids.** *Evol Appl* 2013, **6**(3):486–496.
11. Böhne A, Heule C, Boileau N, Salzburger W: **Expression and sequence evolution of aromatase *cyp19a1* and other sexual development genes in East African cichlid fishes.** *Mol Biol Evol* 2013, **30**(10):2268–2285.
12. Herpin A, Adolfi MC, Nicol B, Hinzmann M, Schmidt C, Klughammer J, Engel M, Tanaka M, Guiguen Y, Schartl M: **Divergent expression regulation of gonad development genes in medaka shows incomplete conservation of the downstream regulatory network of vertebrate sex determination.** *Mol Biol Evol* 2013, **30**(10):2328–2346.
13. Hornung U, Herpin A, Schartl M: **Expression of the male determining gene *dmrt1bY* and its autosomal coorthologue *dmrt1a* in medaka.** *Sex Dev* 2007, **1**(3):197–206.
14. Rossi P, Dolci S, Albanesi C, Grimaldi P, Geremia R: **Direct evidence that the mouse sex-determining gene Sry is expressed in the somatic cells of male fetal gonads and in the germ cell line in the adult testis.** *Mol Reprod Dev* 1993, **34**(4):369–373.
15. Turner ME, Martin C, Martins AS, Dunmire J, Farkas J, Ely DL, Milsted A: **Genomic and expression analysis of multiple *Sry* loci from a single *Rattus norvegicus* Y chromosome.** *BMC Genet* 2007, **8:**11.
16. Veith AM, Schafer M, Kluver N, Schmidt C, Schultheis C, Schartl M, Winkler C, Volff JN: **Tissue-specific expression of *dmrt* genes in embryos and adults of the platyfish *Xiphophorus maculatus*.** *Zebrafish* 2006, **3**(3):325–337.
17. Ijiri S, Kaneko H, Kobayashi T, Wang D-S, Sakai F, Paul-Prasanth B, Nakamura M, Nagahama Y: **Sexual dimorphic expression of genes in gonads during early differentiation of a teleost fish, the Nile tilapia *Oreochromis niloticus*.** *Biol Reprod* 2008, **78**(2):333–341.
18. Herpin A, Schartl M: ***Dmrt1* genes at the crossroads: a widespread and central class of sexual development factors in fish.** *FEBS J* 2011, **278**(7):1010–1019.
19. Lange IG, Hartel A, Meyer HHD: **Evolution of oestrogen functions in vertebrates.** *J Steroid Biochem Mol Biol* 2002, **83**(1–5):219–226.
20. Nakamura M: **The mechanism of sex determination in vertebrates - are sex steroids the key-factor?** *J Exp Zool A Ecol Genet Physiol* 2010, **313**(7):381–398.
21. Angelopoulou R, Lavranos G, Manolakou P: **Sex determination strategies in 2012: towards a common regulatory model?** *Reprod Biol Endocrinol* 2012, **10**(1):13.
22. Morohashi K, Baba T, Tanaka M: **Steroid hormones and the development of reproductive organs.** *Sex Dev* 2013, **7**(1–3):61–79.
23. Heule C, Salzburger W, Böhne A: **Genetics of sexual development – an evolutionary playground for fish.** *Genetics* 2014, **196**(3):579–591.
24. Salzburger W: **The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes.** *Mol Ecol* 2009, **18:**169–185.
25. Kocher TD: **Adaptive evolution and explosive speciation: the cichlid fish model.** *Nat Rev Genet* 2004, **5**(4):288–298.
26. Santos ME, Salzburger W: **How cichlids diversify.** *Science* 2012, **338**(6107):619–621.
27. Salzburger W, Van Bocxlaer B, Cohen AS: **The ecology and and evolution of the African Great Lakes and their faunas.** *Annu Rev Ecol Evol Syst* 2014, **5:**19–45.
28. Baroiller JF, D'Cotta H, Saillant E: **Environmental effects on fish sex determination and differentiation.** *Sex Dev* 2009, **3**(2–3):118–135.
29. Reddon AR, Hurd PL: **Water pH during early development influences sex ratio and male morph in a West African cichlid fish, *Pelvicachromis pulcher*.** *Zoology (Jena)* 2013, **116**(3):139–143.
30. Roberts RB, Ser JR, Kocher TD: **Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes.** *Science* 2009, **326**(5955):998–1001.
31. Ser JR, Roberts RB, Kocher TD: **Multiple interacting loci control sex determination in lake Malawi cichlid fish.** *Evolution* 2010, **64**(2):486–501.
32. Parnell NF, Streelman JT: **Genetic interactions controlling sex and color establish the potential for sexual conflict in Lake Malawi cichlid fishes.** *Heredity* 2013, **110**(3):239–246.
33. Wirtz Ocana S, Meidl P, Bonfils D, Taborsky M: **Y-linked Mendelian inheritance of giant and dwarf male morphs in shell-brooding cichlids.** 2014, *Proc Biol Sci* **281**(1794): 20140253.
34. Baroiller JF, D'Cotta H, Bezault E, Wessels S, Hoerstgen-Schwark G: **Tilapia sex determination: Where temperature and genetics meet.** *Comp Biochem Physiol A Mol Integr Physiol* 2009, **153**(1):30–38.
35. Theis A, Salzburger W, Egger B: **The function of anal fin egg-spots in the cichlid fish *Astatotilapia burtoni*.** *PLoS One* 2012, **7**(1):e29878.
36. Fujimura K, Okada N: **Development of the embryo, larva and early juvenile of Nile tilapia *Oreochromis niloticus* (Pisces: Cichlidae).** *Dev Staging Sys Dev Growth Differ* 2007, **49**(4):301–324.
37. Juntti SA, Hu CK, Fernald RD: ***Tol2*-mediated generation of a transgenic haplochromine cichlid, Astatotilapia burtoni.** *PLoS One* 2013, **8**(10):e77647.
38. Al Hafedh YS, Siddiqui AQ, Al-Saiady MY: **Effects of dietary protein levels on gonad maturation, size and age at first maturity, fecundity and growth of Nile tilapia.** *Aquac Int* 1999, **7**(5):319–332.
39. Brawand D, Wagner CE, Li YI, Malinsky M, Keller I, Fan S, Simakov O, Ng AY, Lim ZW, Bezault E, Turner-Maier J, Johnson J, Alcazar R, Noh HJ, Russel P, Aken B, Alfoldi J, Amemiya C, Azzouzi N, Baroiller JF, Barloy-Hubler F, Berlin A, Bloomquist R, Carleton KL, Conte MA, D'Cotta H, Eshel O, Gaffney L, Galibert F, Gante HF, *et al*: **The genomic substrate for adaptive radiation in African cichlid fish.** *Nature* 2014, **513**(7518):375–381.
40. Pandian TJ, Sheela SG: **Hormonal induction of sex reversal in fish.** *Aquaculture* 1995, **138:**1–22.
41. Mair GC, Abucay JS, Abella TA, Beardmore JA, Skibinski DOF: **Genetic manipulation of sex ratio for the large-scale production of all-male tilapia *Oreochromis niloticus*.** *Can J Fish Aquat Sci* 1997, **54**(2):396–404.
42. Poonlaphdecha S, Pepey E, Canonne M, de Verdal H, Baroiller JF, D'Cotta H: **Temperature induced-masculinisation in the Nile tilapia causes rapid up-regulation of both *dmrt1* and *amh* expressions.** *Gen Comp Endocrinol* 2013, **193:**234–242.
43. Poonlaphdecha S, Pepey E, Huang SH, Canonne M, Soler L, Mortaji S, Morand S, Pfennig F, Mélard C, Baroiller JF, D'Cotta H: **Elevated *amh* gene expression in the brain of male tilapia (*Oreochromis niloticus*) during testis differentiation.** *Sex Dev* 2011, **5**(1):33–47.
44. Fernandino JI, Hattori RS, Kishii A, Strussmann CA, Somoza GM: **The cortisol and androgen pathways cross talk in high temperature-induced masculinization: the 11beta-hydroxysteroid dehydrogenase as a key enzyme.** *Endocrinology* 2012, **153**(12):6003–6011.
45. Haugen T, Almeida FF, Andersson E, Bogerd J, Male R, Skaar KS, Schulz RW, Sorhus E, Wijgerde T, Taranger GL: **Sex differentiation in Atlantic cod (*Gadus morhua* L.): morphological and gene expression studies.** *Reprod Biol Endocrinol* 2012, **10:**47.
46. Duffy TA, Picha ME, Won ET, Borski RJ, McElroy AE, Conover DO: **Ontogenesis of gonadal aromatase gene expression in atlantic silverside (*Menidia menidia*) populations with genetic and temperature-dependent sex determination.** *J Exp Zool A Ecol Genet Physiol* 2010, **313**(7):421–431.
47. Blasco M, Fernandino JI, Guilgur LG, Strussmann CA, Somoza GM, Vizziano-Cantonnet D: **Molecular characterization of *cyp11a1* and *cyp11b1* and their gene expression profile in pejerrey (*Odontesthes bonariensis*) during early gonadal development.** *Comp Biochem Physiol A Mol Integr Physiol* 2010, **156**(1):110–118.
48. Meyer A, Schartl M: **Gene and genome duplications in vertebrates: the one-to-four (–to-eight in fish) rule and the evolution of novel gene functions.** *Curr Opin Cell Biol* 1999, **11**(6):699–704.

49. D'Cotta H, Fostier A, Guiguen Y, Govoroun M, Baroiller J-F: **Aromatase plays a key role during normal and temperature-induced sex differentiation of tilapia *Oreochromis niloticus*.** *Mol Reprod Dev* 2001, **59**(3):265–276.
50. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the 2(−Delta Delta C(T)) method.** *Methods* 2001, **25:**402–408.
51. Matson CK, Zarkower D: **Sex and the singular DM domain: insights into sexual regulation, evolution and plasticity.** *Nat Rev Genet* 2012, **13**(3):163–174.
52. Gautier A, Le Gac F, Lareyre J-J: **The *gsdf* gene locus harbors evolutionary conserved and clustered genes preferentially expressed in fish previtellogenic oocytes.** *Gene* 2011, **472**(1–2):7–17.
53. Rondeau E, Messmer A, Sanderson D, Jantzen S, von Schalburg K, Minkley D, Leong J, Macdonald G, Davidsen A, Parker W, Mazzola R, Campell B, Koop B: **Genomics of sablefish (*Anoplopoma fimbria*): expressed genes, mitochondrial phylogeny, linkage map and identification of a putative sex gene.** *BMC Genomics* 2013, **14**(1):452.
54. Shibata Y, Paul-Prasanth B, Suzuki A, Usami T, Nakamoto M, Matsuda M, Nagahama Y: **Expression of gonadal soma derived factor (GSDF) is spatially and temporally correlated with early testicular differentiation in medaka.** *Gene Expr Patterns* 2010, **10**(6):283–289.
55. Morrish B, Sinclair A: **Vertebrate sex determination: many means to an end.** *Reproduction* 2002, **124**(4):447–457.
56. Rodriguez-Mari A, Yan YL, Bremiller RA, Wilson C, Canestro C, Postlethwait JH: **Characterization and expression pattern of zebrafish *Anti-Mullerian hormone (Amh)* relative to *sox9a*, *sox9b*, and *cyp19a1a*, during gonad development.** *Gene Expr Patterns* 2005, **5**(5):655–667.
57. Mayor C, Brudno M, Schwartz JR, Poliakov A, Rubin EM, Frazer KA, Pachter LS, Dubchak I: **VISTA: Visualizing global DNA sequence alignments of arbitrary length.** *Bioinformatics* 2000, **16:**1046.
58. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I: **VISTA: computational tools for comparative genomics.** *Nucleic Acids Res* 2004, **1**(32):W273–W279.
59. Mohapatra S, Liu ZH, Zhou LY, Zhang YG, Wang DS: **Molecular cloning of *wt1a* and *wt1b* and their possible involvement in fish sex determination and differentiation.** *Ind J Sci Technol* 2011, **4**(S8):85–86.
60. Hale MC, Xu P, Scardina J, Wheeler PA, Thorgaard GH, Nichols KM: **Differential gene expression in male and female rainbow trout embryos prior to the onset of gross morphological differentiation of the gonads.** *BMC Genomics* 2011, **12:**404.
61. Han F, Wang Z, Wu F, Liu Z, Huang B, Wang D: **Characterization, phylogeny, alternative splicing and expression of *Sox30* gene.** *BMC Mol Biol* 2010, **11:**98.
62. Baroiller JF, Guiguen Y, Fostier A: **Endocrine and environmental aspects of sex differentiation in fish.** *Cell Mol Life Sci* 1999, **55:**910–931.
63. Burmeister SS, Kailasanath V, Fernald RD: **Social dominance regulates androgen and estrogen receptor gene expression.** *Horm Behav* 2007, **51**(1):164–170.
64. Böhne A, Sengstag T, Salzburger W: **Comparative transcriptomics in East African cichlids reveals sex- and species-specific expression and new candidates for sex differentiation in fishes.** *Genome Biol Evol* 2014, **6**(9):2567–2585.
65. Bezault E, Clota F, Derivaz M, Chevassus B, Baroiller J-F: **Sex determination and temperature-induced sex differentiation in three natural populations of Nile tilapia (*Oreochromis niloticus*) adapted to extreme temperature conditions.** *Aquaculture* 2007, **272, Supplement**(0):S3–S16.
66. Jameson SA, Lin Y-T, Capel B: **Testis development requires the repression of *Wnt4* by Fgf signaling.** *Dev Biol* 2012, **370**(1):24–32.
67. Yu H, Pask A, Shaw G, Renfree M: **Differential expression of WNT4 in testicular and ovarian development in a marsupial.** *BMC Dev Biol* 2006, **6**(1):44.
68. Kim Y, Kobayashi A, Sekido R, DiNapoli L, Brennan J, Chaboissier M-C, Poulat F, Behringer RR, Lovell-Badge R, Capel B: ***Fgf9* and *Wnt4* act as antagonistic signals to regulate mammalian sex determination.** *PLoS Biol* 2006, **4**(6):e187.
69. Jeays-Ward K, Dandonneau M, Swain A: **Wnt4 is required for proper male as well as female sexual development.** *Dev Biol* 2004, **276**(2):431–440.
70. Barrionuevo FJ, Burgos M, Scherer G, Jiménez R: **Genes promoting and disturbing testis development.** *Histol Histopathol* 2012, **27**(11):1361–1383.
71. Miyamoto Y, Taniguchi H, Hamel F, Silversides DW, Viger RS: **A GATA4/WT1 cooperation regulates transcription of genes required for mammalian sex determination and differentiation.** *BMC Mol Biol* 2008, **9:**44.
72. Klüver N, Herpin A, Braasch I, Driessle J, Schartl M: **Regulatory back-up circuit of medaka Wt1 co-orthologs ensures PGC maintenance.** *Dev Biol* 2009, **325**(1):179–188.
73. Wang R-S, Yeh S, Tzeng C-R, Chang C: **Androgen receptor roles in spermatogenesis and fertility: Lessons from testicular cell-specific androgen receptor knockout mice.** *Endocr Rev* 2009, **30**(2):119–132.
74. Colombo M, Diepeveen ET, Muschick M, Santos ME, Indermaur A, Boileau N, Barluenga M, Salzburger W: **The ecological and genetic basis of convergent thick-lipped phenotypes in cichlid fishes.** *Mol Ecol* 2013, **22**(3):670–684.
75. Laird PW, Zijderveld A, Linders K, Rudnicki MA, Jaenisch R, Berns A: **Simplified mammalian DNA isolation procedure.** *Nucleic Acids Res* 1991, **19**(15):4293.

Perspectives

The work presented in my PhD thesis brought new insights into the topics of egg-spot formation and sexual development in East African cichlid fishes:

First, I investigated the development of anal fin egg-spots with the help of a picture series of *Astatotilapia burtoni*. Using one individual male, four morphological stages of egg-spot formation could be described. I investigated these stages in more fish using several set-ups (i.e., low and high density of fish). When fish were raised under different conditions, it turned out that standard body length is a better predictor for egg-spot formation stage than age.

I then focused on sexual development, which is extremely diverse in fish compared to mammals and birds. It can be seen as a continuum of sex determination and sex differentiation, which are regulated by an interplay of heritable and environmental factors. The identification of various master sex-determining genes, the description of complex polygenic regulations and, importantly, the discovery of new functions of presumably conserved down-stream factors led us to a conceptual re-formulation of the evolution of sexual development cascades.

Empirical evidence for the diversity of down-stream factors implicated in sexual development is given by the lineage-specific shifts in gene expression and sequence evolution of the aromatase *cyp19a1* in cichlids. We found expression in testis in Ectodini (*cyp19a1A*) and Haplochromini (*cyp19a1B*) instead of presumably conserved ovary- and brain-specific expression observed in all other so far studied fish species, respectively.

Additionally, supporting the plasticity of down-stream factors of sexual development is our experiment that used hormonally sex-reversed specimens of *Astatotilapia burtoni*. Estrogen treatment followed by crossing experiments provided evidence for a male heterogametic sex-determining system (XX-XY). XY-females (sex-reversed males) crossed to normal XY-males resulted in a putative YY-super male that only produced male offspring. These all-male broods were used to profile the expression of sexual development genes. We described several genes putatively involved in testis formation. One of them has not been implicated with this function before. Moreover, this points towards less conservation in the sexual development cascade than previously thought.

Being a male or female substantially influences all life stages. Obviously, males and females develop structurally and functionally distinct gonads, which in turn produce different types of germ cells. Not only do males and females show internal differences, but they often also vary in the phenotypic appearance. In East African cichlid fishes sexual dimorphism e.g., in coloration, is very pronounced in a number of species, especially within the most-species rich lineage, the haplochromines. In this lineages the females are rather dull and the males are brightly colored and carry egg-spots on their anal fin, a putative evolutionary key-innovation suggested to be important e.g., in intra-sexual communication (Salzburger *et al.* 2005; Salzburger *et al.* 2007; Santos and Salzburger 2012; Theis *et al.* 2012).

Sexual development, leading to the formation of the male and female phenotype, is regulated by a variety of heritable but also environmental factors. Sex-determining systems (e.g., male- or female heterogamety) and genetic cascades must not be conserved, not even between closely related species. Recent studies in fish have shown that the master sex-determining genes vary between different species, and even between sister species (Matsuda *et al.* 2002; Nanda *et al.* 2002; Hattori *et al.* 2012; Kamiya *et al.* 2012; Myosho *et al.* 2012; Yano *et al.* 2012). Our findings now extend this variety to downstream factors of sexual development (Böhne *et al.* 2013). In my thesis, I investigated a set of candidate genes for sexual development. More specifically, we studied these genes in adult cichlid species and in one of them also during early male development. Differential gene expression patterns, measured in brain and gonad tissue, were observed between cichlid lineages and within the same species over time. Adult Ectodini (a lineage that is endemic to Lake Tanganyika) and Haplochromini (the most species-rich lineage occurring in all three Lakes, Tanganyika, Malawi and Victoria (Salzburger *et al.* 2005; Sturmbauer *et al.* 2011)) showed a shift towards testis-specific expression of the aromatase *cyp19a1A* and *cyp19a1B*, respectively. Interestingly, profiling the expression of these gene copies during early male development in one of the haplochromines, *Astatotilapia burtoni*, revealed the opposite pattern than in adults of the same species. In contrast to the high expression of the B-copy in adult testis, the A-copy seemed to be more important in early testis differentiation. And the B-copy, on the other hand, showed a low expression during early development, which seemed to increase at the end of the investigated experimental period. We also observed analogies between adults and juveniles, e.g., *dmrt1*, which is testis-specifically expressed in adults and shows continuous increase in expression in developing testis. We described an interesting pattern for *wnt4B*, with a peak in expression at an early stage of testis development. In adult *A. burtoni* testis, we only detected weak expression of *wnt4B*. We expect a master sex-determining gene to show high expression levels early in developing testis. I therefore decided to compare the genomic sequence of *wnt4B* in more detail in

males and females. Sequence comparison of the coding and the promoter region showed no differences between the sexes and we therefore excluded *wnt4B* as the master sex-determining gene. Although *wnt4B* is not acting at the top of the sexual development cascade it might be one of the first targets of the master sex regulator. A transcription factor binding site analysis revealed a number of genes, possibly acting further up in the cascade, representing putative candidates for the role of the sex-determining gene in *A. burtoni*.

Outlook

My results, once more, underline the flexibility of sexual development in cichlids and highlight that always all processes acting along the continuum of sexual development should be looked at. Developmental series already give a broader insight into the dynamics of sexual development genes, compared to a gene expression assay in adults reflecting a snapshot in time only. To complete the picture of sexual development in *Astatotilapia burtoni*, similar experiments using all-female broods, would be interesting. On one hand, that would provide important insights into the female cascade and on the other hand it would additionally support the results obtained from the male developmental series: Because we measured gene expression in whole heads and trunks - as proxies for brain and gonad tissue - there is still the possibility that other tissues in the trunk highly express a gene presumably concluded to be a testis-specific gene. Although we think this possibility is not likely for most genes investigated (as discussed in Chapter 4), further support for testis and not trunk expression would be given by low expression in the female trunk. Theoretically, all-female broods can be produced by crossing an XX-male (a sex-reversed female) to a normal XX-female. Our attempts to obtain such sex-reversed fish by feeding fry with testosterone treated food were not successful; the testosterone treatment did – in contrast to the estradiol treatment – not result in 100 % sex-reversal. Further studies should, for example, make use of aromatase-inhibitors in order to sex-reverse fish instead of testosterones, which has recently been shown to function even in adults of different fish species, including the cichlid Nile tilapia (*Orechromis niloticus*) (Paul-Prasanth *et al.* 2013).

In addition to genetic experiments, morphological studies could be conducted. In contrast to the most investigated cichlid in terms of sexual development, the Nile tilapia, there is no scientific data about the morphological gonadal development in *A. burtoni* and other cichlids that belong to the rapidly radiating lineage of the haplochromines. Such a study would allow describing putative gene functions more precisely in relation to morphological differentiation.

Regarding the described differences in adult gene expression between species, I suggest to perform developmental series in more species. This will allow an interspecies comparison of the sexual development cascade over time. In adults it seems that species within a lineage share the same expression pattern to some extent. It would be interesting to see if this conservation in adult gene expression within a lineage holds true for developmental series as well.

Sex-determining mechanisms can vary even between populations of the same species as shown e.g., in a sympatric species-pair in sticklebacks (Kitano *et al.* 2009). Different populations of *A. burtoni* in lake and stream habitats are currently investigated using restriction site-associated DNA (RAD)-tag sequencing data (Egger *et al*. in preparation). This technique was already applied to investigate sexual development of other fishes (Anderson *et al.* 2012; Parnell and Streelman 2012). For further studies on genomic differences between males and females, I suggest to involve next generation sequencing approaches like RAD-tag sequencing, whole genome sequencing as well as transcriptome data.

Concluding remarks

I started my doctoral work with the investigation of morphological development of anal fin egg-spots in *Astatotilapia burtoni*. Due to the lack of a sex-specific marker for this species, my data also included females, unless they were phenotypically distinguishable from males, which is only possible after the formation of egg-spots. These limitations led to the question, how sex specific phenotypes come into existence in the first place. This, to me, opened up the fascinating field of sexual development, and with our experiments, we showed that this cascade of forming males and females is highly plastic in fish – in all different processes involved. One aspect of that flexibility is the possibility to reverse sex by hormonal treatment, which allowed producing all-male broods. It is these broods that enable the investigation of egg-spot development in males only, even though I did not find the sex-specific marker I aimed for.

References

Anderson, J. L., A. Rodríguez-Marí, I. Braasch, A. Amores, P. Hohenlohe *et al.*, 2012 Multiple sex-associated regions and a putative sex chromosome in zebrafish revealed by RAD mapping and population genomics. PLoS One 7: e40701.

Böhne, A., C. Heule, N. Boileau and W. Salzburger, 2013 Expression and sequence evolution of aromatase *cyp19a1* and other sexual development genes in East African cichlid fishes. Mol. Biol. Evol. 30: 2268-2285.

Hattori, R. S., Y. Murai, M. Oura, S. Masuda, S. K. Majhi *et al.*, 2012 A Y-linked anti-Müllerian hormone duplication takes over a critical role in sex determination. Proc. Natl. Acad. Sci. U. S. A. 109: 2955-2959.

Kamiya, T., W. Kai, S. Tasumi, A. Oka, T. Matsunaga *et al.*, 2012 A trans-species missense SNP in *Amhr2* is associated with sex determination in the tiger pufferfish, *Takifugu rubripes* (fugu). PLoS Genet. 8: e1002798.

Kitano, J., J. A. Ross, S. Mori, M. Kume, F. C. Jones *et al.*, 2009 A role for a neo-sex chromosome in stickleback speciation. Nature 461: 1079-1083.

Matsuda, M., Y. Nagahama, A. Shinomiya, T. Sato, C. Matsuda *et al.*, 2002 *DMY* is a Y-specific DM-domain gene required for male development in the medaka fish. Nature 417: 559-563.

Myosho, T., H. Otake, H. Masuyama, M. Matsuda, Y. Kuroki *et al.*, 2012 Tracing the emergence of a novel sex-determining gene in medaka, *Oryzias luzonensis*. Genetics 191: 163-170.

Nanda, I., M. Kondo, U. Hornung, S. Asakawa, C. Winkler *et al.*, 2002 A duplicated copy of *DMRT1* in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*. Proc. Natl. Acad. Sci. U. S. A. 99: 11778-11783.

Parnell, N. F., and J. T. Streelman, 2012 Genetic interactions controlling sex and color establish the potential for sexual conflict in Lake Malawi cichlid fishes. Heredity 110: 239-246.

Paul-Prasanth, B., R. K. Bhandari, T. Kobayashi, R. Horiguchi, Y. Kobayashi *et al.*, 2013 Estrogen oversees the maintenance of the female genetic program in terminally differentiated gonochorists. Sci. Rep. 3: 2862.

Salzburger, W., I. Braasch and A. Meyer, 2007 Adaptive sequence evolution in a color gene involved in the formation of the characteristic egg-dummies of male haplochromine cichlid fishes. BMC Biology 5: 51.

Salzburger, W., T. Mack, E. Verheyen and A. Meyer, 2005 Out of Tanganyika: Genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. BMC Evol. Biol. 5: 17.

Santos, M. E., and W. Salzburger, 2012 Evolution. How cichlids diversify. Science 338: 619-621.

Sturmbauer, C., M. Husemann and P. D. Danley, 2011 Explosive speciation and adaptive radiation of East African cichlid fishes, pp. 333-362 in *Biodiversity Hotspots*, edited by F. Zachos and J. Habel. Springer Berlin Heidelberg, Berlin, Heidelberg.

Theis, A., W. Salzburger and B. Egger, 2012 The function of anal fin egg-spots in the cichlid fish *Astatotilapia burtoni*. PLoS One 7: e29878.

Yano, A., R. Guyomard, B. Nicol, E. Jouanno, E. Quillet *et al.*, 2012 An immune-related gene evolved into the master sex-determining gene in rainbow trout, *Oncorhynchus mykiss*. Curr. Biol. 22: 1423-1428.

Acknowledgments

Many people accompanied me during my work on this PhD thesis and I am most grateful for their support:

My biggest thanks go to **Walter Salzburger** for offering me the possibility to work in his group and always for being a great boss and an appreciated friend.

I am most thankful to **Astrid Böhne** who taught me lab techniques, helped me analyzing data and sacrificed a lot of time to review and thereby improve my scientific writing.

I am very grateful to **Judith Mank** who spontaneously agreed to review this thesis and be the second examiner at my defense.

Furthermore I thank **Maria Emília Santos** for introducing me to world of qPCR.

A big thank you goes to the lab managers **Brigitte Aeschach** and **Nicolas Boileau** for their untiring support in the lab and help with organizational issues.

I would like to thank my office mates **Adrian Indermaur**, **Anya Theis**, **Britta Meyer** for discussions and support (scientifically as well as mentally).

Thanks to **Attila Rüegg** and **Yuri Kläfiger** who helped maintaining my fish and additionally to the sequential fish feeders **Adrian Indermaur** and **Bernd Egger**.

I thank the whole **Salzburger group** for a fantastic time.

Not to forget the former lab members **Eveline Diepeveen**, **Michael Matschiner** and **Moritz Muschick** who were in the group when I started my PhD and will always belong to the Salzburger lab.

A special thank goes to **my parents** who supported me during the whole time of my studies.

I would also like to acknowledge the **European Research Council** and the **Swiss Academy of Sciences** for my salary and the travel grant for the fieldtrip to Africa, respectively.