

Molecular Characterization of Two Endothelin Pathways in East African Cichlid Fishes

Eveline T. Diepeveen · Walter Salzburger

Received: 7 September 2011 / Accepted: 4 January 2012 / Published online: 21 January 2012
© Springer Science+Business Media, LLC 2012

Abstract The adaptive radiations of cichlid fishes in East Africa have been associated with the acquisition of evolutionary novelties as well as the ecological opportunities existing in the East African Great lakes. Two remarkable evolutionary innovations are the pharyngeal jaw apparatus, found in all cichlid species, and the anal fin egg-spots of mouthbrooding cichlids. Based on their conserved functions during the development of both the jaw apparatus and pigmentation, the endothelin ligands and receptors form a putative link between these naturally and sexually selected traits. Here we study the evolutionary history of four members of two endothelin pathways (Edn1/EdnrAa and Edn3b/EdnrB1a) to elucidate their possible roles during the evolution and development of key innovations in East African cichlid species. The analyses performed on partial sequences (ca. 6,000 bp per taxon) show that all four endothelin family members evolved under purifying selection, although both ligands are characterized by an accelerated rate of protein evolution in comparison to the receptors. In accordance with earlier findings, we show that the mature protein sequence of Edn1 and Edn3 are highly conserved, also in cichlids, whereas the preproendothelin parts are variable indicating relaxed selective constraints. In the receptors, nonsynonymous substitutions were mainly found in the ligand-binding domains suggesting functional

divergence. Gene expression assays with Real-Time PCR indeed reveal that the two studied endothelin pathways are expressed in the cichlid pharyngeal jaw and in the haplochromine egg-spot (among other pigment-cell containing tissues), suggesting their involvement during morphogenesis of naturally and sexually selected traits in cichlids.

Keywords Endothelin receptor · Pigmentation · Pharyngeal jaw · Neural crest · Key innovation · Molecular evolution

Background

The spectacular adaptive radiations of cichlid fishes provide a unique model system to study evolutionary processes. No other vertebrate family consists of such a great number of species (at least 3,000 species worldwide), such a degree of phenotypic diversity and such a rapid diversification rate. Above all, the species flocks of cichlids in the East African Great Lakes show the greatest variation in morphology, coloration, ecology, behavior, and reproductive biology (Barlow 2000; Coulter 1991; Fryer and Iles 1972; Keenleyside 1991; Kocher 2004; Kornfield and Smith 2000; Salzburger 2009; Seehausen 2006). Furthermore, the cichlids' occurrence in various geographic regions (Africa, India, Sri Lanka, and Central and South America) and the independent (and repeated) colonization of multiple African lakes (e.g., lakes Malawi, Tanganyika, and Victoria) form the basis for some outstanding cases of parallel evolution within and between lakes and regions (Duftner et al. 2007; Elmer et al. 2010; Kocher et al. 1993).

According to the ecological opportunity hypothesis, fast divergence of lineages can occur when available and underutilized niches are being filled (Schluter 2000). This

Electronic supplementary material The online version of this article (doi:10.1007/s00239-012-9483-6) contains supplementary material, which is available to authorized users.

E. T. Diepeveen (✉) · W. Salzburger (✉)
Zoological Institute, University of Basel, Vesalgasse 1,
4051 Basel, Switzerland
e-mail: eveline.diepeveen@unibas.ch

W. Salzburger
e-mail: walter.salzburger@unibas.ch

scenario seems very plausible for cichlids that colonized and subsequently radiated in large rift lakes with their diversity in habitats (e.g., surge habitat, shallow rocky shores, sandy ground, mud floor, benthic water) and food resources (e.g., algae, plants, small invertebrates, and other fish). New niches can also become accessible when novel traits, so-called evolutionary key innovations, evolve with which so far unaccessible niches can be exploited (Gavrilets and Losos 2009; Schluter 2000). The evolutionary success of cichlids in general, and the explosive radiation of the East African haplochromine lineage in particular, have been associated with the acquisition of such evolutionary innovations (Fryer and Iles 1972; Liem 1973; Salzburger et al. 2005). The pharyngeal jaw apparatus (Fig. 1a), for instance, is found in all cichlids and is thought to underlie diversification with respect to feeding ecology (Fryer and Iles 1972; Hunter 1998; Liem 1973). The diversity in cichlid pharyngeal jaw morphology (and the teeth on there) directly correlates with the diversity in diet, from algae to whole fishes (Liem 1973). Even though other groups of teleost fishes are characterized by the presence of pharyngeal structures that aid the passage of food into the esophagus, cichlid fishes show the greatest diversity in pharyngeal jaw morphology resulting in the efficient manipulation of a great variety of food. Thus, the pharyngeal jaw apparatus is thought to play an important role in niche specialization and, hence, speciation in

cichlids (Liem 1973; Muschick et al. 2011; Salzburger 2009).

Sexual selection is the other major factor that is likely to have contributed to the cichlids' species-richness (see, e.g., Meyer 1993; Salzburger 2009; Seehausen et al. 1997; Turner & Burrows 1995), as indicated by the occurrence of multiple color morphs and sexual color dimorphism. This is particularly obvious in the most species-rich group of cichlids, the haplochromines. The haplochromines exclusively consists of maternal mouthbrooders, which protect their offspring from predation in their buccal cavities (Hert 1989; Mrowka 1987; Fig. 1b). The occurrence of egg-spots on the anal fins of male haplochromines and other C-lineage tribes has been proposed as yet another key innovation (Salzburger et al. 2005, 2007; Fig. 1c). Variation in size, number, and color of the egg-spots has been observed both between and within species (Salzburger et al. 2007). Egg-spots appear to play a prominent role in mating behavior (Wickler 1962), although their exact function remains elusive (Egger et al. 2011; Theis et al. 2012).

The above-mentioned evolutionary key innovations of cichlids share a developmental origin. Both the mouth apparatus (including the pharyngeal jaws) and pigmentation (body coloration and egg-spots) are derived from the neural crest (Albertson and Kocher 2006; Gans and Northcutt 1983; Hall 1999; Slack 2001), which is—by

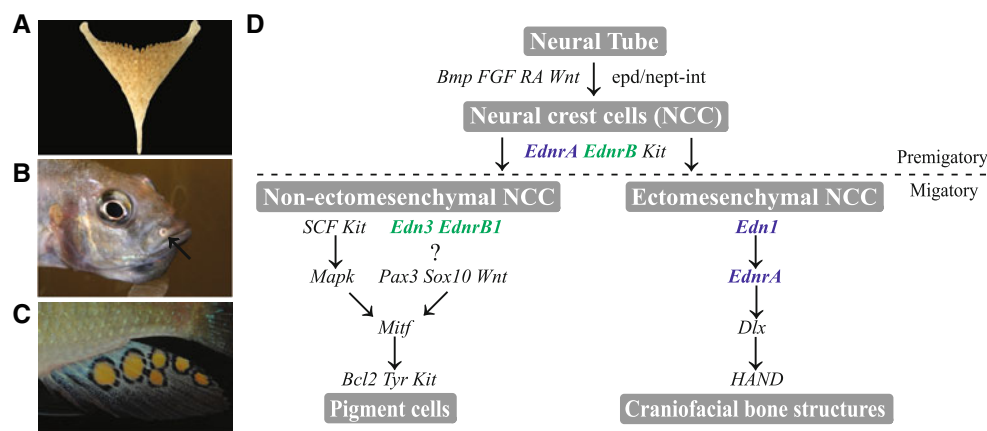


Fig. 1 Three putative key innovations and the genetic pathways involved in neural crest cell differentiation. **a** The lower pharyngeal jaw, which is a fusion of the fifth ceratobranchials, is responsible for grinding the food and is present in all cichlids. **b** A mouthbrooding haplochromine female. The female protects her eggs and later on the fry (as shown) from possible predation by holding them in her mouth. **c** The egg-spot on the anal fin of haplochromine males play an important role in breeding behavior. **d** The endothelin system is involved in the development of (pre)migratory neural crest cells and its derivatives in vertebrates. The *Edn3/EdnrB1* pathway (in green) is involved in pigment-cell development, while the *Edn1/EdnrA* pathway (in blue) is involved in the lower jaw formation (among other craniofacial bone structures). Pathways depicted are simplified

and not exhaustive. *Bcl2* B-cell leukemia/lymphoma 2, *Bmp* bone morphogenetic protein family, *Kit* kit receptor tyrosine kinase, *Dlx* distal-less homeobox, *Edn1/3* endothelin1/3, *EdnrA/B* endothelin receptor A/B, *epd/nept-int* epidermis/neuroepithelium-interaction, *FGF* fibroblast growth factor family, *HAND* heart and neural crest derivatives expressed transcript, *Mapk* mitogen-activated protein kinase, *Mitf* microphthalmia-associated transcription factor, *Pax3* paired box gene 3, *RA* retinoic acid, *SCF* stem cell factor, *Sox10* sex determining region Y-box 10, *Tyr* tyrosinase, *Wnt* wingless-type integration site family. Photographs taken by M. Muschick, A. Theis, and C. Heule. Pathways based on, e.g., Bonano et al. 2008; McGill et al. 2002; Parichy 2000; Pla and Larue 2003; Rawls et al. 2001; Slack 2001; Yanagisawa et al. 2003

itself—an evolutionary innovation of the vertebrate clade (e.g., Gans and Northcutt 1983; Shimeld and Holland 2000). Neural crest cells originate in the neural folds, and migrate away from the neural tube to form a wide variety of cell types across the whole body including neuron cells, pigment cells, craniofacial bones, Schwann cells, and smooth muscle cells (Chung et al. 2009; Slack 2001). The genetic and developmental basis of the neural crest-derived tissues and specifically those involved in evolutionary key innovations is, hence, of great importance for the study of cichlid evolution (Salzburger 2009).

Within the last decade several studies have focused on the genetic basis of naturally and sexually selected traits in cichlids (reviewed in Salzburger 2009). Albertson et al. (2005), for example, showed that *bmp4* is involved in craniofacial diversity and that it has the potential to change the mandible shape of cichlids. It had previously been shown that this gene underwent adaptive protein evolution in East African cichlid species (Terai et al. 2002b). More recently, Fraser et al. (2009) showed that similar genetic pathways underlie the tooth formation on both the oral and the pharyngeal jaws in cichlids. Seven Hox genes and several genes from a previously described dental gene network (Fraser et al. 2008) were shown to be expressed in developing pharyngeal jaw tissue in two Malawi cichlid species (Fraser et al. 2009). Apart from jaw-related genes, several candidate coloration genes have been studied in cichlids. Terai et al. (2003), for example, showed that cichlids express multiple alternatively spliced *hag* variants and that species-rich lineages feature more splice variants than species-poor ones. Also for this gene an accelerated rate of protein evolution was determined (Terai et al. 2002a). The *Pax7* locus appears to be involved in the formation of the orange blotch phenotype (Roberts et al. 2009). Finally, it has been shown that *csflra* is involved in the development of the colorful egg-spots on the male anal fin (Salzburger et al. 2007).

One specific family of candidate genes for neural crest-derived structures and, in particular, two prominent evolutionary innovations of cichlids (i.e., the pharyngeal jaw and coloration) is the endothelin (Edn) family. Endothelin 1 was discovered in 1988 as vasoconstrictor by Yanagisawa et al. (1988). Later, more members of this gene family were found and their functions annotated, so it is now evident that the endothelin system is involved in the formation of craniofacial structures, enteric ganglion neurons, pigmentation, and vasoconstriction in both mouse (Baynash et al. 1994; Clouthier et al. 1998; Hosoda et al. 1994; Kurihara et al. 1994; Ozeki et al. 2004; Yanagisawa et al. 1988) and zebrafish (Miller et al. 2000; Parichy et al. 2000). The active gene product of the endothelin genes is only 21 amino acids long and obtained via two steps of

cleaving from the inactive precursors preproendothelin and big endothelin, respectively (Masaki 2004).

The Edn family arose in the vertebrate lineage from an ancestral signaling peptide and a G protein-coupled receptor. Subsequent expansion took place by three rounds of whole genome duplication (i.e., the two vertebrate-specific whole genome duplications and a fish-specific one), each followed by functional divergence and coevolution of the ligands and receptors (Braasch et al. 2009). Also, the range of interactions between the ligands and the receptors expanded with each round of genome duplication, since all endothelin receptors can—theoretically—bind all ligands, albeit with different affinities (note that this has not yet been examined in teleosts). Braasch et al. (2009) showed that functional changes in the receptors were mainly due to shifts of evolutionary rates and less to radical changes in amino acids. Furthermore, the sites where these changes were observed are mainly found in the extracellular loops or ligand-binding domains of the receptors and not in the intracellular domains where interaction with the G protein occurs. This indicates that with the expanding repertoire of the ligands, the receptors co-evolved by slightly changing the ligand–receptor binding affinity and selectivity. Since the third round of whole genome duplication took place in the lineage leading to the ray-finned fishes (Meyer and van de Peer 2005), the most abundant number of endothelins and receptors are found in today's teleost fishes. Six endothelin genes and five teleost endothelin receptor genes were identified in stickleback, while zebrafish, medaka, spotted green pufferfish, and torafugu are all characterized by a species-specific gene loss, resulting in a lower number of either receptor or ligand (Braasch et al. 2009).

The aim of this study is to elucidate the molecular evolutionary history of the endothelin family in East African cichlids and its possible role during the evolution and development of evolutionary key innovations. Therefore, two components of the endothelin core system (the ligand and receptor of two endothelin signaling pathways) were studied in a phylogenetically representative set of East African cichlids: the Edn1/EdnrA pathway, which is involved in the development of craniofacial structures (Miller et al. 2000; Nair et al. 2007; Pla and Larue 2003; Fig. 1d), and the Edn3/EdnrB pathway, which is involved in the development of pigment cells (Pla and Larue 2003; Fig. 1d). We determined partial DNA sequences of ligands and receptors in 26 cichlid species from Lake Tanganyika (total sequence length per species: ca. 6,000 bp) and performed analyses to determine the evolutionary history of these loci. We show that the endothelin family members evolved under purifying selection and that they are indeed expressed in pharyngeal jaw and egg-spot tissue.

Materials and methods

Tissue Sampling and DNA/RNA Extraction

In this study, we analyzed two components of the endothelin system in 26 species of cichlid fishes from Lake Tanganyika, East Africa, belonging to 14 different tribes, thus representing a great portion of the diversity of East African cichlids (Salzburger et al. 2002, 2005). The samples were collected during two consecutive expeditions to Lake Tanganyika in 2007 and 2008 (see Supplementary Table 1 for specimen information).

Genomic DNA was extracted from white muscle tissue and digested with Proteinase K followed by sodium chloride extraction and ethanol precipitation (Bruford et al. 1998). RNA was extracted from male adult *Astatotilapia burtoni* (laboratory strain) in order to analyze the expression pattern of the studied loci (see below). Animals were kept under standard conditions (12 h light, 12 h dark; 25°C). Twelve types of tissues were dissected (brain, liver, gonads, muscle, pharyngeal jaw, maxilla, mandible, skin, anal fin [both entire anal fin as egg-spot and non-egg-spot tissue independent], caudal fin, eyes, and whole juveniles were used as positive control) and ground using a mortar and a pestle, followed by an RNA extraction procedure with Trizol (Invitrogen). Hereafter, genomic DNA was removed from the samples by DNase treatment following the DNA Free protocol (Ambion Inc, Austin, TX). The RNA concentration was measured with a Nanodrop 1000 spectrophotometer (Thermo Scientific, Wilmington, DE) in order to use the recommended amount of RNA in the subsequent reverse transcriptase (RT) step (see below).

Reverse Transcription and Real-Time PCR

Reverse transcription of the cleaned-up RNA samples was performed with the High capacity RNA-to-cDNA kit (Applied Biosystems) following the manufacturer's protocol. The housekeeping gene *GAPDH* (*Glyceraldehyde-3-phosphat-dehydro-genase*) was used under standard polymerase chain reaction (PCR) conditions with Amplitaq DNA polymerase (see below) to test if the extraction and following steps were successful. Subsequent Real-Time PCRs were conducted on a StepOnePlus Real-Time PCR System (Applied Biosystems) following the manufacturer's protocol. Final cDNA and primer concentrations of 0.5 ng/μl and 900 nM, respectively, were used together with the SYBR Green master (Rox) dye (Roche, Basel, Switzerland). Comparative C_T experiments were used to quantify the level of expression with juvenile tissue as reference tissue and *Actin* (see Supplementary Table 2 for

primer info) as endogenous control. Measurements of expression levels in target tissues were normalized using the endogenous control. Each reaction was replicated three times.

Loci, PCR Amplification, and Sequencing

Four nuclear loci were investigated here: *Endothelin 1* (*Edn1*), *Endothelin 3b* (*Edn3b*), *Endothelin receptor Aa* (*EdnrAa*), and *Endothelin receptor B1a* (*EdnrB1a*). These loci were chosen based on the availability of their teleost orthologous sequences in public databases, the quality of the subsequent alignments and the quality of the PCR products. Also we choose to study only one paralog per endothelin pathway. PCR primers for both the DNA and cDNA samples are listed in Supplementary Table 2 and were initially designed based on teleost orthologs (i.e., zebrafish (*Danio rerio*), medaka (*Oryzias latipes*), Fugu (*Takifugu rubripes*), Tetraodon (*Tetraodon nigroviridis*) and stickleback (*Gasterosteus aculeatus*); see Braasch et al. 2009). Primers were tested on *A. burtoni* DNA. Resultant *A. burtoni* sequences were blasted using a translated nucleotide query. Based on these sequences an additional set of cichlid-specific primers was designed. Standard 12.5 μl PCR reactions were conducted on a 2720 thermal cycler and/or a Veriti 96-well thermal cycler (both Applied Biosystems) with 1× Taq buffer, 0.8 mM dNTP mix, 0.0125 units REDTaq DNA polymerase (all Sigma, Saint Louis, MO), 0.2 μM of each primer, 0.5 mM MgCl₂, and 20 ng of template DNA. For initially unsuccessful amplifications the following, modified, protocol was used: 1× Taq buffer, 0.25 units Amplitaq DNA polymerase (both Applied Biosystems), 0.8 mM dNTP mix (Sigma), 0.2 μM of each primer, 0.1 mM MgCl₂ and 20 ng of template DNA. Amplification consisted of an initial denaturation step at 94°C (2 min), 30–40 cycles of 94°C (45 s), 47–57°C (45 s), and 72°C (60–120 s) followed by a final extension at 72°C (10 min).

All PCR products were visualized with ethidium bromide on a 1.5% agarose gel. Hereafter, 2.5 μl amplified PCR-product was purified with ExoSAP-IT following the manufacturer's instructions (USB, Cleveland, OH). The subsequent sequencing reactions for both forward and reverse sequences separate were performed with 1 μl Terminator Ready Reaction Mix 3.1 (Applied Biosystems), 0.625 μM primer and the purified amplification products. The conditions of the cycle sequencing reaction were as followed: 94°C (1 min), 25 cycles of 94°C (10 s), 52°C (20 s), and 60°C (4 min).

The final purification step was performed with the BigDye X-Terminator system (Applied Biosystems). SAM solution (22.5 μl), X-Terminator beads (5 μl) and H₂O (14.5 μl) was added to each sample and mixed (2,000 rpm)

for 30 min, followed by a final spin down (1,500 rpm, 2 min).

The DNA fragments were visualized on a 3130xl capillary sequencer (Applied Biosystems). Sequences of the DNA samples were used in subsequent analyses (see below), while sequences from the cDNA samples were only used to confirm the respective locus. All sequences have been deposited into GenBank under accession numbers JQ309683–JQ309786 (see Supplementary Table 1).

Cloning

Since one PCR reaction for the *EdnrB1a* locus did not result in a clear signal in the sequencing reaction, PCR products were cloned with a TOPO TA cloning kit (Invitrogen) following the manufacturer's protocol. PCR reactions were performed directly on the clone products with the M13 primer pair that was supplied by the manufacturer. After confirmation of the proper insert size (i.e., again with visualization with ethidium bromide on a 1.5% agarose gel), PCR products were cleaned-up with a Gen Elute PCR clean up kit (Sigma) followed by the normal sequencing protocol (see above). A minimum of six colonies were sequenced for each sample and aberrant nucleotide positions due to PCR errors and only observed in one of these sequences, were removed before the analyses.

Molecular and Phylogenetic Analyses

Sequences were aligned and edited using Codon Code Aligner 3.5.6 (CodonCode Corporation, Dedham, MA). Exon/intron boundaries were determined by homology comparisons with five teleost orthologs (see Braasch et al. 2009).

First of all, we constructed phylogenetic trees consisting of both cichlid and other teleost orthologs to confirm the exact paralog of our sequences. Therefore, cichlid *Edn* and *Ednr* sequences were translated into protein sequences and aligned with the other teleost *Edn* and *Ednr* sequences (i.e., *D. rerio*, *O. latipes*, and *G. aculeatus*; see Braasch et al. 2009 for references) using TCOFFEE (Notredame et al. 2000; Poirot et al. 2003). ProtTest (Abascal et al. 2005; Drummond and Strimmer 2001; Guindon and Gascuel 2003) was used to select the most appropriate model of protein evolution, which was then used in the phylogenetic analysis on the RAxML server (Stamatakis et al. 2008).

Three different data sets were constructed from the cichlid endothelin sequences: a whole data set (wds), a protein-coding data set (cod), and a non-coding data set (noncod). The number of segregating sites (S) and the K_a/K_s values for the wds and the cod were conducted with the software package DnaSP 5.10 (Librado and Rozas 2009). For the cod data sets, sliding window

analyses of the K_a/K_s ratio were also performed with the same program with a window size of 25 and an overlap of 10 bp. This comparison was performed between different groups defined on the basis of existing phylogenies (e.g., outgroup species, and the three large Tanganyikan cichlid clades haplochromines, lamprologines, and ectodines; see, e.g., Salzburger et al. 2002, 2005, 2007) and the remaining samples.

Phylogenetic analyses were performed for three data sets per locus: wds, cod, and noncod along with one concatenated data set. The concatenated data set includes the four endothelin family member data sets plus one additional non-endothelin, autosomal coding gene to increase the total length to 9,451 bp (Diepeveen and Salzburger unpublished data). Due to inconsistent results for the Cyprichromini *Edn1* sequence, this single sequence was omitted from the concatenated data set in an additional analysis. Likelihood ratio tests (LRTs) were carried out with jModeltest 0.1.1 (Guindon and Gascuel 2003; Posada 2008) and Bayesian information criterion and corrected Akaike information criteria were used to determine the best-fitting models of nucleotide substitutions. These models were used in the maximum likelihood (ML) search in PAUP* 4.0 (Swofford 2002) with *T. polylepis* and *O. tanganicae* as outgroup species (see Salzburger et al. 2002). To test the robustness of the resulting topologies, we conducted bootstrap analyses with 100 replicates. Bayesian Inference of phylogeny was conducted in MrBayes 3.2 with 10,500,000 generations in total and *T. polylepis* as outgroup species (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). To determine whether the Markov chain Monte Carlo analysis had run long enough, we applied the Are We There Yet (AWTY) system (Nylander et al. 2008).

To test if the studied loci evolved under non-neutral selection, we analyzed the synonymous (silent; K_s) and nonsynonymous (amino acid-changing; K_a) substitution rates of amino acid sequences within and between the 26 species. First, we ran the data on the Selecton Server (Doron-Faigenboim et al. 2005; Stern et al. 2007). LRTs were conducted between the M8 model (β and $\omega_s \geq 1$), which enables positive selection, and the M8a (β and $\omega_s = 1$; no positive selection) and M7 (only purifying selection) models. Furthermore, we applied a Bayesian approach to predict whether a site is undergoing positive selection, as implemented in Selecton.

Next, a branch test was performed using HyPhy (Kosakovsky Pond et al. 2005) to test whether specific branches showed signs of adaptive sequence evolution. Both the obtained ML phylogeny (see above) and an independent tree based on sequences of the mitochondrial *ND2* gene (Salzburger et al. 2002, 2005) were used for reconstructions of the d_n/d_s ratios on the branches of the

phylogeny. The Suzuki–Gojobori derived adaptation selection tool was then used ($P = 0.05$) to visualize the d_n/d_s ratios on a branch-scaled tree. This method also indicates sites evolving under non-neutral evolution.

The program Codeml from the PAML (Phylogenetic Analysis by ML) 4.2 package (Yang 1997, 2007) was subsequently used to test both site- and branch-specific adaptive evolution of coding sequences. Tests of positive selected sites were conducted by performing LRT of the following model comparisons: M1a (Nearly Neutral) with M2a (Positive Selection), M7 (β) with M8 (β & ω), and M8a (β & $\omega_s = 1$) with M8. The comparison between M0 (one-ratio) and M3 (discrete) was used as a test of variable ω among sites. The naïve empirical Bayes (NEB; Nielsen and Yang 1998; Yang et al. 2000) and the Bayes empirical Bayes (BEB; Yang et al. 2005) criteria were used to calculate the posterior probabilities for site classes and to identify sites under positive selection in case the LRT was significant. In addition to the site models, the branch model (*free-ratios* model with ω allowed to vary) was applied to detect positive selection acting on particular lineages (Yang 1998; Yang and Nielsen 1998).

Finally, we studied the observed nonsynonymous amino acid substitutions in more detail. For this, we compared our data with the information known about the protein domains of both human (*Homo sapiens*) and mouse (*Mus musculus*) sequences of the proteins studied here (source: The Universal Protein Resource database <http://www.uniprot.org/>).

Results

RNA Expression Pattern

The gene expression levels and patterns of the endothelin family members are depicted in Fig. 2 (Real-Time PCR-results) and Supplementary Fig. 1 (PCR-results). Overall the four endothelin family members are widely expressed with *Edn3b* expression not being detectable in three tissues (i.e., liver, muscle, and pharyngeal jaw). The ligand–receptor pair *Edn1* and *EdnrAa* are co-expressed in all craniofacial bone tissues (i.e., maxilla, mandible, and pharyngeal jaw) with a three-fold increase of *EdnrAa* expression in the pharyngeal jaw, although up to three-fold increases are also found in liver, muscle, maxilla, skin, and anal fin (Fig. 2a). The combination of *Edn3b* and *EdnrB1a* expression is found in the pigment containing tissues (i.e., skin and fin tissues), for which also the highest levels of *Edn3b* expression were detected (Fig. 2b). Furthermore, all genes that are expressed in the anal fin are also expressed in both egg-spot tissue and non-egg-spot tissue.

Sequencing Results and Orthology Assessment

The total length of the captured regions for *EdnrAa* was 963 bp (of which 246 bp were protein coding (cod)), 2091 bp (528 bp cod) for *EdnrB1a*, 2449 bp (540 bp cod) for *Edn1*, and 512 bp (195 bp cod) for *Edn3b* (see Table 1).

The ML analyses of the obtained cichlid sequences and the downloaded teleost sequences revealed that cichlid

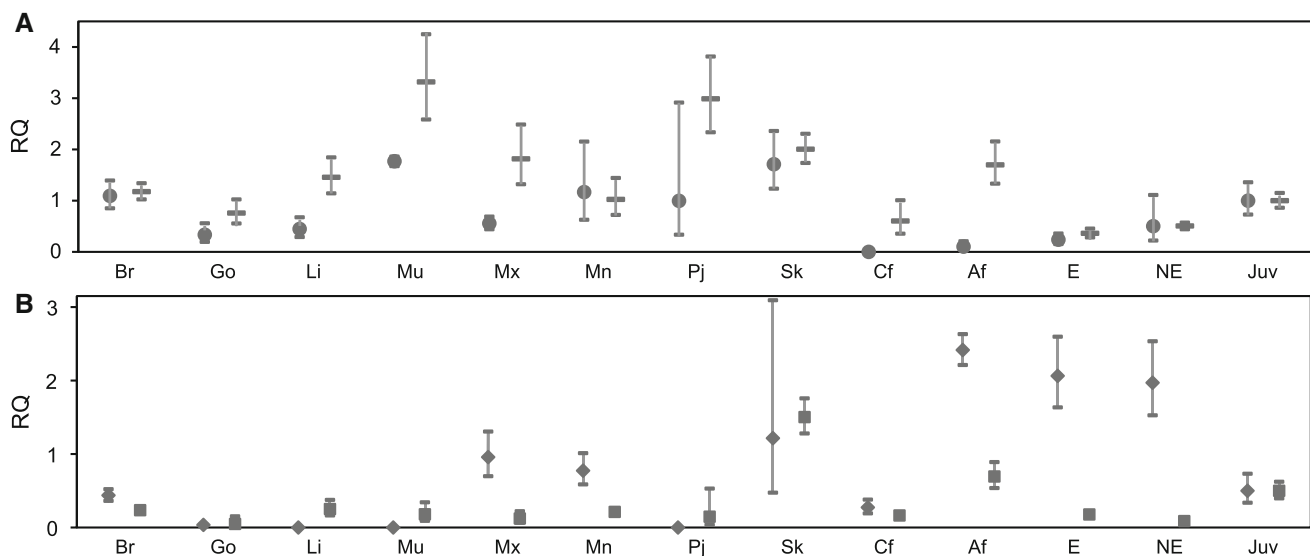


Fig. 2 Gene expression levels of the endothelin family members. **a** Relative quantitation (RQ) plot of the members of the *Edn1* (circles)/*EdnrAa* (bars) pathway. Gene expression was detected for *Edn1* and *EdnrAa* in all tissues except for *Edn1* in the caudal fin tissue. **b** Relative quantitation (RQ) plot for the *Edn3b* (diamonds)/*EdnrB1a* (squares). The two genes appear to be expressed in all tissues (except for *Edn3b* in the liver, muscle, and pharyngeal jaw

tissues), albeit sometimes at low levels. For both analyses, juveniles were used as reference sample (Rf) and *Actin* as endogenous control. Error bars indicate the maximum and minimum expression levels determined by a 95% confidence level. Af anal fin, Br brain, Cf caudal fin, E egg-spot tissue, Go gonads, Li liver, Mn mandible, Mu muscle, Mx maxilla, Ne non-egg-spot tissue, Pj pharyngeal jaw, Sk skin

Table 1 Segregating sites and K_a/K_s values

| Data set | Length ^a | S^b | K_a/K_s^c |
|----------------|---------------------|-------------|-------------|
| <i>Edn1</i> | 2449 (540) | 16.1 (14.7) | 0.2312 |
| <i>Edn3b</i> | 512 (195) | 14.1 (12.9) | 0.5700 |
| <i>EdnrAa</i> | 963 (246) | 16.1 (9.4) | 0.0880 |
| <i>EdnrB1a</i> | 2091 (528) | 17.4 (3.3) | 0.0000 |

^a Total length of the sequenced region and protein coding regions (in parentheses) in base pairs

^b Percentage of segregating sites for total sequenced region and protein coding regions (in parentheses)

^c Nonsynonymous/synonymous substitution rate ratio for protein coding regions

EdnrAa, *EdnrB1a*, *Edn1*, and *Edn3b* each is more closely related to their teleost orthologs than to other teleost paralogues (Supplementary Fig. 2a, b). This indicates that the obtained cichlid sequences are indeed the presumed endothelin genes.

Phylogenetic Analysis

The wds consensus Bayesian phylogenies (BAY) with ML bootstrap values and BAY posterior probabilities are depicted in Supplementary Fig. 3a–e. In most data sets, both ML and BAY methods revealed a similar topology that is largely in agreement with available species trees (Salzburger et al. 2002, 2005, 2007). This topology consists of the Tylochromini (*T. polylepis*), Tilapiini (*O. tanganyicae*), and the endemic Bathybatini, Boulengerochromini, and Trematocarini being placed in a basal position with respect to the substrate spawning Lamprologini, the Eretmodini and the ‘C-lineage’ (Cyphotilapiini, Cyprichromini, Ectodini, Limnochromini, Perissodini, and the Haplochromini/Tropheini clade; Clabaut et al. 2005; Day et al. 2008; Nishida 1991; Salzburger et al. 2002, 2005; Takahashi 2003).

Although, relationships between tribes are generally not recovered by the individual gene trees (e.g., *Edn3b* tree; Supplementary Fig. 3c), the data sets do contain an adequate amount of phylogenetic signal to group members of a tribe together. This is further confirmed by the moderate to high support values for the monophyly of the major lineages, while support was typically low for many interspecific relationships—reflecting a limited number of variable sites and in the case of *Edn3b* a small size of the data set (ca. 500 bp).

As in previous phylogenetic analyses, the position of *Eretmodus cyanostictus* remains unclear (Clabaut et al. 2005; Day et al. 2008; Salzburger et al. 2002); either unresolved as part of a major polytomy (Supplementary Fig. 3b–d) or even at a derived position within the C-lineage (Supplementary Fig. 3a). A similar observation was made for *Cyprichromis leptosoma*; a position as sister

group to the perissodines was alternated with polytymous positions and even a position within the Lamprologines in the *Edn1* phylogeny. Although this clustering was not supported by the cod and noncod analyses (data not shown), the same signal was retrieved when the wds was re-analyzed in parts of 800 bp, indicating that convergence via, e.g., lateral gene transfer or an hybridization event could have occurred.

Segregating Sites and K_a/K_s Values

The number of segregating sites (S) and the K_a/K_s values for each of the studied loci are depicted in Table 1. The lowest value for S was observed for the *EdnrB1a* cod data set, while the *EdnrB1a* wds data set was characterized by the highest value for this parameter. All the endothelin family members were characterized by K_a/K_s ratios <1 indicating purifying selection. Remarkable low K_a/K_s ratios were found for both endothelin receptors. When comparing the K_a/K_s ratios of ligands versus receptors we found significantly higher ratios in the ligands (unpaired T test, $P = 0.02$; Fig. 3a). Furthermore, the K_a/K_s ratios observed in the ligands is significantly higher than the mean K_a/K_s ratios observed by Baldo et al. (2011; Fig. 3b) in a pairwise comparison of 1,098 EST’s between *Ophthalmotilapia ventralis* and *A. burtoni* (unpaired T test, $P < 0.01$).

Selection Pressure on Sites and Lineages

Even though the endothelin family members have conserved functions among all vertebrates, which would be a cause for selection against deleterious mutations, the co-opted functions of these genes during the development

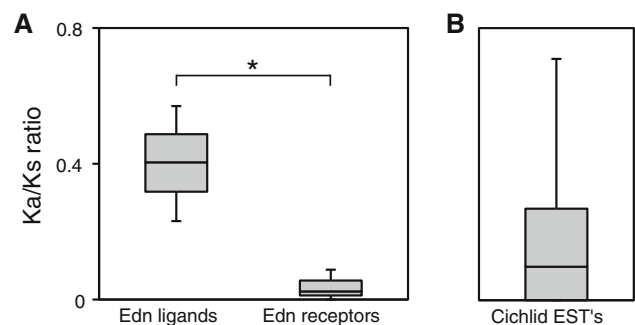


Fig. 3 K_a/K_s comparison between the endothelin ligands, endothelin receptors, and cichlid EST’s. **a** The K_a/K_s values for the Edn ligands and Edn receptors. Higher values were found for the ligands than for the receptors ($P = 0.02$; unpaired T test; $N = 3$). **b** The K_a/K_s values for the 1,055 EST’s derived from a genome wide approach based on a *O. ventralis/A. burtoni* comparison performed by Baldo et al. (2011). In this data set, 43 outliers were detected. Outlier values ranged between 0.74 and 2.66 and were omitted from the figure, but not from the analysis. Depicted are the mean, first and third quartile, and the minimum and maximum K_a/K_s values for each of the three groups

and evolution of two of the cichlids key innovations might be of such importance that this would have left signs in the genome by means of positive selection. If the latter is the case, differences in the strength of the selection pressure between lineages are to be expected, especially for the *Edn3b/EdnrB1a* pathway, since not all species possess egg-spots and/or colorful body patterns.

Table 2 summarizes the results of the various tests that were performed to infer sites and/or lineages under non-neutral evolution. Negatively selected sites were found in all loci. None of the LRT's of the PAML (BEB) and Selecton analyses for the endothelin family members were significant, suggesting that these loci evolved under purifying selection. Trends toward significance were found for *EdnrAa* and *Edn1* with Selecton; while with NEB in PAML 23 and 15 positively selected sites were identified for *Edn1* and *Edn3b*, respectively.

Positive selected branches were found with HyPhy for all loci except *EdnrAa* (Table 2). In *EdnrAa*, only one species (*C. macrops*) showed a slightly elevated K_a/K_s or d_n/d_s ratio ($d_n/d_s = 0.6$ (PAML) – 0.9 (HyPhy)). Most species and clades showed d_n/d_s ratios around zero in *EdnrB1a* except for three Tropheini/Haplochromini species (*L. labiatus*, *S. diagramma*, *T. moori*), one Perissodini species (*P. straeleni*) and *C. frontosa*, for which elevated d_n/d_s ratios were detected. Nonetheless, other species of the Tropheini and Perissodini did not show elevated d_n/d_s ratios. In the branch-specific PAML analysis, no signs of positive selection were found in *EdnrB1a*.

Besides species-specific increases of K_a/K_s and d_n/d_s ratios the analysis of the *Edn1* data set also showed lineage-specific increases. All major branches towards the derived clades showed increased K_a/K_s and d_n/d_s ratios.

Both lineage and species-specific increases of K_a/K_s and d_n/d_s ratios were also found in *Edn3b*. Again several

internal branches showed increased K_a/K_s and d_n/d_s ratios. The results for the internal branches were not unambiguous due to incongruent results of the separate analyses. The Perissodini were the only clade that consistently showed the same strong signal (i.e., positive selection) in all analyses. Furthermore, the *Edn3b* sequences of the following species showed signs of positive selection; *G. permaxillaris*, *N. pulcher*, *B. microlepis*, *T. nigrifrons*, *O. tanganicae* and the clade consisting of *S. diagramma* and *P. curvifrons*.

Sliding Window Analysis

The results of the sliding window analyses are depicted in Fig. 4. For *EdnrAa* there is only one region found with elevated omega values (K_a/K_s) in all analyzed groups. Two regions of weakly elevated omega values were observed for *EdnrB1a*. It seems that the Haplochromini have the highest omega value at the first peak, while the second peak is found in more groups (i.e., the two outgroup species, Haplochromini, Lamprologini, and Perissodini). Nonsynonymous substitutions were found in both the transmembrane and the extracellular and cytoplasmic domains of both the endothelin receptors (see also Supplementary Table 3).

For the *Edn1* and *Edn3b* data sets, several peaks were identified in the preproendothelin for which all groups surveyed show elevated omega values for at least one of these peaks. In all groups, we have found species or lineage-specific nonsynonymous substitutions that coincide with these peaks. For both endothelins, the highest scores are found for the outgroup species *T. polylepis*. Furthermore, the results for the Perissodini show two strong peaks, caused by two lineage-specific substitutions. One remarkable nonsynonymous substitution is found in *E. cyanostictus*; a transversion led to a Threonine–Serine substitution at the second amino acid position of the functional protein. A Serine at this position is also observed in the *Edn1*, 2, and 4 proteins.

Discussion

In this study, we investigated the molecular evolution of two endothelin ligand–receptor pairs (*Edn1–EdnrAa* and *Edn3b–EdnrB1a*) in East African cichlid fishes. We performed a variety of phylogenetic and molecular analyses in a set of 26 species and assessed the expression of the genes under study in twelve different tissues of a reference species. Here we discuss our results with respect to the observed gene expression pattern, selection pressure, and their functional and evolutionary implications.

Table 2 Overview of selection pressure on sites and branches

| Locus | Selecton/ HyPhy n.s.s. ^a | Selecton/ Hyphy p.s.s. ^b | PAML ^c NEB/BEB p.s.s. | HyPhy/ PAML p.s.b. ^d | Lineages |
|----------------|---|---|--|---------------------------------------|---|
| <i>Edn1</i> | 33/5 | – (4)– | 23/– | Yes/yes | All |
| <i>Edn3b</i> | 14/3 | –/– | 15/– | Yes/yes | All |
| <i>EdnrAa</i> | 16/3 | –(1)– | –/– | –/– | – |
| <i>EdnrB1a</i> | 37/8 | –/– | –/– | Yes/– | Cypho, Periss, Troph ^e |

^a Negative selected sites ($P \leq 0.05$)

^b Positive selected sites ($P \leq 0.05$); values in brackets: Selecton $P \leq 0.10$

^c Models M2a and M8; NEB ($P \leq 0.01$); BEB ($P \leq 0.05$); for *Edn1* only model M2a

^d Positive selected branches

^e Cyphotilapiini, Perissodini, Tropheini

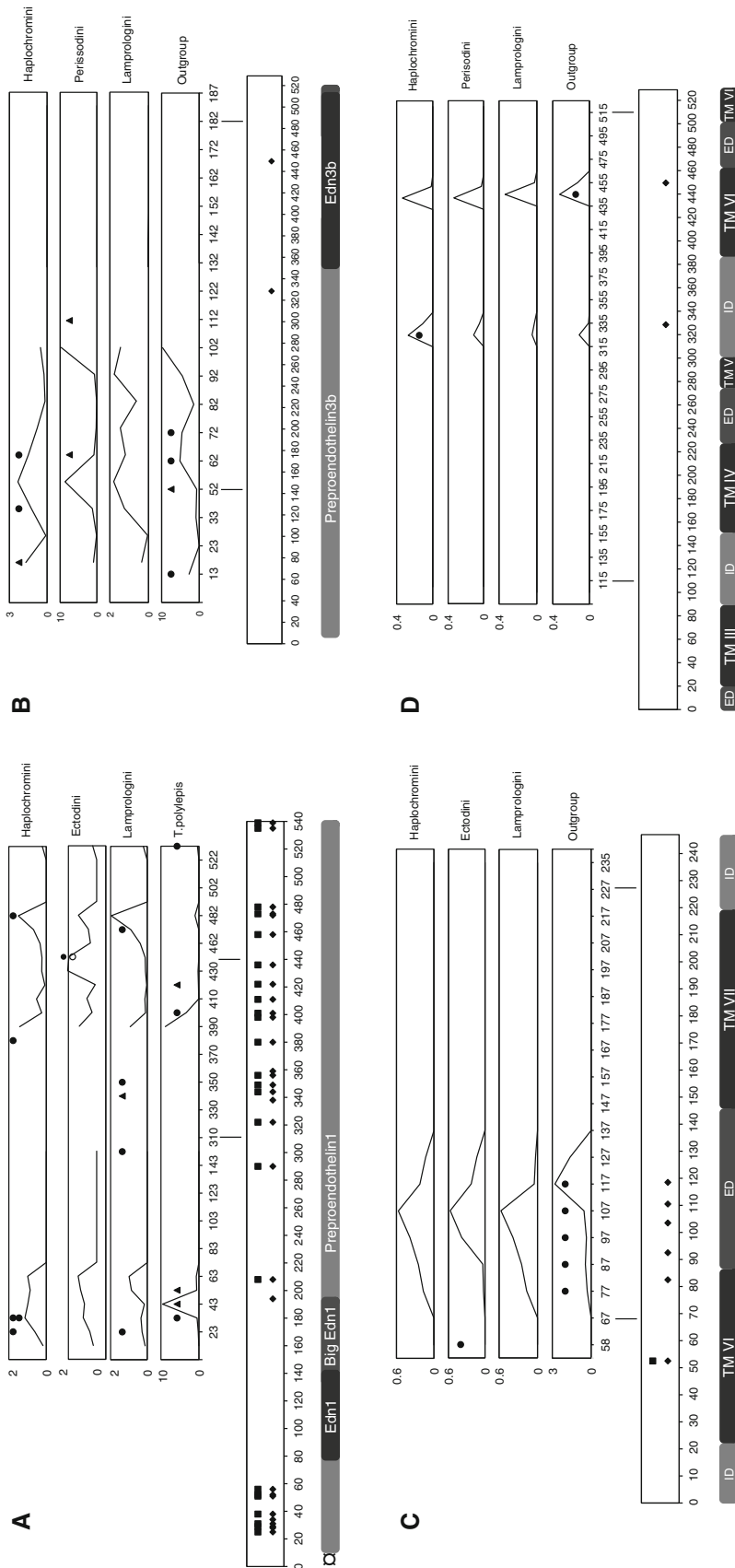


Fig. 4 Sliding window plots of the endothelin ligands or receptors. Species-specific nonsynonymous substitutions are displayed by *filled circles* in the sliding window plots, while *filled triangles* depict lineage-specific nonsynonymous substitutions. Comparisons were performed between the members of the specified group and the remaining samples. Window sizes were 25 bp with 10 bp overlap. *Bars* show corresponding regions between the sliding windows and the overview of putative positive selected sites (*filled squares*) and the nonsynonymous substitutions (*filled diamonds*). Functional domains are depicted as *gray-tone blocks*. **a** Elevated K_a/K_s ratio, positive selected sites (i.e., PAML NEB, $P \leq 0.01$) and nonsynonymous sites are mainly found in the preproendothelin 1, with one nonsynonymous substitution found in big endothelin 1. The sequenced part of the signaling protein is depicted by the *filled circle symbol*. **b** One nonsynonymous substitution was found in the transmembrane and extracellular parts of EdnrAa. The positive selected sites were found with the NEB ($P \leq 0.01$). **c** Nonsynonymous substitutions were found in the transmembrane and extracellular parts of EdnrB1a hold only two nonsynonymous substitutions intracellular and transmembrane parts of EdnrB1a

Gene Expression Assays

Our gene expression assays reveal that the members of the two studied endothelin pathways are expressed in a broad variety of tissues in the cichlid *A. burtoni* (Fig. 2, Supplementary Fig. 1), which is in concordance with the wide range of functions known for the endothelin family members. Most studied and described functions of the endothelin system are related to the blood pressure regulation functions of *Edn1* in adult vertebrates (e.g., Yanagisawa et al. 1988) and the functions during early development of neural crest derivatives (e.g., Miller et al. 2000; Parichy et al. 2000). Yet, the endothelin system is not limited to vasoconstriction or early development, since several neural crest derivatives such as pigmentation and the (dentition on the cichlids') pharyngeal jaws are produced and maintained throughout adulthood. Here we show that the two endothelin pathways under study are differently expressed in the pharyngeal jaw apparatus and anal fin egg-spots. We find *Edn1* and *EdnrAa* co-expression in the pharyngeal jaw, while *Edn3b* and *EdnrB1a* are co-expressed in the anal fin and egg-spot tissue (Fig. 2, Supplementary Fig. 1). This is in agreement with previously described endothelin functions: *EdnrB1/EdnrB2* duplicates and *Edn3* are known to be involved in pigmentation in a range of vertebrates including zebrafish (e.g., Parichy et al. 2000; Pla and Larue 2003); whereas a correlation between *Edn1*, *EdnrA* duplicates and the pharyngeal bone development has been uncovered in zebrafish (Kimmel et al. 2003; Pla and Larue 2003).

Deciphering the Selection Pressure Acting Upon the Endothelin Family Members in Cichlids

Patterns of DNA sequence variation are important indicators of the mode of selection acting upon a locus. In particular, the ratio of non-synonymous substitutions per non-synonymous site (K_a) to synonymous substitutions per synonymous site (K_s) is considered a powerful tool in the study of adaptive evolution (Hughes 1997). Under neutrality synonymous and non-synonymous substitutions are assumed to occur at equal rates, resulting in a ratio close to one. Purifying selection will remove deleterious mutations, thus, leading to a decrease of the K_a/K_s ratio, while positive selection will favor advantageous mutations, which is characterized by a K_a/K_s ratio exceeding one. Theory predicts that both positive and purifying selection can occur after gen(om)e duplication events. Although the fate of most duplicated genes is nonfunctionalization (e.g., Ohno 1970; Zhang 2003), models of adaptive evolution by gene duplication (Hughes 1999) predict that soon after a duplication event positive selection acts on mutations that lead to new or enhanced functions (i.e., neofunctionalization).

Subsequently, once having evolved, these functions are maintained by purifying selection acting at the gene level. Neofunctionalization is, however, not the only outcome that leaves a signature of purifying selection in the genes after a duplication event. Selection against deleterious mutations could also occur subsequent to subfunctionalization. Subfunctionalization is known to have occurred in the zebrafish *EdnrA* paralogs, which have partially redundant functions during lower jaw formation (Nair et al. 2007).

Here we show that the endothelin ligand and receptor family members evolved under purifying selection. Interestingly, substantial variation between loci was observed in the K_a/K_s comparisons. While we detected very low K_a/K_s values for the two endothelin receptors, the two endothelin ligands showed elevated rates (Table 1), even when compared to genome wide comparisons between two mouthbrooding species from Lake Tanganyika (Fig. 3). The accelerated rate of protein evolution observed in the ligands is similar to what has been observed for *bmp4* by Terai et al. (2002b). In both type of genes, the variation accounting for the elevated rate of protein evolution is not observed in the functional protein domain, but rather in the pro-domain/preproendothelin parts. Therefore, the mature proteins are likely to be functional, although posttranslational regulation or cleavage efficiency could have been affected (see next section; Terai et al. 2002b).

Implications on the Functional Level of Endothelins and Their Receptors in Cichlids

As mentioned before, the extracellular section and loops of the endothelin receptors are used for ligand binding, while interaction with the G protein takes place at the intracellular loops (Masaki et al. 1999; Orry and Wallace 2000). The extracellular part can be further divided into an “address” (ligand–receptor selectivity) and a “message” (ligand–receptor binding) part (Masaki et al. 1999; Sakamoto et al. 1993). In both endothelin receptors under study, almost all non-synonymous changes (detected in *T. polylepis*, *C. macrops*, and *O. tanganyicae*) were confined to the “address” region. The additional substitution in *EdnrB1a* (specific for *A. burtoni*) concerns the cytoplasmic part (G protein interaction). Thus, most changes in the endothelin receptors of cichlids are found in the parts responsible for ligand–receptor interaction, supporting the co-evolution theory of Braasch et al. (2009).

The majority of the changes in the ligands affected the preproendothelin part with only two exceptions: one substitution occurred in the big endothelin 1 of *N. furcifer* and one in the *Edn3b* mature protein part of *E. cyanostictus*. This last substitution is found at the second amino acid of the mature protein, a site that is also variable among other

endothelins (Braasch et al. 2009). A serine at this position is commonly observed in the Edn1, 2, and 4 proteins among teleost and other vertebrates, but—so far—not in any of the Edn3 duplicates.

Because of the particular way of protein maturation (preproendothelin > big endothelin > endothelin), it might be speculated that the observed mutations in the preproendothelin part could affect posttranslational regulation. At least for *endothelin 1*, it is known that the regulation is ultimately controlled by the rate of transcription. *Cis*-acting elements both in the promoter and further upstream, epigenetic mechanisms and several other types of regulatory interactions (e.g., calcium levels, extracellular hypertonicity) do play an important role, too (reviewed in Stow et al. 2011). It was further shown that several regulatory interactions take place directly on the *Edn1* mRNA (e.g., through GAPDH, microRNA), which can either be stimulating or repressing (Stow et al. 2011).

The observed changes in the preproendothelin parts could—in principle—also affect the cleaving process. We found that the pairs of basic amino acids, where the actual cleavage takes place, stay unaffected in cichlids. However, substitutions in the preproendothelin part could still affect the efficiency of cleavage and with that the concentration of the mature protein (the mature protein itself stays unaffected and would function normally). Definite conclusions about the functional effects of the nonsynonymous changes can, however, only be made on the basis of functional assays (e.g., Hoekstra et al. 2006).

Evolutionary Implications

The pharyngeal jaw apparatus and anal fin egg-spots are two putative evolutionary key innovations of cichlids. A key innovation is a novel trait in evolution that enables the exploitation of new niches and, hence, could initiate adaptive radiation (Gavrilets and Losos 2009; Schluter 2000; Simpson 1953). In this way, the trait is directly responsible for the evolutionary success of a radiating group. Little is known, however, about how the relationship between trait evolution and organismal diversification is reflected in the genes and genetic pathways underlying such innovations. Often it is assumed that genes underlying (ecologically) important traits show signs of non-neutral evolution. In cichlids, for example, the downstream color gene *csf1ra*, which is expressed in egg-spots, appears to have evolved adaptively in the direct ancestor of the egg-spot bearing haplochromines (Salzburger et al. 2007). An accelerated protein evolution has also been found in *bmp4*, which is involved in (pharyngeal) bone formation (Fraser et al. 2009; Terai et al. 2002b). It is unclear how common such signatures of positive selection are in the genes underlying evolutionary innovations. Obviously, many

more genes are likely to be involved in the evolution of a novel trait than the single/few positively selected genes in the above-mentioned examples. Also, in cases where genes underlying a novel trait are part of a conservative pathway that is co-opted, one would expect a signature of purifying rather than positive selection. This scenario could have happened with the endothelin family members. For instance, when the conserved pathway involved in the development of neural crest-derived tissues was co-opted during the evolution of key innovations in cichlid fishes.

Speciation and organismal diversification can be greatly facilitated by the concurrent action of natural and sexual selection. This is particularly the case when an ecologically important, hence naturally selected trait, is at the same time the target of mate choice. In sticklebacks, for example, it has been shown that body size, which is influenced by the foraging habitat, also acts as trigger in mate preference (Hatfield and Schluter 1999; Nagel and Schluter 1998). Feulner et al. (2009) suggested that the electric organ discharge of weakly electric fish is a magic trait serving for both electrolocation (foraging) and mate recognition (mate choice). Another single trait that is both under disruptive natural selection and acts as a cue for assortative mating is the color patterns of the coral reef fishes of the genus *Hypoplectrus* (Puebla et al. 2007).

Just as natural and sexual selection are both affecting a single trait in these examples, the endothelin family as a whole might be affected by both natural and sexual selection. The endothelin family members, thus, form a link between a naturally selected trait (mouth morphology, including the pharyngeal jaws) and a sexually selected one (coloration)—via their underlying genes and genetic pathways. This makes the endothelin gene family a strong candidate for a more thorough functional, developmental, and evolutionary examination in cichlid fishes.

Conclusions

In this work, we studied the molecular evolution of two ligand-receptor pairs of the endothelin system (Edn1/EdnrAa and Edn3b/EdnrB1a) in a representative set of 26 East African cichlid species. Sequences of the endothelin family members show signs of purifying selection in cichlids, a common force observed after gen(om)e duplication events. The two ligands show an accelerated rate of protein evolution in comparison to the receptors, most likely caused by the highly variable preproendothelin parts. No evidence was found for one of the endothelin pathways to have evolved under positive selection in one or more lineages. While the amino acids of the mature endothelins are rather conserved in cichlids, the variable preproendothelin parts could influence posttranslational regulation.

We also found several changes in the endothelin receptors that could possibly affect the ligand-receptor binding functions. Future functional studies should test these hypotheses. Expression pattern analyses confirm that *Edn1/EdnrAa* are co-expressed at a relatively high expression level in the pharyngeal jaw tissue, whereas *Edn3b/EdnrB1a* at relatively high levels in skin and anal fin tissue (including the anal fin egg-spots). Because of their anticipated function during the morphogenesis of the (pharyngeal) jaw apparatus and coloration and pigmentation, the endothelin gene family members form an important link between these naturally and sexually selected traits in cichlids.

Acknowledgments We would like to thank Brigitte Aeschbach and Nicolas Boileau for help and advice on the lab work, Michael Matschiner for advice on the phylogenetic analyses, Emilia Santos for providing the “egg-spots” and “non-egg-spots” RNA samples and Laura Baldo for providing the K_a/K_s values from the cichlid EST comparison. We are very grateful to Ingo Braasch and the anonymous reviewers for their valuable comments and suggestions on earlier version of this manuscript. This study was supported by the European Research Council (Starting Grant “INTERGENADAPT” to WS) and the Swiss National Science Foundation (Grant 3100A0_122458 to WS).

Conflict of interest The authors declare that they have no conflict of interest.

References

- Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 21:2104–2105
- Albertson RC, Kocher TD (2006) Genetic and developmental basis of cichlid trophic diversity. *Heredity* 97:211–221
- Albertson RC, Streebman JT, Kocher TD, Yelick PC (2005) Integration and evolution of the cichlid mandible: the molecular basis of alternate feeding strategies. *Proc Natl Acad Sci USA* 102:16287–16292
- Baldo L, Santos EM, Salzburger S (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
- Barlow GW (2000) The cichlid fishes: nature’s grand experiment in evolution. Perseus publishing, Cambridge
- Baynash AG, Hosoda K, Giaid A, Richardson JA, Emoto N, Hammer RE, Yanagisawa M (1994) Interaction of endothelin-3 with endothelin-B receptor is essential for development of epidermal melanocytes and enteric neurons. *Cell* 79:1277–1285
- Bonano M, Tribulo C, De Calisto J, Marchant L, Sánchez SS, Mayor R, Aybar MJ (2008) A new role for the endothelin-1/endothelin-A receptor signaling during early neural crest specification. *Dev Biol* 323:114–129
- Braasch I, Volf J-N, Schartl M (2009) The endothelin system: evolution of vertebrate-specific ligand-receptor interactions by three rounds of genome duplication. *Mol Biol Evol* 26:783–799
- Bruford MW, Hanotte O, Brookfield JFY, Burke T (1998) Multilocus and single-locus DNA fingerprinting. In: Hoelzel AR (ed) *Molecular genetic analysis of populations: a practical approach*. Oxford University Press, Oxford, pp 287–336
- Chung I-H, Yamaza T, Zhao H, Chuong P-H, SHi S, Chai Y (2009) Stem cell property of postmigratory cranial neural crest cells and their utility in alveolar bone regeneration and tooth development. *Stem Cells* 27:866–877
- Clabaut C, Salzburger W, Meyer A (2005) Comparative phylogenetic analyses of the adaptive radiation of Lake Tanganyika cichlid fish: nuclear sequences are less homoplasious but also less informative than mitochondrial DNA. *J Mol Evol* 61:666–681
- Clouthier DE, Hosoda K, Richardson JA, Williams SC, Yanagisawa H, Kuwaki T, Kumada M, Hammer RE, Yanagisawa M (1998) Cranial and cardiac neural crest defects in endothelin-A receptor-deficient mice. *Development* 125:813–824
- Coulter GW (1991) *Lake Tanganyika and its life*. British Museum (Natural History) and Oxford University Press, Oxford
- Day JJ, Cotton JA, Barraclough TG (2008) Tempo and mode of diversification of Lake Tanganyika cichlid fishes. *PLOS One* 3:e1730. doi:10.1371/journal.pone.0001730
- Doron-Faigenboim A, Stern A, Bacharach E, Pupko T (2005) Selecton: a server for detecting evolutionary forces at a single amino-acid site. *Bioinformatics* 21:2101–2103
- Drummond A, Strimmer K (2001) PAL: an object-oriented programming library for molecular evolution and phylogenetics. *Bioinformatics* 17:662–663
- 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
- Egger B, Klaefiger Y, Theis A, Salzburger W (2011) A sensory bias has triggered the evolution of egg-spots in cichlid fishes. *PLoS One* 6:e25601
- Elmer KR, Kusche H, Lehtonen TK, Meyer A (2010) Local variation and parallel evolution: morphological and genetic diversity across a species complex of neotropical crater lake cichlid fishes. *Philos Trans R Soc Lond B* 365:1763–1782
- Feulner PGD, Plath M, Engelmann J, Kirschbaum F, Tiedemann R (2009) Electrifying love: electric fish use species-specific discharge for mate recognition. *Biol Lett* 5:225–228
- Fraser GJ, Bloomquist RF, Streebman JT (2008) A periodic pattern generator for dental diversity. *BMC Biol* 6:32. doi:10.1186/1741-7007-6-32
- Fraser GJ, Hulseley CD, Bloomquist RF, Uyesugi K, Manley NR, Streebman TJ (2009) An ancient gene network is co-opted for teeth on old and new jaws. *PLoS Biol* 7:e31. doi:10.1371/journal.pbio.1000031
- Fryer G, Iles TD (1972) *The cichlid fishes of the Great Lakes of Africa: their biology and Evolution*. Oliver & Boyd, Edinburgh
- Gans C, Northcutt RG (1983) Neural crest and the origin of vertebrates: a new head. *Science* 220:268–273
- Gavrilets S, Losos JB (2009) Adaptive radiation: contrasting theory with data. *Science* 323:732–737
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum-likelihood. *Syst Biol* 52:696–704
- Hall BK (1999) *The neural crest in development and evolution*. Springer-Verlag, New York
- Hatfield T, Schluter D (1999) Ecological speciation in sticklebacks: environment-dependent hybrid fitness. *Evolution* 53:866–873
- Hert E (1989) The function of egg-spots in an African mouth-brooding cichlid fish. *Anim Behav* 37:726–732
- Hoekstra HE, Hirschmann RJ, Bunday RA, Insel PA, Crossland JP (2006) A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* 313:101–104
- Hosoda K, Hammer RE, Richardson JA, Baynash AG, Cheung JC, Giaid A, Yanagisawa M (1994) Targeted and natural (piebald-lethal) mutations of endothelin-B receptor gene produce

- megacolon associated with spotted coat color in mice. *Cell* 79:1267–1276
- Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17:754–755
- Hughes AL (1997) Rapid evolution of immunoglobulin superfamily C2 domains expressed in immune system cells. *Mol Biol Evol* 14:1–5
- Hughes AL (1999) Adaptive evolution of genes and genomes. Oxford University Press, Oxford
- Hunter JP (1998) Key innovations and the ecology of macroevolution. *Trends Ecol Evol* 13:31–36
- Keenleyside MHA (1991) Cichlid fishes: behaviour, ecology and evolution. Chapman & Hall, London
- Kimmel CB, Ullmann B, Walker M, Miller CT, Crump JG (2003) Endothelin 1-mediated regulation of pharyngeal bone development in zebrafish. *Development* 130:1339–1351
- Kocher TD (2004) Adaptive evolution and explosive speciation: the cichlid fish model. *Nat Genet* 5:288–298
- Kocher TD, Conroy JA, McKaye KR, Stauffer JR (1993) Similar morphologies of cichlid fish in lakes Tanganyika and Malawi are due to convergence. *Mol Phylogenet Evol* 2:158–165
- Kornfield I, Smith PF (2000) African cichlid fishes: model systems for evolutionary biology. *Annu Rev Ecol Syst* 31:163–196
- Kosakovsky Pond SL, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21:676–679
- Kurihara Y, Kurihara H, Suzuki H, Kodama T, Maemura K, Nagai R, Oda H, Kuwaki T, Cao W-H, Kamada N, Jishage K, Ouchi Y, Azuma S, Toyoda Y, Ishikawa T, Kumada M, Yazaki Y (1994) Elevated blood-pressure and craniofacial abnormalities in mice deficient in endothelin-1. *Nature* 368:703–710
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452
- Liem KF (1973) Evolutionary strategies and morphological innovations: cichlid pharyngeal jaws. *Syst Biol* 22:425–441
- Masaki T (2004) Historical review: endothelin. *Trends Pharmacol Sci* 25:219–224
- Masaki T, Ninomiya H, Sakamoto A, Okamoto Y (1999) Structural basis of the function of endothelin receptor. *Mol Cell Biochem* 190:153–156
- McGill GG, Horstmann M, Widlund HR, Du J, Motyckova G, Nishimura E, Lin Y-L, Ramaswamy S, Avery W, Ding H-F, Jordan S, Jackson I, Korsmeyer S, Golub T, Fisher D (2002) Bcl2 regulation by the melanocyte master regulator Mitf modulates lineage survival and melanoma cell viability. *Cell* 109:707–718
- Meyer A (1993) Phylogenetic relationships and evolutionary processes in East African cichlid fishes. *Trends Ecol Evol* 8:279–284
- Meyer A, van de Peer Y (2005) From 2R to 3R: Evidence for a fish-specific genome duplication (FSGD). *Bioessays* 27:937–945
- Miller CT, Schilling TF, Lee KH, Parker J, Kimmel CB (2000) Sucker encodes a zebrafish endothelin-1 required for ventral pharyngeal arch development. *Development* 127:3815–3828
- Mrowka W (1987) Egg stealing in a mouthbrooding cichlid fish. *Anim Behav* 35:923–925
- Muschick M, Barluenga M, Salzburger W, Meyer A (2011) Adaptive phenotypic plasticity in the Midas cichlid fish pharyngeal jaw and its relevance in adaptive radiation. *BMC Evol Biol* 11:116. doi:10.1186/1471-2148-11-116
- Nagel L, Schluter D (1998) Body size, natural selection, and speciation in sticklebacks. *Evolution* 52:209–218
- Nair S, Li W, Cornell R, Schilling TF (2007) Requirements for endothelin type-A receptors and endothelin-1 signaling in the facial ectoderm for the patterning of skeletogenic neural crest cells in zebrafish. *Development* 134:335–345
- Nielsen R, Yang ZH (1998) Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* 148:929–936
- Nishida M (1991) Lake Tanganyika as an evolutionary reservoir of old lineages of East-African cichlid fishes: inferences from allozyme data. *Cell Mol Life Sci* 47:974–979
- Notredame C, Higgins DG, Heringa J (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302:205–217
- Nylander JAA, Wilgenbusch JC, Warren DL, Swofford DL (2008) AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24:581–583
- Ohno S (1970) Evolution by gene duplication. Springer Verlag, New York
- Orry AJW, Wallace BA (2000) Modeling and docking the endothelin G-protein-coupled receptor. *Biophys J* 79:3083–3094
- Ozeki H, Kurihara Y, Tonami K, Watatani S, Kurihara H (2004) Endothelin-1 regulates the dorsoventral branchial arch patterning in mice. *Mech Dev* 121:387–395
- Parichy DM, Mellgren EM, Rawls JF, Lopes SS, Kelsh RN, Johnson SL (2000) Mutational analysis of the endothelin receptor b1 (rose) during neural crest and pigment pattern development in the zebrafish *Danio rerio*. *Dev Biol* 227:294–306
- Pla P, Larue L (2003) Involvement of endothelin receptors in normal and pathological development of neural crest cells. *Int J Dev Biol* 47:315–325
- Poirot O, O'Toole E, Notredame C (2003) Tcoffee@igs: a web server for computing, evaluating and combining multiple sequence alignments. *Nucleic Acids Res* 31:3503–3506
- Posada D (2008) jModelTest: phylogenetic model averaging. *Mol Biol Evol* 25:1253–1256
- Puebla O, Bermingham E, Guichard F, Whiteman E (2007) Colour pattern as a single trait driving speciation in *Hypoplectrus* coral reef fishes? *Proc R Soc Lond B* 274:1265–1271
- Rawls JF, Mellgren EM, Johnson SL (2001) How the zebrafish gets its stripes. *Dev Biol* 240:301–314
- 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, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574
- Sakamoto A, Yanagisawa M, Sakurai T, Nakao K, Toyo-oka T, Yano M, Masaki T (1993) The ligand-receptor interactions of the endothelin systems are mediated by distinct “message” and “address” domains. *J Cardiovasc Pharmacol* 22:S113–S116
- 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, Meyer A, Baric S, Veheyen 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
- 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. doi:10.1186/1471-2148-5-17
- 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 Biol* 5:51. doi:10.1186/1741-7007-5-51
- Schluter D (2000) The ecology of adaptive radiation. Oxford University Press, Oxford
- Seehausen O (2006) African cichlid fish: a model system in adaptive radiation research. *Proc R Soc Lond B* 273:1987–1998
- Seehausen O, Alphen v JJM, Witte F (1997) Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* 277:1808–1811

- Shimeld SM, Holland PWH (2000) Vertebrate innovations. *Proc Natl Acad Sci USA* 97:4449–4452
- Simpson GG (1953) The major features of evolution. Columbia University Press, New York
- Slack JMW (2001) Essential developmental biology. Blackwell Publishing, Oxford
- Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web-servers. *Syst Biol* 57:758–771
- 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
- Stow LR, Jacobs ME, Wingo CS, Cain BD (2011) Endothelin-1 gene regulation. *FASEB J* 25:16–28
- Swofford DL (2002) PAUP*. Phylogenetic analysis using parsimony (*and other methods). VERSION 4. Sinauer Associates, Sunderland
- Takahashi T (2003) Systematics of Tanganyikan cichlid fishes (Teleostei: Perciformes). *Ichthyol Res* 50:367–382
- Terai Y, Morikawa N, Kawakami K, Okada N (2002a) Accelerated evolution of the surface amino acids in the WD-repeat domain encoded by the hageromo gene in an explosively speciated lineage of East African cichlid fishes. *Mol Biol Evol* 19:574–578
- Terai Y, Morikawa N, Okada N (2002b) The evolution of the prodomain of bone morphogenetic protein 4 (*Bmp4*) in an explosively speciated lineage of East African cichlid fishes. *Mol Biol Evol* 19:1628–1632
- Terai Y, Morikawa N, Kawakami K, Okada N (2003) The complexity of alternative splicing of *hageromo* mRNAs is increased in an explosively speciated lineage in East African cichlids. *Proc Natl Acad Sci USA* 100:12798–12803
- Theis A, Salzburger W, Egger B (2012) The function of anal fin egg-spots in the cichlid fish *Astatotilapia burtoni*. *PloS ONE* 7(1): e29878. doi:10.1371/journal.pone.0029878
- Turner GF, Burrows MT (1995) A model of sympatric speciation by sexual selection. *Proc Biol Sci Lond B* 260:287–292
- Wickler W (1962) Egg-dummies as natural releasers in mouth-breeding cichlids. *Nature* 194:1092–1093
- Yanagisawa M, Kurihara H, Kimura S, Tomobe Y, Kobayashi M, Mitsui Y, Yazaki Y, Goto K, Masaki T (1988) A novel potent vasoconstrictor peptide produced by vascular endothelial cells. *Nature* 332:411–415
- Yanagisawa H, Clouthier DE, Richardson JA, Charite J, Olson EN (2003) Targeted deletion of a branchial arch-specific enhancer reveals a role of *dHAND* in craniofacial development. *Development* 130:1069–1078
- Yang ZH (1997) PAML, A program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 13: 555–556
- Yang ZH (1998) Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol Biol Evol* 15:568–573
- Yang ZH (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586–1591
- Yang ZH, Nielsen R (1998) Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *J Mol Evol* 46:409–418
- Yang ZH, Nielsen R, Goldman N, Pedersen A-MK (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155:431–449
- Yang ZH, Wong WSW, Nielsen R (2005) Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol Biol Evol* 22:1107–1118
- Zhang JZ (2003) Evolution by gene duplication: an update. *Trends Ecol Evol* 18:292–298