Reverse Evolution in RH1 for Adaptation of Cichlids to Water Depth in Lake Tanganyika

Haruka Nagai,†1 Yohey Terai,†1 Tohru Sugawara,1 Hiroo Imai,2 Hidenori Nishihara,1 Michio Hori,3 and Norihiro Okada*1

1Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Midori-ku, Yokohama, Japan
2Department of Cellular and Molecular Biology, Primate Research Institute, Kyoto University, Inuyama, Japan
3Graduate School of Science, Kyoto University, Kyoto, Japan
†These two authors contributed equally to this work.
*Corresponding author: E-mail: nokada@bio.titech.ac.jp.

Abstract

Reverse evolution is a widespread phenomenon in biology, but the genetic mechanism for the reversal of a genetic change for adaptation to the ancestral state is not known. Here, we report the first case of complete reverse evolution of two amino acids, serine and alanine, at a single position in RH1 opsin pigment for adaptation to water depth. We determined RH1 sequences of cichlid fishes from four tribes of Lake Tanganyika with different habitat depths. Most of the species were divided into two types: RH1 with 292A for species in shallow water or 292S for species in deep water. Both types were adapted to their ambient light environments as indicated by the absorption spectra of the RH1 pigments. Based on the RH1 locus tree and ecological data, we inferred the ancestral amino acids at position 292 and the distribution of the depth ranges (shallow or deep) of ancestral species of each tribe. According to these estimates, we identified two distinct parallel adaptive evolutions: The replacement A292S occurred at least four times for adaptation from shallow to deep water, and the opposite replacement S292A occurred three times for adaptation from deep to shallow water. The latter parallelism represents the complete reverse evolution from the derived to the ancestral state, following back adaptive mutation with reversal of the RH1 pigment function accompanied by reversal of the species habitat shift.

Key words: molecular parallelism, adaptation, opsin, reverse evolution.

Introduction

Can an organism retrace a previously traversed evolutionary pathway? Reverse evolution is one of the fundamental questions in evolutionary biology. The reacquisition by derived populations of the same character state as that of ancestral populations is a widespread phenomenon in biology. The genetic mechanisms, however, are still unclear (Porter and Crandall 2003). At the genetic level, only a case of the stickleback armor plate occurring by a reverse shift of the allele frequency has been reported (Kitano et al. 2008). Because this example does not accompany a new mutation and its fixation, the reversal of a fixed genetic change for adaptation to the ancestral state is not known.

African cichlid fishes are becoming a model system for the genetics of vertebrate speciation and adaptive evolution (Kocher 2004). Lake Tanganyika, the oldest of the three African great lakes (Cohen and Soreghan 1993), has about 250 cichlid species (Turner et al. 2001) that have been divided into 12 (Poll 1986) or 16 (Takahashi 2003) tribes. Evolutionary biologists have been fascinated by these fishes because most of them are endemic to this lake (Nishida 1991; Salzburger et al. 2005) and are diversified by ecology and morphology (Fryer and Iles 1972; Poll 1986; Coulter 1991; Kawanabe et al. 1997). In particular, the visual systems of cichlids have been studied extensively (Carleton et al. 2000; Carleton and Kocher 2001; Sugawara et al. 2002, 2005; Terai et al. 2002, 2006; Parry et al. 2005; Spady et al. 2006; Maan et al. 2006; Seehausen et al. 2008) because vision is important for food acquisition (Hori et al. 1993; Kohda and Hori 1993; Kawanabe et al. 1997) and mate choice (Fryer and Iles 1972; Seehausen et al. 1997, 2008; Seehausen and van Alphen 1998). Their visual systems are thought to have adapted to different ambient light conditions that vary with depth, water color, and turbidity (Seehausen et al. 1997, 2008; Sugawara et al. 2005; Maan et al. 2006; Terai et al. 2006).

The visual pigments are located in photoreceptor cells of the retina (Yokoyama 2000) and their absorption spectra determine the range of wavelengths that contribute to vision (Yokoyama 2000; Trezise and Collin 2005). Visual pigments consist of a light-absorbing chromophore (retinal) and a protein (opsin) (Shichida 1999). The absorption spectrum of a visual pigment can be altered by amino acid replacements within opsins (Shichida and Imai 1998; Kochendoerfer et al. 1999). Cichlid species have eight opsins: a single rod opsin and seven cone opsins (Carleton et al. 2000; Carleton and Kocher 2001; Sugawara et al. 2002; Parry et al. 2005). Rod opsin (rhodopsin1 or RH1) makes it possible to form black and white images in dim light (Yokoyama R and Yokoyama S 1990; Yokoyama S and Yokoyama R 1996). Recently, we showed that East African cichlid species have adapted to deepwater ambient light environments by adaptive and
parallel amino acid replacements in the opsin gene RH1
(Sugawara et al. 2002, 2005), namely an amino acid change
from alanine to serine at position 292 (A292S) that shifts
the peak wavelength absorbance \( \lambda_{\text{max}} \) value by 11–14
nm toward blue light. This shift from the longer (about
500 nm) to the shorter \( \lambda_{\text{max}} \) (470–490 nm) is an adaptation
from shallowwater to deepwater light environments.

Here, we focused on four tribes in Lake Tanganyika
(Cyprichromini, Perissodini, Lamprologini, and Ectodini;
Poll 1986; Takahashi 2003) and determined their
RH1 sequences. These tribes have distinct water-depth distributions. We demonstrate two different kinds of parallelism by
amino acid substitutions at a single amino acid position in RH1 for adaptation to water depth. One is the conven-
tional parallelism from an ancestral to a derived state, and the other represents the complete reverse evolution from the derived state back to the ancestral state.

Materials and Methods

Collection of Specimens and Ecological Data

A total of 67 species from four tribes (7, 9, 32, and 19 species from Cyprichromini, Perissodini, Lamprologini, and Ectodini, respectively) and three species from genus Bentrochro-

tomis were used in this study. The specimens were caught during several expeditions to Lake Tanganyika from 1997 to 2007 at the locations shown in supplementary table S1, Supplementary Material online, and figure 1. Ecological data were also collected during the expeditions. Especially, data for depth of habitat were collected with the aid of scuba for shallow species and with a fish finder (HONDEX

DNA Extraction and Sequencing

Genomic DNA extraction and the determination of RH1 sequences were performed as described (Sugawara et al. 2002). DNA sequences were aligned using Genetyx (ver. 5.0) and manual alignment.

Determination of Sequences Upstream and Downstream of RH1

Screening of a bacterial artificial chromosome clone, sub-

cloning, and determination of the sequences of the DNA fragment including RH1 and its flanking regions were per-
formed as described (Terai et al. 2006). Based on the se-
quence, we designed a pair of primers (RH1 long F1 and
RH1 long R1, supplementary fig. S2 and S3, Supplementary
Material online) to amplify RH1 (1 kb), 3 kb of upstream
data, and 1 kb of downstream DNA (5 kb total) from Lake
Tanganyika cichlids. The 5-kb DNA fragment containing RH1 was amplified by long polymerase chain reaction
(PCR) as described (Terai et al. 2006) except that the ex-
tension time was decreased to 7 min. The amplified prod-
uct was then used as a template to amplify the regions flanking RH1, which were determined using the primers shown in supplementary figures S2 and S3, Supplementary
Material online.

Construction of RH1 Locus Tree and Reconstruction of Ancestral RH1 Sequences

Nucleotides that were heterozygous or undetermined and gaps in either species were excluded from the analysis. A total of 2,321 bp upstream and 491 bp downstream were concatenated and subjected to the following three phylo-
genetic analysis with 1,000 bootstrap replications. Neighbor
joining (NJ) analysis (Saitou and Nei 1987) was performed with MEGA 4 software (Tamura et al. 2007) with the K2P
model. Maximum parsimony (MP) analysis was performed
with PAUP* using the branch-and-bound search (Swofford
2002). Maximum likelihood (ML) analysis was performed with RAxML (Stamatakis et al. 2005) under general time
reversible + gamma model. Boulengerichromis microlepis
was used as an outgroup species. The phylogenetic trees constructed by three methods were almost the same and the tree was used for the construction of ancestral se-
quences of RH1 with PAML 4 (Yang 2007).

Construction of ND2 Tree

PCR and determination of NADH dehydrogenase subunit 2
gene (ND2) were performed as described (Brandstätter et al. 2005). A 1,005 bp ND2 was subjected to the phylo-
genetic analysis with 1,000 bootstrap replications. NJ analysis
(Saitou and Nei 1987) was performed with MEGA 4
software (Tamura et al. 2007) with the K2P model. Para-
cyprichromis brieni and Pa. nigripinnis were used as the out-
group species.

Measurement of the Absorption Spectra of RH1 Pigments

Expression, reconstitution, and purification of RH1 pig-
ments were performed as described (Ueyama et al. 2002; Sugawara et al. 2005). Absorption spectra of the RH1 pigment solution before and after photobleaching
(>480 nm for 5–20 min) were recorded using a spectropho-
tometer (UV-2400, Shimadzu, Kyoto, Japan). The measure-
ments were performed in the presence of hydroxylamine

Fig. 1. Map of the southern part of Lake Tanganyika showing the
sampling localities. Numbers indicate localities at which cichlids
were collected: 1, Chipwa; 2, Isanga; 3, Chisanza; 4, Kasenga; 5,
Kasenga off; 6, Wonzye; 7, Kumbula Island; 8, Kasakalawe; 9, Katoto;
10, Nachissa; 11, Cape Kaku; 12, Kasaba Bay; 13, Cape Nangu; 14,
Sumbu N.P.; 15, Ndole Bay; and 16, Chibanga. Location 17, Kigoma,
is located at the northern part of Lake Tanganyika.
The measurements were performed three to five times before and after photobleaching. We determined the mean peak spectral values and standard errors from multiple preparations and measurements for each RH1 pigment.

Accession Number of Sequences
The nucleotide sequences have been deposited in the GenBank database (accession numbers: AB457848–AB458142 and AB588055–AB588119).

Results and Discussion
Molecular Adaptation to Light Environments
We determined the RH1 sequences that correspond to amino acid positions 13–320 of the full-length RH1 (354 residues). Sequences were analyzed from 293 individuals (supplementary table S1, Supplementary Material online) in four tribes and genus Benthochromis from Lake Tanganyika (fig. 1). In each tribe, the variable positions of amino acid vary from 5 to 24 (supplementary fig. S1A–E, Supplementary Material online). At residue 292, both serine (292S) and alanine (292A) were observed in the species from each of the four tribes. In the genus Benthochromis, only 292S was observed in three species.

To identify the functional difference between RH1s containing 292A and those containing 292S, we selected a few representative sequences and measured the absorption spectra of their corresponding proteins in complex with retinal (RH1 pigments). As shown in fig. 2A–I, we used $\lambda_{\text{max}}$ values for RH1 pigments from 14 representative species of tribes. We measured seven of those values and refer to seven of those values that were reported previously (Sugawara et al. 2005). RH1s containing 292S had $\lambda_{\text{max}}$ values from 488 to 494 nm (fig. 2D–G and I), whereas RH1s containing 292A had $\lambda_{\text{max}}$ values from 497 to 504 nm (fig. 2A–C and I) with no overlap between the two groups (fig. 2H). Species living in shallow water tended to have RH1 sequences with 292A (fig. 3A–D, light gray boxes), whereas deepwater species tended to have 292S (fig. 3A–D, dark gray boxes). These absorption spectra showed that the $\lambda_{\text{max}}$ values of the pigments reconstituted from the sequences with 292S (average 491 nm, fig. 2I) were shifted 9 nm toward blue light compared with the pigments with 292A (average 500 nm, fig. 2I) for adaptation to blue-greener light (470–490 nm) that predominates in deeper water (Sugawara et al. 2005).

The Period and Direction of Amino Acid Replacements at Position 292
To determine the period and identity of the two variant residues at position 292 during cichlid evolution, we constructed a high-resolution RH1 locus tree using concatenated sequences upstream (2,321 bp) and downstream (491 bp) of RH1 (supplementary fig. S2, Supplementary Material online) using NJ, MP, and ML methods (fig. 4). We used both 292A and 292S representatives from each tribe. Using all three methods, monophyletic clades of the Cyprichromini, Perissodini, and Lamprologini tribes were supported with high bootstrap values (fig. 4). Although the species in Ectodini formed a monophyletic clade with a relatively low bootstrap value (fig. 4), several phylogenetic
Nagai et al.  


Such multiple replacements might be explained by the persistence and independent fixation of the ancestral polymorphisms or introgression by hybridization over the tribes. If the 292A and 292S alleles were generated just once and have been sorted in different lineages accompanied by the up- and down-stream flanking regions, the flanking region tree should show the monophyly of each of 292A and 292S alleles, respectively. But the flanking region tree did not show such monophyletic clades (fig. 4), suggesting that the persistence of polymorphic 292A and 292S alleles could not explain multiple replacements without recombination. If the 292A and 292S alleles have persisted by ancestral polymorphism and have exchanged the RH1 gene by recombination, the multiple replacements at 292 might be explained. To test this possibility, we analyzed the synonymous substitutions in RH1 gene (supplementary fig. S4, Supplementary Material online). Most of the synonymous substitutions showed distribution specific to tribes or lineages, indicating that RH1 gene has evolved along with species divergences. At six sites (supplementary fig. S4, Supplementary Material online), the synonymous substitutions were shared by 292A and 292S alleles, suggesting that the A292S and S292A replacements occurred independently after the synonymous substitutions at these six sites. At two synonymous sites, substitutions to two different nucleotides were observed (supplementary fig. S4, Supplementary Material online, sites 75 and 639), suggesting multiple substitution at the same site is not a rare evolutionary event. At position 292, the codons for alanine and serine were GCC and UCC, respectively, in all 292A and 292S alleles. The reason why the amino acids at 292 have been coded by only two codons may be that GCC and UCC
can be exchanged by a single nucleotide substitution, whereas the other codons need two or more substitutions. Because the synonymous substitution rate was too low (overall mean ds among four tribes: 0.00645), the synonymous substitution may have not occurred at this third position of these codons. These results supported the conclusion that the inferred parallel replacements occurred independently in different lineages.

Only in the case that the polymorphism/variations of 292A and 292S alleles might has persisted for a long period by balancing selection and that the recombination rate might be high enough so that even the short region including amino acid position 292 has not evolved under balancing selection in most lineages. If the balancing selection might have persisted 292A/292S polymorphism in Cy. coloratus, the high heterozygosity of 292A/292S might be expected in all populations of this species. Therefore, we considered the $\lambda_{\text{max}}$ values of the RH1 pigments are the same as in figure 2f. The ancestral branches that are inferred as shallow water and deep water are highlighted in yellow and blue, respectively.
the inferred parallel replacements were independent rather than a result of fixation of the ancestral polymorphisms or introgression by hybridization over the tribes.

Why are the 292A/292S polymorphisms in the two populations of *Cy. coloratus* (supplementary table S1, Supplementary Material online and fig. 1: Isanga and Chisanza) observed? The flanking sequences of 292A and 292S alleles from *Cy. coloratus* formed a monophyletic clade (fig. 4), suggesting that those alleles have arisen and evolved in this species. To elucidate the population structure in *Cy. coloratus*, we determined mitochondrial ND2 sequences and constructed a phylogenetic tree. The ND2 tree clearly showed the monophyletic clades of each of western populations and eastern populations, respectively (supplementary fig S5, Supplementary Material online). In the western populations, the RH1 was fixed in 292A allele and in the eastern populations, the RH1 was fixed in 292S except two polymorphic populations (supplementary table S1, Supplementary Material online). Thirty thousand years ago, the water level of Lake Tanganyika was 300–600 m lower than the present, and there may have been the long sandy shore (supplementary fig S6, Supplementary Material online, yellow thick lines) between western and eastern rocky shores, being a geographical barrier between the rock-dwelling cichlid species in both of these rocky shores (Kohda et al. 1996, supplementary fig S6A and B, Supplementary Material online). According to this geographic isolation, the western and eastern populations of *Cy. coloratus* may have evolved independently with the different RH1 alleles (supplementary fig S6A and B, Supplementary Material online). As the water level arose, the western and eastern populations may have extended their distribution to the southern end of Lake Tanganyika (supplementary fig S6C, Supplementary Material online), and two populations (Isanga and Chisanza) might have acquired 292A allele by recent introgressive hybridization from western populations (supplementary fig S6D, Supplementary Material online). Hence, the RH1 alleles have evolved not only during species divergence but also during differentiation of populations within species. The 292A and 292S alleles were also observed in *Pa. brieni* and *Pe. paradoxus*, and one of those alleles was fixed within each population (supplementary table S1, Supplementary Material online). These allele differentiations may also be due to different adaptive evolution of RH1 in each population.

Reverse Evolution for Adaptation to Water Depth

To elucidate the adaptive role of the two variant residues at position 292, we inferred the water-depth distribution of the species in each tribe. As shown in fig. 3A and B, species in Cyprichromini and Perissodini distribute at a wide range of shallow water (5–40 m) or deep water (>40 m). Species in the basal lineages of Cyprichromini, Perissodini, and Perissodini/Benthochromis clade are deepwater fish (fig. 4, *Cy. pavo* in Cyprichromini, *Pe. eccentricus* in Perissodini, and *Benthochromis* species; supplementary table S1, Supplementary Material online), and the ancestral amino acid at position 292 of RH1 was inferred to be S (deepwater adaptive, fig. 4). Based on this information, the ancestors of these two tribes and the Perissodini/Benthochromis clade were inferred to have deepwater distribution. Molecular analysis using amplified fragment length polymorphism also supports the deepwater origins of these tribes (Takahashi et al. 2007). On the other hand, most of the species in Lamprologini and Ectodini are shallow-water species (fig. 4), and the ancestral amino acid at position 292 was inferred to be A (shallowwater adaptive, fig. 4). Thus, the ancestors of these two tribes were inferred to have shallowwater distribution. The deep and shallowwater ancestor tribes were assumed to have originated in and adapted to deep and shallow water, respectively.

According to the ancestral types and the periods of the replacements at RH1 position 292, we observed two different parallel adaptations. In the tribes with shallowwater ancestors (Lamprologini and Ectodini), 292S was acquired twice, and the species with 292S were only observed in deep water (fig. 3C and D). The periods of the habitat shifts from ancestral shallow to derived deep water were inferred at the ancestral lineages of 292S species in these tribes (fig. 4), suggesting adaptation to the ambient light environment in deep water. The replacement A292S was also observed in the ancestral lineage(s) of deepwater lineages, Cyprichromini, and Perissodini/Benthochromis (fig. 4). The phylogenetic relationships of these lineages are ambiguous, but it was clearly concluded that A292S occurred at least once (if these lineages form a monophyletic group) or twice (marked by “+” in fig. 4) with the distribution shifts from ancestral shallow to derived deep water (fig. 4). In total, A292S occurred at least four times during the evolution of these tribes. The *Cy. coloratus* from eastern populations distribute in relatively shallow water with 292S (fig. 4 and supplementary fig S5, Supplementary Material online). The ancestor of eastern populations might have undergone the deepwater environment with the acquisition of deepwater adaptive 292S allele (fig. 4).

Surprisingly, the opposite parallelism was also observed. The 292S reverted to three times, and the species with 292A in tribes with deepwater ancestors (Cyprichromini and Perissodini) live in shallow water (fig. 3A and B). The periods of the habitat shifts from ancestral deep to derived shallow water were also coincident with the periods during which the replacement S292A was acquired (fig. 4), suggesting adaptation to ambient light in shallow water. In all four tribes, not all the species that have shifted their water-depth distribution have acquired the adaptive amino acid replacement (fig. 4).

In the ancestral lineages of Cyprichromini and Perissodini/Benthochromis, A292S occurred for shallowwater to deepwater adaptation followed by S292A in these tribes for deepwater to shallowwater adaptation (fig. 4), indicating that reversal of adaptation to the ancestral state was achieved by back amino acid replacement. The absorption spectra of RH1 pigments showed the reversal of the function of RH1 pigments to the ancestral state (shallowwater
adaptive: fig. 2A and B). These results demonstrate reverse evolution by showing reversal of the habitat shift followed by back amino acid replacement with reversal of the function of the RH1 pigment during adaptive radiation in Lake Tanganyika. The driving force of the reverse and repeated evolution might as well be natural selection because molecular convergence and parallelism are thought to be strong evidence of natural selection. The identity of the amino acid residue at position 292 may be less constrained relative to those at other positions during adaptations for ambient light environments, so this may be why the adaptive replacement A292S in RH1 has been observed in many distantly related deepwater fishes (Hunt et al. 1996; Yokoyama et al. 1999; Yokoyama 2000; Bowmaker and Hunt 2006).

In this report, we suggest the importance of a limited number of amino acid positions for adaptive evolution. Thus, if natural selection is strong enough, the evolution from the derived to the ancestral state may occur. Cases of reverse evolution at the morphological level have been reported. Many cases of evolution of sexually selected characteristics have evolved from the derived dimorphic to the ancestral monomorphic state, and the loss and recovery of wings in stick insects are well-known examples (Wiens 2001; Whiting et al. 2003). Studies about the molecular mechanisms of such morphological reverse evolution will show that organisms can retrace a previously traversed evolutionary pathway at the molecular level.

Supplementary Material

Supplementary table S1 and figures S1–S6 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

Grants from the Ministry of Education, Culture, Sports, and Science and Technology of Japan (to N.O.) supported this work. The expedition in Lake Tanganyika was conducted with permission from the Zambian Ministry of Agriculture, Food, and Fisheries and complied with the current laws of Zambia. This study was partly supported by grants-in-aid for Scientific Research in Priority Areas (14087203) and from the Global Center of Excellence Program, Ministry of Education, Culture, Sports, and Science and Technology (MEXT), Japan.

References


