Immunogenetics (2005) 56: 894–908
DOI 10.1007/s00251-005-0767-5

ORIGINAL PAPER

Corine P. Kruiswijk · Trudi Hermsen · Joost van Heerwaarden · Brian Dixon · Huub F. J. Savelkoul · René J. M. Stet

Major histocompatibility genes in the Lake Tana African large barb species flock: evidence for complete partitioning of class II *B*, but not class I, genes among different species

Received: 22 October 2004 / Revised: 7 December 2004 / Published online: 8 February 2005

Abstract The 16 African 'large' barb fish species of Lake Tana inhabit different ecological niches, exploit different food webs and have different temporal and spatial spawning patterns within the lake. This unique fish species flock is thought to be the result of adaptive radiation within the past 5 million years. Previous analyses of major histocompatibility class II *B* exon 2 sequences in four Lake Tana African large barb species revealed that these sequences are indeed under selection. No sharing of class II *B* alleles was observed among the four Lake Tana African large barb species. In this study we analysed the class II *B* exon 2 sequences of seven additional Lake Tana African large barb species and African large barbs from the Blue Nile and its tributaries. In addition, the presence and variability of major histocompatibility complex class I *UA* exon 3 sequences in six Lake Tana and Blue Nile African large barb species was analysed. Phylogenetic lineages are maintained by purifying or neutral selection on non-peptide binding regions. Class II *B* intron 1 and exon 2 sequences were not shared among the different Lake Tana African large barb species or with the riverine barb species. In contrast, identical class I *UA* exon 3 sequences were found both in the lacustrine and riverine barb species. Our analyses demonstrate complete partitioning of class II *B* alleles among Lake Tana African large barb species. In contrast, class I alleles remain for the large part shared among species. These different modes of evolution probably reflect the unlinked nature of major histocompatibility genes in teleost fishes.

Keywords Fish · Species flock · Evolution · MHC · Class I · Class II

Electronic Supplementary Material Supplementary material is available for this article at http://dx.doi.org/10.1007/s00251-005-0767-5.

C. P. Kruiswijk · T. Hermsen · J. van Heerwaarden · H. F. J. Savelkoul · R. J. M. Stet (✉)
Cell Biology and Immunology Group,
Wageningen Institute of Animal Sciences,
Wageningen University,
Marijkeweg 40,
6709 PG Wageningen, The Netherlands
e-mail: rene.stet@wur.nl
Tel.: +31-317-483922
Fax: +31-317-483955

B. Dixon
Department of Biology, University of Waterloo,
200 University Avenue West,
ONT, N2L 3G1, Canada

R. J. M. Stet
Scottish Fish Immunology Research Centre,
University of Aberdeen,
Zoology Building, Tillydrone Avenue,
Aberdeen, AB24 2TZ, Scotland, UK

Introduction

Lake Tana, situated in the north-western highlands of Ethiopia, harbours a unique aquatic ecosystem system formed during 5 million years of isolation. In late Pleistocene times, a volcanic eruption blocked the outlet to the Blue Nile by creating a 40-m-high waterfall (Mohr 1962). This event effectively prevented gene flow among fish species from the lake and the Blue Nile and all other freshwater systems connected to this river. During this period of isolation, the ancestral *Barbus intermedius* fish population underwent speciation due to adaptive radiation into different ecological niches within the lake, resulting in 15 novel African 'large' barb species (Nagelkerke and Sibbing 2000). In contrast, only a single extant species, *B. intermedius*, is found in the Blue Nile and its tributaries. Next to the 15 species, the lake harbours a population of African large barbs occupying the shores of Lake Tana, resembling the *B. intermedius* of the Blue Nile system. During speciation events, selection pressure is hypothesized to be strong on major histocompatibility complex (MHC) class I and class II molecules (Klein 1987), since individuals move to new environments and as a consequence encounter new pathogens.

Dixon et al. (1996) studied MHC class II polymorphism in four (*B. acutirostris*, *B. nedgia*, *B. tsanensis* and *B. truttiformis*) out of 15 species belonging to the Lake Tana

African large barb species flock. The four species under study did not possess identical sequences. This suggested that each species has a unique set of class II *B* sequences derived from a common ancestral suit within a time span of 5 million years due to an expansion into a new ecological environment.

In contrast to all other jawed vertebrates, teleost fish like the African large barbs have MHC class I and class II genes located on different linkage groups (Bingulac-Popovic et al. 1997; Hansen et al. 1999; Sato et al. 2000; Sultmann et al. 2000). This lack of linkage is likely to have influenced selection on these genes during ecological specialization relative to a situation where they are linked (Stet et al. 2003). To test this hypothesis we extended the study on MHC class II *B* encoding sequences with six Lake Tana African large barb species (*B. intermedius*, *B. brevicephalus*, *B. macrophtalmus*, *B. megastoma*, *B. platydorsus* and *B. surkis*) and African large barbs (*B. intermedius*) from the Blue Nile and its tributaries, comprising in total 18 individuals. In addition, to validate whether MHC class I genes evolved in a similar species specific manner, class I sequences from six different Lake Tana African large barb species and from Blue Nile African large barbs, comprising in total 35 individuals, were analysed.

Materials and methods

Sampling of fish and DNA extraction

Muscle or fin clip samples from Lake Tana *B. intermedius* species flock individuals and *B. intermedius* individuals of the Blue Nile system were collected from various locations in the lake and the Blue Nile system. Genomic DNA was isolated using a Wizard Genomic DNA Purification Kit (Promega, Madison, Wis., USA) according to the protocol provided. DNA concentration was determined using the GeneQuant system (Amersham Pharmacia Biotech, Roosendaal, The Netherlands).

PCR conditions

Standard PCR reaction conditions were 1× reaction buffer, 1.5 mM $MgCl_2$, 1 U *Taq* polymerase (Goldstar, Eurogentec, Seraing, Belgium), 0.2 mM dNTPs, 0.2 μM each primer and 100 ng genomic DNA. The cycling profile was one cycle at 94°C for 5 min, followed by 30 cycles consisting of denaturing at 94°C for 30 s, annealing at 55°C for 30 s, polymerization at 72°C for 1 min and a final cycle of 10 min at 72°C.

Amplification of genomic MHC class I exon 3 and class II intron 1 and exon 2 sequences

Genomic class II intron 1 and exon 2 sequences were amplified with the same primers used by Dixon et al. (1996). Alpha-2 domains of class I alpha chains were amplified using primers matching the nucleotide sequence of the start (5′-GGTGTTCACTCAGTCCAG 3′) and the end (5′-CTTTTCTCTCCAGAGAGTCCTT 3′) of known cyprinid class I alpha-2 encoding exons. Amplification of exon 2 sequences was hampered by the inability to design a conserved primer matching the start of exon 2. In addition, introns separating exon 1 and exon 2 and exon 2 and exon 3 have shown to be large in cyprinids (Erp van et al. 1996; Michalova et al. 2000; Kruiswijk et al. 2004). Therefore, only exon 3 sequences were amplified and analysed.

Cloning and DNA sequencing

PCR products were ligated and cloned using the pGEM T-easy kit (Promega), following the manufacturer's description. Plasmid DNA was isolated from bacterial cells, using the QIAprep Spin miniprep kit (Qiagen, Valencia, Calif., USA) according to the protocol provided. Subsequently, plasmid DNA was sequenced using the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kit (PE Applied Biosystems, Foster City, Calif., USA) and analysed using an ABI 377 sequencer (PE Applied Biosystems).

Accession numbers and nomenclature

The new sequences reported here were deposited in the EMBL database under the followings accession numbers: AJ506649–AJ506736. The nomenclature used to assign new sequences or rename existing sequences adheres, in part, to the recommendations described in the HLA Facts Book (Marsh et al. 2000). The first four letters refer to the species names and are followed by a dash for the locus designation (*UA* for class I sequences and *DAB* for class II *B*), an asterisk and four digits. According to the HLA Facts Book (Marsh et al. 2000), the first two digits following the asterisk describe the lineage, and the third and fourth digits that follow assign alleles. However, it is not yet clear whether the class I or class II *B* sequences isolated belong to a single locus or multiple loci. Therefore, the nomenclature given assumes that all class I and all class II *B* sequences belong to a single locus *UA* for class I and *DAB* for class II *B* and represent a single lineage, *01. The following two digits, in general describing the alleles, indicate in this study the order in which a novel sequence was isolated from a species. Sequences that were shared among different African large barb species start with the abbreviation *Baic*, referring to the *B. intermedius* species complex. Fourteen out of 22 previously described class I exon 3 sequences characterized in an earlier study (Stet et al. 1998) were renamed and included in this study. The 14 sequences were isolated from two *B. nedgia* (Lip) individuals, a *B. acutirostris* (Acute, Bob) and a *B. intermedius* (Shore-complex, Flip) individual. The sequences were previously designated *L01* to *L04*, *L06* to *L09, A01* to *A05* and *F02* and renamed to *Bane-UA*0101*, *Baic-UA*0101*, *-*0112*, *-*0111* and *-*0114*, *Bane-UA*0107*, *-*0108* and *-*0109*,

*Baac-UA*0101*, *-*0102*, *-0103*, *-*0104* and *-*0104*, *Bain-UA*0104* (acc. nos. AJ007885 to 88, -90 to -93, AJ007879 to -83, and -99), respectively. All analyses performed only included class I *UA* exon 3 sequences found at least two times in two independent PCRs. Two sequences, designated *L12* and *F01*, that seemed pseudo genes in an earlier study (Stet et al. 1998), were shown to be sequence artifacts. Thorough analyses revealed that these sequences were similar. Subsequently, this sequence was designated *Baic-UA*0102*.

Nucleotide sequence and amino acid sequence analyses

The class I and class II nucleotide sequences used for the analyses were represented by at least two identical clones from two independent PCR amplifications. Sequence data obtained were analysed with Sequencer, version 4.1, software (Gene Code, Ann Arbor, Mich., USA). Multiple alignments were performed using the program Clustal W, version 1.8 (Thompson et al. 1997). Phylogenetic analyses were constructed in Mega, version 2.1, software (Kumar et al. 2001), using the neighbour-joining algorithm (Saitou and Nei 1987), with the Jukes–Kantor method for nucleotide sequences. Synonymous (d_S), non-synonymous (d_N) substitution, nucleotide substitution rates (d) and d_N–d_S were calculated with the Mega software package, using the Jukes–Kantor method. Standard errors were estimated by 1,000 bootstrap replications. We realize that bootstrap tests in some cases may lead to erroneous values when the numbers of d_S and d_N substitutions are small. Nei and Jin (1989) provided a better method for computing variances and co-variances. However, this method is difficult to perform with large data sets.

Results

Analyses of *Barbus* class II *DAB*01* genes

Dixon et al. (1996) amplified class II *B* intron 1 and exon 2 encoding sequences by PCR on genomic DNA of 16 Lake Tana African large barb species flock individuals, using a primer set designed to match the leader sequence and the end of exon 2 of the *Cyca-DAB2*01* gene. The 16 individuals represented four different African large barb species: *B. acutirostris*, *B. nedgia*, *B. tsanensis* and *B. truttiformis*.

We extended the analyses of class II intron 1 and exon 2 encoding sequences in the Lake Tana African large barb species flock with 18 individuals, using a similar approach. Genomic DNA was obtained from one *B. brevicephalus*, one *B. macrophtalmus*, five *B. megastoma*, four *B. platydorsus*, one *B. surkis* and six *B. intermedius* individuals. The latter represented five individuals from the Blue Nile system and one shore complex individual from Lake Tana (see Electronic Supplementary Material, S-1).

All PCRs yielded products of approximately 480 bp for each African large barb individual, the expected product size of African large barb *DAB*01* equivalents (Dixon et al. 1996). Sequence analyses of the 480-bp fragments revealed 29 novel *DAB*01* sequences, 23 identified in individuals of the Lake Tana African large barb species flock and six identified in African large barb individuals from the Blue Nile system. Together with the 40 *DAB*01* sequences identified by Dixon et al. (1996), a total of 70 different *DAB*01* sequences were available for further analyses (Fig. 1). Alignment of the deduced amino acid sequence of all *DAB*01* partial exon 2 sequences confirmed that these sequences were equivalent to the previously identified African large barb *DAB*01* sequences. Based on motifs at variable positions, the 70 African large barb *DAB*01* sequences could be grouped.

Remarkably, no sharing of these *DAB*01* sequences was observed among individuals from the different African large barb species. However, *DAB*01* sequences were shared by individuals of the same species (see Electronic Supplementary Material, S-1). Also, the *B. intermedius* individuals from different locations in the Blue Nile system did not share any *DAB*01* sequences. The number of different *DAB*01* sequences observed in a single African large barb individual varied from one to six.

Although in some cases the DAB*01 amino acid sequences of some species are identical [Basu-DAB*0102 (Zu), Bame-DAB*0104 (Se) and Baac-DAB *0107 (Ac)], they differ in their exon 2 nucleotide sequences or intron sequences. Two pairs of identical partial exon 2 nucleotide sequences [*Bane-DAB*0106 (Li)-Bain-DAB*0101 (Sc)* and *Bame-DAB*0106 (Se)-Bats-DAB*0117 (In)*] were found. However, these sequences all differ in their intron sequences, suggesting that the exon 2 sequences most likely differ in the missing 51 nucleotides. The last 17 amino acid residues of the beta-1 domain contain at least three polymorphic residues (Kruiswijk et al. 2004).

Phylogenetic relationship of *Barbus DAB*01* intron 1 and exon 2 sequences

Analyses of the phylogenetic relationship of the 70 African large barb *DAB*01* partial exon 2 sequences revealed 13 clusters supported by high bootstrap values (above 70%) that comprised several different African large barb species. Eleven out of the 29 novel *DAB*01* sequences clustered within one of these ten clusters (Fig. 2). Three additional clusters (1–3), supported by high bootstrap values, were formed that comprised in total five novel sequences and three sequences identified by Dixon et al. (1996). The African large barb clusters contained up to a maximum of six different species per cluster, while the clusters VIII, IX and X contained sequences from a single species.

A phylogenetic tree constructed using 66 *DAB*01* intron 1 sequences of Blue Nile system large barb (*B. intermedius*) and Lake Tana African large barb species, and common carp *DAB*01* and *DAB*02* intron 1 sequences showed a specific clustering separating the two genera (*Cyprinus*

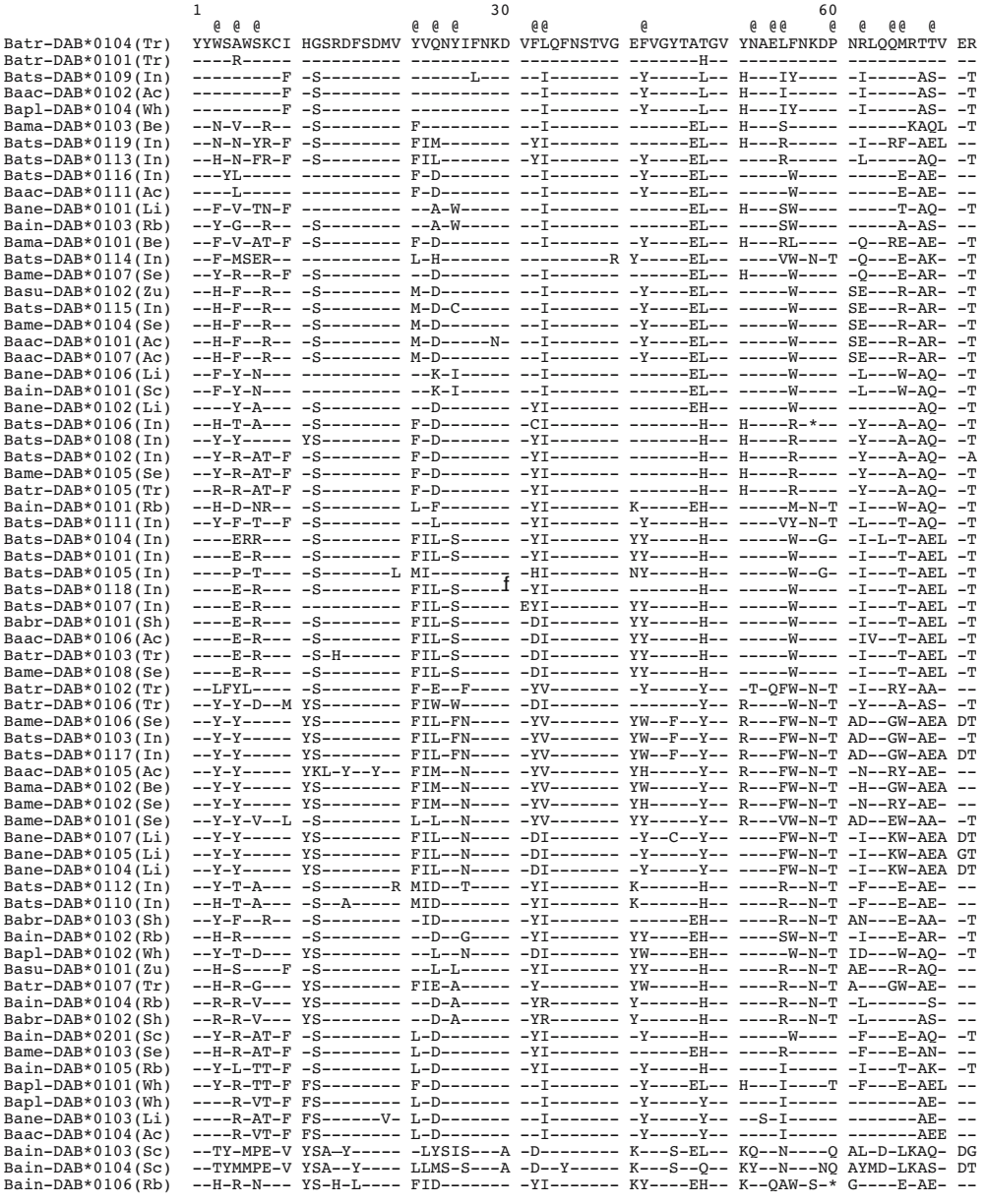

Fig. 1 The deduced African large barb class II *DAB*01* exon 2 amino acid sequences. The *numbers above* denote the amino acid position relative to the mature protein. *Dashes* indicate identity to the Batr-DAB*0107(Tr) sequence, and *asterisks* indicate gaps. Residues known to be involved in peptide binding in mammalian class II B molecules (Brown et al. 1993) are indicated by *ampersands* below numbering. African 'large' barb sequences starting with Bats-, Baac-, Bane- and Batr- are renamed sequences described by Dixon et al. (1996). Previously, the Lake Tana African large barb species were indicated by the morphotype nomenclature. To facilitate identification, morphotype abbreviations are indicated *between brackets*, following the sequence designation. *Ac* Acute, *Se* bigmouth small-eye, *Be* bigmouth big-eye, *In* intermedius, *Li* lip, *Sh* shorthead, *Sc* shore-complex, *Tr* troutlike, *Wh* white hunch, *Zu* Zurki

from *Barbus*) (Fig. 3). Eleven clusters, A–K, were observed when intron sequences were analysed. However, the clusters B, E and I, showed bootstraps value below 50. Each cluster contained up to a maximum of four different species. Four clusters (C, D, I and J) only contained intron sequences from a single African large barb species.

The clusters G and D together comprise intron sequences belonging to the exon 2 sequences of cluster I in the phylogenetic analysis of exon 2 sequences, with the exception of *Batr-DAB*0102*. The intron clusters A, B, C, E, F, H and I comprised similar sequences to the exon clusters VI, IV, X, II, V, III and VIII, respectively, with some exceptions (*Baac-DAB*0107*, *Bane-DAB*0107*, *Bats-DAB*0104*, *-06* and *-08*). No intron cluster was observed that correspond to exon cluster VII.

Nucleotide diversity in *Barbus* class II *DAB*01* intron 1 and exon 2

Clusters in phylogenetic analyses of MHC genes usually represent lineages of alleles. On the basis of the phylogeny of African large barb *DAB*01* exon 2 sequences, 12 such lineages could be identified (Fig. 2). The phylogenetic clusters were used for comparison of nucleotide diversity of intron 1 and exon 2 sequences within and among lineages. Nucleotide diversity of exon 2 sequences was calculated separately for peptide-binding regions (PBRs) and non-PBRs, as they were expected to differ (Table 1). The human MHC class II model (Brown et al. 1993) was used to assign PBR and non-PBR amino acid residues, since *DAB*01* variability corresponded well with *HLA-DRB* polymorphism (Dixon et al. 1996).

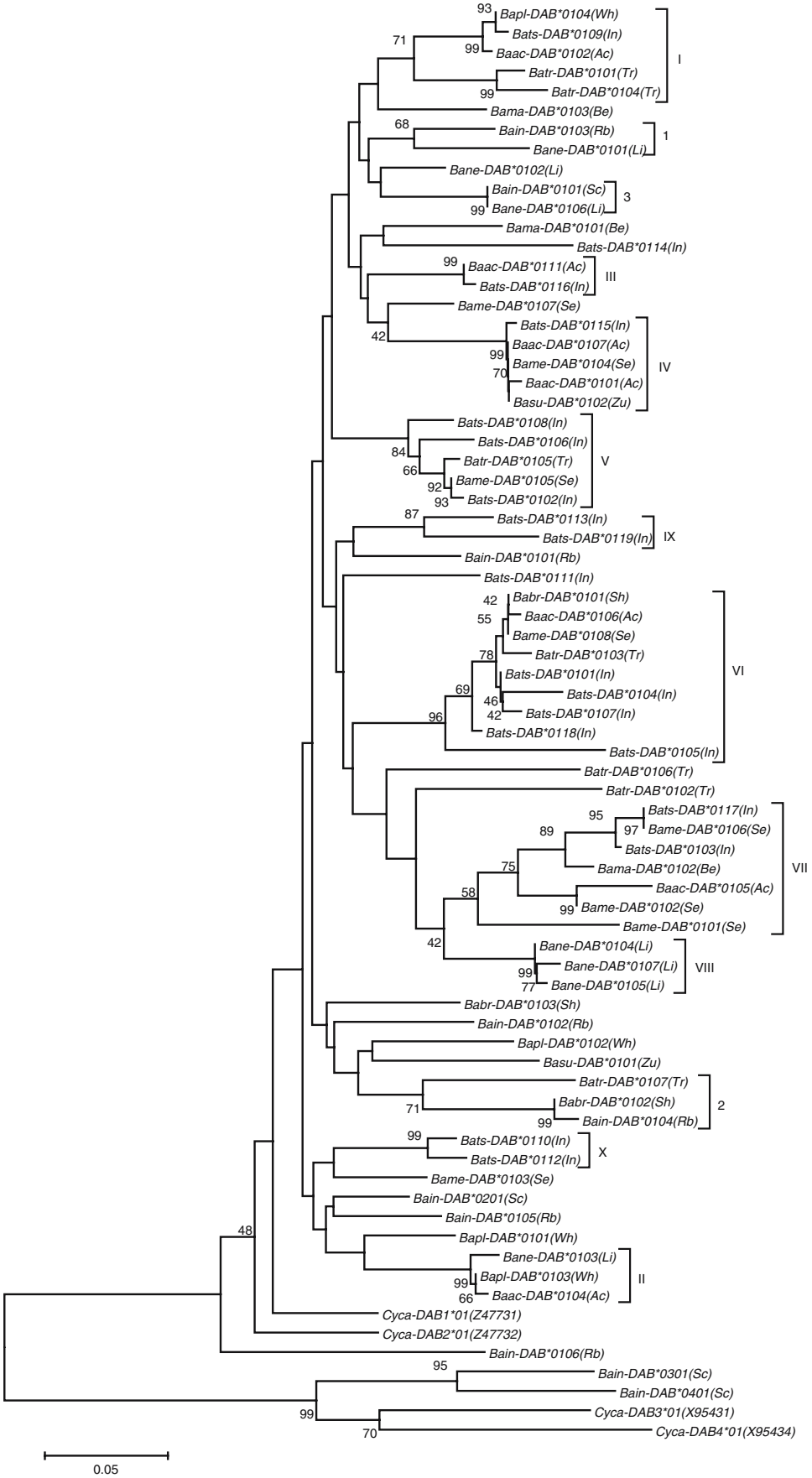

Fig. 2 Neighbour-joining tree of African large barb class II *DAB*01* exon 2 nucleotide sequences. Trees were constructed using the method of Saitou and Nei, based on Jukes–Kantor method in Mega software. *Numbers at branch notes* indicate bootstrap confidence levels of 1,000 bootstrap replications. Only bootstrap values over 40% are shown. African large barb sequence nomenclature is similar to that used in Fig. 1. The accession numbers of common carp (*Cyca*) and zebrafish (*Dare*) class II *B* sequences are shown *in brackets*, following the sequence assignments. *Brackets* indicate denote clusters (I–XI and 1–3) supported by bootstrap values over 70%. The clusters I–XI correspond to the clustering pattern reported by Dixon et al. (1996), while those numbered 1–3 represent novel clusters identified in this study

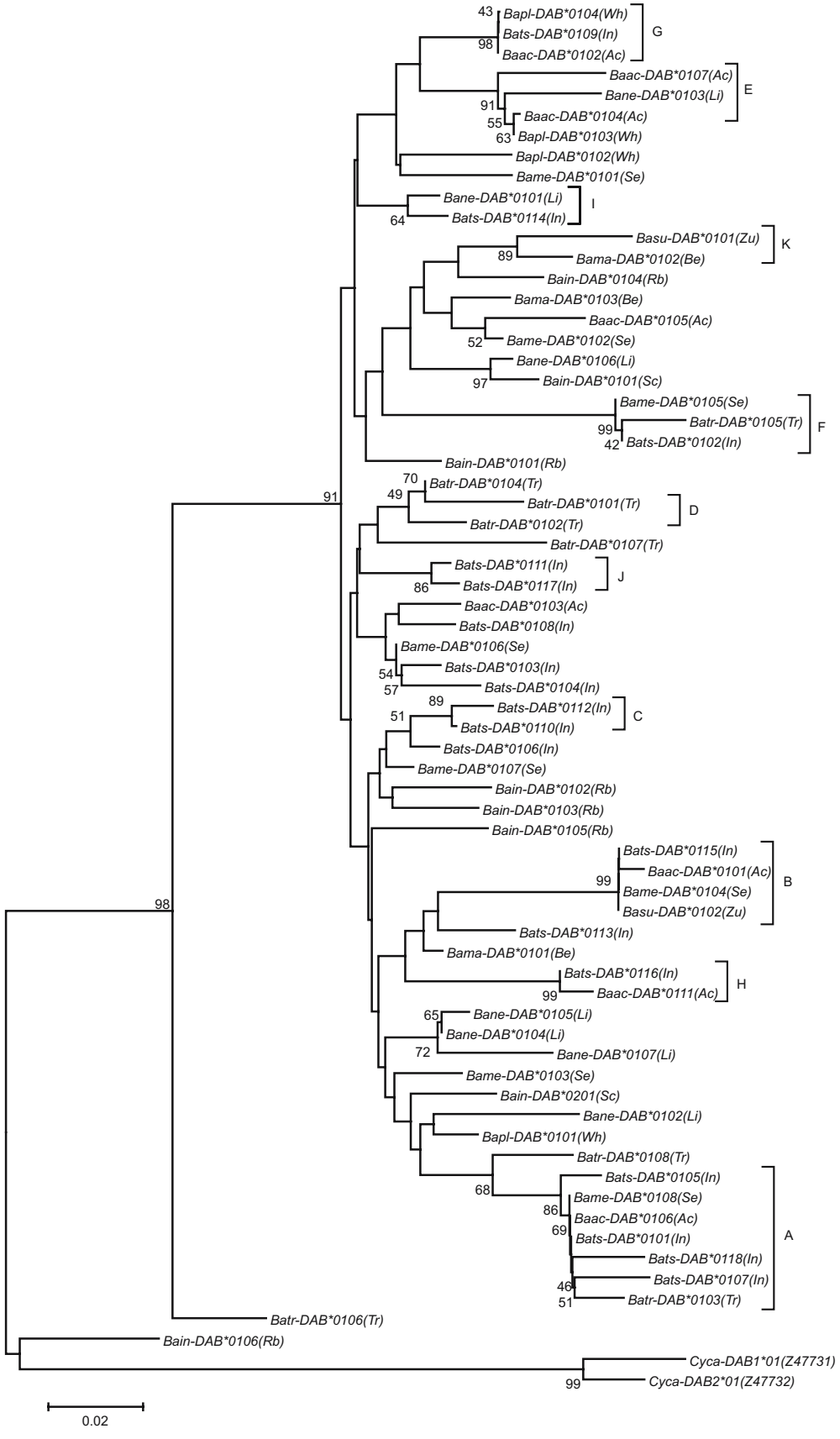

Fig. 3 Neighbour-joining tree of African large barb class II *DAB*01* intron 1 nucleotide sequences. Trees were constructed using the method of Saitou and Nei, based on Jukes–Kantor method in Mega software. *Numbers at branch notes* indicate bootstrap confidence levels of 1,000 bootstrap replications. African large barb sequence nomenclature is similar to that used in Fig. 2. The accession numbers of common carp (*Cyca*) and zebrafish (*Dare*) class II *B* sequences are shown in *brackets*, following the sequence name. *Brackets indicated with the letters A–I* correspond to the clustering pattern reported by Dixon et al. (1996). *Brackets indicated with the letters J, K and I* represent novel clusters observed in this study, which were supported by bootstrap values over 70%

Table 1 Mean±SE of nucleotide substitutions per site (d) in intron 1, synonymous (d_S) and non-synonymous (d_N) substitutions per site in exon 2 within and among groups of African 'large' barb class II *DAB*01* alleles as defined by phylogenetic relationship. *PBR* Peptide-binding region

	intron 1			exon 2					PBR					Non-PBR			
	d	SE	[a]*	d_N	SE	d_S	SE	[b]*	d_N	SE	d_S	SE	[b]*	d_N	SE	d_S	SE
Overall[c]	0.07	0.01	*	0.17	0.03	0.07	0.02	*	0.65	0.13	0.14	0.05	*	0.07	0.02	0.06	0.02
Within[d]	0.04	0.01		0.04	0.01	0.03	0.02		0.11	0.05	0.04	0.03		0.02	0.01	0.03	0.02
Among[e]	0.07	0.01		0.18	0.04	0.08	0.03	*	0.72	0.27	0.15	0.10	*	0.07	0.02	0.06	0.03

[a] d_S exon 2 PBR significantly higher than d intron 1. Tested hypothesis: $d_S > d$, $^*P<0.10$
[b] d_N significantly higher than d_S. Tested hypothesis: $d_N > d_S$, with $^*P<0.05$
[c] Mean±SE of pairwise comparison of 69 *Barbus* class II *DAB*01* sequences
[d] Mean±SE of pairwise comparison within groups of *Barbus* class II *DAB*01* sequences
[e] Mean±SE of pairwise comparison among groups of *Barbus* class II *DAB*01* sequences

Overall d_N and d_N substitution rates of pairwise comparisons among all *DAB*01* exon 2 sequences, and within and among phylogenetic lineages revealed a significantly higher mean d_N value of over two times the mean d_S (Table 1).

Pairwise comparison of PBR residues of phylogenetically defined *DAB* lineages revealed significantly higher mean d_N values of almost five times the mean d_S. Mean d_N and d_S values were essentially similar for overall-, within- and among-lineage pairwise comparison of non-PBR residues (Table 2; Electronic Supplementary Material, S-2).

Mean d_s values of overall-, within- and among-lineage pairwise comparisons, calculated for PBR, non-PBR or all exon 2 codons, were not significantly higher with respect to mean d values in intron 1. Pairwise analyses of the African large barb *DAB*01* intron 1 and exon 2 substitution rates within and among phylogenetic lineage was performed to evaluate whether recombination events might have homogenized intron 1 sequences relative to exon 2 sequences within lineages of alleles, but not among lineages. Plots of d in intron 1 versus d_s in exon 2 of pairwise comparisons within and among phylogenetic lineages, however, did not reveal any correlation (Fig. 4a, b). Although both plots showed some pairwise comparisons with relatively high d values in introns, d_s in exons were low and vice versa.

Calculations of mean values of pairwise comparisons per individual sequence revealed significantly higher d_s rates in exons of some sequences, namely *Bats-DAB*0111*, *Bane-DAB*0101*, *Bats-DAB*0101*, *Bama-DAB*0104*, *-05* and *-06* (Fig. 4c). However, plots of d in intron 1 versus d_S in exon 2 of all pairwise comparisons of these sequences with all others did reveal a high correlation with exon 2, evolving twice as fast as intron 1 (Fig. 4d).

Table 2 Mean±SE of ratios of d_N per d_S substitutions per site and $d_n - d_s$ values of PBRs and non-PBRs within and among groups of African large barb class II *DAB*01* alleles as defined by phylogenetic relationship. *Sig.* Significance

	Non-PBR					PBR				
	d_n/d_s	SE	$d_n - d_s$	SE	Sig.[a]	d_n/d_s	SE	$d_n - d_s$	SE	Sig.[a]
All	1.145	0.022	0.009	0.023	**	4.745	0.139	0.514	0.108	*
I	0.875	0.028	−0.003	0.022	**	1.101	0.124	0.013	0.118	**
II			0.015	0.009	**			0.017	0.017	**
III			0.008	0.008	**					
IV			0.003	0.003	**			0.010	0.010	**
V	1.091	0.014	0.015	0.008	**	2.000	0.080	0.053	0.044	**
VI	0.895	0.013	−0.001	0.012	**			0.019	0.015	**
VII	3.429	0.021	0.034	0.021	**	10.000	0.063	0.126	0.050	*
VIII			−0.008	0.019	**					
IX	0.124	0.066	−0.094	0.060	**			0.260	0.085	*
X	2.429	0.025	0.021	0.018	**			0.025	0.025	**
1	0.523	0.061	−0.039	0.060	**	1.212	0.157	0.033	0.112	**
2	0.267	0.042	−0.053	0.037	**	4.087	0.144	0.284	0.154	*
IV vs V	0.568	0.056	0.056	0.056	**	8.842	0.192	0.448	0.173	*
IV vs VI	1.743	0.035	0.032	0.032	**	4.841	0.409	0.630	0.341	*
IV vs VII	2.870	0.050	0.050	0.050	*	4.755	0.290	0.598	0.287	*
IV vs VIII	3.000	0.042	0.038	0.038	**	3.796	0.300	0.467	0.296	*
IV vs IX	0.449	0.046	0.044	0.044	**	7.472	0.382	0.809	0.334	*
IV vs X	2.343	0.039	0.036	0.036	**	26.154	0.276	0.654	0.294	*
IV vs 1	0.548	0.045	0.043	0.043	**	4.137	0.311	0.480	0.241	*
IV vs 2	0.610	0.060	0.058	0.058	**	4.761	0.521	1.039	0.495	*

[a] *Z*-test: neutral selection, H: $d_N = d_S$, $^{**}P<0.10$; positive selection, H: $d_N < d_S$, $^*P<0.05$

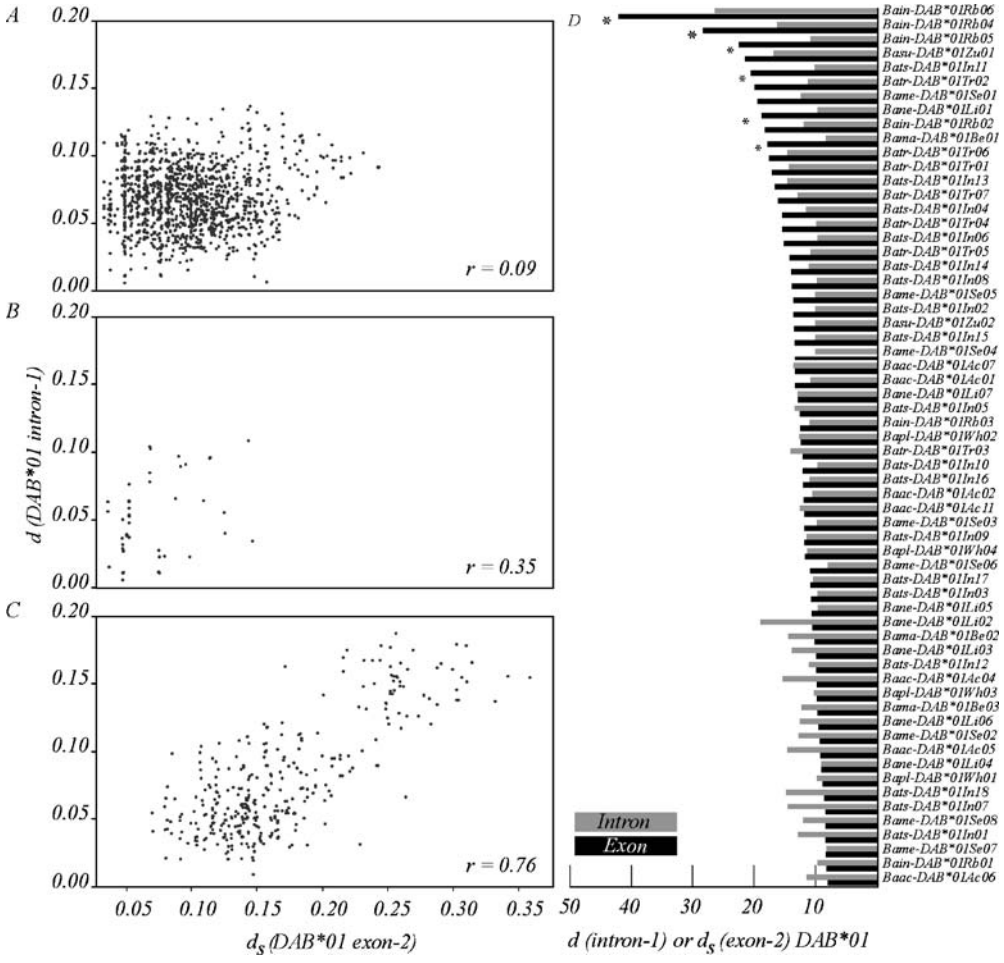

Fig. 4 Pairwise comparisons of substitution rates in *DAB*01* intron 1 nucleotide substitution rates (d) and in *DAB*01* exon 2 synonymous substitution rates (d_S). **a** Plots of among-lineage pairwise comparisons, **b** within-lineage pairwise comparisons and **c** only pairwise comparisons *Bats-DAB*0111(In)*, *Bane-DAB*0101(Li)*, *Bats-DAB*0101 (Be)*, *Bain-DAB*0104(Rb)*, **0105(Rb)* and **0106(Rb)* with all other Lake Tana large barb class II *DAB*01* sequences. Correlation coefficients (r) are indicated. **d** Pairwise comparisons of single *DAB*01* intron 1 or *DAB*01* exon 2 sequences with all other Lake Tana large barb class II *DAB*01* sequences. Test H_0: $d_s=d$, H_1: $d_s>d$, $^*P<0.05$, using the Z-test

Analyses of *Barbus* class I *UA* genes

The genomic DNA isolated from 34 Lake Tana African large barb species flock or Blue Nile *B. intermedius* individuals that served as template in PCRs for *DAB*01* analyses was also used to amplify class I *UA* exon 3 sequences, with the exception of individuals 29, 43 and 62. The class I *UA* analyses included four individuals that were not included in the *DAB*01* analyses (*B. intermedius* no. 08971, no. 32972 *B. megastoma* no. 74 and *B. platydorsus* no. 23). Analyses of 35 individuals by PCR using primers matching the start and the end of class I exon 3 yielded products of approximately 250 bp. In total, 72 different genomic class I *UA* sequences (Fig. 5) could be identified by sequence analyses and could be grouped on the basis of motifs at variable positions. In addition to the 72 African large barb sequences, an identical single pseudogene was detected in four *B. platydorsus* individuals. This pseudogene has two in frame stop codons and was not included in the analyses.

Fourteen out the 72 sequences identified were shared by at least two species (see Electronic Supplementary Material, S-3). All individuals analysed possessed sequences that were shared among species with exception of three individuals (Bob, no. 08971 and no. 08973), which did not possess shared sequences. The number of shared sequences per individual varied from one to six. The remaining 48 sequences were only found in individuals of the same species, i.e. species specific.

Phylogenetic relationship of *Barbus* class I *UA* sequences

Phylogenetic analyses included the African large barb class I *UA* sequences and class I *UA* sequences from zebrafish and common carp (Fig. 6). The sequences clustered in a trans-species manner in 11 sub-clusters, supported by bootstrap levels of at least 70%. Ten clusters contained only African large barb sequences and one cluster comprised a common carp and a zebrafish sequence (cluster VIII). Clustering of small sub-clusters into larger clusters was observed. Bootstrap levels of these larger clusters, however, were below 40%. The African large barb sub-clusters contained sequences from at least two African large barb species, with the exception of cluster X. This cluster comprised only sequences from two Blue Nile system individuals (no. 08971 and no. 08973). All sequences of these individuals were found together in a single clade, separate from all other clusters.

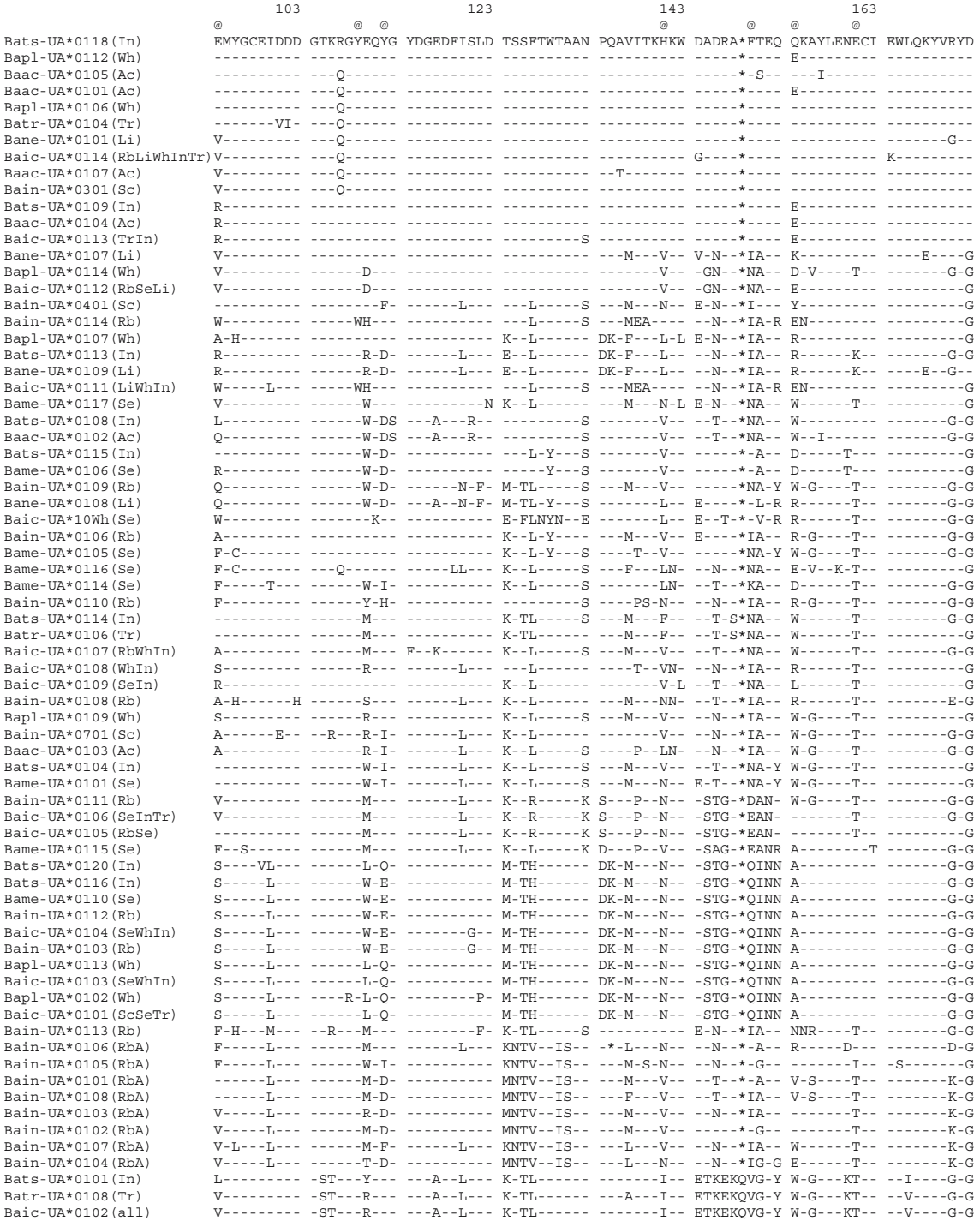

Fig. 5 The deduced African large barb class I *UA* exon 3 amino acid sequences. The *numbers above* denote the amino acid position relative to the mature protein. *Dashes* indicate identity to the Bats-UA*0118(In) sequence, and *asterisks* indicate gaps. Residues known to be involved in peptide binding in HLA-A2 molecule (Saper et al. 1991) are indicated by *ampersands*. African large barb sequence nomenclature is similar to that used in Fig. 1

Selection on *Barbus* class I *UA* allelic sequences

Similar to the situation for class II molecules, the amino acid residues of PBRs of the class I UA molecules are expected to be under selection with d_N/d_S ratios higher than 1. Although residues involved in peptide binding have not yet been identified for fish class I molecules, variability of the African large barb class I *UA* exon 3 sequences corresponds well to that of the HLA class I molecules. We have used the model of HLA-A2 (Saper et al. 1991) to assign PBR and non-PBR amino acid residues.

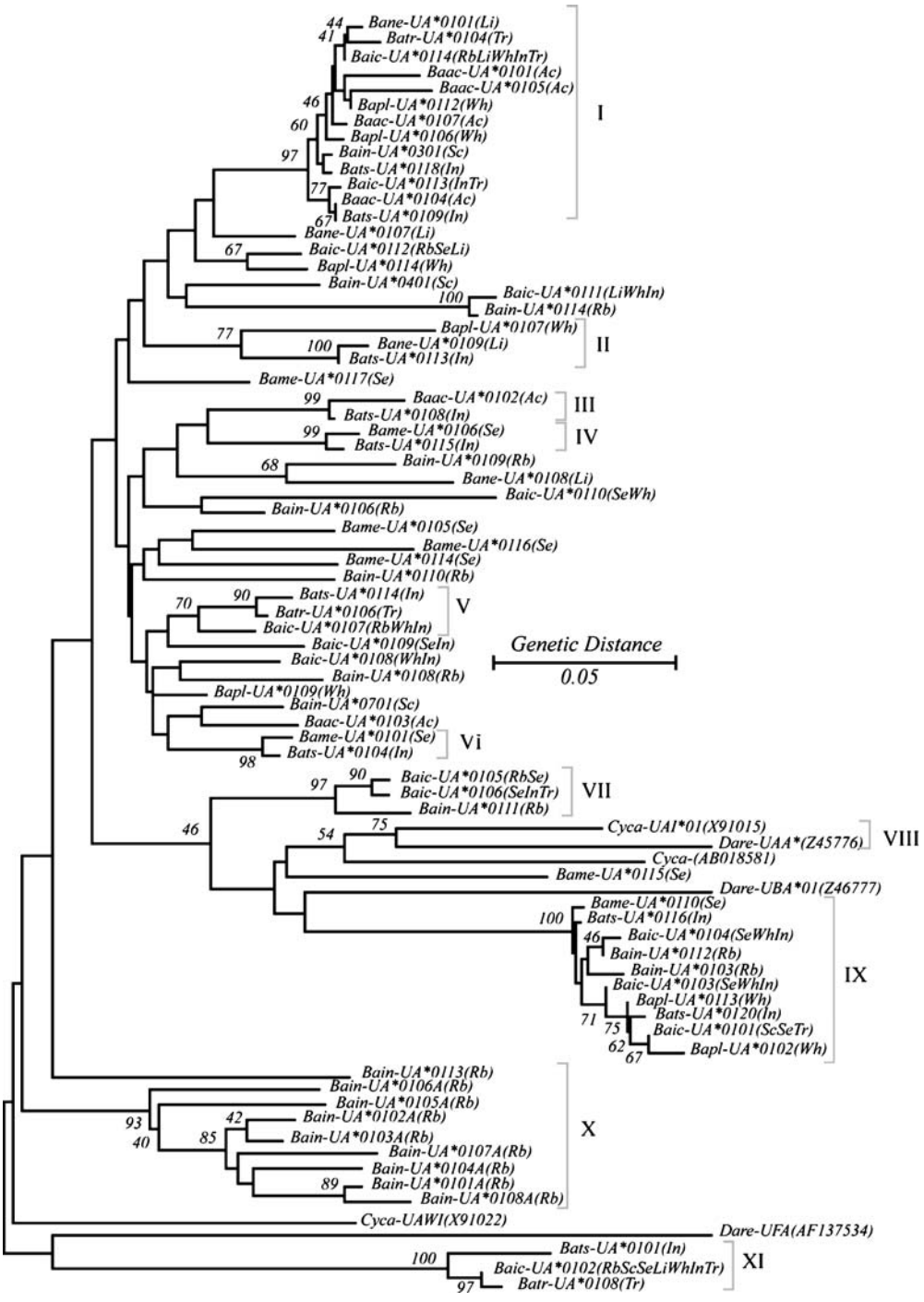

Fig. 6 Neighbour-joining tree of African large barb class I *UA* exon 3 nucleotide sequences. Trees were constructed using the method of Saitou and Nei, based on Jukes–Kantor method in Mega software. *Numbers at branch notes* indicate bootstrap confidence levels of 1,000 bootstrap replications. Accession numbers of common carp (*Cyca*) and zebrafish (*Dare*) sequences are shown in *brackets*, following the sequences designation. African large barb sequence nomenclature is similar to that used in Fig. 1. Sequences in cluster X have an additional 'A' before the allele number. Shared class I *UA* exon 3 sequences are marked by *grey boxes*. *Numbered brackets* denote clusters with bootstrap values over 70%

The d_N/d_S ratios calculated for PBR residues of all *UA* sequences, some within-lineage comparisons (I, II and X) and several among-lineage comparisons were above 3 (Table 3; Electronic Supplementary Material, S-2). Performing a statistical test of the difference between d_N and d_S for PBR residues revealed that d_N values calculated for all *UA* sequences were significantly higher than d_S. The d_N/d_S ratios calculated for non-PBR residues were below 3 with two exceptions (III vs IV, IV vs VI). However, performing statistical tests revealed that all d_N values were not significantly different from d_S values.

Divergence time estimates

In the phylogenetic analyses of *DAB*01* intron 1 and exon 2 sequences, the six riverine African large barb sequences (*Bain-DAB*0101–Bain-DAB*0106*) showed to be more distantly related to the Lake Tana African large barb sequences, as reflected in the genetic distance separating these sequences (Figs. 2, 3). Only one exception is observed, *Bain-DAB*0104* exon 2, which seems closely related to *Babr-DAB*0102* exon 2. Divergence time calculations

Table 3 Mean±SE of ratios of d_N per d_S substitutions per site and d_N–d_S values of PBRs and non-PBRs within and among groups of African large barb class I *UA* alleles as defined by phylogenetic relationship

	Non-PBR					PBR				
	d_N/d_S	SE	d_N/d_S	SE	Sig.[a]	d_N/d_S	SE	d_N–d_S	SE	Sig.[a]
All	0.881	0.030	0.015	0.028	**	3.714	0.146	0.897	0.123	*
I	1.194	0.009	0.003	0.008	**	23.204	0.121	0.304	0.112	*
II	1.505	0.032	0.030	0.029	**	4.965	0.327	0.304	0.253	*
III	0.633	0.024	−0.010	0.019	**			0.062	0.068	**
IV			0.007	0.007	**			0.133	0.165	**
V	1.119	0.019	0.003	0.018	**			0.087	0.064	**
VI			0.007	0.006	**			0.128	0.159	**
VII	0.380	0.018	−0.014	0.018	**	1.101	0.264	0.017	0.282	**
VIII	0.586	0.070	−0.079	0.072	**			0.383	0.192	*
IX	0.196	0.017	−0.026	0.018	**			0.078	0.049	**
X	0.690	0.026	−0.019	0.023	**	3.917	0.206	0.534	0.181	*
XI	0.253	0.023	−0.031	0.002	**			0.221	0.254	**
I vs II	1.175	0.036	0.015	0.037	**	8.022	0.214	0.669	0.169	**
I vs II	1.794	0.034	0.035	0.029	**	36.323	0.535	1.455	0.539	*
I vs IV	1.301	0.030	0.012	0.031	**	23.091	0.379	0.800	0.372	*
I vs V	1.580	0.034	0.029	0.034	**	15.710	0.586	1.667	0.552	*
I vs VI	1.633	0.041	0.036	0.043	**	15.000	0.682	1.644	0.640	*
I vs VII	1.427	0.051	0.038	0.050	**	15.158	0.444	1.154	0.432	*
I vs VIII	0.799	0.071	−0.045	0.067	**	2.116	0.717	0.804	0.868	**
I vs IX	1.036	0.063	0.005	0.064	**	3.837	1.025	1.617	0.732	*
I vs X	0.524	0.066	−0.100	0.062	**	6.296	0.319	0.988	0.269	*
I vs XI	0.720	0.085	−0.077	0.081	**	50832.	0.525	1.605	0.663	*

[a]*Z*-test: neutral selection, H: $d_N=d_S$, $^{**}P<0.10$; positive selection, H: $d_N>d_S$, $^{*}P<0.05$

based on d_S in exons and d in introns might provide information on the origin of these sequences.

Using a substitution rate of 2.85×10^{-9} per site per year (Dixon et al. 1996), divergence time calculations estimated mean values ranging from 7.7 to 41.7 million years for the six river barb *DAB*01* exon 2 sequences when compared to all Lake Tana African large barb sequences (Table 4). Similar calculation based on intron 1 sequences revealed mean values ranging from 9.4 to 26.2 million years for the six river barb *DAB*01* intron sequences. Remarkably, in several cases intron and exon values differed substantially. For example divergence time calculations between *Bain-DAB*0106* and all Lake Tana African large barb *DAB*01* exon 2 sequences suggested that exons diverged on average 41.7 million years ago, while similar calculation based on *DAB*01* intron 1 revealed mean values of 26.2 million years. Similar differences were calculated for *Bain-DAB*0102*, *Bain-DAB*0104* and *Bain-DAB*0105*, with exons estimated to be 1.5 times older than introns. This is in contrast with divergence time calculations for *Bain-DAB*0101* and *Bain-DAB*0103*, for which *DAB*01* exon 2 and intron 1 sequences were estimated to have a similar age.

Such differences were not only observed among comparisons of river barb sequences with Lake Tana African large barb sequences, but also among Lake Tana African large barb sequences (Table 4). The reverse situation, with exons estimated to be much younger than introns, was also observed among some pairwise comparisons of Lake Tana African large barb *DAB*01* sequences and between some pairwise comparisons of Lake Tana African large barb *DAB*01* sequences and river barb *DAB*01* sequences.

Out of all 1,880 pairwise comparisons of African large barb *DAB*01* exon 2 sequences, 86.7% diverged more than 5 million years. The sequences that had diverged less than 5 million years frequently lacked d_S, although they do possess d_N (data not shown). Most pairwise comparisons that lack d_S are within phylogenetic lineage comparisons. Calculations based on *DAB*01* intron 1 rates revealed that 96% of intron 1 sequences were estimated to be older than 5 million years.

Divergence time estimates based on d_S in exon 3 of African large barb class I *UA* sequences, again using a rate of 2.85×10^{-9} substitutions per site per year (Dixon et al. 1996), yielded invariably divergence time estimates above 5 million years for all pairwise comparison. Only a minority (5.8%) of all pairwise comparisons yielded divergence time estimates below 5 million years. Almost half of this minority was hampered by a lack of d_S. Although these sequences lacked d_S, they all possessed d_N (data not shown). Most pairwise comparisons that lack d_S are those within phylogenetic lineages.

Eight sequences unique to two individuals, collected in the Blue Nile River system, clustered together in the phylogenetic analyses. To highlight these sequences in the phylogenetic tree, they are denoted with an additional ‘A’ following the allele number (Fig. 6). Their position seems to suggest that these sequences represent an ancient lineage. Divergence time estimates revealed values between 33.7 and 63.5 million years. The common carp class I *UA* sequences diverged from the African large barb sequences between 45.2 and 51.8 million years ago. These estimates are higher than the estimated divergence time of 30 million years, separating the genera *Cyprinus* and

Table 4 Divergence time estimates based on d_S in *DAB*01* exon 2 or *UA* exon 3 sequences or *d* in *DAB*01* intron 1 sequences

Class II *DAB*01*					
		Time	SE	Time	SE
Vs		Million years		Million years	
*Bain-DAB*0101(Rb)*	All *Barbus* sequences	7.7	3.1	9.4	1.9
*Bain-DAB*0102(Rb)*	All *Barbus* sequences	17.7	3.4	11.7	2.4
*Bain-DAB*0103(Rb)*	All *Barbus* sequences	11.8	4.5	10.9	2.5
*Bain-DAB*0104(Rb)*	All *Barbus* sequences	27.5	4.6	16.4	2.5
*Bain-DAB*0105(Rb)*	All *Barbus* sequences	22.4	3.4	10.6	2.5
*Bain-DAB*0106(Rb)*	All *Barbus* sequences	41.7	4.5	26.2	2.4
*Bain-DAB*0101(Rb)*	*Bats-DAB*0118(In)*	1.0	1.0	13.8	3.6
*Bain-DAB*0102(Rb)*	*Bain-DAB*0113(In)*	7.1	4.6	17.7	4.0
*Bain-DAB*0103(Rb)*	*Bain-DAB*0103(Rb)*	3.7	2.5	18.5	4.3
*Bain-DAB*0101(Rb)*	*Bats-DAB*0118(In)*	40.2	14.9	25.4	5.3
*Bain-DAB*0102(Rb)*	*Bain-DAB*0113(In)*	26.7	10.2	9.6	2.8
*Bain-DAB*0103(Rb)*	*Bain-DAB*0103(Rb)*	27.1	9.5	9.2	2.9
*Baac-DAB*0104(Ac)*	*Bane-DAB*0104(Li)*	4.7	2.8	15.7	3.9
*Baac-DAB*0105(Ac)*	*Bane-DAB*0102(Li)*	4.3	3.8	27.0	5.0
*Bame-DAB*0102(Se)*	*Batr-DAB*0103(Tr)*	6.6	4.7	17.6	4.0
*Bats-DAB*0118(In)*	*Bats-DAB*0104(In)*	7.4	5.4	17.7	3.9
*Bame-DAB*0102(Se)*	*Bats-DAB*0111(In)*	19.8	8.8	8.8	2.7
*Basu-DAB*0101(Zu)*	*Batr-DAB*0102(Tr)*	30.1	9.9	17.2	4.1
*Bats-DAB*0106(In)*	*Bama-DAB*0101(Be)*	26.4	9.8	5.5	2.2
*Bats-DAB*0111(In)*	*Batr-DAB*0102(Tr)*	30.9	9.9	8.3	2.7
Class I *UA*					
*Bain-UA*0101A(Rb)*	All *Barbus* sequences	53.4	6.6		
*Bain-UA*0102A(Rb)*	All *Barbus* sequences	40.4	7.7		
*Bain-UA*0103A(Rb)*	All *Barbus* sequences	44.1	7.4		
*Bain-UA*0104A(Rb)*	All *Barbus* sequences	46.3	6.9		
*Bain-UA*0105A(Rb)*	All *Barbus* sequences	36.8	6.9		
*Bain-UA*0106A(Rb)*	All *Barbus* sequences	33.7	6.7		
*Bain-UA*0107A(Rb)*	All *Barbus* sequences	57.2	8.4		
*Bain-UA*0108A(Rb)*	All *Barbus* sequences	63.5	8.1		
Cyca-(AB018581)	All *Barbus* sequences	45.2	9.9		
*Cyca-UA1*01(X91015)*	All *Barbus* sequences	51.8	8.5		
Cyca-UAW1(X91022)	All *Barbus* sequences	51.4	7.1		
Dare-UAA(Z45776)	All *Barbus* sequences	54.9	6.8		
*Dare-UBA*01(Z46777)*	All *Barbus* sequences	58.2	6.9		
Dare-UFA(AF137534)	All *Barbus* sequences	82.3	9.9		

Barbus as determined by others (Cavender 1991; Dixon et al. 1996; Zardoya and Doadrio 1999). The zebrafish class I sequences diverged between 51.4 and 82.3 million years from the African large barb sequences, which is also longer than the estimated divergence time separating the genera *Danio* and *Barbus* (Cavender 1991; Dixon et al. 1996; Zardoya and Doadrio 1999).

Discussion

It is hypothesized that an ancestral *B. intermedius* population present in the Blue Nile system and in Lake Tana gave rise to the present-day Lake Tana African large barb species flock. During 5 million years of isolation of the lake, individuals of an ancestral population adapted to different ecological niches in the lake. As a result, they underwent speciation, resulting in 15 novel African large barb species (Nagelkerke and Sibbing 2000). Dixon and co-workers provided evidence that the Lake Tana African large barb class II *DAB* sequences were subject to selection. Analyses of d_N and d_S rates calculated for within and among phylogenetic lineages revealed that d_N/d_S ratios in PBR codons were usually higher than d_N/d_S ratios in non-PBR codons. Values calculated for PBR codons were mostly above 3, indicating strong selection. In addition, the four Lake Tana African large barb species studied possessed their own set of species-specific class II *DAB*01*.

The present study on class II *DAB*01* sequences isolated from six additional Lake Tana African large barb species underpins previous observations that each *Barbus* species has its own set of class II *B* alleles (Dixon et al. 1996).

However, the divergence time estimates calculated on the basis of the extended set of *DAB*01* alleles refutes the notion that these alleles have emerged after the ancestral Lake Tana large barb population has been isolated from their riverine counterparts. We were able to include class II *DAB*01* intron 1 and exon 2 sequences obtained from Blue Nile system barbs. Divergence time estimates based on synonymous substitution rates in class II *DAB* exon 2 sequences indicated that the riverine sequences predated those obtained from the Lake Tana. Phylogenetic analyses suggested that one of the riverine sequences (*Bain-DAB* 0106*) emerged before the separation of the genera *Cyprinus* and *Barbus*, 30 million years ago (Cavender 1991; Dixon et al. 1996; Zardoya and Doadrio 1999). The maximum divergence time estimate of approximately 42 million years supports this observation.

Comparisons of pairwise divergence time estimates of *DAB*01* intron 1 and exon 2 sequences revealed remarkable differences. Obviously, differences in divergence time estimates are reflected in d in *DAB*01* intron 1 and d_S in *DAB*01* exon 2. In most genes, d_S in exons and d in introns are essentially identical (Hughes 1999). However, higher d_S rates in exons are hypothesized to occur as a consequence of recombination and subsequent genetic drift of a gene on which selection acts to maintain polymorphism in an exon (Hughes 2000). The observed differences in substitution rates suggests that intron–exon shuffling has taken place in *DAB*01* genes as observed in comparisons of the phylogenies of intron and exon sequences separately.

Homogenization of introns is suggested to occur within allelic lineages (Cereb et al. 1997; Hughes 2000). In the analyses performed by Hughes (2000), comparisons of African large barb *DAB*01* intron 1 and exon 2 sequences revealed that they evolved at similar rates, while between-genera (*Barbus* and *Cyprinus*) comparisons revealed that intron 1 and exon 2 sequences evolved at different rates. The African large barb *DAB*01* may represent unique allelic lineages and thus, only within -lineage comparisons may reveal a correlation between d_S in exons and d in introns, with higher d_S rates than d rates. However, plots of d in intron versus d_S in exons for within and among pairwise comparisons of phylogenetic lineages did not reveal significantly different substitution rates in exons and introns. In contrast, such differences between substitution rates in exons and introns were observed for pairwise comparisons of some relatively older sequences and are therefore more representative for the comparisons between the genera (*Barbus* and *Cyprinus*) analyses presented by Hughes (2000).

It should be taken into account that analyses of substitution rates may be hampered by the fact that the sequences used were amplified from genomic DNA. This does not allow differentiation between functional and pseudogenes. In functional genes polymorphism would be concentrated in PBRs, while sequence variation in pseudogenes is expected to occur at random (Hughes and Nei 1989). Non-functional MHC class II sequences might be present among those isolated. However, with only one exception, the partial *DAB*01* sequences amplified were not pseudogenes, although it can not be excluded that other parts of these genes may have in-frame stops due to point mutations, deletions/insertions or complete deletion of exons, rendering them pseudogenes.

Analyses of the class I *UA* sequences showed that, unlike class II *DAB*01*, class I *UA* sequences were shared among Lake Tana African large barb species. Fourteen out of 72 genomic class I *UA* exon 3 sequences were shared by at least two African large barb species, while the remaining 58 were species specific. It should be taken into account that sharing of sequences is based on exon 3 sequences only. The complete *UA* sequences may differ in exon 1, which in general is the most divergent exon. However, sharing of exon 3 sequences were not only observed among Lake Tana African large barb species, but also between Lake Tana species and African large barbs from the Blue Nile system.

Five out of 14 shared sequences were also present in at least one African large barb individual from the Blue Nile system, strongly suggesting maintenance of class I sequences for more than 5 million years. Such sharing of alleles among closely related species is rarely seen. Comparisons of MHC class I genes among closely related teleostean or mammalian species (Gyllensten et al. 1990; Ono et al. 1993; Bontrop 1994; McAdam et al. 1995; Watkins 1995) did not reveal sharing of alleles, except in two cases. Cooper et al. (1998) isolated an MHC class I sequences shared between two species of chimpanzee, while Evans et al. (1998) identified an identical expressed class I allele in two new world primate species.

In the phylogenetic analyses all African large barb *UA* clusters, except cluster XI, comprised sequences of multiple African large barb species, indicating that these sequences have evolved in a trans-species manner as reported for other teleosts (Sato et al. 1997) and mammals (Mayer et al. 1988). These clusters usually represent allelic lineages present in a common ancestral population, except in some cases in which severe recent bottlenecks have occurred (Miller and Withler 1996). A recent study revealed that highly divergent class I *UBA* sequences in rainbow trout and Atlantic salmon were derived from a single locus (Shum et al. 2001; Aoyagi et al. 2002; Grimholt et al. 2002). In this study, sequences within a cluster, in general, seem to be more related as short genetic distances separated them. Sequences outside these clusters, on the other hand, seem to have relatively longer branch lengths. These included sequences from riverine barbs that were collected in the Blue Nile system. This may indicate that some class I lineages have been maintained in several African large barb species, most likely by a process of natural selection.

Average values of all pairwise comparisons indicated that positive selection acted upon PBRs, while non-PBRs were subject to neutral selection. Several within-lineage comparisons were hampered by a lack of d_S, although d_N were present. This indicates that selection pressure seems to be in favour of d_N. The lack of significant d_n–d_s values for positive selection on PBRs of within-lineage comparisons might be the result of the limited number of sequences in most clusters. In general, positive selection seems to

diversify PBRs among phylogenetic lineages, while non-PBRs are maintained by neutral selection. The generality of this conclusion is supported by analyses of the divergent phylogenetic clusters I to XI (Table 3). The Lake Tana African large barb species population are estimated to be isolated from their riverine counterparts in Blue Nile system in the order of 5 million years. Within this time span, class I alleles present in the ancestral African large barb population are expected to have diversified. Pairwise comparison of all sequences yielded invariably divergence time estimates above 5 million years. This indicated that most sequences were already diversifying within the ancestral African large barb population before the isolation of Lake Tana. Only a minority of 5.8% of all pairwise comparisons yielded divergence time estimates below 5 million years, of which almost half were hampered by a lack of d_S. Although these sequences lacked d_S, they all possessed d_N. Interestingly, these sequences belonged to a single phylogenetic cluster (Fig. 6, I). Furthermore, most clusters comprise sequences that are also present in riverine barbs. Data presented in this study clearly indicate that phylogenetic lineages are maintained in multiple species, most likely by purifying selection. Diversification seems to be located in PBRs, since d_N are mainly restricted to these positions. Alleles within allelic lineages seem to be maintained by purifying selection or neutral selection on non-PBRs, whereas sequences belonging to different lineages have been under positive selection.

However, the most striking observation from the data presented in this study is that the Lake Tana African large barb species investigated showed absolutely no sharing of class II *DAB*01* sequences, while they do share class I *UA* sequences. Similarly, the riverine African large barbs studied did not share class II *DAB*01* sequences with the Lake Tana African large barbs but do share the same class I *UA* sequences. We are aware of the fact that some species are only represented by a single individual. However, even with the small sample sizes, sharing of *UA* exon 3 sequences is easily identified. Our data seem to indicate that class I and class II alleles were subject to episodes of differential selection during ecological specialization. This resulted in maintenance of class I alleles that were present in the African large barb population before the isolation of the Lake and partitioning of class II alleles among the different barb species within a time span of 5 million years.

In mammals MHC genes are found in a complex and haplotypes of class I and class II alleles show a confounding pattern of linkage. In contrast, MHC class I and class II genes in teleosts are found in different linkage groups (Sultmann et al. 2000). The functional class I locus in fish is closely linked to *LMP* and *TAP* [rainbow trout (Hansen et al. 1999), zebrafish (Murray et al. 1999), Atlantic salmon (Grimholt et al. 2002)], whereas class II loci can be found in several linkage groups, either as functional genes or as pseudogenes (Graser et al. 1996; Kuroda et al. 2002), or as a single locus on a single linkage group as demonstrated for Atlantic salmon (Grimholt et al. 2002; Stet et al. 2002). The independent segregation of MHC genes in bony fish may have facilitated the partitioning of class II alleles among the different Lake Tana African large barb species.

Acknowledgements The authors would like to thank Martin de Graaf, Leo Nagelkerke and Sibbing for providing tissue samples of Lake Tana African large barb species flock individuals and Darrell Siebert and Elizabeth Barratt for providing tissue samples of riverine African large barb individuals.

References

Aoyagi K, Dijkstra JM, Xia C, Denda I. Ototake M, Hashimoto K, Nakanishi T (2002) Classical MHC class I genes composed of highly divergent sequence lineages share a single locus in rainbow trout (*Oncorhynchus mykiss*). J Immunol 168:260–273

Bingulac-Popovic J, Figueroa F, Sato A, Talbot WS, Johnson SL, Gates M, Postlethwait JH, Klein J (1997) Mapping of Mhc class I and class II regions to different linkage groups in the zebrafish, *Danio rerio*. Immunogenetics 46:129–134

Bontrop RE (1994) Nonhuman primate *Mhc-DQA* and *-DQB* second exon nucleotide sequences: a compilation. Immunogenetics 39:81–92

Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, Wiley DC (1993) Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature 364:33–39

Cavender TM (1991) The fossil record of the Cyprinidae. Chapman & Hall, London

Cereb N, Hughes AL, Yang SY (1997) Locus-specific conservation of the *HLA* class I introns by intra-locus homogenization. Immunogenetics 47:30–36

Cooper S, Adams EJ, Wells RS, Walker CM, Parham P (1998) A major histocompatibility complex class I allele shared by two species of chimpanzee. Immunogenetics 47:212–217

Dixon B, Nagelkerke LA, Sibbing FS, Egberts E, Stet RJM (1996) Evolution of MHC class II beta chain-encoding genes in the Lake Tana barbel species flock (*Barbus intermedius* complex). Immunogenetics 44:419–431

Erp van SHM, Dixon B, Figueroa F, Egberts E, Stet RJM (1996) Identification and characterization of a new major histocompatibility complex class I gene in carp (*Cyprinus carpio* L.). Immunogenetics 44:49–61

Evans DT, Piekarczyk MS, Cadavid L, Hinshaw VS, Watkins DI (1998) Two different primate species express an identical functional MHC class I allele. Immunogenetics 47:206–211

Graser R, O'Huigin C, Vincek V, Meyer A, Klein J (1996) Trans-species polymorphism of class II *Mhc* loci in danio fishes. Immunogenetics 44:36–48

Grimholt U, Drabløs F, Jorgensen SM, Høyheim B, Stet RJM (2002) The Major histocompatibility class I locus in Atlantic salmon (*Salmo salar* L.): polymorphism, linkage and protein modelling. Immunogenetics 54:570–581

Gyllensten UB, Lashkari D, Erlich HA (1990) Allelic diversification at the class II DQB locus of the mammalian major histocompatibility complex. Proc Natl Acad Sci USA 87:1835–1839

Hansen JD, Strassburger P, Thorgaard GH, Young WP, Du Pasquier L (1999) Expression, linkage, and polymorphism of MHC-related genes in rainbow trout, *Oncorhynchus mykiss*. J Immunol 163:774–786

Hughes AL (1999) Adaptive evolution of genes and genomes. Oxford University Press, New York

Hughes AL (2000) Evolution of introns and exons of class II major histocompatibility complex genes of vertebrates. Immunogenetics 51:473–486

Hughes AL, Nei M (1989) Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proc Natl Acad Sci USA 86:958–962

Klein J (1987) Origin of major histocompatibility complex polymorphism: the trans-species hypothesis. Hum Immunol 19:155–162

Kruiswijk CP, Hermsen T, Fujiki K, Dixon B, Savelkoul HFJ, Stet RJM (2004) Analysis of genomic and expressed major histocompatibility class Ia and class II genes in a hexaploid Lake Tana African 'large' barb individual (*Barbus intermedius*). Immunogenetics 55:770–781

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17:1244–1245

Kuroda N, Figueroa F, O'Huigin C, Klein J (2002) Evidence that the separation of *Mhc* class II from class I loci in the zebrafish, *Danio rerio*, occurred by translocation. Immunogenetics 54:418–430

Marsh SGE, Parham P, Barber LD (2000) The HLA facts book. Academic, San Diego

Mayer WE, Jonker M, Klein D, Ivanyi P, Seventer van G, Klein J (1988) Nucleotide sequences of chimpanzee MHC class I alleles: evidence for trans-species mode of evolution. EMBO J 7:2765–2774

McAdam SN, Boyson JE, Liu X, Garber TL, Hughes AL, Bontrop RE, Watkins DI (1995) Chimpanzee MHC class I A locus alleles are related to only one of the six families of human A locus alleles. J Immunol 154:6421–6429

Michalova V, Murray BW, Sultmann H, Klein J (2000) A contig map of the *Mhc* class I genomic region in the zebrafish reveals ancient synteny. J Immunol 164:5296–5305

Miller KM, Withler RE (1996) Sequence analysis of a polymorphic *Mhc* class II gene in Pacific salmon. Immunogenetics 43:337–351

Mohr PA (1962) The Geology of Ethiopia. University college of Addis Ababa Press, Addis Ababa, Ethiopia

Murray BW, Sultmann H, Klein J (1999) Analysis of a 26-kb region linked to the Mhc in zebrafish: genomic organization of the proteasome component beta/transporter associated with antigen processing-2 gene cluster and identification of five new proteasome beta subunit genes. J Immunol 163:2657–2666

Nagelkerke LAJ, Sibbing FA (2000) The large barbs (*Barbus* spp., Cyprinidae, Teleostei) of Lake Tana (Ethiopia), with a description of a new species, *Barbus osseensis*. Neth J Zool 50:179–214

Nei M, Jin L (1989) Variances of the average numbers of nucleotide substitutions within and between populations. Mol Biol Evol 6:290–300

Ono H, O'Huigin C, Tichy H, Klein J (1993) Major-histocompatibility-complex variation in two species of cichlid fishes from Lake Malawi. Mol Biol Evol 10:1060–1072

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Saper MA, Bjorkman PJ, Wiley DC (1991) Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 A resolution. J Mol Biol 219:277–319

Sato A, Klein D, Sultmann H, Figueroa F, O'Huigin C, Klein J (1997) Class I *Mhc* genes of cichlid fishes: identification, expression, and polymorphism. Immunogenetics 46:63–72

Sato A, Figueroa F, Murray BW, Malaga-Trillo E, Zaleska-Rutczynska Z, Sultmann H, Toyosawa S, Wedekind C, Steck N, Klein J (2000) Nonlinkage of major histocompatibility complex class I and class II loci in bony fishes. Immunogenetics 51:108–116

Shum BP, Guethlein L, Flodin LR, Adkison MA, Hedrick RP, Nehring RB, Stet RJM, Secombes C, Parham P (2001) Modes of salmonid MHC class I and II evolution differ from the primate paradigm. J Immunol 166:3297–3308

Stet RJM, Kruiswijk CP, Saeij JPJ, Wiegertjes GF (1998) Major histocompatibility genes in cyprinid fishes: theory and practice. Immunol Rev 166:301–316

Stet RJM, de Vries B, Mudde K, Hermsen GT, van Heerwaarden J, Shum BP, Grimholt U (2002) Unique haplotypes of co-segregating major histocompatibility class II *A* and II *B* alleles in Atlantic salmon (*Salmo salar*) give rise to diverse class II genotypes. Immunogenetics 54:570–581

Stet RJM, Kruiswijk CP, Dixon B (2003) Major histocompatibility lineages and immune gene function in teleost fishes: the road not taken. Crit Rev Immunol 23:441–471

Sultmann H, Sato A, Murray BW, Takezaki N, Geisler R, Rauch GJ, Klein J (2000) Conservation of Mhc class III region synteny between zebrafish and human as determined by radiation hybrid mapping. J Immunol 165:6984–6993

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

Watkins DI (1995) The evolution of major histocompatibility class I genes in primates. Crit Rev Immunol 15:1–29

Zardoya R, Doadrio I (1999) Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. J Mol Evol 49:227–237