

Phylogeography of the vairone (*Leuciscus souffia*, Risso 1826) in Central Europe

WALTER SALZBURGER,* ANITA BRANDSTÄTTER,† ANDRÉ GILLES,‡ WALTHER PARSON,† MELANIE HEMPEL,* CHRISTIAN STURMBAUER§ and AXEL MEYER*

*Department of Biology, University Konstanz, Universitaetsstrasse 10, 78457 Konstanz, Germany, †Institute of Legal Medicine, University of Innsbruck, Muellerstrasse 44, 6020 Innsbruck, Austria, ‡Laboratoire d'Hybridiologie, Université de Provence, 1 Place Victor Hugo, 13331 Marseille, France, §Institute of Zoology, University Graz, Universitaetsplatz 2, 8010 Graz, Austria

Abstract

The vairone *Leuciscus souffia* is a cyprinid fish that inhabits river systems in and around the Alps. The complete mitochondrial DNA control region (945 bp) was sequenced in 295 vairone from 22 populations in Central Europe. A total of 51 haplotypes were identified with a maximum pairwise distance between haplotypes of 5.6%. Phylogenetic analyses revealed two major clades in *L. souffia*, an 'Italian' clade, and an 'Alpine' clade. Two hybrid zones exist, in the Mediterranean Alps and in the Soca basin. The position of the sister species of *L. souffia*, *L. turskyi*, to the 'Alpine' and the 'Italian' clade could not be resolved unambiguously. However, a linearized tree analysis indicated that *L. turskyi* represents a third lineage, that originated at the same time as the 'Alpine' and the 'Italian' clades of *L. souffia*. In the 'Alpine' clade two groups were resolved, a subclade with haplotypes from the Rhône and Var basins and a cluster with haplotypes from the Danube and Rhine systems. Our data suggest a long history of the vairone in Central Europe, predating Pleistocene glacial cycles. Two main refugia during glaciations must have existed, one in Italy and another one most probably in the Danube system. However, age estimates based on molecular clock calibrations suggest the survival of 'Alpine' haplotypes in several drainages during the last glaciation cycles. The Rhine system was only recently colonized.

Keywords: Cyprinidae, hybrid zone, mismatch analysis, mitochondrial DNA, Pleistocene glaciations, refugia

Received 28 February 2003; revision received 2 June 2003; accepted 2 June 2003

Introduction

Phylogeography, the combination of phylogenetics and population genetics with biogeography, has existed as a formal discipline for over 15 years (Avice *et al.* 1987; Avice 1998, 2000; Bernatchez & Wilson 1998). By integrating past and present biogeographical records and the information provided by molecular data, this conjunction has led to important insights in population genetics, evolutionary biology and ecology (Avice 1994, 1998; Schluter 1997; Bernatchez & Wilson 1998; McCusker *et al.* 2000; Bernatchez 2001). Phylogeographic approaches have also been used to define conservation units within species, and were therefore incorporated in the concept of the evolutionarily significant unit (ESU)

by several authors (Avice 1994; Moritz 1994; Crandall *et al.* 2000) and in that of adaptive evolutionary conservation (AEC; Fraser & Bernatchez 2001). One of the most prominent subjects of phylogeography is the documentation of the impact of historical events on extant faunas and floras. Pleistocene glaciations, for example, have repeatedly shaped the distribution and genetic diversity of many species, and the cyclic climatic changes during the Quaternary were the most significant events in the evolutionary history of many extant lineages (Zink & Slowinski 1995; Avice 1998; Avice 2000; Bernatchez & Wilson 1998; Taberlet *et al.* 1998; Hewitt 1999, 2000). During major glaciations the polar ice shields were considerably more massive, and, in the northern hemisphere, large portions of the continents were covered with ice; mountain chains were glaciated, and the sea-level was remarkably lowered (Webb & Bartlein 1992; Rohling *et al.* 1998; Yokoyama *et al.* 2000).

Correspondence: Walter Salzburger. Fax: + 49-7531-883018; E-mail: walter.salzburger@uni-konstanz.de

Freshwater fish are particularly informative for the study of the effects of glaciations, the localization of refugia during cold periods, and the colonization of new habitats following glacial retreats (Bernatchez *et al.* 1992; Schluter & McPhail 1993; Bernatchez & Osinov 1995; Bernatchez & Wilson 1998; Durand *et al.* 1999; Nesbø *et al.* 1999; Wang *et al.* 1999; Englbrecht *et al.* 2000; Bernatchez 2001; Smith *et al.* 2001). This is the result of the island-like property of freshwater habitats, the relative instability of these habitats, the impacts of climatic changes on rivers and lakes and the availability of geological data for aquatic systems. To date, Pleistocene glaciation effects in the North American ichthyofauna have been studied in some detail, partly because North American glaciations were of greater impact relative to those in Eurasia (Dawson 1992; Bernatchez & Wilson 1998). However, there is a growing body of phylogeographic studies in European freshwater fish, including the brown trout *Salmo trutta* (Bernatchez *et al.* 1992; Bernatchez & Osinov 1995; Aurelle & Berrebi 2001; Bernatchez 2001; Weiss *et al.* 2001; Duftner *et al.* 2003), the grayling *Thymallus thymallus* (Koskinen *et al.* 2000; Weiss *et al.* 2002), the European perch *Perca fluviatilis* (Refseth *et al.* 1998; Nesbø *et al.* 1999), the bullhead *Cottus gobio* (Hänfling & Brandl 1998a; Englbrecht *et al.* 2000; Hänfling *et al.* 2002), and the chub *Leuciscus cephalus* (Hänfling & Brandl 1998b; Durand *et al.* 1999).

The vairone *Leuciscus souffia*, Risso 1826, is a riverine cyprinid that inhabits the upper reaches of rivers in the Alps and surrounding regions, where it feeds on invertebrates, mainly insects. *Leuciscus souffia* lays its eggs in the gravel, its breeding style has been classified as 'nonguarding, broodhiding lithophil', and successful spawning has been observed at temperatures as low as 12 °C (Bless 1996). The vairone is an endangered species. It is listed in Appendix III of the Bern convention (protected fauna) and in the fauna–flora–habitat directive (Natura 2000) of the European Union. Up to three subspecies of *L. souffia* have been described forming two main groups, *L. s. muticellus* in northern Italy as well as *L. s. souffia* and *L. s. agassi* in Alpine regions and north of the Alps (Kottelat 1997; Gilles *et al.* 1998; Machordom *et al.* 1999). Often, the latter are considered as a subspecies *L. s. souffia*. Phylogenetic studies based on mitochondrial DNA have placed *L. souffia* outside the remaining representatives of *Leuciscus*, as sister group to *Chondrostoma*, and it was suggested to include this species in the genus *Telestes* (Zardoya & Doadrio 1998, 1999; Machordom *et al.* 1999).

Here, we present the phylogeography of *Leuciscus souffia*, based on mitochondrial control region sequences of 295 specimens from 22 populations located throughout the distribution range of this species in the area of the Alps. We have chosen *L. souffia* for several reasons. Most importantly, it occurs in the upper reaches of the main Central European river systems draining the Alps (Danube, Rhine, Rhône and Po), which play a key role in our understanding of the pre- and postglacial history of Central European

freshwater fish. Unlike in most other European freshwater fish species studied so far, there is no evidence for human-induced faunal translocations in the vairone, which can be explained by the lack of economic interest for this species and its endangerment. The present distribution of *L. souffia* should thus reflect the natural population history. Our goals were (i) to apply a phylogeographic approach in order to reconstruct the recent phylogeographic history of the vairone as a model system for circum-alpine riverine taxa, (ii) to identify refugia during glaciations and routes of re-colonization of the Alps, (iii) to investigate the importance of postglacial connections between Alpine drainage systems for the present distribution of this species, and (iv) to compare our results with previous phylogeographic studies on European freshwater fish, such as the economically important salmonid species grayling (*Thymallus thymallus*) and brown trout (*Salmo trutta*). The extensive sampling of our study further allowed us to test the hypothesis that populations of formerly glaciated areas display lower levels of genetic diversity than those in ice-free areas (Bernatchez & Wilson 1998). Given that the vairone is an endangered species, our phylogeographic study might also provide the basis for declaring conservation units in this species.

We performed a phylogenetic analysis based on DNA sequences of the mitochondrial control region in order to delineate the major clades in *L. souffia*. A linearized tree approach (Takezaki *et al.* 1995) was applied to estimate divergence times between lineages based on previously published molecular clock calibrations. For each major clade, a haplotype network was constructed. Levels of genetic diversity were tested by a hierarchical analysis of molecular variance (AMOVA; Excoffier *et al.* 1992), and the demographic histories of the lineages were compared by means of mismatch analyses (Schneider & Excoffier 1999).

Materials and methods

Samples

A total of 295 specimens of *Leuciscus souffia* from 22 populations had been collected (Fig. 1 and Table 1). We also sampled its presumed sister species *Leuciscus turskyi* from the Vrba river in Croatia. Fin-clips or muscle tissue preserved in 100% ethanol were used as source of DNA. Voucher specimens are available as life stocks in the Alpenzoo Innsbruck, Austria (Inn, Var, Natisone, Torre, Iudrio), or from the authors (remaining samples).

mtDNA amplification and sequencing

Total DNA was extracted by applying a proteinase K digestion followed by sodium chloride extraction and ethanol precipitation (Bruford *et al.* 1998). Polymerase chain reaction (PCR) amplification was performed according to

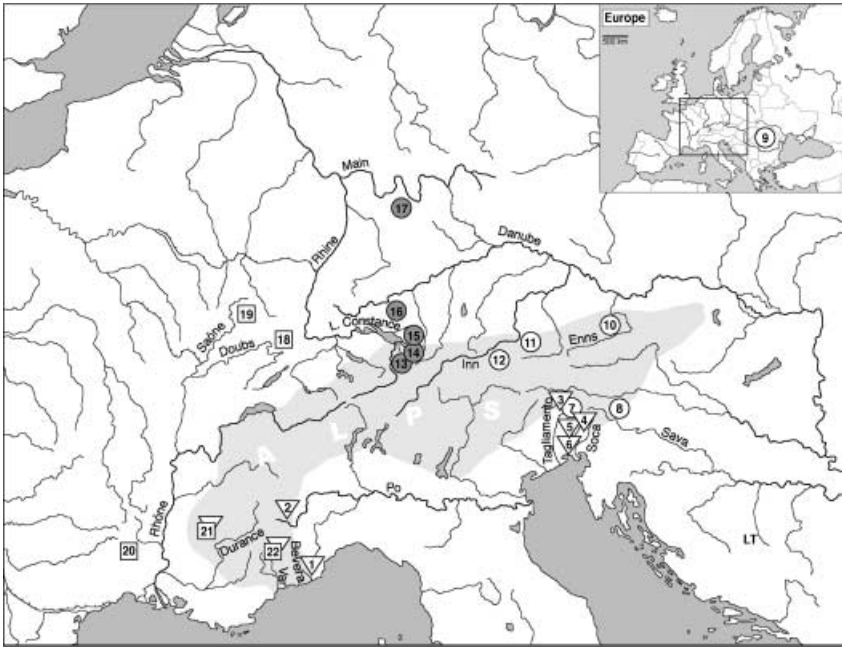


Fig. 1 Map of Central Europe showing the major river systems draining the Alps and the sampling sites of *Leuciscus souffia* for this study. Symbols refer to the affiliation of haplotypes of a population to one of the three groups identified by our phylogenetic analyses (∇ = 'Italian' clade, \square = 'Rhône/Var' clade, \circ = 'Danube/Rhine' clade in Fig. 2). LT indicates the sample locality for *L. turskyi*.

standard methods on an ABI 9700 Cycler using one published (L-Pro-F; Meyer *et al.* 1994) and one newly designed (Leuc-DR 5'-GCT TCA GTG CCG TGC TTT GTT C-3') primer under the following conditions: 35 cycles with denaturation at 94 °C for 30 s, annealing at 52 °C for 30 s, and extension at 72 °C for 90 s (reaction volume: 25 μ L). PCR products were purified using ExoSapIt (Amersham) or the Qiaquick PCR purification kit (Qiagen), sequenced in both directions with the BigDye termination reaction chemistry and analysed with ABI 3100 or ABI 3700 capillary DNA sequencers (Applied Biosystems).

Sequence alignment, phylogenetic and population genetic analyses

Sequences were aligned using the computer program SEQUENCHER (GenCodes, USA), resulting in an alignment of 945 bp (base pairs). Up to six indels, which were coded as gaps, had to be included in the complete alignment. The sequences have been deposited in GenBank under the accession numbers AY301621–AY301925.

For phylogenetic tree reconstruction we applied maximum likelihood, maximum parsimony and neighbour-joining methods using the computer program PAUP* 4.0b10 (Swofford 2002). *Chondrostoma nasus* and *C. genei* were declared as outgroup (see Zardoya & Doadrio 1999; Gilles *et al.* 2001). The optimal model of molecular evolution for the maximum likelihood analysis was determined in a likelihood-ratio testing framework running the computer program MODELTEST 3.06 (Posada & Crandall 1998). A maximum-parsimony analysis using all equal weights was performed under the heuristic search option (50 replicate searches with random addition

of taxa). For neighbour-joining, we used Jukes–Cantor model, because of the close relatedness of the in-group taxa. A bootstrap analysis was performed with 5000 pseudo-replicates for maximum parsimony and neighbour-joining, and under the fast stepwise addition option for maximum parsimony. For maximum likelihood, we performed 100 bootstrap replicates (fast stepwise addition option).

To date tentatively the major cladogenetic events in *L. souffia* we constructed a linearized tree and calculated average pairwise distances between clades at different hierarchical levels. For the construction of the linearized tree we performed the two-cluster-test of Takezaki *et al.* (1995) to test rate constancy for all internal nodes of a neighbour-joining topology with the computer program LINTRE (Takezaki *et al.* 1995). First, the rate of molecular evolution was tested for all internal nodes of the topology. In our case, none of the taxa had to be excluded according to the two-cluster test ($P < 0.01$), so that the linearized tree could be constructed under the assumption of rate constancy among all ingroup taxa. For this analysis we chose two outgroup taxa, one specimen each of *Chondrostoma genei* and *C. nasus*.

We then calculated average uncorrected pairwise distances among the haplotypes of the three major lineages in the linearized tree ('Alpine', 'Italian', *L. turskyi*) in order to estimate the percentage divergence, which separated these lineages. In a second step, we excluded the haplotypes from *L. turskyi* and focused on the divergence between the 'Alpine' and the 'Italian' clade only. This second linearized tree analysis was carried out because the phylogenetic position of *L. turskyi* was not consistent in our analysis. Average pairwise distances were also calculated between haplotypes of the Rhône/Var basins and the Danube/Rhine systems

Table 1 Sample locations, drainage system, numerical code for each population (Pop. no.; see Fig. 1), sample size (*N*) and haplotype frequencies of all sequenced individuals of *Leuciscus souffia*

Population	Pop.	Nr. drainage	N	L01	L02	L03	L04	L05	L06	L07	L08	L09	L10	L11	L12	L13	L14	L15	L16	L17	L18	L19	L20	L21
Bevera	1	Bevera	1									1												
Carmagnola	2	Po	10																		1			
Tagliamento	3	Tagliamento	5															2				1		
Iudrio	4	Soca	12												3		1	2	1			3		
Natisone	5	Soca	10															6	3			1		
Torre	6	Soca	22												5	1	4	4	2	1	4			
Cosizze	7	Soca	9																				5	3
Sava/Jevscek	8	Danube	5																				1	
Romania	9	Danube	14																					
Enns/Reichraming	10	Danube	30	17																				
Inn/Kirchbichl	11	Danube	26	3	1																			
Inn/Schwaz	12	Danube	1																					
Schwarzach	13	Rhein	12			11																		
Bregenzerach	14	Rhein	20			20																		
Leiblach	15	Rhein	8			7	1																	
Argen	16	Rhein	23			19	4																	
Main	17	Rhein	7			2																		
Doubs	18	Rhone	11						8	1														
Saone	19	Rhone	7						4															
Chassezac	20	Rhone	13					3					5											
Durance	21	Rhone	11			1	2	1		3														
Var	22	Var	38					27	4	2	1		1		2	1								
			295	20	1	59	5	31	18	2	3	1	4	5	2	9	1	5	14	6	2	9	6	3

Note that haplotype codes L33 and L34 were assigned to *L. turskyi* and haplotypes L49 and L50 were assigned to *Chondrostoma genei* so these haplotypes do not appear in the table.

as well as within each lineage. The obtained pairwise distances were used to estimate the timing of major cladogenetic events in the vairone based on previously published molecular clock estimates for freshwater fish, ranging between 1% and 2% substitutions per million years (control region: Nesbø *et al.* 1999; Englbrecht *et al.* 2000; Bernatchez 2001; Koskinen *et al.* 2002; cytochrome *b* gene: Zardoya & Doadrio 1999; Dowling *et al.* 2002).

Then, a minimum spanning network was constructed for both the North Italian populations (corresponding to subspecies *L. s. muticellus*) and the Alpine populations (corresponding to subspecies *L. s. souffia/agassi*). For the construction of the network we used the computer program TCS (Clement *et al.* 2000). In case of alternative branching orders in the TCS-generated network, we chose those connections between haplotypes that were favoured by a maximum likelihood search in PAUP*. Therefore, the optimal maximum likelihood topology was translated into an unrooted network. The general topology of the minimum spanning network thus remained unaffected. However, in case of alternative connections between haplotypes, only the most likely one was depicted. As further evaluation of the minimum spanning networks, we assigned the consistency index (CI; Kluge & Farris 1969) for each mutation under maximum parsimony using PAUP* and depicted this information on the network trees.

Levels of genetic diversity within and among the major lineages of *L. souffia* were tested by a hierarchical analysis of molecular variance (AMOVA; Excoffier *et al.* 1992) using ARLEQUIN 2.0 (Schneider *et al.* 1999) and 1000 permutations. The Bevera sample (locality 1) was excluded from this analysis, since there was only one specimen available. For AMOVA, we grouped the populations by drainages [Danube, Rhine, Rhône, Po, Var, Soca (includes Cosizze, Iudrio, Natisone, Torre), Tagliamento]. Finally, a mismatch analysis (according to Schneider & Excoffier 1999) was performed with ARLEQUIN 2.0 in order to compare the demographic history of the major lineages and to test for recent population expansion. We computed the moment estimator of time to the expansion (τ), and the mutation parameters before ($\theta_0 = 2\mu N_0$) and after ($\theta_1 = 2\mu N_1$) the expansion using a parametric bootstrap approach (1000 simulations), where μ is the mutation rate and N is the female effective population size. We also computed the raggedness index of Harpending (Harpending 1994) and used the sum of square deviations (SSD) between the observed and the expected mismatch as a test statistic for the validity of the estimated stepwise expansion model (Schneider & Excoffier 1999). Finally, we used the equation $\tau = 2\mu t$ to infer a time-scale for the demographic expansion (where t is the expansion time in generations). We assumed an average generation time of 2 years (G. Pechlaner, personal observation).

Table 1 (Continued)

L22	L23	L24	L25	L26	L27	L28	L29	L30	L31	L32	L33	L35	L36	L37	L38	L39	L40	L41	L42	L43	L44	L45	L46	L47	L48	L51	L52	L53	L54	L55	
						3						2															1	1	2		
1															1																
1							1																								
	1																														
		4																													
		5																													
		22													1																
							1																								
1	3	27	4	1	3	3	1	1	3	2	2	2	2	3	1	1	10	4	4	2	1	1	1	1	1	1	1	1	1	1	2

Results

Phylogenetic analyses

Fifty-one different haplotypes were detected in the 295 specimens of *Leuciscus souffia*. Haplotypes differed from each other between 1 and 53 mutations in the 945-bp-long alignment, which translated to a maximum uncorrected pairwise distance of 5.6%. The largest genetic difference, however, was found between *L. turskyi* and two haplotypes of *L. souffia* (haplotypes L09 and L17), which differed from each other by 63 mutational steps (6.7%). The nucleotide composition was A: 31.4%, C: 21.9%, G: 15.6%, T: 31.1%, the overall transition : transversion ratio was 2.1. By running MODELTEST 3.06 (Posada & Crandall 1998), the HKY model with a gamma substitution correction (shape parameter $\alpha = 0.96$; estimated from the data-set) and a proportion of invariable sites of 0.75 was suggested as best fitting model.

Phylogenetic analyses of the data set including all 51 different haplotypes of *Leuciscus souffia*, the two specimens of *L. turskyi*, and the outgroup taxa *Chondrostoma nasus* and *C. genei* revealed consistent results. In all phylogenetic trees obtained, three main groups were identified among the in-group taxa, the *L. turskyi*-clade, the 'Italian'-clade with haplotypes from Northern Italy (except the haplotypes of

the Cosizze river), one haplotype that also occurred in the Var basin and one haplotype of the Durance river, and the 'Alpine' clade with all remaining haplotypes. In maximum likelihood ($-\ln 2631.4548$; tree not shown), *L. turskyi* was resolved as sister group to the 'Alpine'-clade, both being sister to the 'Italian' clade of *L. souffia*. However, after enforcing a molecular clock, *L. turskyi* was placed as sister to both the 'Italian' and the 'Alpine' clade of *L. souffia*. The monophyly of the 'Italian' clade was supported by a bootstrap value of 95, that of the 'Alpine' clade by a bootstrap value of 76. In the neighbour-joining tree (Fig. 2), *L. turskyi* was resolved as the most ancestral split, sister group to the 'Italian' and the 'Alpine' clade of *L. souffia*, however, with low bootstrap support. In the 'Italian' clade, no clear phylogeographic pattern could be derived from the tree. Haplotype L18, which was found in one specimen of the Carmagnola population (locality 2) and in one of the Torre population (locality 6), was resolved as most ancestral split in the 'Italian' clade, but supported by a low bootstrap value. Within the 'Alpine' clade two main groups could be observed: a clade including the representatives of the Rhône drainage and the Var basin, and the Danube-clade including the five haplotypes found in the Rhine system, which formed a monophyletic group. Due to the large number of closely related haplotypes, maximum parsimony analysis was limited

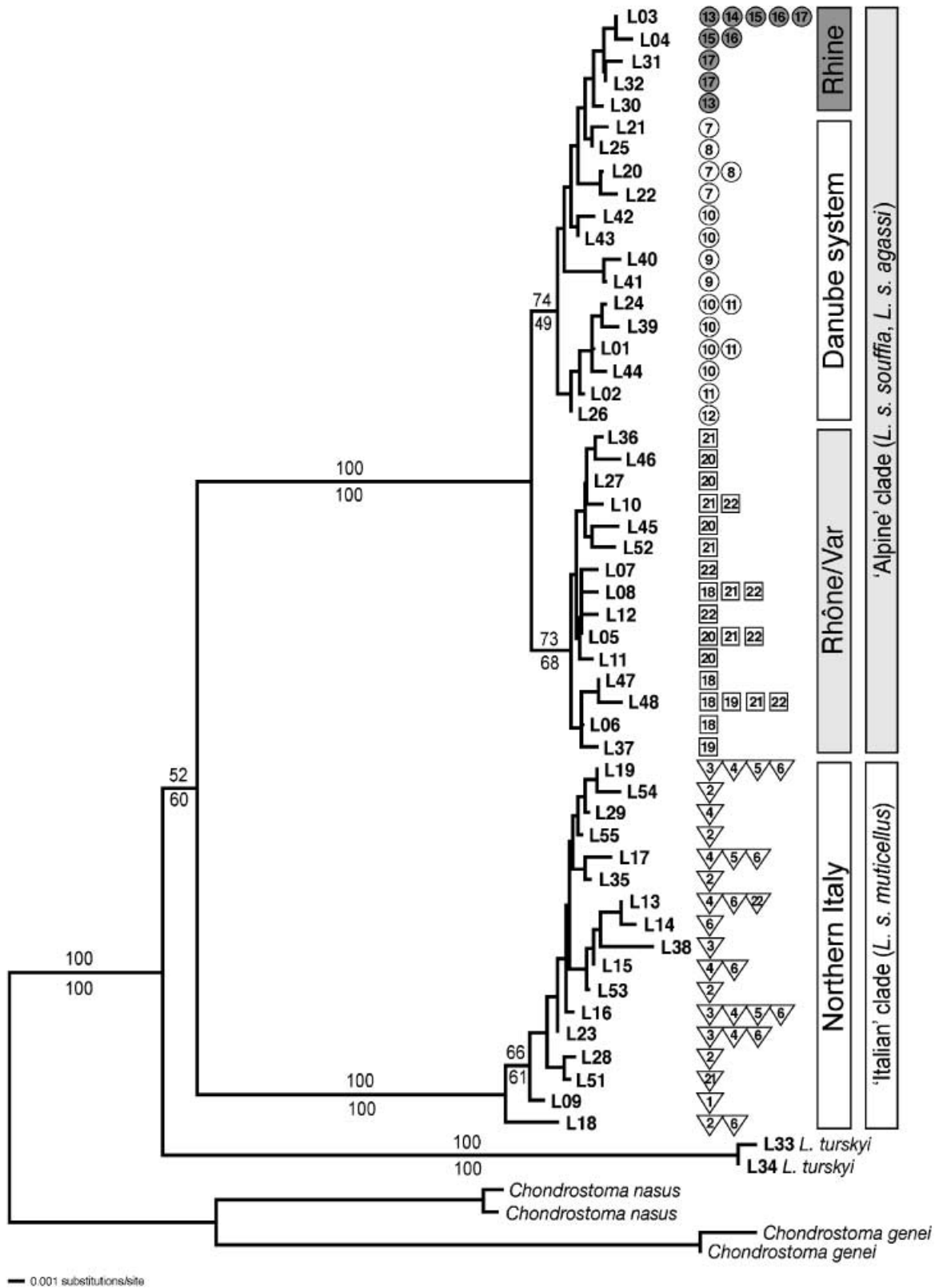


Fig. 2 Neighbour-joining tree of all 51 haplotypes identified in *Leuciscus souffia* and the two specimens of *L. turskyi*. *Chondrostoma nasus* and *C. genei* were used as outgroup taxa (see Zardoya & Doadrio 1999). Bootstrap values are depicted for branches of the main clades, with neighbour-joining bootstraps above and maximum parsimony bootstraps below the corresponding branch. Numbers of the haplotypes refer to those in Table 1.

Table 2 Average uncorrected pairwise distances between and within groups resolved by the linearized tree analysis

Comparison	% divergence	SD	$t_{\text{divergence}}$ (1% rate)	$t_{\text{divergence}}$ (2% rate)
<i>L. turskyi</i> : 'Alpine': 'Italian'	5.294	0.597	5.29 ± 0.60 Myr	2.64 ± 0.29 Myr
'Alpine': 'Italian'	5.089	0.351	5.09 ± 0.35 Myr	2.54 ± 0.17 Myr
Danube/Rhine:Rhône/Var	0.969	0.150	969 ± 150 kyr	484 ± 74 kyr
'Italian'	0.410	0.205	410 ± 205 kyr	205 ± 102 kyr
Danube/Rhine	0.389	0.184	389 ± 184 kyr	195 ± 92 kyr
Rhône/Var	0.328	0.138	328 ± 138 kyr	164 ± 69 kyr
Rhine	0.196	0.086	196 ± 86 kyr	98 ± 43 kyr

The obtained values were translated into divergence times ($t_{\text{divergence}}$) with a rate of 1% and 2% substitutions per million years

to 10^9 rearrangements (~10 000 trees saved; unweighted tree length: 224 steps). In the strict consensus topology (not shown), *L. turskyi* was placed as most ancestral split, followed by the 'Italian' clade, sister to the 'Alpine' clade of *L. souffia*. Again, two groups were identified in the 'Alpine' clade, the Rhône/Var clade and the Danube clade including the Rhine-system haplotypes. In all analyses, one haplotype found in the Var river (locality 22; haplotype L13) and one haplotype found in the Durance river (locality 21; haplotype L51) were grouped into the 'Italian' clade. The haplotypes of the small Cosizze-river in the border region of Italy and Slovenia (locality 7; haplotypes L20, L21, L22) were consistently grouped into the 'Alpine' clade, although the Cosizze river is connected via the Natisone river (locality 5) to rivers Torre (locality 6) and Iudrio (locality 4), all of which comprise exclusively haplotypes assigned to the 'Italian' clade.

The linearized tree (not shown) indicated that the split between *L. turskyi* and the 'Alpine' and 'Italian' clade of *L. souffia* happened contemporaneously. It was also suggested that the split between the Rhône/Var and the Danube/Rhine subclade predated the oldest splits in any of the clusters from Italy, Danube/Rhine, or Rhône/Var. The average pairwise distance among haplotypes of the three main lineages was 5.29% (SD 0.598), and that between the 'Alpine' and the 'Italian' clade was 5.09% (SD 0.351; see Table 2). Among the haplotypes assigned to the Rhône/Var and to the Danube/Rhine subclade we found an average pairwise distance of 0.97% (SD 0.150). Average distances were low within the major groups: Danube/Rhine = 0.39% (SD 0.184), Rhône/Var = 0.33% (SD 0.138), Italy = 0.41% (SD 0.205), Rhine = 0.20% (SD 0.086). Average pairwise distances and estimated divergence times with two different molecular clock rates are depicted in Table 2.

Population level analyses

In the minimum-spanning network of the 'Italian' clade (Fig. 3a) haplotype L55 from Carmagnola in the upper reaches of the Po river was resolved as central haplotype, through which the most abundant haplotypes L13, L16, and L19 were connected. The haplotype diversity was largest

in the Carmagnola population (locality 2) and in the Torre river (locality 6). Also, in these two populations the most ancestral haplotype according to the phylogenetic analyses, haplotype L18, was found.

The unrooted network tree including the haplotypes of the 'Alpine' clade is depicted in Fig. 3(b). In the Rhône/Var subclade, the two most abundant haplotypes L05 and L06 occupied a central position in the network and they were found in a large portion of the sampling area. The two populations in the upper reaches of the Rhône drainage, the Saône and Doubs, displayed fewer haplotypes than the three southern populations. Also, the haplotypes found in rivers Saône and Doubs were restricted to one part of the network tree, while the haplotypes of the remaining populations were scattered across the network tree. The Rhône/Var subclade connected to the Danube subclade through haplotype L06, which connected to the westernmost haplotype L26 of the Danube system. In the Danube subclade, haplotype L01 occupied a pivotal position since it connected to the remaining haplotypes of rivers Inn and Enns, to a southern clade including the haplotypes found in Jevscek (Slovenia), Cosizza (Italy) and Romania, and to the haplotypes found in the Rhine drainage. All Rhine haplotypes shared one synapomorphic mutation (a C-T change in position 197 in our alignment; Fig. 3b). The most abundant haplotype in the Rhine-system (haplotype L03), which occurred in all five populations of the Rhine drainage, connected to the central haplotype of the Danube subclade (haplotype L01) via the haplotype L32 found in river Main.

AMOVA results and mismatch analysis

To quantify population genetic structuring within and among populations AMOVA was performed (Table 3), in which all but one of the 22 sampled populations were grouped according to drainage systems (Danube, Rhine, Rhône, Po, Var, Soca, Tagliamento). AMOVA revealed that 70.08% of the variation was found between drainages, 22.03% within drainages and only 7.90% within populations (global ϕ_{ST} = 0.92).

A mismatch analysis was performed to compare the demographic history of the major lineages in *L. souffia*. The

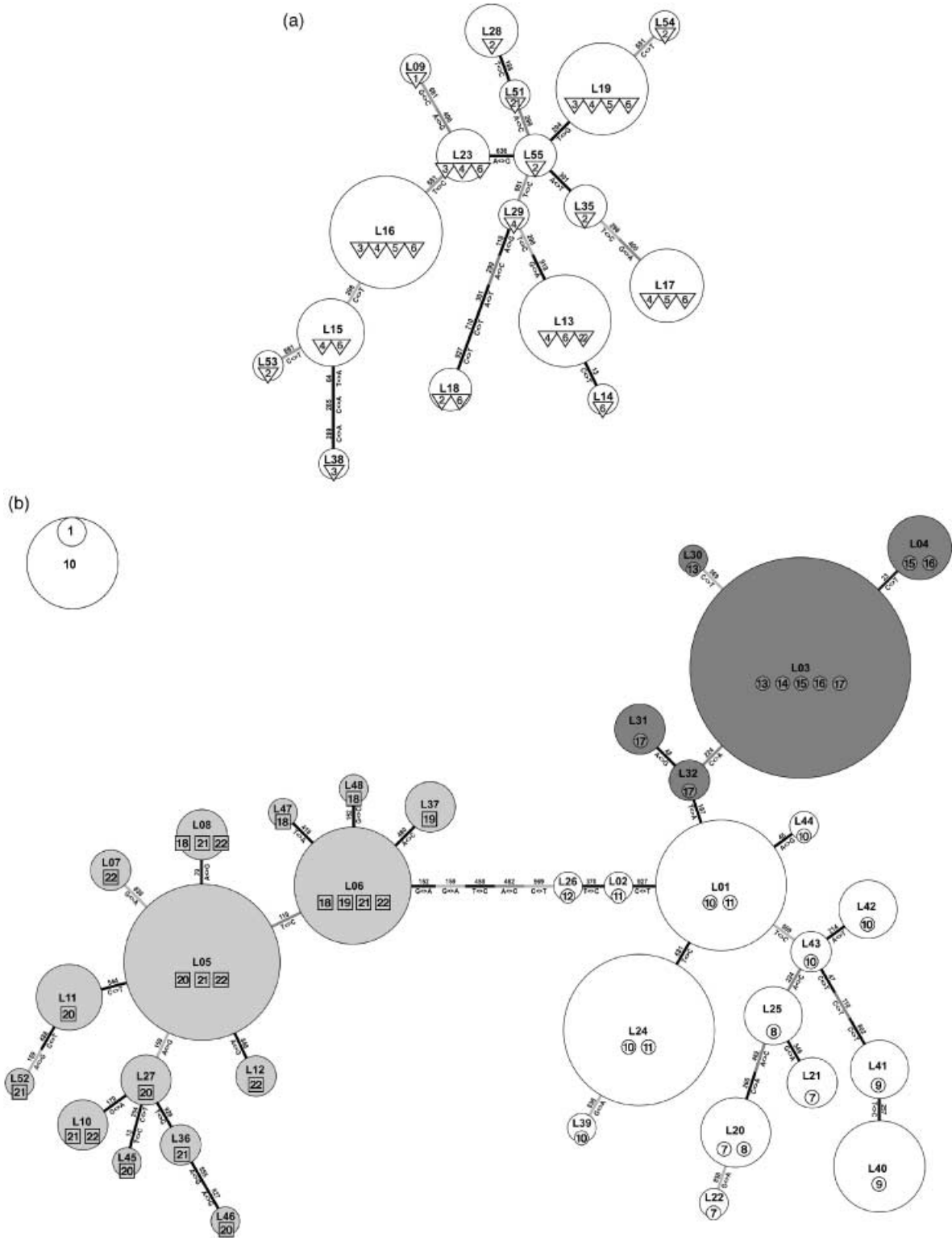


Fig. 3 Unrooted haplotype networks for the 'Italian' (a) and the 'Alpine' (b) clade generated with TCS (Clement *et al.* 2000). In case of alternative connections between haplotypes, those connections were chosen, that were also found by a maximum likelihood analysis of the same data-set in PAUP* (Swofford 2000). As evaluation for the minimum spanning networks, we depicted the consistency index (CI; Kluge & Farris 1969) for each mutational step on the network trees, together with the information about the mutation itself (base change and positional information in the alignment). Branches were shaded in grey according to the value of the CI. Symbols for locations and haplotype numbers refer to Figs 1, 2, and Table 1, the sizes of the circles resemble the number of individuals per haplotype.

Table 3 Results from the hierarchical analysis of molecular variance (AMOVA)

Analysis structure	Source of variation	Sum of squares (d.f.)	Variance component	<i>P</i>	Fixation index	% of variation
by drainage	among groups	2235.454 (7)	Va = 8.3159	< 0.001	$\Phi_{CT} = 0.70$	70.08
	among populations	464.078 (14)	Vb = 2.6140	< 0.001	$\Phi_{SC} = 0.74$	22.03
	within populations	253.980 (271)	Vc = 0.9372	< 0.001	$\Phi_{ST} = 0.92$	7.90

Populations have been grouped by drainage systems [Danube, Rhine, Rhône, Po, Var, Soca (includes Cosizze, Iudrio, Natisone, Torre), Tagliamento].

Table 4 Results of the mismatch analyses

Clade	τ	Obs. mean		θ_1	Ragged.	$P_{(HARP)}$	$P_{(SSD)}$	$t_{\text{expansion}}$ (kyr)
		Obs. mean	θ_0					
Italy	3.93 (1.71–6.36)	3.256	0.0 (0.0–1.7)	13.2 (6.5–4803)	0.025	0.57	0.43	371 (162–601)
Danube	6.597 (2.34–13.01)	3.270	0.0 (0.0–5.1)	5257 (1848–39090)	0.041	0.56	0.39	624 (221–1230)
Rhine	1.873 (0.20–3.12)	0.386	0.0 (0.0–0.6)	0.4 (0.0–3072)	0.280	0.59	0.58	177 (19–295)
Rhône/Var	1.552 (0.69–1.95)	1.484	0.0 (0.0–1.1)	540.3 (6.0–6735)	0.045	0.53	0.67	147 (65–184)
Danube/Rhine	4.874 (1.67–9.25)	3.151	0.0 (0.0–1.7)	6.5 (3.0–2022)	0.051	0.28	0.23	461 (158–874)
Danube/Rhine/Rhône	10.000 (4.11–16.50)	5.868	0.0 (0.0–4.8)	11.2 (5.1–56577)	0.030	0.21	0.18	945 (388–1560)

The expansion parameter (τ), the mismatch observed mean (obs. mean), the mutation parameter before (θ_0) and following (θ_1) expansion, and the raggedness index (Harpending 1994) (Ragged.) were calculated for all major groups separately and for the groups of haplotypes according to the obtained phylogenies. Values in brackets refer to 95% consistency intervals. $P_{(HARP)}$ is the probability of getting a higher value of the raggedness index of Harpending (1994) by chance. $P_{(SSD)}$ (sum of squared deviations) is the probability of observing a less good fit between the model and the observed distribution by chance. The timing of the most important demographic expansion ($t_{\text{expansion}}$) in each group was determined on the basis of the equation $\tau = 2\mu t$, assuming a mutation rate of 1% per million years and an average generation time t of 2 years. See Fig. 4 for graphs.

mismatch distribution differed substantially among the lineages (Fig. 4), with the lowest mean number of differences in the Rhine population (0.386; age expansion parameter $\tau = 1.873$; see Table 4 for more details) suggesting a very recent expansion. For the Rhône/Var subclade a mean number of difference of 1.484 was detected ($\tau = 1.552$) and for the 'Italian' clade a mean number of difference of 3.256 was found ($\tau = 3.930$). Within the Danube drainage only (mean number of difference = 3.270; $\tau = 6.597$) two modes were observed in the pairwise comparisons. When combining the Danube and Rhine populations – which seemed more appropriate according to our phylogenetic analysis and the network trees – a mean number of difference of 3.151 ($\tau = 4.874$) was observed, the distribution being again bimodal. The mismatch distribution of the entire 'Alpine' clade (mean number of difference = 5.868; $\tau = 10.000$) showed three clear peaks. A model of sudden demographic expansion was supported for all analysed groups. The mismatch distribution including all lineages was multimodal, with one mode corresponding to the number of differences between the two major groups ('Italian', 'Alpine'; ~50 mutations) and the other to differences within lineages. The estimated timing of the most important demographic

expansion in each group, based on a mutation rate of 1% per million years, is depicted in Table 4.

Discussion

Phylogenetic reconstruction

For *Leuciscus souffia*, up to three subspecies have been described, *L. s. souffia* from the Var, *L. s. agassi* from the Rhône, Rhine and Danube basins, and *L. s. muticellus* from Italy (reviewed in Kottelat 1997). In a study based on morphometric analyses, allozyme data and 16S rRNA sequences, Gilles *et al.* (1998) found no significant difference between the Var population and other populations in southern France and they suggested combining *L. s. souffia* and *L. s. agassi* to form just one subspecies *L. s. souffia*. Our phylogenetic analysis clearly demonstrated the existence of two main clades in *L. souffia*, which resembled the morphologically distinct subspecies *L. s. souffia* (+ *agassi*) from the Var, Rhône, Rhine and Danube basins ('Alpine'), and *L. s. muticellus* from Italy ('Italian'). Within the 'Alpine' clade, two distinct groups were recognized, the Rhône/Var subclade and the Danube/Rhine group. The five haplotypes detected in the Rhine drainage formed a

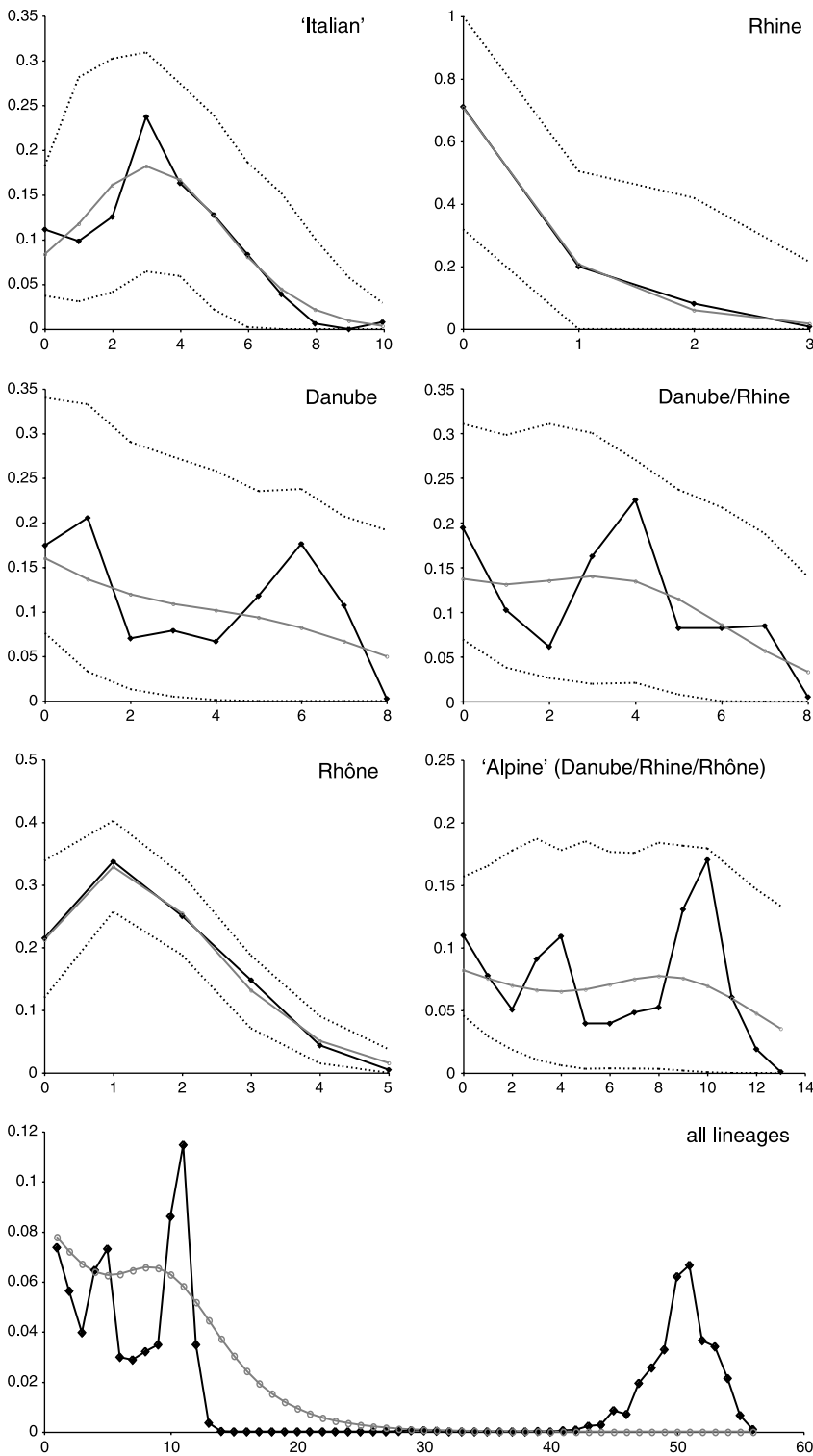


Fig. 4 Frequency distribution of pairwise mutational differences between individuals of *Leuciscus souffia* obtained by mismatch analyses. We analysed each subclade separately ('Italian', Rhine, Danube, Rhône/Var), and we also grouped the various haplotypes according to our phylogenetic analyses (Danube/Rhine, Danube/Rhine/Rhône) and analysed the entire data set. The black curve delineates the observed values, the grey curve is the model fitted to the data, and the dashed lines represent the 2.5 and 97.5 percentile values (1000 simulations). Relative numbers of pairwise comparisons are depicted on the *y*-axis, the number of mutational differences on the *x*-axis (see Table 4 for further details).

distinct clade within the Danube/Rhine group. The three groups identified by our analyses reflect the subspecies suggested by Banarescu (1992), with *L. s. souffia* from the Var and Rhône basins, *L. s. agassi* from the Rhine and Danube systems, and *L. s. muticellus* from Italy (see also Machordom *et al.*

1999). In our phylogeny, also *L. turskyi* from Vrba river in Dalmatia (Croatia) was placed in close relationship to the two main lineages of *L. souffia*. Furthermore, the linearized tree indicated that the split between *L. turskyi* and the 'Alpine' and 'Italian' clade of *L. souffia* occurred contemporaneously.

The timing of cladogenetic events in *L. souffia*

The application of a molecular clock for the estimation of divergence time is problematic for several reasons (see, e.g. Gillespie 1991; Page & Holmes 1998). However, a molecular-clock-based time estimate provides an approximate time frame for evaluating phylogeographic hypotheses. Several molecular clock calibrations exist for the mitochondrial DNA of fishes ranging from less than 1% substitutions per million years for the cytochrome *b* gene of cyprinids (Zardoya & Doadrio 1999), 3.6% in the control region of percoids (Donaldson & Wilson 1999), to more than 6% in the 5'-end of the control region of cichlids (Sturmbauer *et al.* 2001). For freshwater fishes of the northern hemisphere a rate between 1 and 2% substitutions per million years is generally applied in the control region (Nesbø *et al.* 1999; Englbrecht *et al.* 2000; Bernatchez 2001; Koskinen *et al.* 2002). For our data set we applied both, a 1% substitution rate per million years as well as a 2% rate. However, a substitution rate of 1% per million years seemed more adequate for our data, since the obtained results more closely matched previous calibrations based on other genes (Gilles *et al.* 1998; Zardoya & Doadrio 1999).

We applied a linearized tree approach to estimate divergence times between lineages. We translated the average pairwise distances between groups indicated in the linearized tree into divergence times. For the dating of the divergence of the three major clades, 'Alpine', 'Italian' and *L. turskyi* we thus obtained a divergence time of 5.29 ± 0.59 million years (Myr) (5.09 ± 0.35 Myr without *L. turskyi*; see Table 2) when applying the 1% rate. Gilles *et al.* (1998) already connected the split between *L. s. souffia* ('Alpine') and *L. s. muticellus* ('Italian') to the Messinian salinity crisis (~5 Myr ago) using 16S rRNA sequences. The divergence time of roughly 5 Myr is also consistent with the cytochrome *b* based estimates of Zardoya & Doadrio (1999) for the split between *L. (Telestes in their study) souffia* and its sister species *T. beoticus* and *T. pleurobipunctatus*. A rate of 2% would translate into a divergence time of 2.64 ± 0.29 Myr (2.54 ± 0.17 Myr without *L. turskyi*). For the split between the Danube/Rhine populations and the Rhône/Var group we estimated 969 ± 149 thousand years (kyr) (1% rate) or 484 ± 74 kyr (2% rate). We also determined the timing of the most important demographic expansion in different groups, based on the mismatch analyses (Table 4). The timing of the demographic expansion does not necessarily match the age estimates based on average pairwise distances, which are more likely to resemble maximum ages of lineages. While for the 'Italian' and the 'Alpine' clade, the Danube/Rhine haplotypes, and the Rhine populations the divergence times based on average pairwise comparisons (see Table 2) closely matched to the respective timing of the most important demographic expansion (see Table 4), the Rhône/Var group seemed to have undergone

demographic expansion at a later stage (147 kyr) compared to its maximum age (328 kyr). Note that the 95% consistency intervals around the expansion parameter τ generally lead to larger categories of time estimates in the mismatch analysis based approach in comparison to the standard deviations of the average pairwise distance method.

Glacial refugia and colonization routes for *L. souffia*

The postglacial re-colonization of the European biota has been extensively studied and for a series of plant and animal species the glacial refugia as well as the re-colonization routes have been identified based on molecular data (reviewed in Taberlet *et al.* 1998; Hewitt 1999, 2000). The three main refugial habitats during Quaternary cold periods in Europe were the Iberian peninsula, Italy as well the Balkans and Greece; almost all studied taxa used these three potential refugia (see Taberlet *et al.* 1998; Hewitt 1999). While for refugia in the Balkans and Greece no major barrier existed for their northward colonization, the Pyrenees and especially the Alps posed strong barriers to the expansion of refugial populations from the Iberian peninsula and Italy (Hewitt 1999).

For the vairone, the long branches leading to both the 'Italian' and the 'Alpine' clade suggest two major glacial refugia. For the 'Italian' clade, a refugium in the Italian peninsula seems highly likely, since all haplotypes assigned to the 'Italian' clade were restricted to the area south of the Alps. The refugium for the 'Alpine' clade cannot unambiguously be derived from our data. However, it seems plausible that this second refugium was located in the lower Danube drainage in the Balkans. This is supported by the greater diversity of the Danube population compared to its sister group in the Rhône and Var basins. An Iberian refugium for *L. souffia* is out of question, since the westernmost distribution of this species is the Rhône drainage. While for the 'Italian' clade the Alps seem to have been a strong enough barrier preventing northward expansion, the 'Alpine' group spread over a large area and is presently distributed from the Balkans over the whole Alpine area to the Mediterranean Alps. The most likely scenario is an expansion from an initial refugium in the lower Danube over the Danube system to Rhine and Rhône/Var. Based on our estimates the split between the Danube/Rhine and the Rhône/Var group happened about 0.5–1 Myr ago (Ma). Thus, two refugia must have existed in the 'Alpine' clade for the subsequent glacial maxima, one in the lower Danube system and the second southwest of the Alps for the present day Rhône system and Var populations. Also, *L. turskyi* survived several glacial maxima in an isolated basin comparable to the Var and at roughly the same geographical latitude.

A clear-cut phylogeographic pattern in the Danube and Rhine populations can be derived from the unrooted network tree (Fig. 3b). In the Danube system, common haplotypes between populations are rare and only occur between

adjacent sample localities. In general, only haplotypes of adjacent populations are connected. The most abundant Rhine haplotype (L03), which occurs in all populations, is connected to the central Danube haplotype (L01) via a haplotype found in the Main river only (L32), suggesting a colonization of the Rhine basin from North of Lake Constance. Our data demonstrate that the Rhine basin was the last river to be founded by *L. souffia*. We thus propose an initial spread of vairone populations via a connection from north of the Alps to the Mediterranean Alps in the early Pleistocene and a second late Pleistocene expansion when populations of the Danube system colonized the Rhine basin. In our phylogenies and network trees it is also indicated that the Rhine was founded by one relatively small vairone population, since all Rhine haplotypes shared the diagnostic C-T mutation in position 197 (Fig. 3b). The Rhine system could have been founded by vairone via the past drain of the Main into the Danube system. However, the connection existed only before 0.6 Ma (Hantke 1993). Our molecular clock estimates assume a more recent expansion of *L. souffia* into the Rhine, possibly via periglacial lakes or rivers during deglaciation. Other populations in formerly glaciated regions (e.g. populations 11, 21, 22) were founded by more than one haplotype. Although the Rhône/Var group is older and hence more diverse than the Rhine group, both lineages underwent the most important demographic expansion at roughly the same time (147 000 vs. 177 000 years ago), which seems to be correlated to the last interglacial maximum.

Aside from the fact that populations of the northernmost drainage system, the Rhine, show a smaller genetic divergence, there is no indication that populations of formerly glaciated areas show lower levels of genetic diversity than populations in ice-free areas (Bernatchez & Wilson 1998). This might be explained by the habitat of the vairone and the generally high degree of isolation for all populations and especially for the drainage systems as is reflected by the large global Φ_{ST} value (Table 3). The vairone occurs throughout its distribution range in smaller rivers or the upper reaches of larger river systems around the Alps. This habitat type, also inhabited by brown trout, grayling and bullhead, might have shifted up or down along the course of the rivers during glaciation cycles but it might not have changed dramatically during cold periods. The refugia during the last glaciation cycle(s) might thus have been located in the respective basin, not too far away from the present day locations. The Enns population (10) and the Sava and Cosizza (7, 8) fish, for example, do not share common haplotypes, suggesting that they have not been connected recently and that two founder populations outlived at least the last glacial maximum in two different branches of the Danube system. The gradual movement of the populations alongside the courses of the rivers according to the habitat changes during glaciation might thus have shifted the genetic diversity from past refugia to present

habitats. Then, no postglacial dispersal over longer distances occurred, which might explain why the genetic diversity is similar in all present populations.

Southern richness due to sea level low-stand?

In many studies it has been recognized that southern populations reveal greater diversity in species or subspecies number and genetic variation than northern ones (see Hewitt 1999 and references therein). In our sample, the northernmost populations from the Rhine basin displayed the lowest diversity, in terms of both haplotype number and haplotype diversity. Shared haplotypes among populations were rare in the Danube basin. However, shared haplotypes were common, in the Rhône / Var basins – especially in the three more southern localities, Chassezac, Durance and Var (localities 20–22; Fig. 3b) – and in the ‘Italian’ clade (Fig. 3a). The existence of shared haplotypes between distant populations, e.g. haplotype L06 in localities 18, 19, 21, 22 or haplotype L18 in localities 2 and 6, suggests a relatively recent connection between Var and Rhône as well as the Po and Soca basins. The occurrence of haplotype L13 in localities 4, 6, 22 suggests a recent connection between the Var basin and a drainage inhabited by *L. s. muticellus* (‘Italian’). The estuaries of Rhône and Var as well as of Po and Soca are distant at present sea level, which makes it impossible for a freshwater species like the vairone to cross these distances through the open sea. However, sea level was considerably lower during glacial maxima. For the last glacial maximum between 22 000 and 19 000 years ago sea level was about 120 m lower than it is today (Rohling *et al.* 1998; Yokoyama *et al.* 2000). Such a lowstand associated with changes in the course of rivers and possibly their confluence might have permitted exchange between river systems and thus explain the observed haplotype distribution in the southern populations of *L. souffia*. Another explanation for postglacial contact could be the exchange of individuals via periglacial rivers or melt-water lakes during deglaciation. Such a scenario has, for example, been proposed for the postglacial expansion of the European perch (Nesbø *et al.* 1999). A human-induced faunal translocation seems highly unlikely, given the large distances between populations with shared haplotypes and the relative rarity of *L. souffia*.

Two contact zones exist southwest and southeast of the Alps

Another interesting aspect of postglacial re-colonization is the occurrence of hybrid zones at areas where different lineages meet after expansion. As a result of the colonization of the North through populations from the South and the existence of barriers to expansion, major contact zones between different lineages in Europe are situated in the

Pyrenees between Spanish and French taxa, in the Alps between Italian and Northern biota, and in Central Europe for Western (Iberian) and Eastern (Balkans) lineages (Hewitt 1999, 2000). In an extensive study of *L. souffia* populations of the Mediterranean Alps in the border region between France & Italy, Gilles *et al.* (1998) already found introgression of *L. s. muticellus* ('Italian') genomes into the *L. s. souffia* population of the Var basin. In the small river Bevera, the same authors also found both kinds of haplotypes. However, the *L. s. muticellus* genomes prevailed in the Bevera sample and these fish were morphologically identified as *L. s. muticellus*. In our independent sampling of the Var population, we confirmed the results of Gilles *et al.* (1998) in that we found one 'Italian' haplotype in the Var population (haplotype L13). In the present study, we identified a second contact zone between two different lineages of the vairone in the border region between Italy and Slovenia. In the sample of the small river Cosizze, which is connected to rivers Iudrio, Natisone and Torre (localities 4–6; Soca basin), we found haplotypes identical or derived from those of the Jevscek population (locality 8) of the Danube drainage. Both areas, the Mediterranean Alps and the Italian/Slovenian border region, are actually the only possible contact zones between the 'Italian' and the 'Alpine' lineages, since the water sheds of the Alps are obviously a too strong barrier for dispersal of freshwater fish. It is interesting, however, that hybridization occurred between the 'Italian' and both subclades of the 'Alpine' lineage in the westernmost and easternmost parts of the distribution range of the 'Italian' haplotypes.

European freshwater fish phylogeography

In the past few years, a series of studies on the phylogeography of European freshwater fish have been published (Bernatchez *et al.* 1992; Bernatchez & Osinov 1995; Hänfling & Brandl 1998a,b; Refseth *et al.* 1998; Durand *et al.* 1999; Nesbø *et al.* 1999; Englbrecht *et al.* 2000; Aurelle & Berrebi 2001; Bernatchez 2001; Weiss *et al.* 2001, 2002; Hänfling *et al.* 2002; Weiss *et al.* 2002). One of the best studied species is the brown trout, for which an impressive data set exists (Bernatchez *et al.* 1992; Bernatchez & Osinov 1995; Aurelle & Berrebi 2001; Bernatchez 2001; Weiss *et al.* 2001; Duftner *et al.* 2003). Five major lineages of brown trout have been identified in Europe, with a nucleotide divergence among the groups between 1.21 and 2.19% in 1250 nucleotides of mtDNA including a segment of the control region (Bernatchez 2001). While the dating for the expansion of the Danube lineage lies in the range of the Danube/Rhine clade expansion in *L. souffia*, the remaining trout lineages seem to have expanded more recently than the *L. souffia* lineages in the respective basins. This is especially the case for the *marmoratus* lineage of brown trout, which averages the distribution of the 'Italian' clade of *L. souffia* (*muticellus*).

Another well studied salmonid species is the grayling *Thymallus thymallus* for which up to eight different lineages have been identified. Three distinct lineages were resolved in Koskinen *et al.* (2000) and eight in Weiss *et al.* (2002). In the phylogeny of Weiss *et al.* (2002), two Danube lineages were placed at the basis, sister to the clade containing the remaining haplotypes. Within this clade, an ancestral 'Adriatic' clade was resolved (with a similar distribution to the 'Italian' clade of the present study), sister to a 'Mixed Rhine/Danube' clade and a cluster comprised by a 'Mixed Central European' clade, the 'Scandinavian' clade and the 'Slovenian Danube system (Sava)' clade. While in our study such ancestral Danube lineages are not present, the branching order of the remaining lineages in *Thymallus thymallus* is similar to that of *L. souffia*, with an ancestral Adriatic ('Italian' in our study) clade sister to a Central European one ('Alpine'). In contrast to the present results, the structuring within the Central European clade of grayling is not that clear-cut. For example, two groups of haplotypes have been identified in the Rhine drainage, one from a 'Mixed Rhine/Danube' clade (comparable to our Danube/Rhine group) and another one from the 'Mixed Central Europe' clade. The genetic divergence among the main lineages in the grayling (4%; Weiss *et al.* 2002) is similar to that of the vairone.

With the chub (*Leuciscus cephalus*) also a cyprinid was studied (Hänfling & Brandl 1998b; Durand *et al.* 1999). Durand *et al.* (1999) suggested the existence of four refugial habitats in Europe, from which lineages of pre-Pleistocene origin recently colonized Central Europe. They identified two ancestral lineages in the Aegaeis and in Eastern Europe, sister to an Adriatic lineage (with representatives from Greece, Albania and Italy) and the Western lineage. Interestingly, Durand *et al.* (1999) found a close relatedness between Rhône and Rhine haplotypes (see also the 'Mixed Central Europe' clade of grayling in Weiss *et al.* 2002), which is in clear contrast to the affiliation of the Rhine system haplotypes to the Danube clade in *L. souffia* (see also the 'Mixed Rhine/Danube' lineage of grayling in Weiss *et al.* 2002).

Four lineages were also identified for the European perch (Nesbø *et al.* 1999), with a southern European lineage sister to a North European, an Eastern European and a Western European clade. Nesbø *et al.* (1999) suggested a refugia north of the Alpine ice-shield from which all Central European lineages re-colonized their present habitats after the last glacial maximum.

A pre-Pleistocene origin of major Central European populations was suggested for the bullhead *Cottus gobio* (Englbrecht *et al.* 2000 based on mtDNA, Hänfling *et al.* 2002 based on microsatellites). Like the vairone, the bullhead is a more cold-water-adapted species, not being subject to faunal translocation or stocking, and both species are sympatric in most parts of their overlapping distribution. One

haplotype lineage in the bullhead (group I of Englbrecht *et al.* 2000) has a similar distribution to the 'Alpine' clade of *L. souffia*, although it reaches north to Scandinavia. With $0.75 \pm 0.36\%$ the average pairwise distance within this group resembles the results found for the 'Alpine' group of our study. Between the major lineages, an average pairwise distance of up to 4.3% has been found for the bullhead (Englbrecht *et al.* 2000), which again is similar to our results (see also Weiss *et al.* 2002 for the grayling). An interesting pattern can be found for the Rhine system, where Englbrecht *et al.* (2000) found two distinct lineages in the upper and lower Rhine, which are substantially older than the vairone lineage of the Rhine basin. Hänfling *et al.* (2002) showed that northern bullhead populations most probably survived the last glaciation cycle on the British Isles and in several river drainages. However, phylogeographic studies of the bullhead never included a large enough sample from the Alpine region to allow a more detailed comparison to the present study.

Conclusion

The phylogeographic analysis of *Leuciscus souffia* populations in and around the Alps provides a clear scenario for the colonization history of Central Europe. Two main lineages exist, which colonized the Alps from the south ('Italian') or from the East, North and West ('Alpine'). Two hybrid zones exist where both lineages met, in the Mediterranean Alps and in the Soca basin. Major drainages like the Rhine and Rhône basins have been founded only once by *L. souffia*, even though several interglacial connections existed between these drainages and the Danube system, as is reflected in the phylogeography of other freshwater fish species (see e.g. Englbrecht *et al.* 2000; Weiss *et al.* 2002). The colonization history of *L. souffia* shows similarities to one clade of the grayling (Weiss *et al.* 2002). The 'Alpine' clade of *L. souffia* roughly resembles the situation found in the Danube clade of the bullhead, including the existence of several refugia for the last glaciation cycle(s) in different basins (Englbrecht *et al.* 2000; Hänfling *et al.* 2002). Such concordant phylogeographic patterns and similar age estimates suggest that the same geological and palaeoclimatological events shaped the European freshwater fish faunal distributions and population genetic structuring. We found no indications that populations in formerly glaciated areas display lower levels of genetic diversity. However, southern populations show increased numbers of shared haplotypes, suggesting recent connections between rivers like Var and Rhône or Po and Soca, possibly induced by a low sea-level stand during the last glacial maximum. The vairone is endangered in large parts of its distribution range. Therefore, however, this species has not been subject to faunal translocations or stocking, so that a clear-cut natural population structuring exists. Conservation efforts after renaturation of rivers,

which are clearly eligible to retain the present diversity, should bear this population structuring in mind, i.e. stocking with nonindigenous individuals should be prevented.

Acknowledgements

We thank the following colleagues who kindly provided samples for this study: B. Grabherr (Austria), U. Groß (Germany), M. Povz (Slovenia) and W. Hauer/Bundesanstalt für Wasserwirtschaft, Mondsee (Austria). We are especially grateful to E. Miotti (for organizing sampling in Italy and the Balkan area), to G. Pechlaner/Alpenzoo Innsbruck, and to N. Medgyesy for providing specimens, for discussions and for help during fieldwork. We thank E. Hespeler and H. Niederstätter for technical support. We also acknowledge M. Barluenga, S. Weiss, three anonymous reviewers, and L. Bernatchez (Subject Editor) for valuable comments on the manuscript. This study was supported by the University of Innsbruck ('Förderungsstipendium der Naturwissenschaftlichen Fakultät') to W.S., and grants of the Deutsche Forschungsgemeinschaft to A.M.

References

- Aurelle D, Berrebi P (2001) Genetic structure of brown trout (*Salmo trutta*, L.) populations from south-western France: data from mitochondrial control region variability. *Molecular Ecology*, **10**, 1551–1561.
- Avice JC (1994). *Molecular Markers, Natural History and Evolution*. Chapman & Hall, New York.
- Avice JC (1998) The history and purview of phylogeography: a personal reflection. *Molecular Ecology*, **7**, 371–379.
- Avice JC (2000) *Phylogeography – the History and Formation of Species*. Harvard University Press, Cambridge.
- Avice JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 459–522.
- Banareescu P (1992) *Zoogeography of Freshwaters, 2. Distribution and Dispersal of Freshwater Animals in North America and Eurasia*. Aula, Wiesbaden, Germany.
- Bernatchez L (2001) The evolutionary history of brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analyses of mitochondrial DNA variation. *Evolution*, **55**, 351–379.
- Bernatchez L, Osinov A (1995) Genetic diversity of trout (genus *Salmo*) from its most eastern native range based on mitochondrial DNA and nuclear gene variety. *Molecular Ecology*, **4**, 285–297.
- Bernatchez L, Wilson CW (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology*, **7**, 431–452.
- Bernatchez L, Guyomard R, Bonhomme F (1992) DNA sequence variation of the mitochondrial control region among geographically and morphologically remote European brown trout *Salmo trutta* populations. *Molecular Ecology*, **1**, 161–173.
- Bless R (1996) On the spawning behavior and the ecology of early life stages of the fourie (*Leuciscus souffia* Risso, 1826). *Fischökologie*, **10**, 1–10.
- Bruford MW, Hanotte O, Brookfield JFY, Burke T (1998) Multi-locus and single-locus DNA fingerprinting. In: *Molecular Genetic Analysis of Populations* (ed. Hoelzel AR), pp. 283–336. Oxford University Press, New York.

- Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.
- Crandall KA, Bininde-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution*, **15**, 290–295.
- Dawson AG (1992) *Ice Age Earth*. Routledge Press, London.
- Donaldson KA, Wilson RR (1999) Amphipanamic geminates of snook (Percoidae: Centropomidae) provide a calibration of the divergence rate in the mitochondrial DNA control region of fishes. *Molecular Phylogenetics and Evolution*, **13**, 208–213.
- Dowling TE, Tibbets CA, Minckley WL, Smith GR (2002) Evolutionary relationships of the plagiopeterns (Teleostei: Cyprinidae) from cytochrome *b* sequences. *Copeia*, **2002**, 665–678.
- Duftner N, Weiss S, Medgyesy N, Sturmbauer C (2003) Enhanced phylogeographic information about Austrian brown trout populations derived from complete mitochondrial control region sequences. *Journal of Fish Biology*, **62**, 427–435.
- Durand JD, Persat H, Bouvet Y (1999) Phylogeography and post-glacial dispersion of the chub (*Leuciscus cephalus*) in Europe. *Molecular Ecology*, **8**, 989–997.
- Englbrecht CC, Freyhoff J, Nolte A, Rassmann K, Schliewen U, Tautz D (2000) Phylogeography of the bullhead *Cottus gobio* (Pisces: Teleostei: Cottidae) suggests a pre-Pleistocene origin of the major central European populations. *Molecular Ecology*, **9**, 709–722.
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.
- Fraser DJ, Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular Ecology*, **10**, 2741–2752.
- Gilles A, Chappaz R, Cavalli L, Lörtscher M, Faure E (1998) Genetic differentiation and introgression between putative subspecies of *Leuciscus souffia* (Teleostei: Cyprinidae) of the region of the Mediterranean Alps. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 2341–2354.
- Gilles A, Lecointre G, Miquelias A, Loertscher M, Chappaz R, Brun G (2001) Partial combination applied to phylogeny of European cyprinids using mitochondrial control region. *Molecular Phylogenetics and Evolution*, **19**, 22–33.
- Gillespie JH (1991) *The Causes of Molecular Evolution*. Oxford University Press, Oxford.
- Hänfling B, Brandl R (1998a) Genetic variability, population size and isolation of distinct populations in the freshwater fish *Cottus gobio* L. *Molecular Ecology*, **7**, 1625–1632.
- Hänfling B, Brandl R (1998b) Genetic and morphological variation in a common European cyprinid, *Leuciscus cephalus*, within and across Central European drainages. *Journal of Fish Biology*, **52**, 706–715.
- Hänfling B, Hellemans B, Volckaert FAM, Carvalho GR (2002) Late glacial history of the cold-adapted freshwater fish *Cottus gobio*, revealed by microsatellites. *Molecular Ecology*, **11**, 1717–1729.
- Hantke R (1993) *Flußgeschichte Mitteleuropas*. Ferdinand Enke-Verlag, Stuttgart.
- Harpending H (1994) Signature of ancient population growth in a low resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591–600.
- Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.
- Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.
- Kluge AG, Farris JS (1969) Quantitative phyletics and the evolution of anurans. *Systematic Zoology*, **18**, 1–32.
- Koskinen MT, Ranta E, Piironen J, et al. (2000) Genetic lineages and postglacial colonization of grayling (*Thymallus thymallus*, Salmonidae) in Europe, as revealed by mitochondrial DNA analyses. *Molecular Ecology*, **9**, 1609–1624.
- Koskinen MT, Knizhin I, Primmer CR, Schlötterer C, Weiss S (2002) Mitochondrial and nuclear DNA phylogeography of *Thymallus* spp. (grayling) provides evidence of ice-age mediated environmental perturbations in the world's oldest body of fresh water, Lake Baikal. *Molecular Ecology*, **11**, 2599–2611.
- Kottelat M (1997) European freshwater fish. *Biologia*, **52** (suppl. 5), 1–271.
- Machordom A, Nicolas Y, Crivelli AJ (1999) Genetic variability and differentiation in *Leuciscus (Telestes) souffia*. Taxonomic and conservation inferences. *Les Comptes Rendus de l'Académie Des Sciences Paris Série Biologies III-Vie*, **322**, 15–28.
- McKusker MR, Parkinson E, Taylor EB (2000) Mitochondrial DNA variation in rainbow trout (*Oncorhynchus mykiss*) across its native range: testing biogeographical hypotheses and their relevance to conservation. *Molecular Ecology*, **9**, 2089–2108.
- Meyer A, Morrissey JM, Schartl M (1994) Recurrent origin of a sexually selected trait in *Xiphophorus* fishes inferred from a molecular phylogeny. *Nature*, **368**, 539–542.
- Moritz C (1994) Defining 'evolutionary significant units' for conservation. *Trends in Ecology and Evolution*, **9**, 373–375.
- Nesbø CL, Fossheim T, Vøllestad LA, Jakobsen KS (1999) Genetic divergence and phylogeographic relationships among European perch (*Perca fluviatilis*) populations reflect glacial refugia and postglacial colonization. *Molecular Ecology*, **8**, 1387–1404.
- Page RDM, Holmes EC (1998) *Molecular Evolution: a Phylogenetic Approach*. Blackwell Sciences, Oxford.
- Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.
- Refseth UH, Nesbø CL, Vøllestad LA, Fjeld E, Stacy JE, Jakobsen KS (1998) Genetic evidence of different migration routes of freshwater fish into Norway revealed by analysis of current perch (*Perca fluviatilis*) populations in Scandinavia. *Molecular Ecology*, **7**, 1015–1027.
- Rohling EJ, Fenton M, Jorissen FJ, Bertrand P, Ganssen G, Caulet JP (1998) Magnitude of sea-level lowstands of the past 500 000 years. *Nature*, **394**, 162–165.
- Schluter D (1997) Ecological speciation in postglacial fishes. In: *Evolution on Islands* (ed. Grant PR), pp. 114–129. Oxford University Press, Oxford.
- Schluter D, McPhail JD (1993) Character displacement and replicate adaptive radiation. *Trends in Ecology and Evolution*, **8**, 197–200.
- Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079–1089.
- Schneider S, Roessli D, Excoffier L (1999) *Arlequin, Version 2.0: a software for genetic data analysis*. Genetics and Biometry Laboratory, University of Geneva, Geneva.
- Smith CT, Nelson RJ, Wood CC, Koop BF (2001) Glacial biogeography of North American coho salmon (*Oncorhynchus kisutch*). *Molecular Ecology*, **10**, 2775–2785.
- Sturmbauer C, Baric S, Salzburger W, Rüber L, Verheyen E (2001) Lake level fluctuations synchronize genetic divergences of

- cichlid fishes in African lakes. *Molecular Biology and Evolution*, **18**, 144–154.
- Swofford DL (2002) *PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods)*. Sinauer Associates, Sunderland.
- Taberlet P, Fumagalli L, Wust-Saucy AG, Cossons JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.
- Takezaki N, Rzhetsky A, Nei M (1995) Phylogenetic test of the molecular clock and linearized trees. *Molecular Biology and Evolution*, **12**, 823–833.
- Wang HY, Tsai MP, Yu MJ, Lee SC (1999) Influence of glaciation on divergence patterns of the endemic minnow, *Zacco pachycephalus*, in Taiwan. *Molecular Ecology*, **8**, 1879–1888.
- Webb T, Bartlein PJ (1992) Global change during the last 3 million years: climatic controls and biotic responses. *Annual Reviews in Ecology and Systematics*, **23**, 141–173.
- Weiss S, Schlötterer C, Waidbacher H, Jungwirth M (2001) Haplotype (mtDNA) diversity of brown trout *Salmo trutta* in tributaries of the Austrian Danube: massive introgression of Atlantic basin fish – by man or nature? *Molecular Ecology*, **10**, 1241–1246.
- Weiss S, Persat H, Eppe R, Schlötterer C, Uiblein F (2002) Complex patterns of colonization and refugia revealed for European grayling *Thymallus thymallus*, based on complete sequencing of the mitochondrial DNA control region. *Molecular Ecology*, **11**, 2599–2611.
- Yokoyama Y, Lambeck K, De Deckker P, Johnston P, Fifield LK (2000) Timing of the last glacial maximum from observed sea-level minima. *Nature*, **406**, 713–716.
- Zardoya R, Doadrio I (1998) Phylogenetic relationships of Iberian cyprinids: systematic and biogeographical implications. *Proceedings of the Royal Society London Series B*, **265**, 1365–1372.
- Zardoya R, Doadrio I (1999) Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. *Journal of Molecular Evolution*, **49**, 227–237.
- Zink RM, Slowinsky JB (1995) Evidence from molecular systematics for decreased avian diversification in the Pleistocene epoch. *Proceedings of the National Academy USA*, **92**, 5832–5835.

The research of Walter Salzburger in the lab of Axel Meyer focuses on the understanding of speciation and adaptive radiation (particularly in East African cichlids). Melanie Hempel was sequencing some vairone-populations as part of an undergraduate course in the Meyer-lab. Anita Brandstätter is a PhD student in the group of Walther Parson, where part of the sequence data have been produced. André Gilles is interested in the phylogeny and phylogeography of European freshwater fish. Christian Sturmbauer and his group work on phylogenetic and population genetic questions in European taxa and Lake Tanganyika cichlids.
