Phylogeny and biogeography of the genus *Pseudobarbus* (Cyprinidae): Shedding light on the drainage history of rivers associated with the Cape Floristic Region

Ernst R. Swartz a,b,*, Paul H. Skelton a, Paulette Bloomer b

a South African Institute for Aquatic Biodiversity, Private Bag 1015, Grahamstown 6140, South Africa
b Molecular Ecology and Evolution Programme, Department of Genetics, University of Pretoria, Pretoria 0002, South Africa

**A R T I C L E   I N F O**

**Article history:**
Received 14 December 2007
Revised 29 September 2008
Accepted 7 October 2008
Available online xxxx

**Keywords:**
Biogeography
Phylogenetics
Drainage evolution
*Pseudobarbus*
Cyprinidae
South Africa
Lesotho
Cape Floristic Region

**A B S T R A C T**

Relationships among the historically isolated lineages of *Pseudobarbus* were reconstructed using molecular and morphological data. Contradictions between the molecular and morphological phylogenies suggest convergent evolution and homoplasy in some morphological characters. The earliest divergence in *Pseudobarbus* was between *P. quathlambae* in Lesotho and the rest of the genus associated with the Cape Floristic Region in South Africa. A close relationship between *P. phlegethon* from the Olifants River system on the west coast of South Africa and a lineage of *P. afer* from small river systems in Afrotemperate Forests on the south coast, can only be explained through previous occurrence and subsequent extinction of ancestral populations in the Gourits River system. Several river systems had confluences before reaching lower sea levels, most notably during the last glacial maximum about 18,000 years ago, explaining closely related populations across different river systems. Mainly river capture explains shared lineages across river systems that did not share a common confluence during lower sea levels.

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

The Cape Floristic Region (CFR) is the only floral kingdom that is located entirely within the geographical confines of one country and known for its high levels of plant diversity (Goldblatt and Manning, 2002). In contrast, the freshwater fish diversity is low (Skelton, 1986), despite the region being characterised by many river systems of various sizes. The evolutionary history of primary freshwater fish often reflects drainage evolution, due to limited opportunities for dispersal between different river systems (Brito et al., 1997; Waters and Wallis, 2000; Mesquita et al., 2001). However, since there are only two to 10 primary freshwater fish species per river system and only three widespread genera (*Pseudobarbus*, *Galaxias* and *Sandelia*) across the CFR, prospects of using freshwater fish to study the biogeography of this region’s rivers are limited. Of the three widespread CFR fish genera, *Pseudobarbus* is the most species rich with six currently described species occurring in the CFR and one in Lesotho compared to one species each for *Galaxias* and *Sandelia*. The distribution of the *Pseudobarbus* species is relatively well known and is also the freshwater genus in the CFR that has been revised taxonomically most recently (Skelton, 1988), making it the ideal group with which to investigate the drainage evolution of the CFR.

*Pseudobarbus* occur in all the major river systems and most of the smaller river systems of the CFR and in the neighbouring Orange River system. The latter river system has its origins in the highlands of Lesotho, is the largest river system in South Africa and once shared a mouth with the Olifants River system in the north-western CFR (De Wit, 1993). The Olifants, Berg, Breede, Gourits, Gamtoos and Sundays River systems are the larger systems in and bordering the CFR and drain from interior regions of the Western Cape and Eastern Cape provinces associated with the semi-arid Karoo region. The Keurbooms and Swartkops River systems penetrate coastal mountain ranges, but originate within the CFR. There are also several small coastal river systems that do not penetrate coastal mountain ranges, especially along the southern coastline (Fig. 1).

Several river captures are evident from the current drainage patterns of rivers in the CFR. During lower sea levels most of the rivers of the CFR would have had common confluences before reaching the sea, forming palaeoriver systems that are currently drowned. During the last glacial maximum (LGM) about 18,000 years ago, the sea level along the south coast of the CFR was about 130 m below present levels (Tankard, 1976; Rogers, 1985; Ramsay and Cooper, 2002). Probably, only about nine to 10 major palaeoriver systems that have not been completely drowned in subsequent
Pseudobarbus originate from the southern African region with historical links to the rest of the African continent. *Pseudobarbus* species are all tetraploid (Naran et al., 2006) and closely related to the tetraploid ‘*Barbus*’ species that are endemic to southern Africa (Machordom and Doadrio, 2001b; Tsigenopoulos et al., 2002). *Pseudobarbus* and the southern African tetraploid ‘*Barbus*’ are the sister group to the pan-African diploid ‘*Barbus*’ (Machordom and Doadrio, 2001b; Tsigenopoulos et al., 2002). When Skelton (1980, 1988) reviewed the taxonomy and established morphological relationships within *Pseudobarbus*, he found that several morphological characters are shared between the Maloti Minnow (*P. quathlambae*) in the upper reaches of the Orange River system in Lesotho and *P. tenuis* from southern parts of the CFR (Gourits and Keurbooms River systems), which suggested a sister species relationship. The latter was surprising given the large geographic distance between these two species. Based on the morphological characters, *P. phlegethon* from the Olifants River system in the western parts of the CFR was inferred as the sister species to *P. quathlambae* and *P. tenuis*. *Pseudobarbus afer*, which is the most widespread redfin species (southern CFR), and the Karoo adapted sea level transgression, based on reconstructions by Swartz et al. (2007 and unpublished), are relevant to the biogeography of *Pseudobarbus* (Figs. 1 and 6). It cannot, however, be assumed that known river captures, confluence of rivers during lower sea levels or other climatic and geological events influenced the evolution of *Pseudobarbus* without testing congruence with the latter’s population history.

The *Pseudobarbus* species are all tetraploid (Naran et al., 2006) and closely related to the tetraploid ‘*Barbus*’ species that are endemic to southern Africa (Machordom and Doadrio, 2001b; Tsigenopoulos et al., 2002). *Pseudobarbus* and the southern African tetraploid ‘*Barbus*’ are the sister group to the pan-African diploid ‘*Barbus*’ (Machordom and Doadrio, 2001b; Tsigenopoulos et al., 2002). When Skelton (1980, 1988) reviewed the taxonomy and established morphological relationships within *Pseudobarbus*, he found that several morphological characters are shared between the Maloti Minnow (*P. quathlambae*) in the upper reaches of the Orange River system in Lesotho and *P. tenuis* from southern parts of the CFR (Gourits and Keurbooms River systems), which suggested a sister species relationship. The latter was surprising given the large geographic distance between these two species. Based on the morphological characters, *P. phlegethon* from the Olifants River system in the western parts of the CFR was inferred as the sister species to *P. quathlambae* and *P. tenuis*. *Pseudobarbus afer*, which is the most widespread redfin species (southern CFR), and the Karoo adapted to the tetraploid ‘*Barbus*’ for the large geographic distance between these two species. Skelton (1980) suggested that an ancestor of the latter species must have occurred in Karoo tributaries of the Orange River system (north of the CFR in Fig. 1). These populations would have given rise to *P. tenuis* in the Gourits River system where they now occur in sympatry with *P. asper*. Sympatry between *P. tenuis* and *P. asper* was therefore likely the result of secondary contact. Redfins do not currently occur in the area of the Orange River system where Skelton (1980) proposed an exchange between the Orange and Gourits River systems. Extinction of the proposed ancestral populations in Karoo tributaries of the Orange River system therefore had to be assumed.

Skelton (1980, 1988) noted morphological variation between populations within species, but the differences were not clear or consistent enough to warrant full species status for these unique populations. However, mtDNA evidence suggests that 15 historically isolated lineages can be identified within the seven species that have only one pair of barbels.

These morphological relationships were easily explained through differentiation between adjacent river systems, except for the large geographic distance between *P. quathlambae* and *P. tenuis*. Skelton (1980) suggested that an ancestor of the latter species must have occurred in Karoo tributaries of the Orange River system (north of the CFR in Fig. 1). These populations would have given rise to *P. tenuis* in the Gourits River system where they now occur in sympathy with *P. asper*. Sympatry between *P. tenuis* and *P. asper* was therefore likely the result of secondary contact. Redfins do not currently occur in the area of the Orange River system where Skelton (1980) proposed an exchange between the Orange and Gourits River systems. Extinction of the proposed ancestral populations in Karoo tributaries of the Orange River system therefore had to be assumed.

Skelton (1980, 1988) noted morphological variation between populations within species, but the differences were not clear or consistent enough to warrant full species status for these unique populations. However, mtDNA evidence suggests that 15 historically isolated lineages can be identified within the seven
Pseudobarbus species (Fig. 1): Mohale and Eastern lineages were identified within P. quathlambae (Swartz et al., unpublished), four historically isolated lineages (Forest, Krom, St. Francis and Mandel-a) were found within P. afer (Swartz et al., 2007), Gourits and Keurbooms lineages were found within P. tenuis (Swartz et al., unpublished), P. burchelli consists of the Breede, Heuningnes and Tradou lineages (Swartz et al., unpublished) and Bloomer and Imp-son (2000) identified the Berg and Verlorenvlei lineages within P. burgi. Swartz et al. (2004) found fixed allelic differences at seven allozyme loci between the Olifants and Doring populations of P. phlegeton, but these differences were not as clearly reflected in the mtDNA sequences (Swartz et al., 2007). This discrepancy may reflect complex population histories within the Olifants River system, but for the purpose of the present study P. phlegeton will be referred to as a single lineage. Minor differentiation was also found within the Forest lineage of P. afer (Swartz et al., 2007) that could have important phylogeographic implications, but these lineages were not divergent enough to alter the topology of phylo- genetic reconstructions. Only one mutational step distinguished the Gamtoos populations of P. asper from some Gourits alleles of the same species, and therefore this species was also considered as a single lineage (Swartz et al., unpublished).

Following the documentation of these historically isolated lin- eages, intraspecific phylogeographic investigations suggested that the confluence of rivers during lower sea levels and river capture were the main routes for dispersal to explain current geographic distribution of genetic diversity (Swartz et al., 2007, unpublished). The wide distribution of the Forest, St. Francis and Mandela lin- eages of P. afer (Swartz et al., 2007) and the Breede lineage of P. burchelli (Swartz et al., unpublished) across currently isolated river systems (Fig. 1) were explained by confluence of river systems during the last glacial maximum (LGM) around 18,000 years ago, when the coastline was about ~130 m below current levels in some places (Tankard, 1976; Rogers, 1985; Ramsay and Cooper, 2002). There are also inland examples of dispersal, such as the occurrence of P. asper in both the Gourits and Gamtoos River sys- tems (Swartz et al., unpublished) that was explained by dispersal across low gradient areas between these two river systems, espe- cially during wet periods. However, isolation as an evolutionary process, has dominated genetic differentiation within Pseudobarbus species.

The intraspecific studies exposed the regional evolutionary pro- cesses and how it potentially linked to climatic and geological events, but a broader geographic context is needed to understand Pseudobarbus biogeography and how it is related to drainage evo- lution in the CFR. This can be partly accomplished by assessing the phylogenetic relationships among the extant historically iso- lated lineages across the CFR and in relation to the populations in Lesotho, since these lineages are more informative with regards to phylogeny reconstruction than the currently described species and are also more appropriate units for understanding the biogeog- raphy of the genus.

To achieve this, representatives of all 15 historically isolated lin- eages were subjected to morphological and mtDNA phylogenetic analyses to test hypotheses about the biogeography of the genus that have previously been based on morphology alone. We test whether phylogenetic relationships based on morphology mirror those based on genetics. Based on combined and separate morpho- logical and molecular phylogenetic analyses, we test whether Pseudobarbus speciation in the CFR have mainly been driven by iso- lation in neighbouring river systems, except for the relationship between P. quathlambae and P. tenuis where another explanation is required. For the latter relationship, we hypothesize that isolation occurred between the Orange and Gourits River system, with subsequent extinction of ancestral P. quathlambae/P. tenuis popula- tions in Karoo tributaries of the Orange River system to explain current distribution patterns. Relationships between newly discov- ered historically isolated lineages are established to infer bioge- ographic patterns that could lead to a better understanding of drainage evolution in the CFR.

2. Materials and methods

2.1. Sampling

Sequences and specimens from Bloomer and Impson (2000), Machordom and Doadrio (2001b), Tsigenopoulos et al. (2002), Swartz et al. (2004, 2007 and unpublished), under the GenBank Accession Nos. AF180848–AF180851 and AF287449–AF287454 were included in the study. Additional P. burgi, P. phlegeton and outgroup specimens were collected by snorkelling with a handnet or with a 3 m seine net. Representatives from all the extant south- ern African serrated tetraploid ‘Barbus’ species were included as outgroups. Whole fish samples were stored in liquid nitrogen in the field and transferred to a −70 °C freezer upon returning to the laboratory or muscle, fin-clips or whole fish samples were placed in 98% EtOH (Department of Genetics, University of Pretoria). The remaining carcasses and/or additional samples were fixed in formalin and deposited in the South African Institute for Aquatic Biodiversity (SAIAB) collection (Grahamstown) as voucher specimens. South African National Parks, Cape Nature, the Eastern Cape’s Department of Economic Affairs, Environment and Tourism approved collecting methods and granted permits. The University of Pretoria’s ethics committee and SAIAB also approved collecting methods.

2.2. DNA extraction, amplification and sequencing

Standard protocols of chemical digestion and phenol/chloroform extraction were used to isolate genomic DNA (Sambrook et al., 1989). Most of the 5’ end and part of the 3’ end of the mitochondrial DNA control region was amplified (PCR) with the primers L16S500 (5’-CCAAAGCAGAATTCTTAC-3’) and H677 (5’-GTGCGCAGAAACC AAAG-3’) (Swartz et al., 2007). The primers GluF (5’-AACACGGCTT GTATTCACATCAA-3’) and ThrR (5’-ACCTCGGTCCTCGATTACAG ACCG-3’) from Machordom and Doadrio (2001a) were used to amplify almost the entire mitochondrial cytochrome b gene. The widely used vertebrate primers 16Sar and 16Sbr (Palumbi et al., 1991) were used to amplify part of the 3’end of the 16S mtDNA ribosomal gene. Reagents (apart from the primers) and conditions for amplification, purification and cycle sequencing were the same for control region, cytochrome b and 16S. Amplification was performed in 50 μl volumes containing 1x buffer, 2 mM MgCl₂, 0.2 mM of each of the four nucleotides (Promega), 25 pmol of each primer, 1.5 U of Super-Therm DNA polymerase (ABGENE) and 100–200 ng of purified DNA as template. PCR conditions for amplification involved an initial denaturation of 2 min at 94 °C, followed by 35 cycles of 30 s at 94 °C, 30 s at 54 °C and 45 s at 72 °C and a final extension of 5 min at 72 °C.

PCR products were purified using the High Pure™ PCR Product Purification Kit (Boehringer Mannheim), followed by elution in ddH₂O. Cycle sequencing was performed in 10 μl volumes, with 2 μl ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Mix (Applied Biosystems), 1.6 pmol of a single primer (L16S500 or H677 for control region, GluF or ThrR for cytochrome b or 16Sar or 16Sbr for 16S) and 100 ng of purified DNA as template. PCR cycling and cycle sequencing were performed in a Geneamp® PCR System 9700 (Applied Biosystems). Nucleotide sequences were determined through ABI 377 or ABI 3100 automated sequencers. Consensus sequences were obtained from the forward and reverse sequences of each specimen and by comparing these to sequences from other individuals through alignment and inspection in...
Sequence Navigator 1.01 (Applied Biosystems). The consensus sequences were aligned using Clustal X (Thompson et al., 1997) and checked manually.

2.3. Morphology

External and internal morphological characters from Skelton (1980) were coded for comparison to the molecular analysis. Morphological data were not available for three of the historically isolated lineages, namely the Heuningnes and Tradou lineages of *P. burchelli* and the Mohale lineage of *P. quanthlambe*. There were also no clear morphological characters investigated by Skelton (1980) that distinguish between the *P. afer* or *P. tenuis* lineages. Characters were ordered according to character states in the outgroups *B. calidus* and *B. erubescens*.

2.4. Phylogenetic analysis

Control region, cytochrome *b* and 16S sequence datasets were first analysed separately and then were combined to assess the robustness of relationships across datasets. Congruence between these datasets were not rejected with a partition homogeneity test (*P = 0.410*) as implemented in PAUP* (Swofford, 2002). The morphological dataset was also analysed separately and finally, combined with the combined molecular dataset. Only representatives of the historically isolated *Pseudobarbus* lineages were used for more computationally intensive phylogenetic analyses. To assess the phylogenetic relationships among the historically isolated lineages of *Pseudobarbus*, neighbour joining (NJ; Saitou and Nei, 1987), maximum likelihood (ML; Felsenstein, 1981) and maximum parsimony (MP; based on “parsimony” of Hennig, 1966) analyses were performed in PAUP*. In addition, Bayesian analyses were performed in MrBayes 3.0b4 (Huelsenbeck and Ronquist, 2001).

The nucleotide substitution model that best fits the molecular data was selected from 56 such models with the Akaitek test in MODELTEST version 3.06 (Posada and Crandall, 1998). Ti/Tv ratio, proportion of invariable sites (I) and the a value of the gamma distribution (rate variation among sites) were also estimated using MODELTEST. NJ, ML and Bayesian analyses were based on these models. Starting trees for the ML analyses were obtained through the NJ method. The optimal tree was obtained through a heuristic search with 10 random sequence additions. Nodal support was assessed with 10,000 non-parametric bootstrap replicates (Felsenstein, 1985), each with 10 replicates of random taxon addition. For control region and 16S, single gaps were treated as a 5th character state. The sequence regions where adjacent gaps occurred were coded as different character states of a single character.

3. Results

3.1. Sequence analysis

The different datasets and the analyses to which they were subjected are summarised in Table 1. From Swartz et al. (2007, unpublished) and additional sequences in the present study, analysis of 259 individuals from 25 river systems and 102 localities yielded 115 control region alleles. The locality information, number of individuals and number of alleles analysed for cytochrome *b* (63 individuals, 51 alleles, 45 localities, 19 river systems), 16S (49 individuals, 31 alleles, 39 localities, 19 river systems) and the additional control region sequences (17 individuals, 17 alleles, 12 localities, 5 river systems) are given in Table 1 and Supplementary data S1 and S2. The same mtDNA regions were consistently used, which were positions 17–615 for control region, 15,350–16,429 for cytochrome *b* and 2959–3521 for 16S with reference to the mtDNA genome sequence of *Cyprinus carpio* (Chang et al., 1994). This yielded 609–614, 1080 and 568 base pairs for control region, cytochrome *b* and 16S, respectively, and a combined analysis of 2257 base pairs (total base pairs include gaps and refers to the number of base pairs after alignment).

The number of parsimony informative characters for the molecular datasets are given in Table 1. Forty-eight of the cytochrome *b* amino acids varied within *Pseudobarbus*. Amino acid differences were 0–8 within *Pseudobarbus* lineages, 1–25 between *Pseudobarbus* lineages, 17–31 between *Pseudobarbus* lineages and the southern African tetraploid *‘Barbus’* outgroups and 1–27 among the latter outgroups. Genetic distances based on the HKY substitution model with *I = 0* and equal rates among sites varied between 1% and 10.3% for control region, between 1.6% and 15.8% for cytochrome *b* and between 0.2% and 4% for 16S between lineages, compared to 0–2% for control region, 0–2.5% for cytochrome *b* and 0–0.7% for 16S within lineages.

3.2. Model based molecular analysis

All the control region sequences from the 15 historically isolated *Pseudobarbus* lineages that has been identified by Swartz et al. (2007, unpublished), re-analysed samples from all the *P. burgi* populations that Bloomer and Impson (2000) investigated, an additional *P. burgi* population and five of the seven *P. phlegethon* populations that were investigated by Swartz et al. (2004) were re-analysed and included in a single neighbour joining phylogram (Supplementary data S3) to confirm the existence of only 15 historically isolated lineages. The analysis was based on the HKY85 substitution model (Hasegawa et al., 1985) with a Ti/Tv ratio of 4.222, *I = 0.681* and *g = 0.844* found in MODELTEST. Only representatives from the 15 historically isolated lineages presented in Supplementary data S3 were included in the ML, Bayesian and ML phylogenetic analyses.
The ML analysis on the combined dataset was based on the GTR substitution model as found in MODELTEST with $\alpha = 0.631$, $\alpha = 0.998$ and with base frequencies being $A = 0.314$, $C = 0.221$, $G = 0.162$ and $T = 0.303$. From the ML phylogram in Fig. 2 (summarised in cladogram A of Fig. 5) with only single representatives of each of the 15 Pseudobarbus lineages, it is evident that the lineages of *P. burchelli* and *P. quathlambae* form two monophyletic groups with 100% bootstrap and Bayesian posterior probability support in both cases. However, *P. phlegethon* groups with the Forest lineage of *P. afer* with moderate bootstrap and significant Bayesian posterior probability support, suggesting that *P. afer* as a species may not be monophyletic. In addition, the two *P. burgi* lineages are also not monophyletic. Instead, the Berg lineage of *P. burgi* groups with the *P. burchelli* lineages with low bootstrap support, but with significant Bayesian posterior probability support. There was also a lack of bootstrap or Bayesian posterior probability support for the two *P. tenuis* lineages being monophyletic.

Deeper relationships were well resolved, but with mostly little support. There was moderate bootstrap support for the deepest split within *Pseudobarbus* being between the two *P. quathlambae* lineages and all the other lineages of *Pseudobarbus*. However, there was strong support for all the *Pseudobarbus* lineages with two pairs of barbels being monophyletic and for all the *Pseudobarbus* lineages being monophyletic compared to *B. calidus* and *B. euryclesens*.

### 3.3. Parsimony based molecular analysis

MP analysis was based on all the taxa for which the combined dataset was available and on only representatives of the 15 *Pseudobarbus* lineages (same taxa as Fig. 2). The latter was carried out to make direct comparisons to the model based analyses. The two datasets (24 ingroup taxa versus 15 ingroup taxa) did not differ from each other in terms of relationships that were inferred from the MP analyses. The strict consensus tree of three most parsimonious trees with all the taxa for which the combined dataset was available is shown in Supplementary data S4. The tree scores were 1338 steps with CI = 0.504, RI = 0.734 and RC = 0.370. From this consensus tree, it is evident that the lineages within *Pseudobarbus* are unresolved. There was strong support for the clade with *P. asper* and the two *P. tenuis* lineages as the sister group to the *P. afer* and *P. phlegethon* complex, for the deepest split within *Pseudobarbus* being between the two lineages of *P. quathlambae* and all the other lineages of *Pseudobarbus* and that all the *Pseudobarbus*
lineages were monophyletic compared to five of the six species of southern African serrated tetraploid barbs. For comparative purposes, the separate MP analyses for control region, cytochrome \(b\) and 16S are available as Supplementary data S5–S7, respectively.

### 3.4. Parsimony based morphological analysis

When characters were ordered according to characters states in *B. calidus* and *B. erubescens*, there were 40 parsimony informative characters in the morphological dataset. Three of these characters were left unordered due to uncertain transformation between character states. Twenty-three characters were synapomorphic for *Pseudobarbus* (Supplementary data S8) and only 16 characters were parsimony informative within *Pseudobarbus* (Supplementary data S9). Two most parsimonious trees were found with \(CI = 0.857, RI = 0.592\) and \(RC = 0.331\) with a length of 849 steps. The consensus of these trees is shown in Fig. 4 (summarised in cladogram I of Fig. 5). A clade with *P. afer* lineages associated with *P. phlegethon*, one with the two *P. tenius* lineages grouping with *P. asper* and a clade with all the lineages with two pairs of barbels were strongly supported. A clade with all the lineages with a single pair of barbels (excluding *P. quathlambae*) had moderate bootstrap support. There was also moderate bootstrap support for *P. phlegethon* grouping with the Forest lineage of *P. afer* and high bootstrap support for the two *P. tenius* lineages being monophyletic.

### 3.6. Summary cladograms

Simple phylograms and cladograms that summarise the results of the present phylogenetic study for comparative purposes are presented in Fig. 5. All the branches with bootstrap support <60% and Bayesian posterior probability support <95% were collapsed. There was only one conflict in relationships in the molecular analyses (A–G in Fig. 5). It concerned the position of the Berg lineage of *P. burgi*. In the Bayesian analysis (B in Fig. 5) it groups with the *P. burchelli* lineages, but in the MP analysis (E in Fig. 5) the two *P. burgi* lineages group together.

There were no conflicts between the combined morphological and DNA analysis (cladogram I in Fig. 5) and the model based or MP molecular analyses, since the position of the Berg lineage of *P. burgi* were unresolved in relation to the two other lineages with *P. afer* and *P. tenius* lineages were not collapsed into just two taxa in the former, because of mtDNA divergence. The combined molecular and morphological analysis produced 384 parsimony informative characters (40 morphological and 344 molecular). Four of the morphological characters were unordered, whilst the other morphological characters were ordered according to character states in the outgroups (*B. calidus* and *B. erubescens*). The MP analysis produced two most parsimonious trees with \(CI = 0.559, RI = 0.592\) and \(RC = 0.331\) with a length of 849 steps. The consensus of these trees is shown in Fig. 4 (summarised in cladogram I of Fig. 5). A clade with *P. afer* lineages associated with *P. phlegethon*, one with the two *P. tenius* lineages grouping with *P. asper* and a clade with all the lineages with two pairs of barbels were strongly supported. A clade with all the lineages with a single pair of barbels (excluding *P. quathlambae*) had moderate bootstrap support. There was also moderate bootstrap support for *P. phlegethon* grouping with the Forest lineage of *P. afer* and high bootstrap support for the two *P. tenius* lineages being monophyletic.

**Fig. 3.** Strict consensus of two equally most parsimonious trees based on 40 parsimony informative morphological characters that were investigated by Skelton (1980). Characters were ordered according to character states in *B. calidus* and *B. erubescens*. The broken line indicates a relationship that was recovered from the MP analysis based on the unordered dataset. Bootstrap support values based on the ordered and unordered datasets (latter in brackets) are shown above branches.

**Fig. 4.** Maximum parsimony strict consensus tree based on a combined dataset of control region, cytochrome \(b\), 16S and morphological characters, showing the relationships among lineages of *Pseudobarbus* for which both genetic and morphological datasets were available. Bootstrap support values are shown above branches. Symbols refer to the distribution of lineages in Fig. 1.
two pairs of barbels that were analysed. However, the MP analysis based only on the morphological characters (H in Fig. 5) differed substantially from the molecular analysis, mainly because of the close relationship that was inferred between *P. quathlambae* and *P. tenuis*. Further important differences were that the *Pseudobarbus* lineages with two pairs of barbels were not recovered as a monophyletic lineage and *P. phlegethon* did not group with *P. afer* as in the molecular analyses. Cladograms that summarise results across the model based molecular analyses (C and D in Fig. 5) and across all the molecular analyses (F and G in Fig. 5) are presented to show consistency of relationships across methods.

### 4. Discussion

#### 4.1. Phylogenetic relationships

The morphological and molecular phylogenies are not entirely congruent. The most important difference is the position of *P. quathlambae*. The morphological data set places this species as the sister species to *P. tenuis*. However, the molecular data strongly support a close relationship between *P. tenuis* and *P. asper*. **Skelton (1980)** suggested that *P. asper* was the sister species to *P. afer*. There has even been considerable confusion between the latter two species. Western populations of *P. afer* that mainly fall within the Forest lineage in the present study, were initially placed with *P. asper* (Barnard, 1943; Jubb, 1965). Surprisingly, however, the molecular analysis suggests a close relationship between the Forest lineage of *P. afer* and *P. phlegethon*. In the model based molecular analyses there were support for the St. Francis and Krom lineages being sister groups. The relationship between the Mandela lineage compared to the other *P. afer* lineages and *P. phlegethon* was unresolved in all of the analyses.

The molecular data and several morphological characters support a monophyletic group of all the lineages with two pairs of barbels. Within this group, the three *P. burchelli* lineages were clearly monophyletic. However, the relationship of the two *P. burgi* lineages was either monophyletic or unresolved although their sister relationship to *P. burchelli* was well supported. Other genes will have to be analysed to resolve the phylogenetic position of the Berg lineage of *P. burgi*. Apart from the single lineages of *P. asper* and *P. phlegethon*, only the three lineages of *P. burchelli* and the two lineages of *P. quathlambae* were clearly monophyletic in terms of the currently described species.

Deeper relationships within *Pseudobarbus* were not well resolved by the Bayesian analysis of the combined molecular dataset or the MP analysis of the morphological characters. However, both the ML and MP analyses based on the combined molecular dataset suggested that the lineages with a single pair of barbels are not monophyletic. This is because the earliest divergence within *Pseudobarbus* is between the two *P. quathlambae* lineages and all the other *Pseudobarbus* lineages, which include the lineages with two pairs of barbels. The next divergence was between the lineages with a single pair of barbels (excluding *P. quathlambae*) and the lineages with two pairs of barbels. The ML and MP analyses therefore suggest that the *P. afer–P. phlegethon* complex and the lineage with *P. asper* and *P. tenuis* are sister groups.

The combined morphological and molecular MP analysis showed the best resolution of deeper relationships compared to the morphological or molecular analyses on their own. It is also not in conflict with the cladograms that were constructed from...
comparisons across methods (F and G in Fig. 5). Including the morphological characters in the molecular dataset therefore did not reduce the resolution of the tree where there was conflict between the morphological and molecular characters. Instead, it seems to have contributed towards resolving relationships where there was not strong support across all the molecular analyses.

4.2. Biogeography and evolutionary history

The cladogram labelled F in Fig. 5 that summarises the results across the molecular analyses, presents a good working hypothesis of the relationships within Pseudobarbus, for comparison to existing biogeographic hypotheses. It presents all the clades for which there was support in any of the molecular analyses, as long as there were no conflicts, in which case the relationships were unresolved. It also presents all the lineages that were investigated in the present study and does not conflict with the combined morphological and molecular analysis (I in Fig. 5).

From the cladogram labelled F in Fig. 5, relationships can be mapped onto the regions currently occupied by the Pseudobarbus lineages (Fig. 6). It is reasonable to assume that the ancestor of Pseudobarbus diverged from the southern African serrated tetraploid barbs within the region now occupied by the temperate ichthyofauna, since all the taxa occur within this region. The current phylogenetic investigation challenges the biogeographic conclusions made by Skelton (1986, 1994) in one major respect. Because the morphological characters showed that P. quathlambae and P. tenuis were sister species, Skelton (1980, 1988) had to suggest a link through the Great Karoo and a link between the Gourits and Orange River systems with subsequent extinction of redfins in the Karoo tributaries of the Orange River system. Based on the molecular analyses, such an explanation is unnecessary.

The molecular evidence suggests that the earliest divergence within Pseudobarbus was between P. quathlambae and the common ancestor of all the other lineages of Pseudobarbus, probably in the mid-late Miocene (A in Fig. 6). It is unclear how or between which river systems this initial divergence occurred. The most likely connection that can explain the current distribution of Pseudobarbus species in both the Orange River system and river systems of the CFR, is through connection between the current Orange River system and the Olifants River system (see Dingle and Hendey, 1984; De Wit, 1993). However, De Wit (1993) suggests that this connection occurred no later than the early Tertiary, which would require a mutation rate in cytochrome b for Pseudobarbus that is slower than 0.39/million years compared to a rate of 1.05–1.31/million years (therefore at least 70% slower) than that commonly used for cyprinids (Dowling et al., 2002; Doadrio and Carmona, 2004; Mesquita et al., 2005). It is therefore likely that dispersal between Lesotho and CFR species of Pseudobarbus occurred after the capture of the Upper Orange by the palaeo Augrabies River that isolated the Olifants River system and shifted to Orange River mouth to its current location.

Just after the isolation of P. quathlambae in Lesotho, but also during the mid-late Miocene, divergence occurred between Pseudobarbus lineages with two pairs of barbels that currently occur in the south-western river systems of the CFR (excluding the Olifants River system) and the lineages with a single pair of barbels that are associated with most of the other river systems of the CFR (including the Olifants River system) (B in Fig. 6). Differentiation within the Pseudobarbus lineages with two pairs of barbels is
relatively old, with divergence between three deeper lineages (Verlorenvlei, Berg and the currently described *P. burchelli*) occurring in the late Miocene (C in Fig. 6). Within the lineages that have a single pair of barbels in the CFR, differentiation occurred in the late Miocene to early Pliocene between a group that currently occurs in the Gourits, Gamtoos, Keurbooms (including Bitou) River systems (*P. asper* and *P. tenuis*) and a complex group that occur in eastern river systems of the CFR and in the Olifants River system (D in Fig. 6).

The earliest divergence in the complex group of CFR lineages with a single pair of barbels that include *P. phlegethon* and the four lineages of *P. afer*, gave rise in the late Miocene to early Pliocene to a lineage associated with Nelson Mandela/Algoa Bay, one associated with St. Francis Bay and a puzzling lineage associated with several coastal river systems in the southern Afrotemperate Forest area of the CFR and the Olifants River system on the west coast (E in Fig. 6). The final steps in the evolution of the *P. afer* and *P. phlegethon* complex, would have been differentiation between the populations of *P. afer* from the Afrotemperate Forests (Forest lineage) and *P. phlegethon* and between the Krom River system (Krom lineage) and the Gamtoos and associated river systems (St. Francis lineage) in the early to mid Pliocene (F in Fig. 6). The Pliocene also saw the differentiation within the Breede River system between a Breede–Heuningnes lineage and one that is currently restricted to the Tradou catchment, which is a tributary of the Breede River system (F in Fig. 6).

During the mid to late Pliocene, differentiation occurred between two lineages of *P. quathlambae* between the Mobale catchment and eastern rivers of Lesotho and between *P. tenuis* and *P. asper* (G in Fig. 6). This was soon followed by differentiation between a *P. tenuis* lineage from the Keurbooms and Bitou River systems and a *P. tenuis* lineage from the Gourits River system in the late Pliocene to the Pleistocene (H in Fig. 6). *Pseudobarbus asper* and *P. tenuis* occur in sympatry in the Gourits River system. Their close genetic relationship raises the possibility that ecological speciation occurred or if secondary invasion occurred, at least ecological displacement and selection for different habitat preference. The divergence between *P. asper* and *P. tenuis* being relatively recent, implies that rapid morphological differentiation occurred between them compared to, for example, the *P. afer* lineages that seem to be older, but with less morphological differentiation. Finally, differentiation occurred between the Breede and Heuningnes River systems to give rise to the currently recognised Breede and Heuningnes lineages also during the late Pliocene to the Pleistocene (H in Fig. 6).

These time estimates based on cytochrome *b* are much older than previous estimates based on control region with a mutation rate of about 3% per million years that has been suggested for the salmonid mtDNA control region (Bernatchez and Danzmann, 1993). For example, here we estimate a late Miocene to mid Pliocene time of divergence compared to an estimate of mid Pliocene to Pleistocene by Swartz et al. (2007) for the differentiation of *P. afer* lineages and *P. phlegethon*. The age estimate of these lineages based on control region would be significantly more recent if one assumes the much faster rates of control region mutation that has been suggested for other fish (Brown et al., 1993). The difference in mutation rate between cytochrome *b* and control region therefore appears to be more similar in *Pseudobarbus* than what has been suggested in the literature for other cyprinids or saturation has affected control region more than cytochrome *b* at this level of differentiation.

Whereas the relationships suggested by the present molecular analysis provides biogeographic scenario’s that are much simpler in terms of the occurrence of *P. quathlambae* and *P. tenuis* to those proposed by Skelton (1980, 1988), it suggests novel phylogenetic relationships that are difficult to explain. All of the relationships based on the molecular analyses can be explained through proximality of river systems and through processes of river capture or influence of different river systems during lower sea levels, except for the close relationship between *P. phlegethon* and the Forest lineage of *P. afer*. The only logical link between these lineages is through the Gourits River system. An ancestor of *P. phlegethon* and the Forest lineage must have occurred in the Gourits River system, but most likely has subsequently been extirpated. Apart from possible sympatric speciation between *P. asper* and *P. tenuis* and possible parapatric differentiation between the Breede and Tradou lineages of *P. burchelli* (Swartz et al., unpublished), all other speciation and differentiation events within *Pseudobarbus* seem to be the result of isolation and allopatric divergence between different river systems.

Sea level transgressions would also have played a role in the differentiation of *Pseudobarbus* lineages. The last major transgression, however, occurred during the early Pliocene (about 3.4–5.2 MYA) and reached levels of around +200 m (Butzer and Helgren, 1972) to over +300 m (Siesser and Dingle, 1981) along the south coast of South Africa. During this time many river systems would have been drowned, but ever since the early Pliocene transgression, sea levels have not risen more than +30 m above present sea levels (Butzer and Helgren, 1972; Rogers, 1985) and thus would only have affected the smallest and lower altitude river systems. More samples will have to be investigated to test whether sea level transgressions left a signal on *Pseudobarbus* mitochondrial DNA diversity.

4.3 Drainage evolution in the Cape Floristic Region

Without a fossil record or clear calibration events with which to date the *Pseudobarbus* phylogeny, inferences that can be made about the drainage evolution of the CFR remains limited. The importance of palaeorivers in allowing recent dispersal between currently isolated river systems, however, is clear from the widespread distribution of the Forest, St. Francis and Mandela lineages of *P. afer* (Swartz et al., 2007) and the Breede lineage of *P. burchelli* (Swartz et al., unpublished). There were probably only five to six major palaeoriver systems along the south coast of the CFR during the last glacial maximum (LGM) about 18,000 years ago. These are referred to as the Mandela, St. Francis, Plettenberg, Wilderness, Gourits and Breede systems (Fig. 6), with Swartz et al. (2007) suggesting that the Plettenberg and Wilderness systems possibly formed one system.

It is also evident that the Gourits River system played an important role in the diversification of the redfins, providing inland opportunities for dispersal through river captures or low-gradient marches during wetter periods. This accounts for the occurrence of *P. asper* in both the Gourits and Gamtoos River systems, *P. tenuis* in the Gourits and Keurbooms/Bitou River systems and is the only logical pathway to explain the close relationship between *P. phlegethon* on the west coast and the Forest lineage of *P. afer* on the south coast. Except for the Plettenberg palaeoriver system that has two lineages due to a river capture that allowed *P. tenuis* to invade the Keurbooms River system, all the other palaeoriver drainages each only have a single lineage. In contrast, the central Gourits and Breede palaeoriver systems each carry three lineages. The occurrences of all these lineages are due to differentiation and invasions before the recent connections during the LGM, except for the occurrence of the Breede lineage of *P. burchelli* in the Goukou River system that was part of the Gourits palaeoriver system. The Breede lineage of *P. burchelli* was not able to spread to the current Gourits River system, suggesting that river capture must have occurred between the Duivenhoks and Goukou River systems (between the Breede and Gourits palaeoriver systems) after isolation of the Goukou and Gourits River systems during post-LGM sea level transgressions, suggesting a river capture event that is less than 18,000 years old.
On the west coast of the CFR, the Berg and Verlorenvlei River systems follow the eastern CFR pattern of having only a single lineage per river system. Much older connections must have occurred between the major river systems and isolation in neighbouring river systems is evident across the CFR for *Pseudobarbus*. Linking these, however, with specific geological or climatic events with confidence will require advances in dating geomorphological features, discovery of *Pseudobarbus* fossils and/or a better understanding of the *Pseudobarbus* molecular clock through comparative phylogenetic analysis with other CFR genera.

Acknowledgments

M. Cunningham, I.M. Russo, W. Delport and K. Tolley are thanked for assistance with the phylogenetic analyses. D. Naran provided *B. trevelyani* samples through the SAIAAB collection. N.D. Impson collected most of the *P. burgi* specimens. G. Gouws, S. Daniels and F. Gordon assisted in the collection of *B. hospes* specimens and G. Kempen assisted in the collection of *P. burgi* samples. Several other collaborators and volunteers, all of whom has been acknowledged (Swartz et al., 2004, 2007) or will be acknowledged in subsequent papers that deal with individual species complexes, assisted in field surveys and the collection of genetic material. WWF-SA (Table Mountain Fund) and the National Research Foundation (South Africa) funded the research.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2008.10.017.

References


