Molecular Ecology (2008) **17**, 4812–4826 doi: 10.1111/j.1365-294X.2008.03959.x

Significant population structure and asymmetric gene flow patterns amidst expanding populations of *Clinus cottoides* (Perciformes, Clinidae): application of molecular data to marine conservation planning in South Africa

SOPHIE VON DER HEYDEN,* KIM PROCHAZKA†§ and RAURI C. K. BOWIE*‡
*Evolutionary Genomics Group, Department of Botany and Zoology, Stellenbosch University, Private Bag XI, Matieland 7602, South Africa, †International Ocean Institute Southern Africa, Department of Biodiversity and Conservation Biology, University of the Western Cape, Private Bag X17, Bellville 7535, South Africa, ‡Museum of Vertebrate Zoology and Department of Integrative Biology, 3101 Valley Life Science Building, University of California, Berkeley, CA 94720, USA

Abstract

***Clinus cottoides* is a fish endemic to the coast of South Africa, predominantly inhabiting rock pools. All South African clinids are viviparous, but probably breed throughout the year; as such, their dispersal may be limited, unlike species with pelagic larval stages. We analysed 343 fish from 14 localities on the west, south and east coasts using two mitochondrial genes and the second intron of the S7 ribosomal gene. Mitochondrial DNA analyses recovered significant genetic differentiation between fish populations from the east coast and other sampling locations, with a second break found between Gansbaai and Cape Agulhas on the south coast. Nuclear DNA recovered shallower, but significant, levels of population structure. Coalescent analyses suggested remarkably asymmetrical gene flow between sampling locations, suggesting that the cold Atlantic Benguela Current and Indian Ocean Agulhas counter-current play important roles in facilitating dispersal. There was no gene flow between the east coast and the other sites, suggesting that these populations are effectively isolated. Divergence times between them were estimated to at least 68 000 years. Neutrality tests and mismatch distributions suggest recent population expansions, with the exception of peripheral western and eastern populations (possibly a consequence of environmental extremes at the edge of the species distribution). Analyses of the current South African marine protected areas network show that it is not connected and that De Hoop, one of South Africa's largest marine reserves, appears to be an important source population of recruits to both the south and southwest coasts.**

Keywords: Clinidae, gene flow, marine protected areas, mitochondrial DNA, nuclear DNA, phylogeography, South Africa

Received 11 June 2008; revision received 27 August 2008; accepted 8 September 2008

Introduction

Phylogeographical analyses play an increasingly important role in determining the presence of evolutionary significant units, as well as 'genetic diversity hotspots' (Moritz 2002; Rocha *et al*. 2007). With increasing pressures on marine ecosystems from exploitation, climate change and invasive species, it is becoming increasingly important to evaluate areas offering protection using not only biological, but also genetic data (Moritz 2002; Rocha *et al*. 2007). As many marine species also have wide and often disjunct populations, it is crucial to understand the directionality and intensity of gene flow, which is shaping the populations and species of tomorrow (Rocha *et al*. 2007; von der Heyden *et al*. 2007), and to use this data to make informed decisions regarding the placement of marine protected areas (MPA). It is

Correspondence: S. von der Heyden, Fax: +27 (0)21 8082405; E-mail: svdh@sun.ac.za
§Present address: Department of Environmental Affairs and Tourism, Branch: Marine and Coastal Management, Private Bag X2, Roggebaai, Cape Town 8012, South Africa

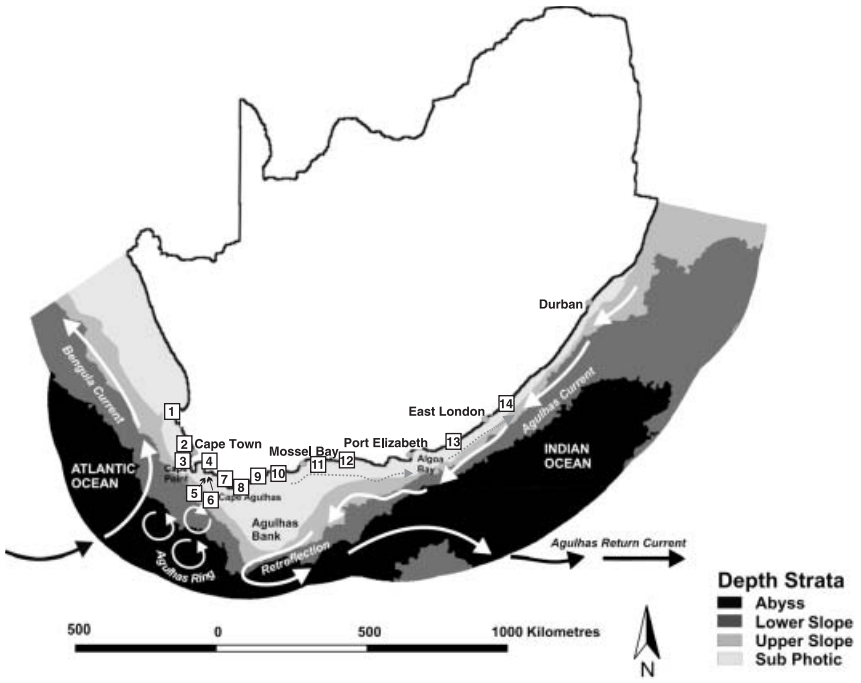

Fig. 1 Map showing the distribution (dashed arrows) and sampling localities of *Clinus* cottoides. 1, Jacobsbaai; 2, Mouille Point; 3, Kommetjie; 4, Wooley's Pool; 5, Rooiels; 6, Betty's Bay; 7, Gansbaai; 8, Cape Agulhas; 9, De Hoop; 10, Cape Infanta; 11, Herrold's Bay; 12, Knysna; 13, Port Alfred; 14, Haga Haga/Gonubie. Major physical oceanographic features around the South African coastline showing the Benguela and Agulhas Currents that are likely to influence gene flow patterns in *C. cottoides* and other South African marine organisms are highlighted (e.g. abalone and crustacea; see text for details). Agulhas ring eddies and the Agulhas counter-current are also depicted: modified after Lutjeharms & Ballegoyen (1988) and Lutjeharms & Ansorge (2001).

important to understand that MPAs do not only act to retain species and their propagules, but also act as a 'reproductive reservoir' which supplies recruits to adjacent areas which are not protected (Bird *et al*. 2007). This is especially important in South Africa where marine protection is minimal and often poorly enforced. Importantly, although 23% of the South African coast is protected by MPA status, only 9% is designated as 'no-take' zones, with the remaining coastline under open access for commercial, artisanal and recreational exploitation (A.T. Lombard (Conservation Systems), T. Strauss (Conservation Systems), D.J. Harris (Ezemrelo KZN Wildfire), K. Sink (African Coelocanth Project), C. Attwood (Marine Coastal Management), L. Hutchings (Marine Coastal Management), unpublished data).

The South African coastal marine environment is characterised by very high species richness as a consequence of its long coastline and variable oceanic conditions (Awad *et al*. 2002). Three main biogeographical provinces are recognised: the 'cool-temperate Namaqua Province', 'warm-temperate south coast Agulhas province' and 'the subtropical east coast province' (Emanuel *et al*. 1992; Figs 1 and 3). Due to the dynamic nature of the intertidal zone, it is not possible to determine the boundaries of each biogeographical province, yet it is clear that there are significant faunal differences between the southeast and northwest coasts separated by Cape Point (Fig. 1; Branch *et al*. 1994). In contrast, the transition between the south and east coasts remains unclear (Turpie *et al*. 2000; Harrison 2002).

Whether there are also genetic boundaries that relate to the ecologically determined coastal biogeographical boundaries in South Africa is poorly understood. For the caridean shrimp *Palaemon peringueyi*, the abalone *Haliotis midae*, and three estuarine invertebrate species, significant genetic structure with a discontinuity in the Cape Agulhas region was detected (Evans *et al*. 2004; Teske *et al*. 2006). Additional breaks on the west and east coasts have also been detected for the isopod, *Exosphaeroma hylocoetes* and the cumacean, *Iphinoe truncata*, respectively (Teske *et al*. 2006). In contrast, for the rockpool inhabiting goby, *Caffrogobius caffer*, no population structure was recovered over the genetic barrier at Cape Agulhas (Neethling *et al*., in press).

Fishes of the family Clinidae (klipfish) have a disjunct distribution globally, with representatives in North and South America, the Mediterranean, Australia, New Zealand and at some localities in the Indo-Pacific. Southern Africa has the largest number of clinid species, with most of these being restricted to the South African coasts, although some species with large distributional ranges are also found in southern Namibia. They are prominent members of the intertidal and are found both in rock pools, and in the shallow subtidal environment. Prochazka & Griffiths (1992) in a study of South African intertidal rock pool communities reported that clinids constituted between 88% to 98% of fishes collected. Southern African clinids are viviparous and 'give birth' to well-developed postflexion larvae, as are clinid species found in Australia and New Zealand. Clinids from North and South America are oviparous (Stepien 1992), although some oviparous clinids also exhibit nesting and guard behaviour by male fish (Coyer 1982).

Clinus cottoides (bluntnose klipfish) is an abundant, small (maximum length 15 cm; Smith 1995), carnivorous clinid species, endemic to South Africa with a distribution along

the west, south and southern east coasts of South Africa (Fig. 1). Their preferred habitats are tide pools on the rocky shore, although rarely, individual fish have been seen to around 8 m depth (Zsilavecz 2005). As for all clinid post-flexion larvae, those of *C. cottoides* are rarely encountered in plankton surveys, in contrast to a number of other South African endemic fishes (N. Strydom, personal communication).

Most marine fishes spawn directly into the water column and have pelagic larval stages that disperse freely. A result of this dispersal strategy is that many marine species show little or no genetic structure over large areas (Grant & Bowen 1998; Waples 1998). Vivipary may therefore reduce the potential for free dispersal, especially in organisms utilising rocky shore habitats. Fish without pelagic larval stages, for example the black surfperch (*Embiotoca jacksoni*, Bernardi 2000), the damselfish (*Acanthochromis polyacanthus*, Planes *et al*. 2001) and the Banggai cardinal fish (*Pterapogon kauderni*, Bernardi & Vagelli 2004) show strong genetic differentiation among sampled localities. Therefore, viviparous fish may have more pronounced population genetic structuring compared to species with pelagic life-history stages, although a number of papers have shown that active larval behaviour may influence population structure by limiting dispersal via larvae 'homing' in on specific oceanographical features (Leis *et al*. 2005; Cowen *et al*. 2006).

Given that *C. cottoides* has a potentially abbreviated life history as a result of 'live-born' larvae and hence putatively has low dispersal abilities, we use this fish as a model species in order to investigate population structure and gene flow patterns along the South African coast. Specifically, we aim to test the hypothesis that the major dispersal agents of *C. cottoides*, their larvae, and therefore the associated gene flow patterns show directionality with respect to two major ocean currents, the Benguela and Agulhas (Fig. 1). In addition to our work, we use published research in order to carry out a preliminary assessment of the MPA system currently in place in South Africa. In particular, we focus on the De Hoop marine reserve, the largest MPA in South Africa, located just east of Cape Agulhas, the dominant genetic break found for marine organisms in the southern Africa intertidal to date. Furthermore, we test the hypothesis that De Hoop contributes significant recruits to populations located to the west, with dispersal predominantly occurring by means of the Agulhas Current and associated Agulhas ring eddies (Lutjeharms & Ballegooyen 1988). We use sequence data from two mitochondrial gene regions [the 5′ variable part of the control region (D-Loop) and NADH subunit 2], as well as one nuclear intron (the second intron of the S7 ribosomal gene) for 343 fishes spanning the entire distribution range of *C. cottoides*. This approach additionally allows us to critically compare the extent of population structure recovered with maternally vs. biparentally inherited marker systems.

Materials and methods

Sample collection and laboratory procedures

Fish were collected exclusively from tide pools (Fig. 1) using handnets and the anaesthetic ethyleneglycolmonophenylether (Merck). Muscle tissue was dissected from each fish and total genomic DNA extracted with the Puregene kit (Gentra Systems) with an overnight proteinase K digestion at 55 °C. All fish will be deposited in the collection of the South African Institute for Aquatic Science, Grahamstown, South Africa.

Polymerase chain reaction (PCR) primers to amplify the NADH subunit 2 gene (ND2) in clinids were designed from an alignment of mitochondrial gene regions of other perciform fish. The primers Clinus_Nd2F: 5′-AAGAGATCAAAACTCTTAGTGCTTC-3′; Clinus_Nd2R: 5′-CTTTGAAGGCCCTTGGTC-3′ PCR amplified a fragment of *c.* 1100 bp, which encompasses the entire ND2 gene. All sequences were checked to ensure that no pseudogene was amplified. The variable 5′ part of the mitochondrial (D-Loop) region was amplified using the PCR protocol and primers CR-A and CR-E as described by Lee *et al*. (1995). The second intron of the S7 ribosomal gene was amplified as in Chow & Hazama (1998), but additional clinid-specific S7 primers were designed to facilitate easier amplification; Clinid_S7F: 5′-GTTCGAGTCCGGGATCTCCCAGGTAAG-3′ and Clinid_S7R: 5′-GCTCACTGTCCCTTTGTAGCTCAG-3′. PCR cycling conditions were as follows for ND2; an initial denaturation of 3 min at 94 °C, 35 cycles of denaturation at 94 °C for 30 s, annealing at 54 °C for 45 s, extension at 72 °C for 1.5 min and a final extension at 72 °C for 5 min. The cycling conditions for the clinid-specific PCR primers were as for ND2 but with an annealing temperature of 62 °C and an extension time of 1 min. All PCR products were gel purified using the Gel Extraction kit from QIAGEN. Purified products were sequenced using BigDye terminator chemistry (Applied Biosystems) and run on an ABI 3100 automated sequencer. ND2 fragments were sequenced with the forward and reverse primers. In order to check for *Taq* or sequencing error in the S7 alleles, 10% of fish with single haplotypes were PCR amplified and sequenced twice. All sequences have been deposited with GenBank under the following accession nos: DQ525278–DQ525291 (ND2), DQ525292–DQ525332 (S7 intron) and DQ525333–DQ525352 (D-Loop).

Data analyses

Sequences were edited and aligned in BioEdit (Hall 1999). Four data sets were generated; data sets 1 and 2 contained 343 fish and 420 bp of the D-Loop and 1021 bp of the ND2 gene, respectively. A third data set combined the mtDNA genes for a total of 1441 bp. Data sets were analysed

separately, but as mtDNA is inherited as a single linkage block, analyses were also conducted with a combined D-Loop and ND2 data set. Data set 4 contained 334 fish (but 668 sequences) and 659 bp of the second intron of the S7 ribosomal gene. In order to separate the alleles in the intron data for each individual, we used the program Phase2.1 (Stephens *et al*. 2001; Stephens & Scheet 2005). We considered the haplotypic phase of alleles to be resolved if recovered with a degree of confidence of 90% or greater from a run of 1 million iterations. All data sets were collapsed to haplotypes using the program Collapse1.2 (http://darwin.uvigo.es).

Data sets were analysed in two ways; the first combined all samples from all localities (i.e. a panmixia model); in the second analysis, the 14 sampling localities were ungrouped (see Fig. 1). Standard diversity indices [haplotype (h) and nucleotide (π) diversity], and the number of polymorphic sites were calculated for each data set using Arlequin 3.1 (Excoffier *et al*. 2005).

In order to investigate possible genetic structuring between sampling locations, analyses of molecular variance (AMOVA) and pairwise Φ_{ST} (an analogue of F_{ST} that considers not only haplotype frequency, but also the extent of differentiation among haplotypes) were conducted in Arlequin 3.1. Significance levels for AMOVA were obtained using a nonparametric permutation approach with 10 000 iterations. In addition, SAMOVA analyses were carried out on all data sets. SAMOVA maximises the proportion of the total genetic variation between groups of populations, as well as identifying possible genetic barriers between them, without pre-defining populations as is necessary for AMOVA (Dupanloup *et al*. 2002).

Due to the limited sequence divergence among sampled haplotypes (0.74 ± 0.36% for D-Loop, 0.23 ± 0.06% for ND2, 0.68 ± 0.65% for S7), a network approach was adopted rather than attempting to build a bifurcating tree, which given the low divergence values would almost certainly invalidate the underlying assumptions of any bifurcating tree-building algorithm (for review see Posada & Crandall 2001). A statistical parsimony network was constructed among haplotypes of the combined mtDNA data set and among alleles of the S7 intron using the program TCS version 1.21 (Clement *et al*. 2000).

To investigate the demographic history of *Clinus cottoides*, pairwise mismatch distributions for all the sampling locations combined and each locality separately, and a test of selective neutrality, Fu's F_S (Fu 1997) were calculated in Arlequin 3.1. This was used as: (i) a test whether populations are in mutation–drift equilibrium under an infinite sites model, and (ii) as an indicator of recent demographic change, which could have resulted from either population expansion or contraction, background selection or genetic hitch-hiking (Rogers & Harpending 1992; Harpending 1994).

To estimate past migration rates, as well as the directionality of gene flow between sampling localities, Migrate-n version 2.4 was used (Beerli & Felsenstein 1999, 2001). Estimates of directional gene flow are especially important in understanding whether oceanographical features, that is, the diverse current systems and upwelling cells affect gene flow patterns, despite *C. cottoides* having an abbreviated life history. In order to maximise the statistical power of our gene flow analyses, we sought to minimise unnecessary parameters (see, e.g. Bowie *et al*. 2006) and constructed a stepping-stone model with asymmetrical gene flow. Our study system is particularly suited to this model, given the linear nature of the southern African coastline. For the stepping-stone migration model, two analytical runs were conducted, an initial short run, followed by a second longer run. For both runs, the starting values of the population mutation parameter and the ratio between the immigration rate and the respective population and mutation rate per generation were estimated from F_{ST} values (Beerli & Felsenstein 1999). For the long-run, 10 short-chains, each with a total of 25 000 generations and a sampling increment of 20 generations, and two long-chains each with a total of 250 000 generations and a sampling increment of 50 generations were run twice. A total of 50 000 and 12 500 000 genealogies (recorded steps multiplied by the sampling increment) were visited by the short and long chains, respectively. For both, the short and long chains, the first 10 000 genealogies were discarded (the burn-in). An adaptive heating scheme with four chains (starting values of 5.00, 2.50, 1.50, 1.00) and a swapping interval of one was used to ensure that efficient mixing occurred. For the other settings, default values were implemented.

We also used the program MDIV (Nielsen & Wakeley 2001) to calculate the divergence time between the two most genetically divergent populations of *C. cottoides*, those from the east coast (Port Alfred and Haga Haga) and all other sampling localities. MDIV was used to estimate a variety of parameters: theta ($\theta = 2N_{ef}\mu$), migration rate ($M = 2N_{ef}m$), time of population divergence ($T = t/2N_{ef}$) and time to the most recent common ancestor (TMRCA = $t\mu$), where N_{ef} is the female effective population size, t is the generation time and μ is the mutation rate. Three simulations were run with 2×10^6, 5×10^6 and 10×10^6 generations, all had a 10% burn-in. A finite-sites model with upper bounds of 10 for the scaled migration rate and time of population divergence of 5 U were set. Scaled divergence time was converted to years where $T_{DIV} = T\theta/2\mu$, with T and θ being estimated from MDIV, and μ being calculated by multiplying with the estimated 3.6% per million year mutation rate for the teleost control region (Donaldson & Wilson 1999). It is uncertain at what age *C. cottoides* begins to reproduce, but from specimen length/reproductive data, we estimated that a generation time of 0.5 years is appropriate.

Table 1 Diversity indices for all sampling locations combined for all data sets

Data set	*n*	*h*	π	*S*/percentage of variation	Fu's *Fs*
D-Loop	343	0.66	0.0026	35/8.3%	–30.3, $P < 0.001$
ND2	343	0.42	0.0005	57/5.6%	–34.0, $P < 0.001$
Combined	343	0.79	0.0011	91/6.3%	–27.7, $P < 0.001$
S7 intron	334 (668)	0.89	0.0046	41/6.2%	–25.6, $P < 0.001$

n, the number of fish sampled; *h*, haplotype/allele diversity; π, nucleotide diversity; *S*, number of polymorphic sites. The percentage of variable sites are also given. The number in brackets for the S7 intron is the number of alleles sampled.

Results

Sequence variation and diversity estimates

In total, 343 fish were analysed for the mtDNA data. We excluded nine fish from the S7 data set, as the program Phase2.1 could not resolve two or more alleles with 90% confidence; this left 334 fish for the intron analyses. Overall, the nDNA S7 data set had both the highest haplotype/allelic and nucleotide diversity, with NADH2 being least diverse for all measures of diversity (Table 1). The D-Loop data set included 420 positions with 35 polymorphic sites (8.3%); 41 haplotypes were recovered. The most common haplotype was shared by 57.1% of fish sampled. The combined D-Loop and NADH2 data set consisted of 1441 nucleotide positions and included 91 variable sites (6.3%); 91 haplotypes were recovered with the most common haplotype shared by 35.1% of fish. For both mtDNA data sets, there were no shared haplotypes between the east coast samples (Fig. 1) and any of the other sampling locations Table 2), with one fixed nucleotide difference (data not shown). Furthermore, for the mtDNA data set, haplotype one was restricted to fish collected only from east of Cape Agulhas, haplotype 30 to fishes from Wooley's Pool to Gansbaai (southwest coast) and haplotype 72 to fishes from Cape Agulhas, De Hoop and Cape Infanta. Samples from the south coast (De Hoop to Herrold's Bay) were the most genetically diverse.

The 659 nucleotide positions of the S7 intron included 41 variable sites (6.2%). One hundred and five haplotypes were recovered; 25.3% of fish shared the most common haplotype. Of the 334 fish sampled, 71 (21.3%) were homozygous and 263 were heterozygous. Many alleles (65.7%) were restricted to single sampling sites, with De Hoop on the South Coast being the most diverse (Table 2). As for the mtDNA, no consistent geographical pattern was apparent in either the allelic or nucleotide diversity estimates for the S7 intron (Table 3), although fishes from the south and southwest coasts were more diverse when compared to the west coast.

Analyses of molecular variance

AMOVA revealed significant overall Φ_{ST} values among the 14 sampling sites for both D-Loop ($\Phi_{ST} = 0.26$, $P < 0.001$), as well as for the combined mtDNA data set ($\Phi_{ST} = 0.12$, $P < 0.001$), and the S7 intron ($\Phi_{ST} = 0.05$, $P < 0.001$). Pairwise Φ_{ST} analyses on the mtDNA data showed that the highest Φ_{ST} values were between the east coast (Haga Haga/Gonubie and Port Alfred; Fig. 1) and all other sampled populations and ranged between 0.48 to 0.77, suggesting the presence of a major gene flow barrier between Knysna and Port Alfred (Table 4). A second, less pronounced genetic break mirrors that of other South African molecular studies, in that it lies around Cape Agulhas. There is some shallow, but significant structure between sampling localities from Jacobsbaai to Gansbaai ($\Phi_{ST} = 0–0.14$) but all west and southwest coastal sampling sites are significantly different to Cape Agulhas, De Hoop and other more easterly locations. The Φ_{ST} values between Cape Agulhas and sampling localities to the west are generally higher ($\Phi_{ST} = 0.1–0.22$), suggesting the presence of a genetic barrier in the region. This suggests that any barriers to gene flow must lie on the western side of Cape Agulhas, between Cape Agulhas and Gansbaai (Table 4; Fig. 3).

Interestingly, S7 intron analyses are not analogous to the mtDNA data. Most notably, the east coast populations are not as strongly differentiated from those on the west and south coast and there are no fixed nucleotide differences between them. Port Alfred is significantly different to all other sampling sites, whereas Haga Haga (east of Port Alfred) showed little differentiation, possibly due to the smaller sample size from that locality. Φ_{ST} values between the east coast and other sampled sites are not significantly higher as they are for the mtDNA data, and overall the Φ_{ST} values range from 0.01 to 0.22. De Hoop is also significantly different from all other sampling localities (Table 4), possibly because of the large number of diverse alleles recovered from fishes there.

SAMOVA was run for all three data sets. For the mtDNA data sets, F_{CT} was maximised at two groups, the two east coast locations (Port Alfred and Haga Haga) and all other sampling stations (D-loop: $F_{CT} = 0.51$, $P < 0.02$; combined mtDNA: $F_{CT} = 0.40$, $P < 0.02$), providing further evidence that the east coast populations are isolated from the south and west coast fishes. However, for the S7 data F_{CT} was maximised at four groups ($F_{CT} = 0.065$, $P < 0.01$), but this grouping was not biologically meaningful, as each of the four groups included sampling localities seemingly at random from the west, south and east coast of South Africa. If sampling localities are pooled in accordance with the mtDNA results from the SAMOVA analyses and modelling

Table 2 Sampling localities, sample size, and frequencies of mtDNA haplotypes and nDNA alleles for *Clinus cottoides*

Locality	*n*	Haplotype combined mtDNA	*n*	Alleles nDNA
West coast				
Jacobsbaai (1)	13	2(8) 34(3) 35[1] 36[1]	13 (26)	3(19) 20(2) 28(01) 29(1) 30(2) 35(1)
Mouille Point (2)	33	2(23) 26(1) 31(2) 32[1] 33[1] 37[3] 38[1] 79[1]	33 (66)	3(28) 7(1) 8(4) 9(13) 19(1) 20(5) 27(1) 28(3) 30(2) 33(1) 34[2] 35(3) 101[1] 102[1]
Kommetjie (3)	18	2(10) 31(3) 42(1) 50(1) 69[1] 70[1] 71[1]	18 (36)	3(16) 8(2) 9(3) 20(3) 27(3) 28(3) 30(1) 31(3) 42(1) 63[1]
Southwest coast				
Wooley's Pool (4)	24	2(7) 30(2) 42(2) 46(2) 49[2] 50(2) 51[1] 52[1] 53[1] 54[1] 55[1] 56[1] 57[1]	24 (48)	3(14) 8(2) 9(7) 10(1) 12(1) 20(5) 27(1) 28(5) 30(1) 31(4) 45[1] 46[3] 47(2) 48(1)
Rooiels (5)	42	2(22) 24[1] 25[1] 26(1) 27(6) 30(1) 34(1) 39(3) 40[3] 41(2) 42(2) 43[1]	42 (84)	3(35) 6(4) 9(9) 12(9) 20(5) 25[1] 26[1] 27(3) 28(7) 31(3) 36(3) 37(1) 38[1] 39[1] 40[1]
Betty's Bay (6)	28	2(16) 22(1) 23[1] 27(2) 28[1] 29[1] 30(1) 44[1] 45[1] 46(1) 47[1] 48[1]	27 (54)	3(12) 7(2) 8(3) 9(5) 12(2) 13(1) 19(1) 20(5) 21[1] 22[1] 23[1] 24[1] 27(1) 28(5) 29(2) 30(1) 31(2) 32[1] 33(1) 35(2) 41[1] 42(1) 43[1] 44(1)
South coast				
Gansbaai (7)	26	2(17) 22(1) 30(4) 39(1) 41(1) 67[1] 68[1]	26 (52)	3(8) 6(1) 7(4) 8(2) 9(7) 12(1) 19(1) 20(7) 27(1) 28(4) 31(3) 35(1) 36(2) 37(1) 42(1) 48(1) 53(1) 60(2) 61(2) 62[1] 97[1]
Cape Agulhas (8)	18	1(1) 2(11) 4(1) 5(1) 72(2) 73[1] 74[1] 86[1]	17 (34)	3(6) 4(1) 5(2) 6(1) 7(1) 8(2) 9(2) 17(1) 19(2) 20(3) 27(1) 42(2) 47(1) 58(1) 59[1] 61(2) 64[1] 65[1] 98[1] 99[1] 100[1]
De Hoop (9)	35	1(2) 2(12) 3[1] 4(8) 5(1) 17[1] 18[1] 19[1] 20[1] 72(2) 81[1] 82[1] 83[1] 84[1] 85[1]	30 (60)	1(1) 2[1] 3(7) 4(1) 5(1) 6(5) 7(5) 8(5) 12(3) 15(1) 16[1] 17(1) 18[1] 19(1) 20(3) 31(1) 52(1) 60(1) 61(1) 78(1) 80[1] 81[1] 82[1] 83[1] 84[1] 85[1] 86[1] 87[1] 88[1] 89[1] 90[1] 91[1] 92[1] 92[1] 94[1] 95[2] 96[1]
Cape Infanta (10)	20	1(3) 2(7) 4(2) 72(1) 75[1] 76[1] 77[2] 78[1] 87[1] 88[1]	20 (40)	1(1) 3(5) 4(1) 6(6) 7(3) 8(2) 9(1) 19(1) 20(5) 27(5) 28(3) 44(1) 53(2) 58(1) 61(1) 66[1] 67[1] 78(1)
Herold's Bay (11)	32	1(5) 2(13) 4(2) 26(2) 58[1] 59[1] 60(1) 61[1] 62[1] 63[1] 64(1) 65[1] 66[1] 80[1]	31 (62)	3(19) 4(2) 6(3) 7(1) 8(3) 9(9) 10(1) 12(3) 15(1) 19(1) 20(2) 27(3) 28(1) 49[1] 50[1] 51(1) 52(1) 53(1) 54[2] 55[1] 56[1] 57(2) 77[1] 78(1) 79[1]
Knysna Heads (12)	17	1(3) 2(9) 60(1) 64(1) 89[1] 90[1] 91[1]	16 (32)	3(9) 4(1) 7(3) 9(1) 19(4) 20(6) 28(1) 30(1) 35(1) 42(1) 58(1) 103[1] 104[1] 105[1]
East coast				
Port Alfred (13)	26	6(18) 10[2] 11[1] 12[1] 13[1] 14[1] 15[1] 16[1]	26 (52)	3(8) 6(3) 8(4) 9(5) 10(1) 12(5) 13(1) 20(8) 27(1) 42(1) 51(3) 57(2) 69[1] 70[1] 71[1] 72[3] 74[2] 75[1] 76[1]
Gonubie/Haga Haga* (14)	11	6(8) 7[1] 8[1] 9[1]	11 (22)	3(8) 9(4) 10(2) 11[1] 12(3) 13(1) 14[1] 51(1) 68[1]

n, number of haplotypes sampled for mitochondrial and nuclear DNA markers (and alleles). [ ], haplotypes unique to a specific population; (), haplotypes shared between sampling areas. The nos 1–14 refer to the sampling localities depicted in Fig. 1 (*three samples from Gonubie and eight from Haga Haga).

of gene flow (see below) among localities then four groups or 'populations' can be recognised: the west coast (Jacobsbaai, Mouille Point, Kommetjie), southwest coast (Wooley's Pool, Rooiels, Betty's Bay, Gansbaai), south coast (Cape Agulhas, De Hoop, Cape Infanta, Knysna) and east coast (Port Alfred, Haga Haga). Φ_{ST} values are significant ($P < 0.001$) in all pairwise comparisons among these four groups and range from 0.02–0.51 with an increase in Φ_{ST} from west to east.

Statistical parsimony network

Evaluation of the limits of statistical parsimony suggests that topologies connecting haplotypes by eight steps (mtDNA) and S7 alleles (nDNA) by 11 steps or fewer have a cumulative probability of greater than 95% of being correct. The mtDNA network recovered a star-shaped pattern with one central common haplotype, to which several more recently derived haplotypes are connected

Table 3 Diversity indices for the fourteen sampling locations for the combined mtDNA data and the nDNA S7 intron

Gene			Combined mtDNA			S7 intron		
Locality	*n*	Θ	*h*	π	Fu's F_S	*h*	*π*	Fu's F_S
West Coast								
Jacobsbaai (1)	13 (13)	0.0022	0.60	0.0007	NS	0.47	0.0020	NS
Sea Point (2)	33 (33)	0.0057	0.51	0.0005	–5.1*	0.77	0.0033	–4.3*
Kommetjie (3)	18 (18)	0.0054	0.67	0.0009	–2.9*	0.78	0.0030	NS
South-West Coast								
Wooley's Pool (4)	24 (24)	0.0061	0.90	0.0016	–7.2*	0.87	0.0037	–4.6*
Rooiels (5)	42 (42)	0.0062	0.73	0.0010	–5.9*	0.80	0.0035	–4.2*
Betty's Bay (6)	28 (27)	0.0103	0.68	0.0011	–7.4*	0.93	0.0066	–10.1*
South Coast								
Gansbaai (7)	26 (26)	0.0042	0.56	0.0006	–3.8*	0.93	0.0042	–12.0*
Cape Agulhas (8)	18 (17)	0.0228	0.67	0.0010	–3.4*	0.96	0.0054	–13.6*
De Hoop (9)	35 (30)	0.0078	0.84	0.0013	–9.4*	0.97	0.0072	–25.6*
Cape Infanta (10)	20 (20)	0.0079	0.86	0.0010	–6.3*	0.94	0.0040	–10.0*
Herold's Bay (11)	32 (31)	0.0056	0.82	0.0009	–12.2*	0.89	0.0043	–16.4*
Knysna Heads (12)	17 (16)	0.0094	0.71	0.0007	–4.1*	0.88	0.0036	–6.8*
East Coast								
Port Alfred (13)	26 (26)	0.0035	0.52	0.0005	–5.6*	0.93	0.0052	–8.1*
Gonubie/Haga Haga* (14)	11 (11)	0.0104	0.49	0.0010	NS	0.84	0.0039	NS

n, the number of fish sampled (numbers in brackets are fish sampled for S7 intron); *h,* haplotype allelic diversity; π, nucleotide diversity; *significant Fu's F_S test; NS, not significant. 1–14 refers to the sampling localities on Fig. 1.

(Fig. 2a). The most divergent haplotype was from De Hoop, and there is a subnetwork of 11 haplotypes restricted to the east coast, although with the presence of an internal loop it cannot be resolved whether these restricted haplotypes represent a radiation from a single ancestral lineage (Fig. 2a).

In stark contrast, the S7 network does not show the same pattern as for the combined mtDNA network. Instead, there is no central allele, but rather five dominant ones. Notably, the sub-network consisting only of east coast haplotypes is absent, with these individuals having alleles spread across the network. However, a second highly divergent sub-network consists primarily of alleles from fish in De Hoop, as well as from Betty's Bay and Rooiels (Fig. 2b).

Tests for neutrality and mismatch distributions

For both the mtDNA and nuclear intron data Fu's F_S revealed significant negative values suggesting that overall (all sampling localities combined) *Clinus cottoides* populations are not in genetic equilibrium, possibly as a consequence of population expansion. At the individual level, all sites show demographic expansion, except for the two peripheral populations at the edges of the distribution of *C. cottoides* (Table 3). The mismatch distributions for the mtDNA and the intron, showed a strong signal of population expansion, with unimodal distributions (data not shown).

Coalescent analyses of migration

The coalescent stepping-stone model revealed asymmetrical gene flow patterns around the South African coast (Fig. 3; Table 5). There is northwards gene flow between Kommetjie and Sea Point to Jacobsbaai on the west coast (*Nm* = 28–40.5), but no gene flow southwards from Jacobsbaai or Sea Point to Kommetjie (*Nm* = 0). Gene flow around Cape Point, a significant break in faunal and floral distribution, is also low between Kommetjie on the west coast and Wooley's Pool in False Bay *Nm* = 0–4). There also is no gene flow in an eastwards direction between Wooley's Pool and Rooiels (*Nm* = 0), but significant gene flow in the opposite direction (*Nm* = 24). On the southeast and south coast, between Rooiels and Gansbaai, fishes appear to be a little more mixed with gene flow occurring in both directions, although there is stronger gene flow from west to east. Significantly, there is no gene flow from Gansbaai to Cape Agulhas and only little in return (*Nm* = 3.5). Cape Agulhas receives a high number of migrants from De Hoop (*Nm* = 63), but does not contribute many in return (*Nm* = 3). Interestingly, De Hoop appears to be the source population for all other sampling localities along the east coast with gene flow only occurring east along the coast, but does not receive migrants from localities to the east of it. Populations at Port Alfred (13) and Haga Haga (14) are effectively isolated, with little (*Nm* = 0.2) gene flow to them from the west; they also do not provide migrants to areas west of

Table 4 Pairwise Φ_{ST} values among sampling localities for *Clinus cottoides*. Values below the diagonal are for the mtDNA control region data set, above the diagonal for the nDNA S7 intron data set. Those in bold are significant

Location	Jacobsbaai	Mouille Point	Kommetjie	Wooley's Pool	Rooiels	Betty's Bay	Gansbaai	Cape Agulhas	De hoop	Cape Infanta	Herold's bay	Knysna	Port Alfred	Gonubie/Haga Haga
Jacobsbaai	–	**0.11**	**0.05**	**0.11**	**0.05**	**0.10**	**0.10**	**0.09**	**0.14**	**0.16**	**0.11**	**0.06**	**0.22**	**0.19**
Mouille Point	**0.14**	–	0.01	0.00	**0.02**	**0.02**	0.14	**0.03**	**0.09**	**0.05**	0.00	**0.03**	**0.08**	0.02
Kommetjie	**0.09**	0.006	–	0.00	0.00	**0.03**	0.01	0.03	**0.08**	**0.06**	0.01	0.02	**0.11**	0.05
Wooley's Pool	**0.08**	**0.11**	0.03	–	0.01	0.02	0.01	0.03	**0.08**	**0.05**	0.00	0.04	**0.08**	0.02
Rooiels	0.05	**0.07**	0.04	**0.04**	–	**0.04**	**0.02**	**0.05**	**0.10**	**0.08**	0.02	**0.04**	**0.12**	**0.04**
Betty's Bay	**0.06**	**0.07**	0.01	0.00	0.01	–	0.01	0.02	0.01	0.02	0.02	**0.03**	**0.05**	0.02
Gansbaai	**0.11**	**0.08**	**0.07**	0.04	**0.06**	0.02	–	0.00	**0.04**	**0.01**	0.00	0.01	**0.04**	0.02
Cape Agulhas	**0.15**	**0.22**	**0.14**	**0.1**	**0.15**	**0.13**	**0.18**	–	**0.03**	0.00	0.01	0.00	**0.04**	0.04
De Hoop	**0.16**	**0.20**	**0.16**	**0.13**	**0.17**	**0.16**	**0.18**	0.00	–	**0.03**	**0.06**	**0.04**	**0.06**	**0.08**
Cape Infanta	**0.11**	**0.12**	**0.08**	**0.09**	**0.10**	**0.10**	**0.12**	0.00	0.05	–	0.02	0.02	**0.06**	**0.08**
Herold's Bay	**0.09**	**0.08**	**0.07**	**0.09**	**0.10**	**0.07**	**0.08**	0.05	**0.08**	0.00	–	0.03	**0.04**	0.00
Knysna	**0.16**	**0.15**	**0.11**	**0.10**	**0.10**	**0.08**	**0.13**	0.07	**0.13**	0.00	0.00	–	**0.10**	**0.08**
Port Alfred	**0.71**	**0.76**	**0.66**	**0.54**	**0.56**	**0.58**	**0.72**	**0.58**	**0.54**	**0.61**	**0.66**	**0.73**	–	0.02
Gonubie/ Haga Haga	**0.69**	**0.77**	**0.63**	**0.48**	**0.52**	**0.53**	**0.71**	**0.52**	**0.49**	**0.57**	**0.64**	**0.74**	0.01	–

Table 5 Relative migration rate values (*Nm*) between each population pair for the stepping-stone migration model along the South African coast. Values in brackets represent the 95% confidence values

From population	To population	*Nm*
1 (Jacobsbaai)	2 (Kommetjie)	0 (0–0)
2 (Kommetjie)	1 (Jacobsbaai)	28 (14.5–54)
2 (Kommetjie)	3 (Sea Point)	0 (0–0)
3 (Sea Point)	2 (Kommetjie)	40.5 (16–64)
3 (Sea Point)	4 (Wooley's Pool)	0 (0–0)
4 (Wooley's Pool)	3 (Sea Point)	4 (2–22)
4 (Wooley's Pool)	5 (Rooiels)	0 (0–0)
5 (Rooiels)	4 (Wooley's Pool)	24 (16–101)
5 (Rooiels)	6 (Betty's Bay)	23.5 (12–42)
6 (Betty's Bay)	5 (Rooiels)	0.5 (0.06–8)
6 (Betty's Bay)	7 (Gansbaai)	5.5 (0.5–27)
7 (Gansbaai)	6 (Betty's Bay)	5.5 (4–59)
7 (Gansbaai)	8 (Cape Agulhas)	0 (0–55)
8 (Cape Agulhas)	7 (Gansbaai)	3.5 (1.2–26)
8 (Cape Agulhas)	9 (De Hoop)	3 (1.8–13)
9 (De Hoop) 8	8 (Cape Agulhas)	63 (14–201)
9 (De Hoop) 10	10 (Cape Infanta)	12.5 (8–21)
10 (Cape Infanta)	9 (De Hoop)	0 (0–0)
10 (Cape Infanta)	11 (Herold's Bay)	18 (12–65)
11 (Herold's Bay)	10 (Cape Infanta)	1.5 (0.9–4.5)
11 (Herold's Bay)	12 (Knysna)	102.5 (61.5–1138)
12 (Knysna) 11	11 (Herold's Bay)	0 (0–2.5)
12 (Knysna) 13	13 (Port Alfred)	0.2 (0.01–8)
13 (Port Alfred)	12 (Knysna)	0 (0–0)
13 (Port Alfred)	14 (Haga Haga)	16.5 (8–241)
14 (Haga Haga)	13 (Port Alfred)	0 (0–0)

them. Notably, gene flow is again asymmetrical on the east coast, with migrants going from Port Alfred to Haga Haga, but not in the opposite direction (Fig. 3). Using the 10×10^6 generation run from MDIV, results clearly support that the east coast population is effectively isolated (Fig. 4). Using $T_{div} = T\theta/2\mu$, we estimate the divergence time between the east coast populations and all others to at least 68 000 years ago. TMRCA was estimated to at least 165 000 years ago.

Discussion

Population structure and gene flow patterns of Clinus cottoides *in South Africa*

Understanding population structure relative to the distribution and habitat selection of marine organisms can help identify potential barriers to gene flow or 'genetic breaks'. South Africa has an extensive coastline where various biotic and abiotic factors combine to form variable and dynamic ecosystems, yet whether there are any effective barriers to gene flow on the southern African coast remains unclear.

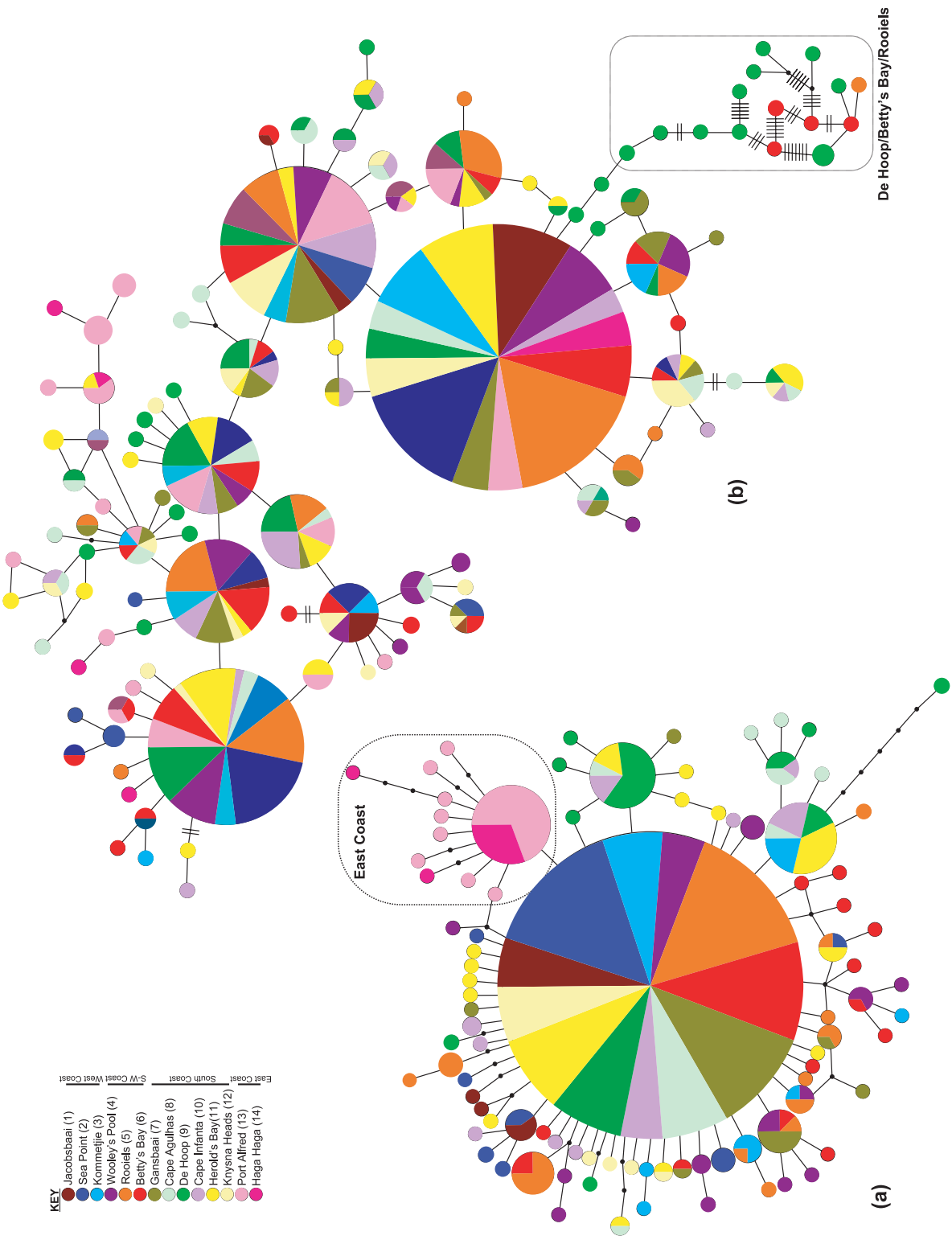

Fig. 2 Haplotype network for the combined control region and ND2 mitochondrial DNA data set (a) and the S7 intron data set (b). The size of the circles is proportional to the frequency of each haplotype; the smallest circles represent unsampled or extinct haplotypes. Each line represents one mutational step. Dotted lines represent alternative one-step mutations. A sub-network of east coast haplotypes is outlined (2a) and a divergent sub-network of alleles from De Hoop, Cape Hangklip and Betty's Bay are highlighted (2b).

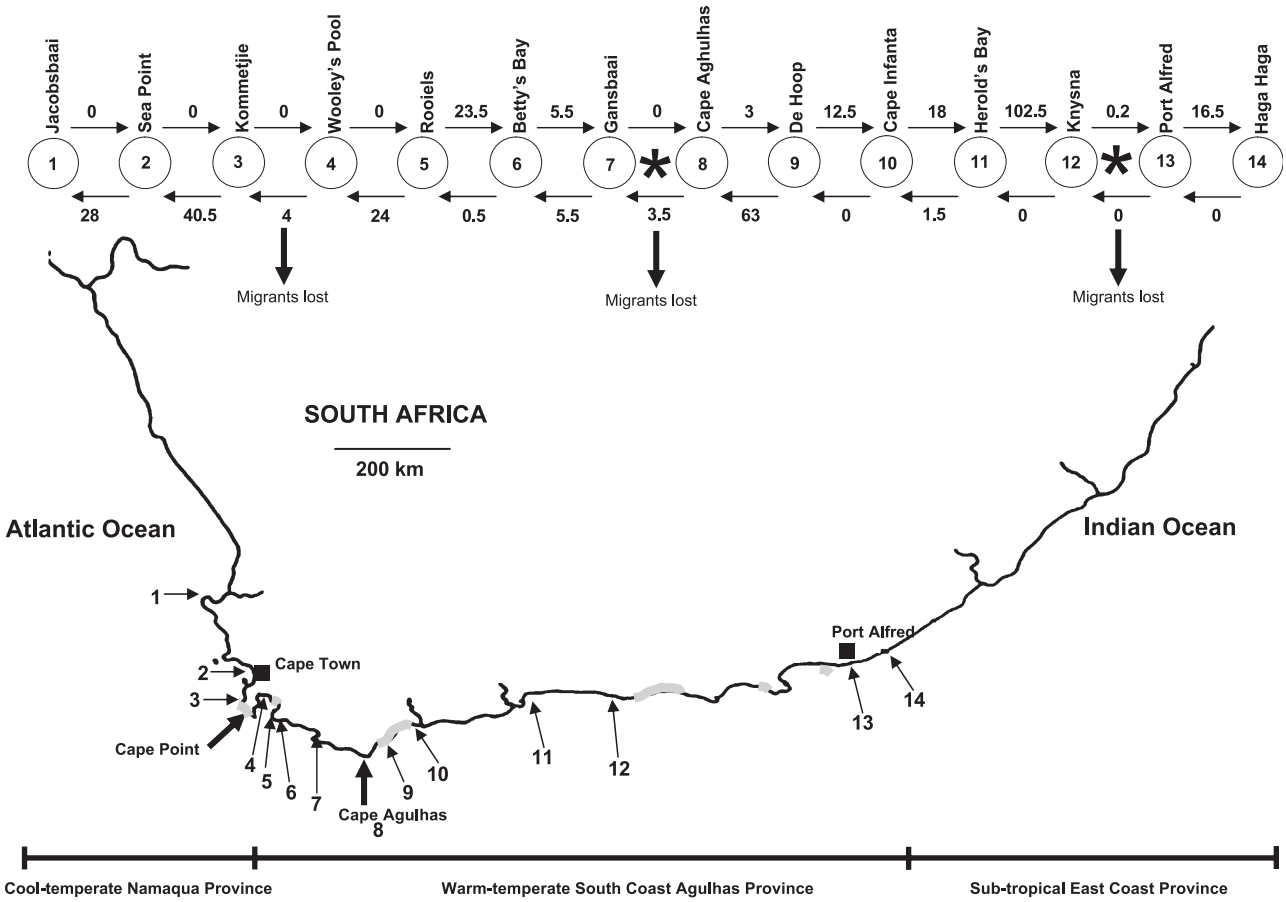

Fig. 3 Map of the South African coastline showing directionality of gene flow for *Clinus cottoides*. The three major marine biogeographical regions of the South African coast (see text) are indicated, as are all the 'no-take' MPA's within the distributional range of *C. cottoides* (solid grey lines). There are some minor 'no-take' zones in False Bay (near Wooley's Pool; sample point 4) and around Cape Point; these have not been shown because of their small size. Relative directionality of gene flow is indicated between populations by arrows. Numbers above and below the arrows denote the relative migration rates. Asterisks denote the three putative barriers to gene flow identified in this study.

AMOVA results of all gene data sets comparing the 14 sampled sites for *Clinus cottoides* clearly show significant structuring among localities (data not shown). If one measures a barrier to gene flow numerically using Φ_{ST} values, then the strongest barrier identified in this study lies between Knysna and the east coast samples ($\Phi_{ST} = 0.74$). SAMOVA analyses also group the east coast samples separately to all others ($F_{CT} = 0.51$). This is further supported by coalescent analyses of migration that suggests there is no gene flow from Knysna east, and zero gene flow from the east coast to areas west of there (Nm = 0–0.2; Fig. 3, Table 5). Moving from west to east along the South African coastline the intertidal substrate becomes progressively less rocky and is interspersed with numerous sandy beaches. The presence of these large tracts of sandy beach limits the available habitat of rock pool inhabiting *C. cottoides*, thereby preventing effective dispersal. Additionally, Maree *et al*. (2000) suggest that for estuarine fish there is a temperature related biogeographical break in the Algoa Bay region (Fig. 1). In this region, the Agulhas Current begins to deflect away from the coast and warm water is replaced by cooler south coast water (Lutjeharms & Ansorge 2001; Fig. 1). Furthermore, isolated upwelling events in the Algoa Bay region and along the south coast also occur, causing dramatic changes in water temperature and affecting marine biota distributions (Schumann *et al*. 1988), thereby providing further barriers to gene flow in the region. Coalescent analyses suggest that divergence of the east coast populations has been occurring for at least 68 000 years.

Aside from significant genetic structuring between east and south/west coast fishes, mtDNA analyses also recovered a second barrier between all sampled sites west of Gansbaai and those to the east of Cape Agulhas, suggesting that the area and oceanic conditions to the west of Cape Agulhas may act as an effective barrier to the free dispersal of *C. cottoides*. Coalescent analyses suggest that there is little gene flow between Cape Agulhas and Gansbaai to the west, compared to gene flow between Cape Agulhas and

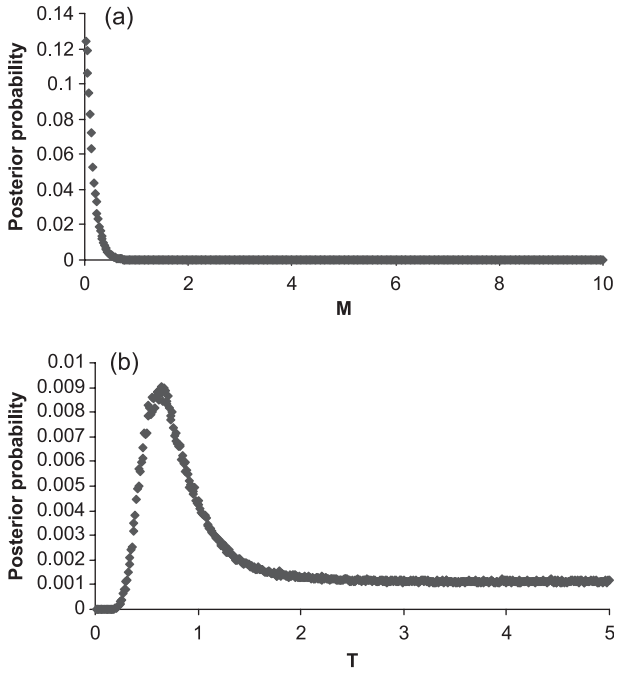

Fig. 4 Posterior probabilities for migration (a; M) and divergence time (b; T) between the east coast and all other populations of *C. cottoides*.

De Hoop to the east (Fig. 3; Table 5). The presence of some shared haplotypes between Gansbaai and Cape Agulhas does imply some, if little, migration between sites. However, there are two haplotypes only found in sampling locations to the east of Cape Agulhas (excluding those from Port Alfred and Haga Haga), and one haplotype that was only found in fish sampled from Wooley's Pool to Gansbaai. Interestingly, biological data also supports Cape Agulhas as a region of discontinuity; Harrison (2002) found that for estuarine fish, Cape Agulhas was a significant biogeographical boundary.

As the Agulhas current retroflects away from the South African coast, it may also represent a formidable barrier to the eastward movement of marine organisms (Lutjeharms & Ansorge 2001; Fig. 1), whereas those organisms carried westward by the current could be moved away from the coast into open water. However, results suggest that dispersal of *C. cottoides* is not affected to a large extent by the large-scale Agulhas current. Instead, they probably utilise smaller counter-currents that develop inshore of the Agulhas Current, which are generated by clockwise eddies. These are known to carry marine animals northwards, which for example happens during the Natal Sardine run (Shillington 1995). Given that gene flow from De Hoop occurs only from west to east (Fig. 3, Table 5), it provides good evidence that *C. cottoides* utilises the smaller inshore currents for dispersal and that such a counter-current probably originates in the De Hoop area on the south coast. Notably, this effect continues throughout the eastern distributional range of *C. cottoides*; gene flow between the east coast sampling sites is also asymmetrical with migrants from Port Alfred to Haga Haga, but none vice versa.

Cape Point (Fig. 1) has also been shown to be a major zone of transition for coastal fishes (Turpie *et al*. 2000) and for invertebrates (Branch *et al*. 1994). There is evidence for population genetic structuring between samples from the west coast and Wooley's Pool (Table 4). Coalescent analyses of migration again reveal that gene flow is asymmetrical, with no migrants going to Wooley's Pool and into False Bay, but some migration to the west coast from Wooley's Pool, which is probably influenced by the Benguela Current. Migration rates are similar in magnitude to gene flow between Cape Agulhas and Gansbaai, which do show differentiation and it is likely that the west coast sites are slowly becoming more isolated by means of genetic drift. The extensive rock habitat in False Bay and along the west coast provides ample habitat for *C. cottoides* and hence, population sizes are likely to be large. Consequently, any fixation of neutral alleles is likely to take a long time even when gene flow is limited.

One possible mechanism that could explain how limited gene flow around Cape Point may occur relates to the formation of warm and cold core-eddies due to volume flux of the Agulhas current that subsequently drift into the southern Atlantic Ocean (Lutjeharms & Ballegooyen 1988). This leads to a substantial transfer of water (and possibly juvenile fishes) from the Indian to the Atlantic Ocean (Fig. 1). Another possible mechanism of gene flow around Cape Point is reported by Roy *et al*. (2001): where upwelling normally associated with the Cape coastline was adversely affected by a Pacific Ocean La Nina event. If cold upwelled water provides a suitable barrier to the dispersal of *C. cottoides*, then the absence of upwelling could potentially allow movement of these fishes on rare occasions.

Contradictory to the mtDNA data, many S7 intron alleles are shared between sites sampled from the east and west coast, as evidenced by the haplotype network and the Φ_{ST} and SAMOVA analyses. These shared S7 intron alleles could be explained by ancestral polymorphism and incomplete lineage sorting between populations due to the high number of single and diverse alleles. Under null expectations, lineage sorting in the mtDNA should be completed first, as mtDNA mutates faster and because mtDNA loci represent only a quarter of the effective population size of biparentally inherited loci. As a consequence, genetic drift will occur more rapidly relative to the nuclear genome (see Zhang & Hewitt 2003; Zink & Barrowclough 2008). Furthermore, selection on mtDNA has also been shown to be an important force in shaping population structure and genetic diversity (Moyer *et al*. 2005; Bazin *et al*. 2006; Grant *et al*. 2006). An alternative explanation would be that male and female post-flexion larvae or adult fish have different

dispersal strategies with male larvae/fish being more dispersive, although there is no evidence for this.

Demographic history of C. cottoides

Grant & Bowen (1998) interpret four basic scenarios of population history based on values of haplotype (h) and nucleotide diversities (π) of marine fishes. They propose that fish with high h and low π probably underwent population expansion after a period of low effective population size. To test for mutation–drift equilibrium, Fu's F_S test was carried out for all four data sets. For each of the data sets, significant negative values were obtained, which strongly suggests recent demographic change ($F_S = -25.6$–34, $P < 0.001$). To further investigate demography in *C. cottoides*, mismatch distributions were calculated for the combined mtDNA and the intron data set (data not shown), which suggest recent population expansion. The starlike haplotype network for the combined mtDNA network (Fig. 2) also reflects the expectations of population expansion. The centrally placed and most common haplotype 2 is probably the haplotype ancestral to the many low frequency haplotypes to which it is closely related (Castelloe & Templeton 1994).

The parsimony network for the intron data is more complex than the starlike mtDNA network, being comprised of a number of high frequency alleles, with a number of alternative connections between alleles. Interestingly, some alleles from De Hoop and a few from the Rooiels/Betty's Bay region form a divergent sub-network that is at least six steps away from the main network, which suggest that *C. cottoides* at De Hoop are genetically more diverse than at other sampling localities, which is also supported by De Hoop's high haplotype and nucleotide diversity (Table 3).

Interestingly, Fu's F_S analyses for both mtDNA and intron data (Table 3) for every single sampling locality suggest recent population expansions, except for Jacobsbaai on the west coast and Haga Haga on the east coast, both of which are located at the edge of the distribution of *C. cottoides* (Fig. 1). Explanations for this include the lack of suitable habitat and/or environmental stress (temperatures that are sub-optimal for population growth), which may result in an allee effect limiting population growth further. Alternatively, *C. cottoides* may be expanding into new localities; gene flow analyses show gene flow to the two localities, but not away from them, suggesting a possible founder effect. However, founder populations are often characterised by low h and π (Grant & Bowen 1998) and neither of these peripheral populations have genetic diversity values significantly lower than from the other sampled areas. It is therefore unlikely that the populations (Jacobsbaai, Haga Haga) at the edge of the distribution of *C. cottoides* are the result of a founder effect, but rather represent populations at the limit of environmental endurance and hence possible sinks. As the west coast is predominantly a rocky shore environment, it is less likely that lack of suitable habitat plays a large role in limiting population growth and distributional limits are probably set as a consequence of the cooling temperature of the cold Benguela Current, coupled to seasonal cold-water upwelling. On the east coast, lack of suitable rocky shore habitat, as well as temperature stress from warmer waters could be important in limiting the distribution of *C. cottoides*.

Can we apply genetic data as useful tools in southern African MPA management?

The decision-making process as to the placement of marine protected areas and their level of protection is often based on a variety of considerations, including biological, physical, economic and socio-political factors, and as such all theoretically contribute towards conservation success (Lubchenco *et al*. 2003; Lourie & Vincent 2004). Several genetic papers examining exploited marine resources have made management recommendations based on population genetic structure and gene flow patterns (e.g. Bird *et al*. 2007; Ridgway *et al*. 2008), yet it appears that only a few marine protected areas are utilising this information (see Lourie & Vincent 2004).

In South Africa, management decisions on the placement of marine protected areas have predominantly taken into account the preservation of coastal environments and fisheries (Attwood *et al*. 1997). Although areas of high species richness were especially important in deciding the placement of MPAs, many were established to protect exploited linefish species. A preliminary assessment by Lombard *et al*. (unpublished data) concluded that with the addition of new data sets conservation prioritisation may differ (i.e. seven new areas were identified to be in need of further protection). Should population genetic analyses complement biological and oceanographical data in helping to improve marine conservation planning? We would argue yes, because molecular analyses have the benefit of recovering structure where morphology or habitat changes are unable to detect changes in populations, which is especially important in South Africa where many organisms span at least 3600 km of coastline.

Our *C. cottoides* data set spans most of South Africa's three major marine bioregions and in combination with other studies provides a preliminary assessment of the suitability of the MPA network currently in place. Interestingly, our genetic data clearly highlights the importance of the De Hoop MPA as a source to both west and east coast areas, although it receives few migrants (Fig. 3). The area around Cape Agulhas, which is currently unprotected, clearly plays an important role as a barrier to gene flow, although this differs between organisms. For *C. cottoides*,

the break lies to the west of Cape Agulhas, whereas for another clinid species *Muraenoclinus dorsalis*, Cape Agulhas is the contact zone between east and south coast fishes and probably the source population for a recent east coast invasion (von der Heyden *et al.*, unpublished data). Abalone *Haliotis midae* also show a founder event to the east of Cape Agulhas, with a recent range expansion (Evans *et al*. 2004).

It appears from published data that Cape Agulhas is more important for genetic partitioning than is Cape Point, although the latter is accepted as the major faunistic break between the cool-temperate Namaqua biogeographical province (west coast) and the warm-temperate south coast Agulhas Province (Fig. 1, Emanuel *et al*. 1992; Branch *et al*. 1994; Awad *et al*. 2002). We show here that Cape Point plays an important role in *C. cottoides* by limiting gene flow between sites on the west coast and False Bay (Fig. 3).

The mtDNA and intron data also show that the south coast area to the east of Cape Agulhas, between De Hoop and Herold's Bay, is the genetically most diverse, with haplotype and nucleotide diversities being the highest of any of the regions. Lombard *et al*. (unpublished data), highlight the coastal stretch in the vicinity of Mossel Bay as an area requiring protection, which falls into this region of high diversity. Given this and Teske *et al*.'s (2007) results that recovered a genetic break for the estuarine *Iphinoe truncata* in this area, it may, with additional genetic data become apparent that some of this coastline will require protection. This area is a popular holiday destination and experiences seasonal increases in not only fishing and exploitation, but also recreational damage and pollution. Thus, this section of coastline warrants protection.

Another such area is that between Rooiels, Cape Hangklip and Betty's Bay. Although a marine reserve exists at Betty's Bay, it does allow recreational fishing and is often used for fishing competitions. Our data suggest that Rooiels is a source population for both Betty's Bay to the east and Wooley's Pool to the west, although like De Hoop, it receives few migrants. This area is an important fishing ground for rock lobster and Matthee *et al*. (2007) showed that *Jasus lalandii*, sampled at Cape Hangklip had much higher genetic diversity indices in relation to other sampling areas. Interestingly, Lombard *et al*. (unpublished data), based on distributional data also highlighted this zone as in need of protection and current discussion on the proposed Kogelberg Marine Park is therefore warranted.

In the light of this and previous studies, genetic analyses show that along the South African coastline, the intertidal MPA network is unconnected, and that several genetic barriers exist. An exception to this is the goby, *Caffrogobius caffer*, for which no structure was detected (Neethling *et al*., in press) for mtDNA between Wooley's Pool and Haga Haga.

With the addition of molecular data for other intertidal marine species and in conjunction with available biological data, it may become possible to identify areas of importance that could potentially be included in a connected MPA network to safeguard not only exploited species, but whole ecosystems. One region that has been identified as requiring urgent protection is the west coast, where north of Cape Town no MPA exists (other than the Langebaan Lagoon MPA). This area has been highlighted as an urgent conservation concern (Lombard *et al*. unpublished data) and research priority (Marine Coastal Management 2008). Interestingly, data for three species of clinid fishes (von der Heyden *et al*. unpublished data), suggest significant population structuring along the west coast and it is hoped that molecular data will contribute significantly to the protection of this and other southern African regions.

Acknowledgements

SvdH was supported by a University of Stellenbosch Postdoctoral Fellowship and a Free-standing Postodoctoral Fellowship from the South African National Research Foundation. This project was funded by a grant to KP, RCKB and W. Holleman from the NRF-SA (GUN2053548) and a Promising Young Scientist's award from Stellenbosch University to RCKB. We thank Marine and Coastal Management for permits to collect fish along the South African coast and thank C.J. von der Heyden, V. Bowie, C. and J. Green, G. Love and G. Zsilavecz for help with collecting samples. We are especially indebted to W. Holleman for logistical support whilst sampling on the east coast. Thank you also to G. Bernardi for providing some tissue samples, to V. Bowie with help in generating the maps, and to V. Almada and an anonymous reviewer for their comments, which helped improve an earlier version of this manuscript.

References

Attwood CG, Mann BQ, Beaumont J *et al*. (1997) Review of the state of marine protected areas in South Africa. *South African Journal of Marine Science*, **18**, 341–367.

Awad AA, Griffiths CL, Turpie JK (2002) Distribution of South African marine benthic invertebrates applied to the selection of priority conservation areas. *Diversity and Distributions*, **8**, 129–145.

Bazin E, Glémin S, Gaultier N (2006) Population size does not influence mitochondrial genetic diversity in animals. *Science*, **312**, 570–572.

Beerli P, Felsenstein J (1999) Maximum likelihood estimation of a migration matrix and effective population sizes in two populations using a coalescent approach. *Genetics*, **152**, 763–773.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences, USA*, **98**, 4563–4568.

Bernardi G (2000) Barriers to gene flow in *Embiotoca jacksoni*, a marine fish lacking a pelagic larval stage. *Evolution*, **54**, 226–237.

Bernardi G, Vagelli A (2004) Population structure in Banggai cardinalfish, *Pterapogon kauderni*, a coral reef species lacking a pelagic larval phase. *Marine Biology*, **145**, 803–810.

Bird CE, Holland BS, Bowen BW, Toonen RJ (2007) Contrasting phylogeography in three endemic Hawaiian limpets (*Cellana* spp.) with similar life histories. *Molecular Ecology*, **16**, 3173–3186.

Bowie RCK, Fjeldså J, Hackett SJ, Bates JM, Crowe TM (2006) Coalescent models reveal the relative roles of ancestral polymorphism, vicariance and dispersal in shaping phylogeographical structure of an African montane forest robin. *Molecular Phylogenetics and Evolution*, **38**, 171–188.

Branch GM, Griffiths CL, Branch ML, Beckley LE (1994) *Two Oceans — A Guide to the Marine Life of Southern Africa*. David Philip, Cape Town & Johannesburg.

Castelloe J, Templeton AR (1994) Root probabilities for intraspecific gene trees under neutral coalescent theory. *Molecular Phylogenetics and Evolution*, **3**, 102–113.

Chow S, Hazama K (1998) Universal PCR primers for S7 ribosomal gene introns in fish. *Molecular Ecology*, **7**, 1247–1263.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.

Cowen RK, Paris CB, Srinivasan A (2006) Scaling of connectivity in marine populations. *Science*, **311**, 522–527.

Coyer JA (1982) Observations on the reproductive behavior of the giant kelpfish, *Heterostichus rostratus* (Pisces: Clinidae). *Copeia*, **1982**, 344–350.

Donaldson KA, Wilson RW (1999) Amphi-panamaic germinates of snook (Percoidei: Centropomidae) provide a calibration of the divergence rate in the mitochondrial DNA control region of fishes. *Molecular Phylogenetics and Evolution*, **13**, 208–213.

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, **11**, 2571–2581.

Emanuel BP, Bustamante RH, Branch GM, Eekhout S, Odendaal FJ (1992) A zoogeographic and functional approach to the selection of marine reserves on the west coast of South Africa. In: *Benguela Trophic Functioning* (eds Payne AIL, Brink KH, Mann KH, Hilborn R), vol. 12, pp. 341–354. South African Journal of Marine Science, Grahamstown, South Africa.

Evans BS, Sweijd NA, Bowie RCK, Cook PA, Elliott NG (2004) Population genetic structure of the perlemoen *Haliotis midae* in South Africa: evidence of range expansion and founder effects. *Marine Ecology Progress Series*, **270**, 163–172.

Excoffier L, Laval G, Schneider S (2005) Arlequin version 3: an integrated software package for population genetics data analysis. *Evolutionary Biology Online*, **1**, 47–50.

Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Grant WS, Bowen BW (1998) Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity*, **89**, 415–126.

Grant WS, Spies I, Canino MF (2006) Biogeographic evidence for selection on mitochondrial DNA in North Pacific Walleye Pollock *Theragra chalcogramma*. *Journal of Heredity*, **97**, 571–580.

Hall TA (1999) BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

Harpending RC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591–600.

Harrison TD (2002) Preliminary assessment of the biogeography of fishes in South African estuaries. *Marine Freshwater Research*, **53**, 479–490.

Lee W-J, Conroy J, Howell WH, Kocher TD (1995) Structure and evolution of teleost mitochondrial control regions. *Journal of Molecular Evolution*, **41**, 54–66.

Leis JM, Hay AC, Trnski T (2005) In situ ontogeny behaviour in pelagic larvae of three temperate, marine, demeral fishes. *Marine Biology*, **148**, 655–669.

Lourie SA, Vincent ACJ (2004) Using biogeography to help set priorities in marine conservation. *Conservation Biology*, **18**, 1004–1020.

Lubchenco J, Palumbi SR, Gaines SD, Andelman S (2003) Plugging a hole in the ocean: the emerging science of marine reserves. *Ecological Applications*, **13**, S3–S7.

Lutjeharms JRE, Ansorge IJ (2001) The Agulhas return current. *Journal of Marine Systems*, **30**, 115–138.

Lutjeharms JRE, Ballegooyen RC (1988) Anomalous upstream retroflection in the Agulhas Current. *Science*, **240**, 1770–1772.

Maree RC, Whitfield AK, Booth AJ (2000) Effect of water temperature on the biogeography of South African estuarine fishes associated with the subtropical/warm temperate subtraction zone. *South African Journal of Science*, **96**, 184–188.

Matthee CA, Cockroft AC, Gopal K, von der Heyden S (2007) Mitochondrial DNA, variation of the west coast rock lobster, *Jasus lalandii*: marked genetic diversity differences among sampling sites. *Marine Freshwater Research*, **58**, 1130–1135.

Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology*, **51**, 238–254.

Moyer GR, Winemiller KO, McPhee MV, Turner TF (2005) Historical demography, selection, and coalescence of mitochondrial and nuclear genes in *Prochilodus* species of northern South America. *Evolution*, **59**, 599–610.

Neethling M, Matthee CA, Bowie RCK, von der Heyden S (in press) Evidence of successful dispersal across a major oceanographic barrier in the southem African endemic, Caffrogobius caffer (Teleostei: Gobiidae). *BMC Evolutionary Biology*.

Nielsen R, Wakeley JW (2001) Distinguishing migration from isolation: an MCMC approach. *Genetics*, **158**, 885–896.

Planes S, Doherty PJ, Bernardi G (2001) Strong genetic divergence among populations of a marine fish with limited dispersal, *Acanthochromis polyacanthus*, within the Great Barrier Reef and the Coral Sea. *Evolution*, **55**, 2263–2273.

Posada D, Crandall KA (2001) Intraspecific gene genealogies: tree grafting into cladograms. *Trends in Ecology & Evolution*, **16**, 37–45.

Prochazka K, Griffiths CL (1992) The intertidal fish fauna of the west coast of South Africa — species, community and biogeographic patterns. *South African Journal of Zoology*, **27**, 115–120.

Ridgway T, Riginos C, Davis J, Hoegh-Guldberg O (2008) Genetic connectivity patterns of *Pocillopora verrucosa* in southern African marine protected areas. *Marine Ecology Progress Series*, **354**, 161–168.

Rocha LA, Craig MT, Bowen BW (2007) Phylogeography and the conservation of coral reef fishes. *Coral Reefs*, **26**, 501–512.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Roy C, Weeks S, Rouault M *et al*. (2001) Extreme oceanographic events recorded in the southern Benguela during the 1999–2000 summer season. *South African Journal of Science*, **97**, 465–471.

Schumann EH, Ross GJB, Goschen WS (1988) Cold water events in Algoa Bay and along the Cape south coast, South Africa, in March/April 1987. *South African Journal of Science*, **84**, 579–584.

Shillington FA (1995) Oceanography of the southern African region. In: *Smiths' Sea Fishes* (eds Smith MM, Heemstra PC), pp. 22–28. Struik Publishers, Cape Town.

Smith MM (1995) Clinidae. In: *Smiths' Sea Fishes* (eds Smith MM, Heemstra PC), pp. 758–769. Struik Publishers, Cape Town.

Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing data imputation. *American Journal of Human Genetics*, **76**, 449–462.

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978–989.

Stepien CA (1992) Evolution and biogeography of the Clinidae (Teleostei: Blennioidei). *Copeia*, **1992**, 375–392.

Teske PR, McQuaid CD, Froneman PW, Barker NP (2006) Impacts of marine biogeographic boundaries on phylogeographic patterns of three South African estuarine crustaceans. *Marine Ecology Progress Series*, **314**, 283–293.

Turpie JK, Beckley LE, Katua SM (2000) Biogeography and the selection of priority areas for conservation of South African coastal fishes. *Biological Conservation*, **92**, 59–72.

von der Heyden S, Groeneveld JC, Matthee CA (2007) Long current to nowhere? – Genetic connectivity of *Jasus tristani* in the southern Atlantic. *African Journal of Marine Science*, **29**, 491–497.

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, **89**, 438–450.

Zhang D-X, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, **12**, 563–584.

Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology*, doi:10.1111/j.1365-294X.2008.03737.x.

Zsilavecz G (2005) *Coastal Fishes of the Cape Peninsula and False Bay: A Divers' Identification Guide*. Southern Underwater Research Group, Cape Town.

Sophie von der Heyden is currently a postdoctoral fellow with a strong interest in the phylogeny and phylogeography of southern African marine organisms. She is especially interested in how reproductive strategies and oceanography interplay to shape the population genetic structure of intertidal species. Kim Prochazka currently oversees fisheries-related research and has an enduring interest in fish ecology and the ecology of rocky shores. Rauri Bowie is an assistant professor in the Department of Integrative Biology and Curator of Birds at the Museum of Vertebrate Zoology at the University of California-Berkeley. His interests include biogeography, systematics and speciation processes in birds and marine organisms, with a particular emphasis on African taxa.
