Chapter 2

Mitochondrial DNA variation in South African dusky kob, Argyrosomus japonicus (Temminck & Schlegel, 1843)

Abstract

Argyrosomus japonicus is abundant from Cape Agulhas to northern KwaZulu-Natal in South African coastal waters. The species is also found along Australia's southern seaboard, from Hong Kong northwards in the South China Sea to southern Korea and Japan and in the northern Indian Ocean off Oman, Pakistan and India. In South Africa *A. japonicus* juveniles are estuarine, while subadults and adults inhabit inshore waters from the surf zone to a depth of 100 meters. Juveniles remain close to their natal estuaries until they reach sexual maturity. Indications are that spawning commences in KwaZulu-Natal from August to November and continues south to the Cape region. Based on otolith dimensions and fish length relationships, it has been suggested that the South African population of *A. japonicus* consists of several allopatric juvenile subpopulations and a single offshore migratory adult population. Analysis of mitochondrial DNA shows that the genetic diversity (mean $h = 0.95$, $\pi = 0.89\%$) compares well with other marine fish species, with 35 haplotypes detected from the 133 control region sequences analysed. Haplotype relationships showed no clear geographic pattern. Analysis of Molecular Variance allocated more than 99% of the variance to within localities and no significant F_{ST} values were found. Lack of genetic differentiation and high maternal gene flow between localities would be in concordance with morphometric and otolith data, which indicate a single South African intermixing adult stock. On the other hand, the isolation-by-distance observed between fish from the southern and Eastern Cape localities and those of KwaZulu-Natal cannot be ignored and deserves further investigation.

Introduction

The dusky kob *Argyrosomus japonicus* (Temminck & Schlegel, 1843) is a large sciaenid with a fragmented global distribution. Along South Africa's coast it is found from Cape Point to southern Mozambique, although only abundant from Cape Agulhas to northern KwaZulu-Natal (KZN) (Fig. 2.1). The species also occurs off the Australian west and south coast and along the South China Sea from Hong Kong to southern Korea and Japan (the type locality). It is also known from Oman, Pakistan and India (Heemstra & Heemstra, 2004). In South African waters, *A. japonicus* is a coastal species that can be found in estuaries, the surf and nearshore zone, with a maximum recorded depth of 100 meters (Griffiths, 1996b; Griffiths & Heemstra, 1995). This species utilise estuaries throughout their range, especially the larger systems such as Lake St. Lucia (Whitfield, 1994). Estuaries with turbid waters are the preferred habitat and, as an ambush predator, the species can effectively forage in murky waters using olfactory and lateral line senses (Van der Elst, 1988). Juveniles are dependent on estuaries as nursery grounds and recruitment of juveniles into estuaries occurs at 20–30 mm TL (approximately one month old). Tag-recapture data suggest that juveniles do not undertake substantial coastal movements and are likely to remain close to their natal estuaries, forming localised populations until sexual maturity is reached (Bullen & Mann, 2004a; Griffiths, 1996b; Whitfield, 1998). Interpreting tagging data for the species has been complicated by misidentification prior to the mid-1990s and, because of similarities in the morphology of *A. japonicus* and *A. inodorus*, misidentification still occurs (Bullen & Mann, 2004a).

In South African waters, prior to the mid-1990s the species was misidentified as *A. hololepidotus* (Lacepède, 1801), and was also not distinguished from the sympatric congeneric, *A. inodorus* (Griffiths & Heemstra, 1995), that co-occurs with *A. japonicus* from Algoa Bay in the Eastern Cape Province westwards to False Bay. Correct identification has proved crucial for management purposes, since the two species have very different life histories (Griffiths, 1996b, 1997c). *Argyrosomus japonicus* is a long-lived species, that can reach an age of 42 years, although individuals older than 27 years are uncommon (Griffiths & Hecht, 1995). Sexual maturity is attained at seven years of age for males and eight years for females, with both sexes having a total length of more than a meter at maturity. In contrast, *A. inodorus* reaches maturity between one and two years of age (310–375 mm TL) and rarely exceeds an age of 13 years (Griffiths, 1996a; Griffiths & Hecht, 1995).

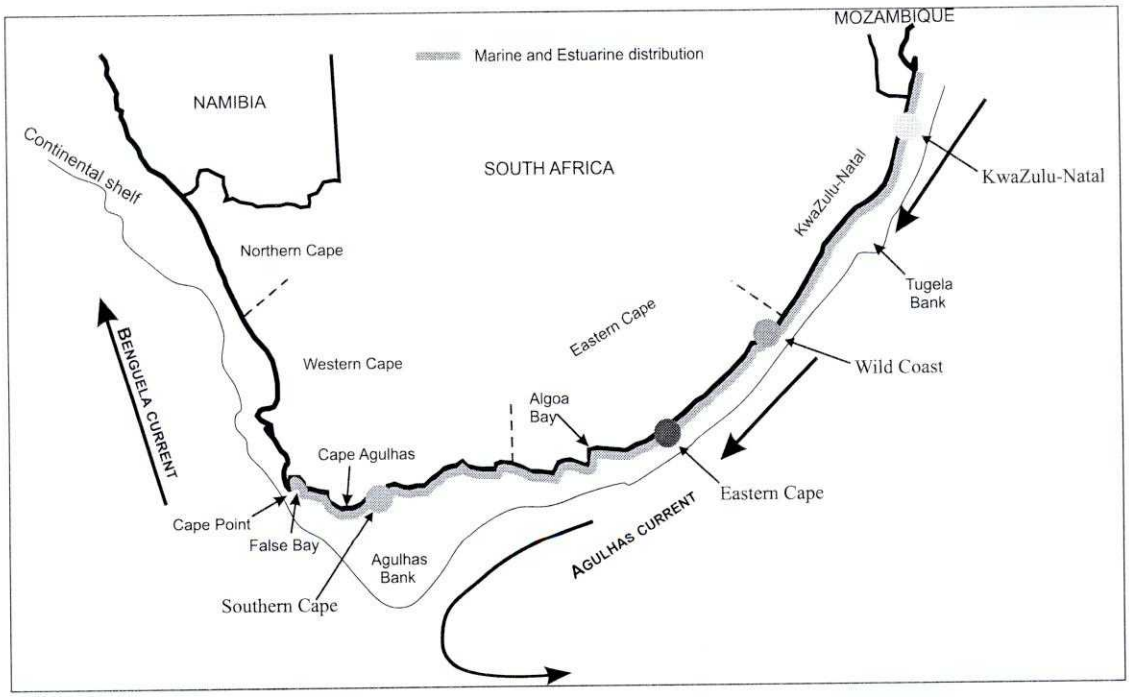

Fig. 2.1 Distribution of *A. japonicus* along South Africa's eastern and southern coast. The four sampling localities where fish were collected are indicated (adapted from Whitfield, 1998).

Increasing numbers of reproductively active *A. japonicus* in linefish catches off the KZN coast during August to November suggests that adult fish from the southeastern and southern Cape migrate to KZN waters to spawn. Spawning from October to January also occurs in the Cape region as far west as Cape Infanta (Griffiths, 1996b). This suggests that breeding commences in KZN and continues as a portion of the spawning population returns to southern waters. A general absence of individuals with spent gonads in KZN, and their presence in southeastern and southern Cape waters, also implies a return migration once spawning has been completed (Griffiths, 1996b; Griffiths & Heemstra, 1995).

Argyrosomus japonicus is one of South Africa's most important angling species and constitutes approximately 18% of the total catch (Lamberth & Turpie, 2003). Per recruit stock assessment models estimate that the South African population has been overexploited and that the spawner biomass is between 1.0 and 4.5% of pristine levels (Griffiths, 1997d; Lamberth & Turpie, 2003). Because of its life history characteristics, namely, age at sexual maturity, longevity and having evolved with relatively low rates of natural mortality, *A. japonicus* is unlikely to sustain high levels of exploitation (Griffiths 1997). There are at least three phases where it is particularly vulnerable to human activities. First, recreational and commercial lineboat fishermen target the dense spawning aggregation off the KZN coast. Second, recreational anglers have targeted juveniles that inhabit the surf zone and estuaries since the 1960s. On the Eastern Cape coast, catches are dominated by *A. japonicus* and they constitute 48% of catch

weight, 35% in KZN and 6% along the southern Cape coast. Third, although early juveniles are not exploited directly, they are estuarine dependent and most of South Africa's estuaries are degraded to various degrees (Griffiths, 1996b, 1997d; Mander *et al.*, 2001).

Based on otolith dimensions and fish length relationships, Griffiths and Hecht (1995) suggested that three or more South African *A. japonicus* stocks might exist for at least part of the species' life cycle. They concluded that the South African population of *A. japonicus* consists of a single offshore migratory adult population with several allopatric juvenile subpopulations in their natal estuaries and adjoining surf zone. Because of the migratory adults, differences among juveniles would reflect environmental conditions rather than genetic differences and it was recommended the species should be managed on a national rather than regional level (Griffiths, 1997d).

The use of the mitochondrial DNA (mtDNA) control region in detecting population structure in marine fish has been successfully applied to many species (Alvarado Bremer *et al.*, 2005; Bernardi *et al.*, 2003; Bernardi & Talley, 2000; Bernardi & Vagelli, 2004; Chenoweth & Hughes, 2003; Fauvelot *et al.*, 2003; Gold & Richardson, 1991; Graves, 1998; Ovenden *et al.*, 2002; Ovenden *et al.*, 2004; Ravago-Gotanco & Juinio-Menez, 2004; Smith *et al.*, 2002; Stefanni & Thorley, 2003). In some cases it has been of limited use due to large intra-oceanic differences and high levels of genetic diversity, e.g. in *Mugil cephalus* (Rocka-Olivares *et al.*, 2000). In the case of *Merluccius merluccius*, analysis of several mtDNA regions was not sensitive enough to reveal structure, whereas microsatellite markers were successful (Lundy *et al.*, 2000). The control region is the section of the mitochondrial genome with the highest substitution rate in most vertebrates, making it useful for population studies. Because of its haploid nature and absence of recombination, the signal obtained from genetic drift is strong and, on average, mtDNA variation is expected to be greater between populations and less within populations (Birky *et al.*, 1989; Ferris & Berg, 1987; Gold & Richardson, 1991; Meyer, 1993; Shaklee & Bentzen, 1998; Waples, 1998; Wilson *et al.*, 1985).

In this study mtDNA control region sequences were analysed to examine the levels of genetic diversity and spatial differentiation among juvenile *A. japonicus* sampled from South African waters. Sampling of estuarine dependent juveniles was important to investigate the possibility that isolated spawning events occur in a specific region or regions, to test if an isolation-by-distance effect exists or whether the South African *A. japonicus* population consist of a single breeding population.

Materials and Methods

Samples

Juvenile *A. japonicus* were collected in estuaries or the adjacent surf zones by seine netting, gillnet or hook and line. Fin clips or gill filaments were taken, placed in 80% ethanol and later stored at -20°C. Length measurements were recorded, as well as locality and date of collecting. Sampling was done at four sites across the species' distribution in South African waters (Fig. 2.1). Samples from the southern Cape were collected in the surf zone at De Hoop Nature Reserve. Eastern Cape samples were collected from the Great Fish River estuary, while the Wild Coast samples were obtained from two adjacent estuaries (Mngazi and Mngazana) with mouths four kilometres apart (samples from these two estuaries were subsequently combined for analysis). All KZN samples were collected from the St. Lucia estuary or the adjacent surf zone. The total distance along the coast between De Hoop in the southern Cape and St Lucia in KZN is 1 542 km. In addition, samples from specimens off Australia's west coast were obtained for comparative purposes. A list of all samples analysed is given in Appendix I.

DNA Extraction

Three different methods for total genomic DNA extraction were used during the course of this study.

1. Phenol-chloroform extraction: A small piece of material (0.01–0.02 g) was placed in 500 µl amniocyte buffer (50 mM Tris, pH 7.6, 100 mM NaCl, 1M EDTA, pH 8.0, 0.5%SDS) with 0.5 mg proteinase K (Roche Diagnostics). The mixture was incubated at 55°C for a minimum of 2.5 hours or until the material was digested, 1.25 µg RNAse A (Roche Diagnostics) was added and incubated at 37°C for 30 min. Next the samples were extracted twice with phenol and then twice with a 24:1 chloroform:isoamyl alcohol solution. Total DNA precipitation was done overnight at -20°C with 0.1 volumes 3M NaAc and two volumes 96% ethanol. The DNA was pelleted through centrifugation at 21 000 *g* for 30 min. The pellet was washed with 2 volumes 70% ethanol, collected through another 30 min at 21 000 *g* centrifugation step and dried. The dried DNA was resuspended in 100 µl Sabax® (Adcock Ingram) water and stored at -20°C (Sambrook *et al.*, 1989).
2. DNA was extracted with the Qiagen DNeasy® tissue kit (Qiagen Inc). A volume of 180 µl ATL buffer and 0.5 mg proteinase K was added to a small piece of material

(0.01–0.02 g) and incubated at 55°C for three hours. This was followed by a RNA digestion with 0.1 mg RNAse A for one hour at 37°C. Next 200 µl AL buffer was added to the solution and incubated at 70°C for 30 min. An equal volume of 100% ethanol was added to this and the solution pipetted into the DNeasy column. The column was centrifuged at 8 000 *g* for one minute. Five hundred microlitres of AW1 buffer was added and the column centrifuged at 8 000 *g* for one minute. Then 500 µl AW2 buffer was added and centrifuged for three minutes at 24 000 *g*. A volume of 50 µl Sabax® water was pipetted onto the membrane and incubated at room temperature for 30 min before being centrifuged for one minute at 8 000 *g*. This procedure was then repeated and the DNA collected in a new centrifuge tube. The DNA solutions were stored at -20°C.

3. The last method of DNA extraction followed Estoup *et al.* (1996). Five hundred microlitres of 10% Chelex resin (Sigma-Aldrich) solution at 60°C was added to a small piece of material (0.01–0.02 g). This was incubated at 100°C for 15 min. After adding 0.5 mg proteinase K, the solution was incubated at 55°C for one hour with constant shaking. The same solution was once again incubated for 15 min at 100°C and then left to cool down before storage at 4°C. Before the Chelex resin extractions were used in a reaction, the tubes were vortexed and centrifuged at 24 000 *g* for 20 s (Estoup *et al.*, 1996).

PCR Amplification

In each Polymerase Chain Reaction (PCR) (Saiki *et al.*, 1988), 80–100 ng of the total genomic DNA was used as template. The full length of the mtDNA control region was amplified using oligonucleotide primers Per L16565 (5' AGC GCC GGT CTT GTA AAC CG 3') and H1033 (5' CAT CTT AAC AGC TTC AGT G 3') (Fig. 2.2).

Primer Per L16565 was designed specifically for this study. Amplification and sequencing of a fragment that included the 3' end of cyt *b*, the two tRNAs, threonine ($tRNA^{thr}$) and proline ($tRNA^{pro}$), as well as the 5' end of the control region was done to achieve this. The obtained sequence was then aligned with sciaenid sequences obtained from GenBank. A conserved sequence in the $tRNA^{thr}$ was identified and selected as template for primer Per L16565. The reverse primer H1033 was developed in our laboratory based on published salmonid sequences (P. Bloomer, pers. comm.). Both primers were numbered according to corresponding nucleotide positions on reference sequence *Salmo salar* (GenBank sequence number NC 001960).

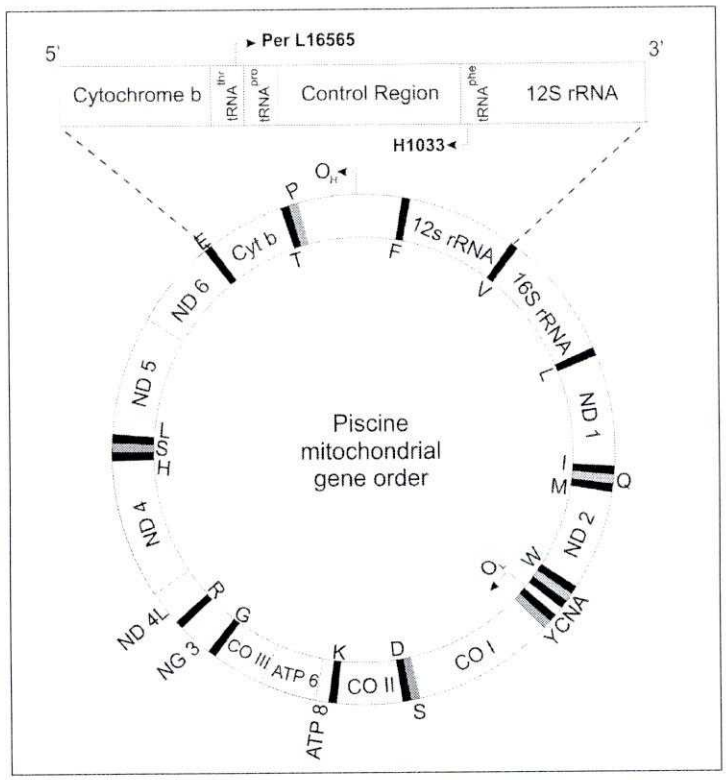

Fig. 2.2 Map of the piscine mitochondrial DNA showing the annealing positions of primers for the amplification of *A. japonicus* complete mtDNA control region (modified from Meyer, 1993).

Each double-stranded amplification reaction contained the following: 1x reaction buffer, 2.0 mM $MgCl_2$, 2 mM of each of the four nucleotides (dNTP's; Promega), 1.5 U Super-therm® DNA Polymerase (Southern Cross Biotechnology) and 10 pmols of each primer. Each reaction was made up with Sabax® water to a final volume of 50 µl. Negative controls were included each time a PCR was set up to screen for possible DNA contamination of solutions. Amplifications were done in a Geneamp® PCR System 9700 (Applied Biosystems) and started with a five minute 94°C denaturing step. This was followed by 35 cycles of denaturing at 94°C (30 s), primer annealing at 58°C (30 s) and extension at 72°C (40 s). The final extension step was done at 72°C for seven minutes. Four microlitres of the PCR product was run through a 1% agarose gel (Invitrogen), stained with ethidium bromide and visualised under UV light to verify successful amplification.

Successfully amplified PCR products were purified directly using a modified sodium acetate precipitation method (Applied Biosystems). This was done by adding 10 µl Sabax® water, 2 µl NaAc (3M) and 90 µl 100% sequencing grade ethanol to the PCR product. The mixture was left to stand at room temperature for 15 min before being centrifuged at 24 000 *g* for 20 min. The ethanol supernatant was removed and the pellet washed by adding 90 µl 70% ethanol (sequencing grade). The DNA was again pelleted by a 10 min centrifugation at 24 000 *g*. All

ethanol was removed and the DNA pellet dried before being resuspended in 15–30 µl Sabax® water. A quality check was done by running some of the resuspended DNA through a 1% agarose gel, stained with ethidium bromide and visualised under UV light.

Sequencing

For all the samples two sequencing reactions were performed, one with Per L16565 and the other with the H1033 primer. Sequencing reactions were done using the BigDye® Terminator Cycle Sequencing Kit (Applied Biosystems). Approximately 100–120 ng purified PCR template, 3.2 pmol of the specific primer and 2 µl BigDye® were mixed and made up with Sabax® water to a final volume of 10 µl and cycled in a Geneamp® PCR System 9700 using the BigDye® program. Purification of cycle sequencing products was done with the same protocol as for PCR products (see PCR amplification above). The one exception was that the pellet was not resuspended in water, but with Hi-Di™ formamide (Applied Biosystems) before loading on the ABI 3100 automated sequencer (Applied Biosystems). Electropherograms of both sequencing reactions per sample were imported into Sequence Navigator™ version 1.0.1 (Applied Biosystems, 1994), aligned and checked for ambiguous bases. The resulting consensus sequence obtained from the alignment was used in multiple sequence alignments performed in Clustal X version 1.8 (Thompson *et al.*, 1997). All substitutions and inferred insertion/deletion (indels) events were double-checked on the original electropherograms to ensure that all bases were scored accurately.

Population Genetic Analysis

Juveniles analysed in this study varied in size and were collected over an extended period from the various sampling localities. Most of the samples were obtained from fish tagged and released as part of other studies. Due to the reported high variability in length-at-age, *A. japonicus* age cannot be estimated accurately from size and should be based on otolith analysis (Griffiths & Hecht, 1995). Since otoliths were not collected, the fish could not be accurately aged. The result was that temporal stability of observed patterns of genetic variability between cohorts at the different localities could not be tested with any certainty and was thus not addressed in the present study. Juvenile fish, irrespective of size class or date collected from a given locality, were grouped to test for geographic variation among localities.

DNASP version 4.0 (Rozas *et al.*, 2003) was used to compute a sliding window of the polymorphisms observed in the analysed fragment. Aligned sequences of all individuals were used to construct an input file for ARLEQUIN 2.0 (Schneider *et al.*, 2000). Only sequences

with less than 5% missing data were used for analysis. Several genetic diversity indices were calculated. First, within sampling locality diversity was estimated by calculation of haplotype diversity (h) (Nei, 1987) and nucleotide diversity (π) (Tajima, 1983, 1993). Second, between locality diversity was calculated as net nucleotide sequence divergence (Reynolds *et al.*, 1983; Slatkin, 1995). Haplotypes and their frequencies at localities were used to construct a Minimum Spanning Network (Excoffier *et al.*, 1992) based on pairwise differences obtained from ARLEQUIN 2.0 to visually examine their genealogical relationship. The number of haplotypes that were only found at a single locality, regardless of frequency, is expressed by the endemism index. This is calculated by dividing the number of 'private' haplotypes at a locality by the total number of haplotypes at that locality.

Before any conclusions on population structure can be drawn from the data, it is important to show that the marker is evolving according to neutral expectations and thus not affected by selection. Population differentiation is assumed to be a result of reproductive isolation, but in the presence of selection this assumption becomes invalid. A mixture of test of neutrality can serve as indicators of demographic parameters like population growth/decline in the evolutionary history of the taxa (Rand, 1996). To test for neutrality of the locus, Tajima's D, Fu's F_s and the R_2 test were used. Tajima (1989) test statistic D is the difference between the average number of nucleotide differences (π) between sequences drawn from random samples and the number of segregating sites (θ). Under the neutral model, in a population at equilibrium, $\pi = \theta$. A negative value for D may either indicate a selective sweep, a population expansion after a bottleneck, or a small founder event. Positive D-values are consistent with models of positive and balancing selection, or an admixture of distinct isolated populations (Rand, 1996). Fu's F_s is a test that works well in detecting population growth for larger sample sizes (Fu, 1997; Ramos-Onsins & Rozas, 2002). F_s tends to be negative when there is an excess of recent mutations (thus an excess of rare alleles). A large negative value is taken as evidence against neutrality of mutations (Fu, 1997). Lastly, the R_2 neutrality test works well for smaller sample sizes. This test evaluates the number of singletons where lower values of R_2 are expected after a recent population expansion (Ramos-Onsins & Rozas, 2002).

Mismatch distributions for all pairwise number of differences observed were calculated and compared to those expected under a population growth and decline model for all localities and for the overall South African population (Harpending, 1994; Rogers, 1995). Homogeneity of mtDNA haplotype frequency was tested via a randomisation procedure (Monte Carlo) (Roff & Bentzen, 1989), with 10 000 permutations.

Analysis of Molecular Variance, AMOVA (Excoffier *et al.*, 1992) was executed in ARLEQUIN 2.0 (Schneider *et al.*, 2000), to determine which part of the variance can be attributed to variance between regions, and variance among individuals within regions. F_{ST} (Wright, 1951) values were used to determine the extent of genetic differentiation between localities and to estimate patterns of spatial genetic structuring. Population pairwise F_{ST} values were calculated and significance at the 5% level determined.

To determine the relationship between genetic differentiation and geography, pairwise F_{ST} values were plotted against geographic distance. A Mantel test (10 000 random permutations) (Smouse *et al.*, 1986) was used to determine the correlation between genetic differentiation (pairwise F_{ST}) and geographic distance between sampling localities, using ARLEQUIN 2.0.

Results

Control region

A fragment of 950 base pair (bp) was obtained through sequencing, including 819 bp from the control region as well as some bases from the flanking tRNA's where the primers annealed. No evidence of heteroplasmy was found since no double peaks were observed in the sequences and a consistent pattern of variability was found across all sequences. From the sliding window it is clear that the distribution of variation in the mtDNA control region of *A. japonicus* is comparable to observations from other vertebrates. Most of the variation was found within the hypervariable region I followed by hypervariable region II, with very conserved 5' (first 72 bp) and 3' (last 84 bp) terminal ends and a relatively conserved central region (Fig. 2.3).

At the 5' end of the control region, one individual had three tandem repeats of a ten bp segment, while all other individuals had only a single copy. Tandem repeats in the 5' end is commonly observed in fish species (Lee *et al.*, 1995). The repeat occurred 18 bp from the 5'end of the control region, and as this falls within the first 72 bp conserved segment, the insert and the first 18 bp of all samples were excluded from all analyses.

By mapping the mutations onto the initial Minimum Spanning Network containing 56 haplotypes, a total of 22 homoplasy characters were identified and removed before analysis to exclude any ambiguity in the relationship among the haplotypes (Avise & Lansman, 1983).

Six sites with insertion/deletions (indels) were also removed. The final fragment analysed was 735 bp in length. Nucleotide composition of the fragment was C 22%, T 30%, A 32% and G 16%, with a total of 58 polymorphic sites. The polymorphic sites contained 61 substitutions, 56 transitions at 55 sites and five transversions at five sites. An alignment of all the variable positions is shown in Table 2.1 (a complete alignment is available on request).

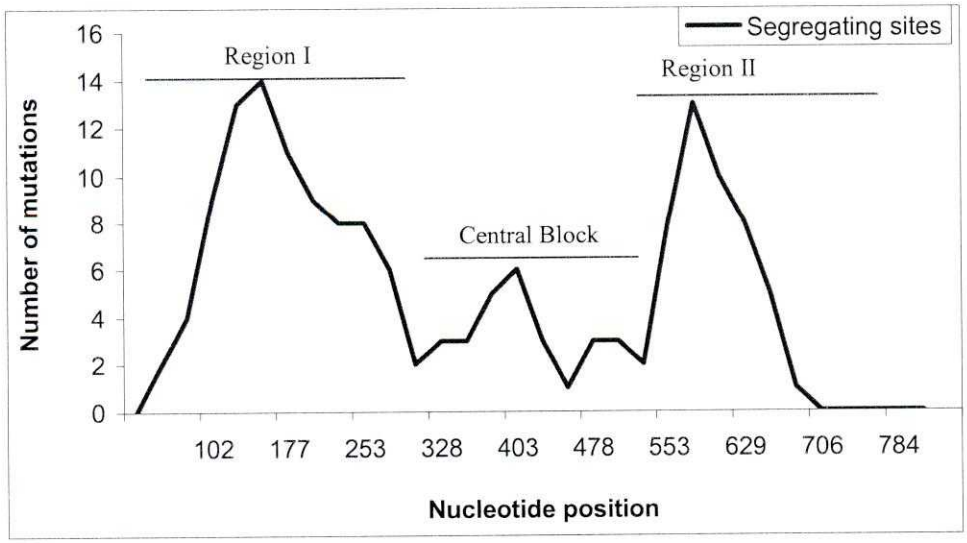

Fig. 2.3 Sliding window of substitutions in the complete mtDNA control region of South African *A. japonicus* indicating the distribution of variable sites; window length 50, step size 25 bp.

Haplotype variation and distribution

The 58 polymorphic sites defined 35 unique haplotypes among 133 individuals analysed from the four South African localities (Table 2.1). Genetic distances among haplotypes ranged from 0.14% (one bp difference) to 2.04% (15 bp difference) with a mean divergence of 1.1% (Appendix II). The most common haplotype (#11) was shared by 14 samples, followed by haplotype numbers three and eight with eleven samples each. Fourteen haplotypes (40%) were represented by a single sample (Table 2.1).

Samples analysed from Australia did not share any haplotypes with those from South Africa. The percentage divergence between the two groups varied between 7.46 and 9.03%. Because of this, the Australian samples were excluded from all the analyses performed on the South African samples.

From the Minimum Spanning Network (Fig. 2.4) it is clear that there is no relationship between haplotype genealogy and geographic location. There are a number of links that contain five or six mutational steps in the network. These large steps are found throughout the network and split the network into several smaller clusters. There is also no apparent

geographic pattern within the smaller haplotype groups created by the large mutational differences.

Table 2.1 Variable sites defining 35 South African *A. japonicus* haplotypes. The number of individuals that shared a specific haplotype is indicated under the "frequency" column. Base one corresponds to base number one in the reference sequence (*Salmo salar* GenBank NC 001960)

Haplotype nr	Frequency	1111111111111111122222222233333333445555555555566666 5578800111233444555679000122459024788890733444567778901334 6761801014046157478295156036204805624868247156122369167355
1	5	AAACATATTCCTCTCTGATTTGCAACATCGCGTTCATACCTAGTATATTAAAAAGAGG
2	1	....G.....
3	11	.....C.....
4	1	.....G.....C.....
5	9	.....A....C.....
6	4	.....T.....A....C.....
7	5	.....C.....
8	11	.....C.....C.....
9	3	.....C..A.....C.....
10	1	.G....C..A.....C.....
11	14	.....C.....A.....A...
12	4	.....C.....A.....C.....A...
13	1	.....C.....A.....G.....A...
14	1	.....C.....A.....G...A...
15	1	.....C.....A.....A.....A...
16	4	.....C....C.....A.....T.....
17	1	.....A....C.....A.....T.....C.....
18	1	.....C..C....C.....A.....T.....G....
19	5	.....C....C..A....AT..T.....T.....C.....
20	1	.....C.C.....A.....C.....C.....
21	1	.....G.....C.....T.....G.....GA.
22	1	.....C.....T.....G.....G...GA.
23	4	....T.....C.....CA.....G.....G.....
24	2	....T.....CTC.....CA.....G.....G.....
25	7	....T.....T..C.....CA.....TCG.....
26	9	...TT.....C.....CA.....A.....
27	1	...TT.....C...G..CA.....A.....
28	2	...TT.....C.....CA.....A..C.....
29	1	G..TT.....C....C.CA.....A.....
30	3	..GTT.....C.....CA..G.....
31	1	...TT.....C.....CCA.....G.....GG.....
32	8	....T....T...C.....CA.....C.....G.C.....
33	2	....T....T...CA.....CA.....C.....G.C.....
34	3	....T....T...C.....CA.....C.....G.C.....A
35	4	....T....T...C.....CA....G.....C.T.....G.C.....

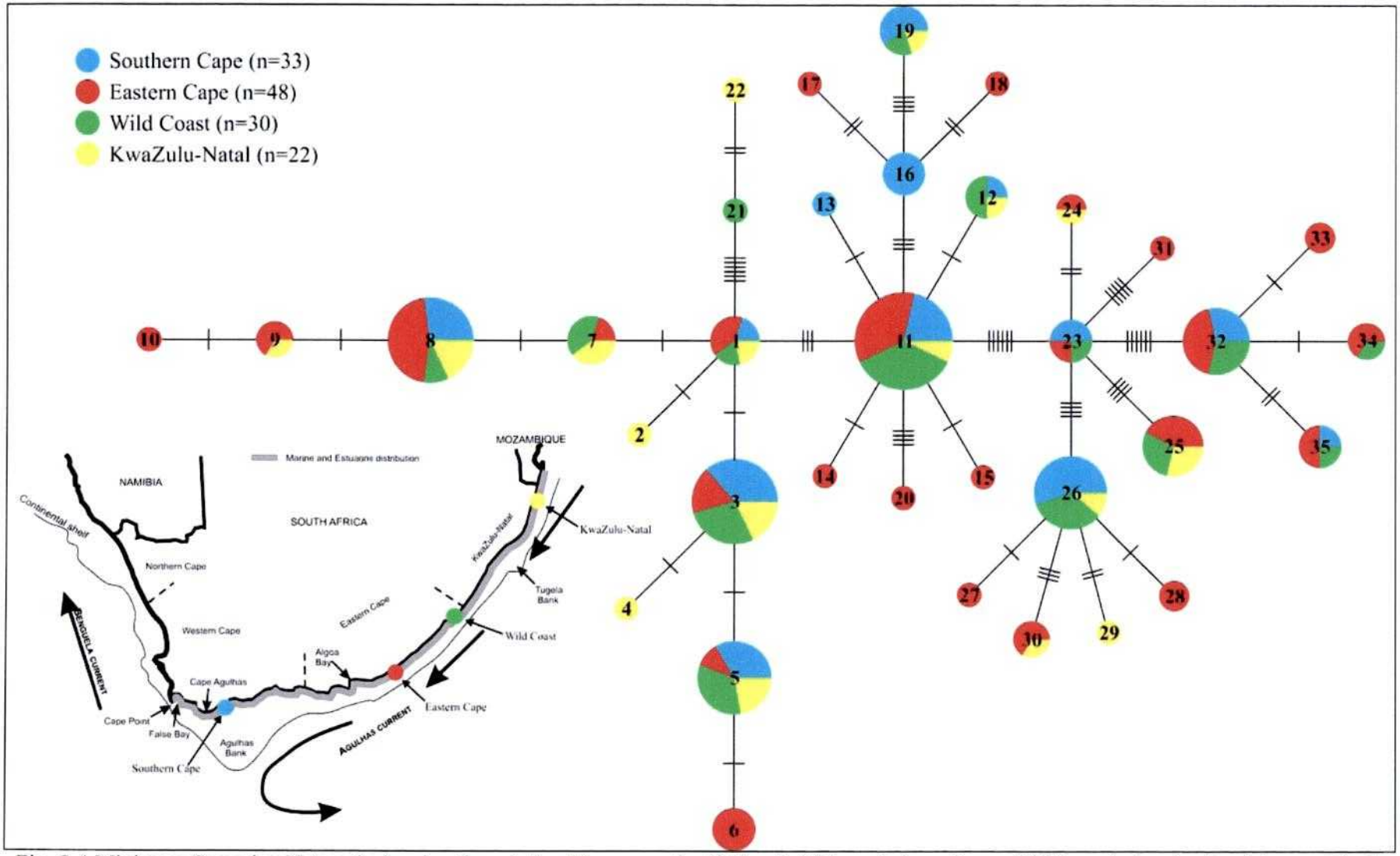

Fig. 2.4 Minimum Spanning Network showing the relationship among the 35 South African *A. japonicus* mtDNA control region haplotypes. Circle diameter is proportional to haplotype frequency. Pie charts display the localities where each haplotype was found and numbers indicate haplotype numbers as in Table 2.1. All haplotypes are separated by at least one mutational step, with hatch marks across the connecting lines indicating the number of mutations.

The mismatch distributions for the four localities (Appendix III) and for the South African population overall (Fig. 2.5) were unimodal. Mismatch distributions from the four localities all have similar distributions that are comparable with the overall mismatch distribution, with that obtained for the southern Cape and KZN being the most different. A relative good fit to a model of sudden expansion was observed in all cases as reflected in the S.S.D. values (Appendix III), although this is not reflected in the *P*-value. The overall ruggedness index was low (r = 0.001) indicating a good fit, but it was not significant (*P* = 0.910).

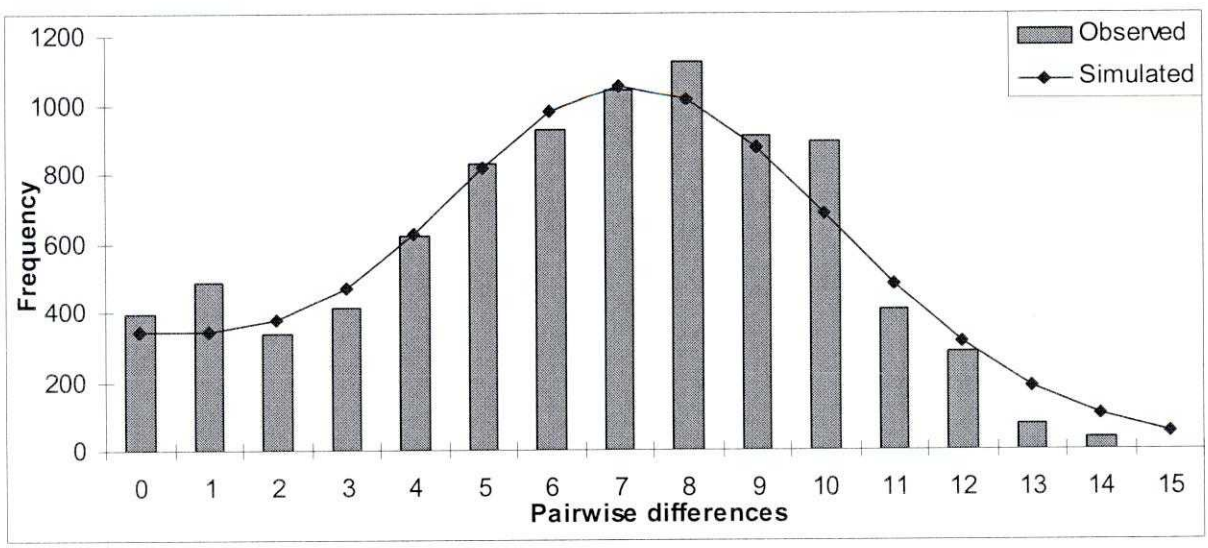

Fig. 2.5 Mismatch distribution of pairwise differences between 133 *A. japonicus* individuals based on 739 bp of the mtDNA control region. The expected frequency is based on a population growth-decline model (S.S.D. = 0.003 *P* = 0.627, r = 0.001 *P* = 0.910, θ_0 = 0.001 and θ_1 = 25.000) determined using ARLEQUIN 2.0.

Tajima D test for neutrality and Fu's F_s test indicated that the mtDNA control region of *A. japonicus* is evolving according to neutral expectations at all the localities, as well as for the South African population overall (Table 2.2). Although R_2 values were significant, both within the localities, as well as for the South African population, the values are not small enough to indicate that the populations have undergone historical growth or that the marker is under selection (Table 2.2).

Table 2.2 Comparison of several statistical tests of neutrality for the South African *A. japonicus* mtDNA control region, analysed per locality and the overall population

Locality	**Tajima D**	***P***	**Fu's F_s**	***P***	**R_2**	***P***
Southern Cape	-0.0699	0.5373	0.3278	0.4554	0.1624	0.0000
Eastern Cape	-0.0650	0.5302	0.3417	0.4465	0.1619	0.0000
Wild Coast	-0.0813	0.5402	0.3296	0.4470	0.1622	0.0000
KwaZulu-Natal	-0.0794	0.5381	0.3518	0.4403	0.1618	0.0000
SOUTH AFRICA	-0.0862	0.5465	0.3458	0.4382	0.1617	0.0000

Genetic diversity within and among sample localities

Eighteen of the 35 haplotypes were only found at a single locality. The Eastern Cape had by far the highest number of private haplotypes, with 44% of the haplotypes from the Eastern Cape not found at other localities (endemism index, Table 2.3). Haplotype diversity at localities were similar to that of the population overall (Table 2.3) and comparable with other marine species. Variation in nucleotide diversity within localities was more variable, with that from the Eastern Cape being the highest (Table 2.3). Comparison of nucleotide divergence between localities indicated that the Eastern Cape and KZN were the most divergent. Corrected pairwise divergence values less than zero indicate that within locality diversity is greater than between locality diversity (Table 2.4).

Table 2.3 Genetic diversity estimates for *A. japonicus* from four South African localities and the South African population overall

Locality	n	No of haplotypes	Endemism index	h (±SD)	π (%)	Mean pairwise differences (bp)
Southern Cape	33	13	0.15	0.9318 ± 0.0178	0.8122	5.9697
Eastern Cape	48	25	0.44	0.9645 ± 0.0111	0.9700	7.1294
Wild Coast	30	15	0.07	0.9425 ± 0.0207	0.8523	6.2644
KwaZulu-Natal	22	17	0.24	0.9784 ± 0.0189	0.8623	6.3377
SOUTH AFRICA	***133***	***35***		***0.9551 ± 0.0057***	***0.8883***	***6.5287***

Table 2.4 Nucleotide divergences between localities. Above diagonal: average percentage pairwise difference between localities. Below diagonal: corrected average pairwise difference

	Southern Cape	Eastern Cape	Wild Coast	KwaZulu-Natal
Southern Cape		0.9002	0.8261	0.8467
Eastern Cape	0.0672		0.8984	0.9294
Wild Coast	-0.0453	-0.0934		0.8644
KwaZulu-Natal	0.0695	0.0979	0.0520	

Geographic variation among localities

Testing for spatial homogeneity of mtDNA haplotype frequencies among the four localities was non-significant ($\chi^2 = 98.72$, $P = 0.6515$). This indicates that no significant structure exists between the four sampling localities.

Results from AMOVA revealed that most of the variation (99.79%) was found among samples within localities and that little could be attributed to variation among localities (0.12%). The overall low F_{ST} value ($F_{ST} = 0.0021$, $P = 0.3537$) also suggested homogeneity among the localities. Similarly pairwise F_{ST} analysis indicated very low differentiation between localities (Table 2.5). Plotting pairwise F_{ST} values against geographic distance indicate that there might

be some isolation-by-distance since a positive correlation (Mantel r = 0.357, P = 0.291) was found between the two variables, this was however not significant (Fig. 2.6).

Table 2.5 *Argyrosomus japonicus* pairwise F_{ST} and associated probability (P) values between the four South African localities. F_{ST} values are below the diagonal and corresponding P values above the diagonal

	Southern Cape	**Eastern Cape**	**Wild Coast**	**KwaZulu-Natal**
Southern Cape		0.17871	0.58984	0.24316
Eastern Cape	0.00917		0.89746	0.19238
Wild Coast	-0.00740	-0.01476		0.26660
KwaZulu-Natal	0.01168	0.01262	0.00827	

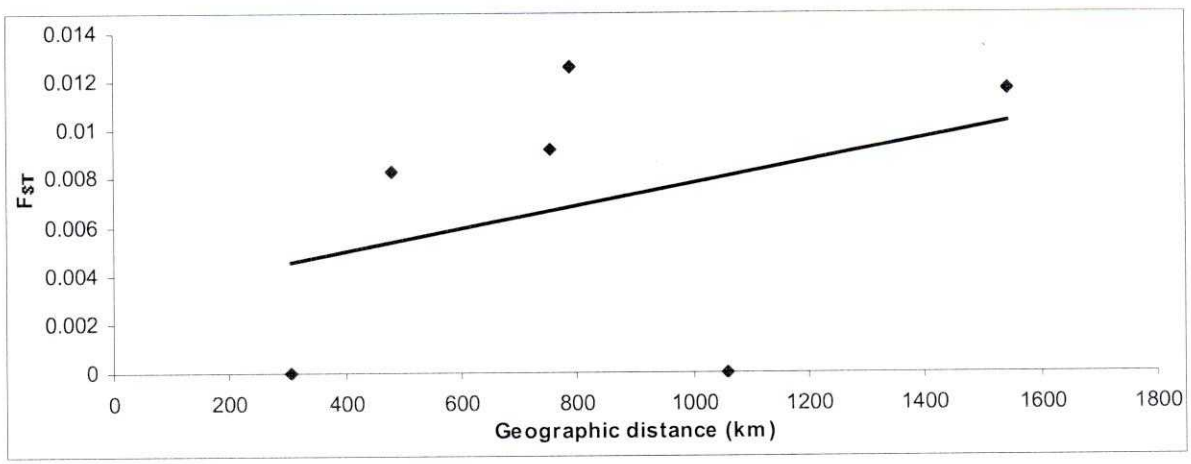

Fig. 2.6 Pairwise F_{ST} values plotted against geographical distance between four South African localities of *A. japonicus*. Mantel correlation value r = 0.413, P = 0.295.

Discussion

This is the first in depth study to determine the genetic variation within any of the *Argyrosomus* species. Allozyme data have been used to distinguish between two sympatric species, *A. inodorus* and *A. coronus*, found off the Namibian coast. The allozymes was also use to determine if any differentiation exits within the Namibian *A. inodorus* population (van der Bank & Kirchner, 1997). In the present study the complete mtDNA control region was shown to be selectively neutral, and thus a suitable molecular marker for a population study. High levels of mtDNA genetic diversity (h = 0.95, π = 0.89%) were found in *A. japonicus* from South Africa. These high levels compare well with that observed in other marine fish species (Beheregaray & Sunnucks, 2001; Carvalho & Hauser, 1998; Gold & Richardson, 1991; Grant & Bowen, 1998; Graves, 1998; Hauser & Ward, 1998; Rocka-Olivares *et al.*, 2000; Zardoya *et al.*, 2004). According to Grant and Bowen (1998) species typically exhibit both high h and π diversity due to either secondary contact of previously differentiated allopatric lineages or as a

result of a long evolutionary history in a large stable population. From the *A. japonicus* Minimum Spanning Network it seems that the high h and π could be due to previously differentiated allopatric lineages as this is evident by a number of steps in the network containing five or six mutations. These divergent lineages could be the result of previous isolated groups that have come into secondary contact and are now integrated in the South African population. The unimodal mismatch distribution (Fig. 2.5) shows a high frequency of large mutational differences (at ±8 steps), which may be a signature left by these divergent lineages. The mismatch distribution obtained from the original 819 bp fragment (Appendix II, Fig II.2) was multimodal, potentially a clearer indication of the possible integration of divergent lineages in the population although population subdivision has a very small effect on the mismatch distribution or the raggedness index (Harpending, 1994; Rogers, 1995).

High h and π diversity is more likely to be expected of a large population since the loss of genetic diversity due to drift is slower and therefore mutations accumulate over time (Hauser & Ward, 1998; Kimura, 1983). Large populations tend to lose more alleles than smaller populations when population size is reduced dramatically and this can result in a considerable change in genetic makeup of the population. But a population do need to reach a new equilibrium after a "crash" before the impact can be determined and in populations with large sizes, such as in many marine fish species, this can take many generations to obtain (if ever). Analysing a small number of individuals might not detect the genetic loss and change in genetic makeup that occurred in the population (Hauser & Ward, 1998; Ryman *et al.*, 1995). It is important to note that species with large population sizes (e.g. *A. japonicus*) are likely to have slow genetic drift at neutral loci, thus taking a long time to reflect differentiation (Hauser & Ward, 1998). *Argyrosomus japonicus* is long-lived, only reaching sexual maturity at an age of 7–8 years, has overlapping generations and a large population size. Consequently, a neutral locus such as the mtDNA control region will be slow in sorting lineages and slow to reach equilibrium.

Very little difference in genetic variation within the four South African localities was observed. Haplotype and nucleotide diversity from the localities is comparable with the overall values for the South African population. The exception is the high nucleotide diversity observed in the Eastern Cape, as well as the high endemism index compared to the other localities. This does not appear to be a result of the larger than average number of samples analysed for this locality, π drops to comparable levels when the samples size is randomly reduced to that of the other localities but the endemism index remains higher. The near identical percentage nucleotide

divergence values within and among *A. japonicus* sampling localities indicates that, on average, any two individuals drawn at random from a given locality will be as different for their mtDNA haplotypes as any two individuals drawn from different localities. Gold and Richardson (1991) also observed this in *Sciaenops ocellatus*, an estuarine dependent sciaenid with a similar life history, from the Gulf of Mexico and the southern United States Atlantic coast. Despite this, they have shown that isolation-by-distance does exist in the Gulf, with gene flow more likely between adjacent estuaries. This also means that conservation should not be focused on individual estuary level but on a wider geographic scale (Gold *et al.*, 2001).

Homogeneity of haplotype frequencies (spatial homogeneity test P=0.65) across the localities, coupled with a low non significant overall F_{ST} value (F_{ST}=0.002, P=0.35), also indicates that no population structure exists among juveniles found at the four South African localities. However, F_{ST} values between the southern Cape and KZN as well as between the Eastern Cape and KZN, although not significant, were five to six times greater than the overall F_{ST} value. This, and the positive correlation found between geographic distances and pairwise F_{ST} values, indicates that some isolation-by-distance does exist.

Sampling of estuarine dependent juveniles is essential to determine if localised spawning occurs within a region, or to detect an isolation-by-distance effect (Gold *et al.*, 1999). This will only give a true reflection if juveniles stay close to their natal estuaries and this appears to be the case for juvenile *A. japonicus* (Bullen & Mann, 2004a; Griffiths, 1996b). Most juveniles found in an estuary is likely to be the result of the proximate spawning events and would thus be expected to reflect the genetic composition of adults from the region. Based on the genetic evidence obtained and life history data, it seems that spawning off KNZ and that observed later in the season off Cape Infanta do not involve the same adult aggregation. The similar differentiation between the two warm temperate localities and KZN show that the juveniles recruiting into estuaries off the Eastern Cape are as isolated from KZN as those off the southern Cape. This indicates that recruitment of juveniles in the Eastern Cape is not merely a result of spawning adults off KZN returning south but that isolated spawning probably occurs within the region (Fig. 2.7). It has been suggested that larvae spawned off South Africa's east coast can be transported by shelf currents southwestwards and that they can reach nursery areas along the south coast in three to four weeks (Beckley & Connell, 1996). However, it would be reasonable to assume that estuaries further from a particular spawning site would receive fewer recruits than estuaries closer to the breeding event. With spawning reported as far south as Cape Infanta, juveniles found in southern Cape estuaries are more likely to be recruited from a

spawning event in this region than from similar events off KZN (Fig. 2.7). Passive larval dispersal would however result in gene flow predominantly in a southwestward direction. This would result in samples from the southern Cape having the highest diversity whereas this study revealed that sample from the Eastern Cape where the most divergent. It might be that the southwestward dispersal of larvae is counter balanced by the migration of adults up the coast to spawn. Total lack of isolation-by-distance between the southern Cape and the Wild Coast as well as between the Eastern Cape and the Wild Coast cannot be explained, specially given that some differentiation is observed between the southern and Eastern Cape. Current life history data indicate that adult *A. japonicus* along the South African coast form a single stock. This cannot be verified due to the lack of tagging information for adult *A. japonicus* (Bullen & Mann, 2004a). The potential isolation-by-distance observed in the current investigation could indicate that not all adult *A. japonicus* travel to spawn off KZN but that most of them tend to spawn close to their natal estuaries.

The general absence of definitive trends in the various tests conducted in this study might be due to the lack of structure in the South African *A. japonicus* population. Another possible reason could be that recent separation occurred therefore not allowing enough time for the accumulation of differences at the neutral locus due to genetic drift. In addition, the ability of tests to detect low levels of differentiation from small sample sizes, like that of the current study, is very limited and therefore run the risk of failing to identify any structure that might exist (Hauser & Ward, 1998; Ward, 2000). Further, individuals analysed in this study were a mixture of juveniles from various cohorts collected over an extended period. This may potentially add a lot of 'noise' to the analysis that could obscure any signal that does exist. Gene flow of only a few migrants per generation among spawning groups would make the isolated populations genetically indistinguishable yet this would not be enough to help overexploited populations to recover (Carvalho & Hauser, 1994; Hauser & Ward, 1998). Thus, due to the lack of power in all the population analysis tests conducted, it would be premature to make any definitive conclusions regarding *A. japonicus* population structure. The marker is not without its limitations. It only gives a reflection on maternal patterns and female dispersal and it reflects historical patterns rather than that of the current conditions (Ferris & Berg, 1987). An increase in sample sizes, as well as better structuring thereof, will be needed to detect if any genetic differentiation does exist within the South African *A. japonicus* population. The issue of temporal stability at localities, although important, is secondary to spatial structure. If spatial differentiation had been detected, then an analysis of temporal

stability would have been important to indicate the stability of breeding populations within a region.

Final conclusions should not be based on only one marker and it would be advisable to make use of a different marker (e.g. microsatellite markers) that is more appropriate at detecting fine-scale genetic population differentiation (Ward, 2000). Given appropriate sampling, analysis of microsatellites will also reflect on the current population gene flow rather than historical genetic information as is the case with mtDNA (Hauser & Ward, 1998). Analyses of spawning adults along the coast might give us a better understanding of the processes operating in the adult *A. japonicus* population along the South African coast. If genetic differentiation is detected between the spawning aggregations the data could be used as baseline information to detect the possible mixing of these stocks as juveniles at the various estuaries.

The mtDNA diversity observed between samples of *A. japonicus* from South Africa and Australia was more than seven times greater than the observed differences within the South African population (data not shown). Studies on other marine taxa with fragmented global distributions have revealed that populations found off South Africa and Australia are either closely related as is the case for *Sardinops* (Bowen & Grant, 1997) but, in general it seems that the South African and Australian populations are not closely related as observed in *Engraulis* (Grant & Bowen, 1998) and *Pomatomus* (Graves, 1998). Similar to the divergence observed in *A. japonicus* from South Africa and Australia was reported by Alvarado *et al.* (2005) that found 8.7% divergence in the variable region I of the mtDNA control region of Atlantic bluefin tuna (*Thunnus thynnus*). This high divergence was attributed to the introgression of maternal lineages from two sister species, *T. orientalis* and *T. alalunga*. The relationship between the two isolated populations (South Africa and Australia) of *A. japonicus* was beyond the scope of this study, but it deserves further investigation. Samples from the South China Sea off Japan, the source locality for the type specimen of this species, will have to be included in such an investigation to unravel the fragmented global distribution of *A. japonicus*.

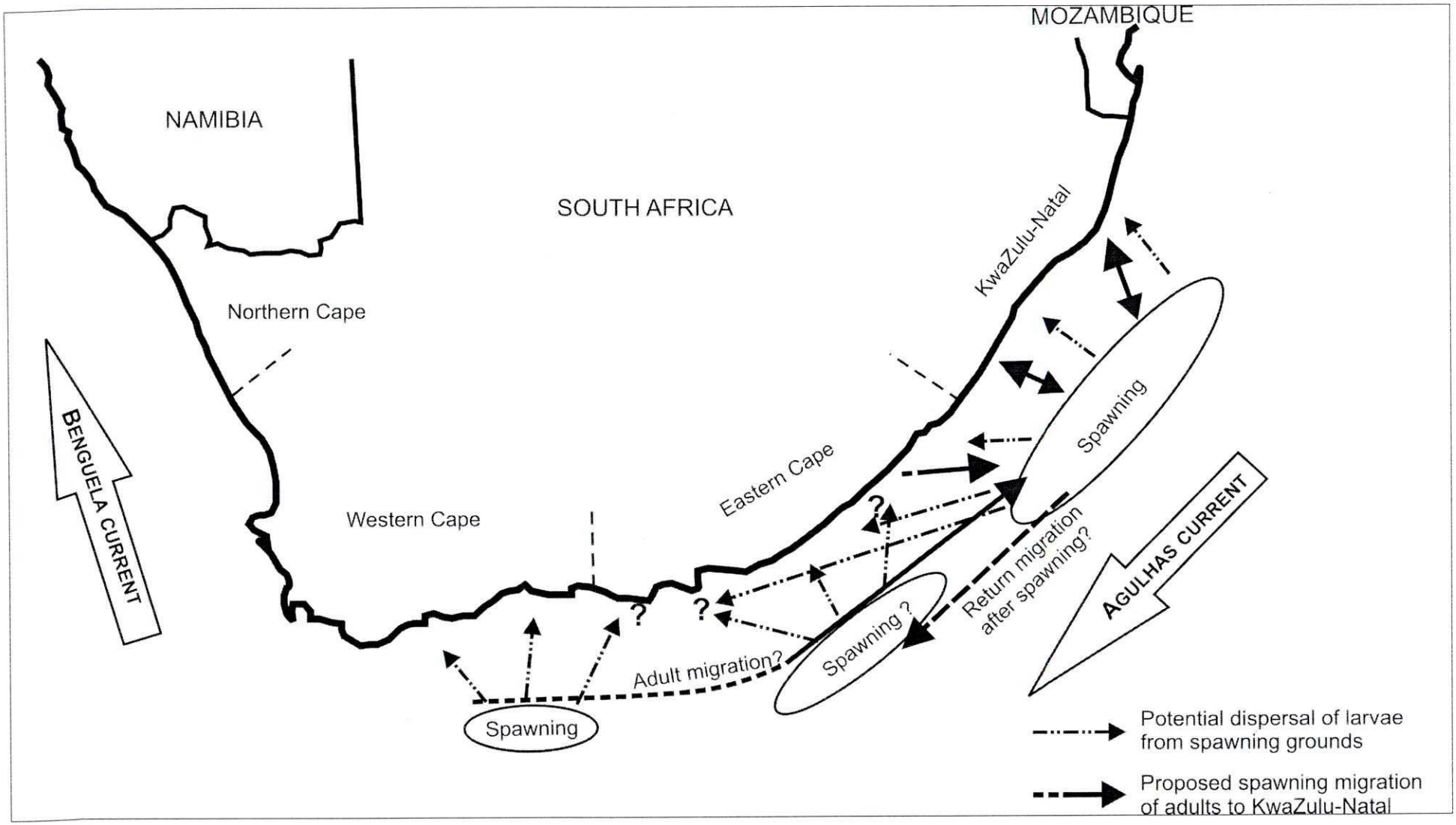

Fig.2.7 Hypotheses regarding maternal gene flow in South African *A. japonicus*. Based on life history data, adult *A. japonicus* undergoes a spawning migration to KZN, but spawning has also been recorded along the south coast later in the season. Influenced by the Agulhas Current, dispersal of larvae would predominantly be in a southwesterly direction whereas adults migrate up the coast to spawn.