BIOCHEMICAL AND MITOCHONDRIAL POPULATION STUDIES OF SOUTHERN AFRICAN HAKE, *MERLUCCIUS CAPENSIS* AND *MERLUCCIUS PARADOXUS*

by

Inga Isabel Becker

A dissertation submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the Faculty of Science, University of Cape Town.

Cape Town, October 1987

CERTIFICATION OF SUPERVISORS

In terms of paragraph 9 of "General regulations for the Degree of Ph D" we as supervisors of the canidate I.I. Becker, certify that we approve of the incorporation in this thesis of material that has already been published or submitted for publication.

Signed

Associate Professor F.T. Robb

Department of Microbiology

Signed

Dr. R. Kirby

Department of Microbiology

FÜR

MEINE ELTERN

CONTENTS

DECLARATION

This thesis reports the results of original research which I carried out in the Fishing Industry Research Institute, and in the Department of Microbiology, University of Cape Town. None of it has been submitted in whole or in part for any other degree and all technical assistance is fully acknowledged. The work presented in Chapter 2 was carried out in collaboration with Dr W. S. Grant.

I.I. Becker

ACKNOWLEDGEMENTS

I am most grateful to my supervisors: Stewart S. Grant for his guidance, criticism, his help with the statistical analysis of the results, and for introducing me to this field of study; Frank T. Robb for his never ending encouragement and advice; and Ralph Kirby for his intense interest and innumerable suggestions during this study, and especially for his help with the analysis of the DNA sequence data. A word of appreciation goes to Dave Woods who granted me the opportunity to do this study.

I would like to thank all my colleagues, who were always kind and willing to help; I am especially deeply indepted to Shelly Deane for all the effort she put into correcting my English, Joe Santangelo for lending me his computer, and Wyn Jones for her technical assistance. Last, but not least, I would like to thank Eric O'Neill for his inexhaustable help, his excellent technical assistance with DNA sequencing, the many amusing fights we had throughout these years, and the effort he put into consuming all the hake.

For arranging the collection of samples, I thank Rob Leslie from Sea Fisheries Research Institute, Earl Moore from Irving and Johnson Fish Company, Cape Town, the Fisheries Research Division of New Zealand, Carl Wicht, and Vernon Edwards.

Many thanks go to Werner Kobi for helping me with the graphs, Peter Becker for being so helpful with the printing, and Ulf Hitzeroth for showing so much understanding and patience during the last months of this study.

This research was funded mainly by the Sea Fisheries Research Institute and I acknowledge the financial support which I received from the Foundation for Research and Development of the Council for Scientific and Industrial Research.

ABSTRACT

ABSTRACT

Two sympatric species of hake, *Merluccius capensis* and *Merluccius paradoxus*, which are morphologically very similar, occur along the coasts of southern Africa. The amount of genetic divergence was estimated between these two species using electrophoretic analysis of 31 protein encoding loci and restriction endonuclease analysis of mitochondrial DNA (mtDNA). No hybrid genotypes were observed with protein electrophoresis suggesting that no interbreeding occurs between the two southern African species of hake. Using protein electrophoresis, the genetic distance was estimated to be 0.583 ($\pm$0.160) between the two species. This is typical of evolutionary divergence between well differentiated congeneric species. The time of divergence, estimated from the genetic distance, was between 7 and 13 million years.

Nucleotide sequence variation in mtDNA was determined between and within the two species of hake using 11 restriction endonucleases. The length of the mitochondrial genomes was 16.9 and 16.7 kb for *M. capensis* and *M. paradoxus*, respectively and this was typical for fish. Fourteen different composite genotypes were observed for 26 individuals of *M. capensis* and for *M. paradoxus* 6 composite genotypes were observed in 24 individuals. A parsimony network connecting the composite genotypes of

either species did not correspond with the geographies of the samples. In *M. capensis* insertions of about 400 bp and deletions of 200 bp were detected in the restriction fragment patterns of *Ava* I, *Xba* I, and *Xho* I. Mapping of three different *Xho* I genotypes confirmed these insertions and deletions and sequencing parts of the mtDNA indicated that these length polymorphisms are most likely situated on or close to the D-loop region of the mtDNA.

Cloning parts of the mtDNA from *M. capensis* and *M. paradoxus* facilitated estimation of sequence divergence between the two southern African hakes and *M. australis*, from New Zealand. The same 11 restriction endonucleases were used to digest the mtDNA of the New Zealand hake and the cloned mtDNA from *M. capensis* and *M. paradoxus* was used as a probe to detect the fragment patterns of *M. australis* mtDNA. This revealed that the salmon sperm present in prehybridisation solutions interfered with the hybridisation of a mtDNA probe. The length of the mtDNA of *M. australis* was 16.9 kb. Only three composite genotypes were observed for seven individuals of *M. australis*.

Merluccius paradoxus and *M. australis* showed a low level of nucleotide site polymorphism relative to *M. capensis*. The low level of nucleotide polymorphism was reflected in the number of composite genotypes and the complexity of the parsimonous trees connecting the composite genotypes with each other. These results suggest that *M. paradoxus* and *M.*

australis may have experienced a population bottleneck in the past.

The amount of sequence divergence between *M. capensis* and *M. paradoxus* was 11.6% (±0.036) and is typical for well differentiated species. The present sympatric distribution of the two species of hake may be the result of repeated dispersal of ancestral North Atlantic species of hake to southern Africa.

The amount of sequence divergence between the New Zealand hake and *M. capensis* and *M. paradoxus* was 17.9% and 11.5%, respectively. No composite mtDNA genotypes were shared between the three species of hake. These results showed that all three species are well separated congeneric species. The time of divergence could be calculated from from the sequence divergence and it was estimated that *M. capensis* and *M. paradoxus* diverged 5.8 million years ago, *M. australis* and *M. capensis* separated approximately 7 to 9 million years ago, and *M. australis* and *M. paradoxus* diverged 5.7 million years ago. This confirmed the hypothesis, based on morphological studies, that the two deep water hake *M. australis* and *M. paradoxus* are genetically closely related and may be biogeographically related by dispersal of ancestral *M. paradoxus* from east to west in the southern Atlantic.

Parts of the cloned mtDNA from both *M. capensis* and *M. paradoxus* were sequenced. The following genes were identified on the mtDNA from *M. paradoxus*: cytochrome oxidase I (COI), COII, lysine tRNA, ATPase8, and ATPase6. Genes identified on the mitochondrial genome of *M. capensis* were ND1 (component of the respiratory-chain NADH dehydrogenase), ND2, isoleucine tRNA, glutamine tRNA, f-methionine tRNA, tryptophan tRNA, and alanine tRNA. The sequences revealed that the hake mitochondrial protein genes were similar in conservation, genetic code, overlapping of reading frames, and termination of reading frames to the equivalent protein genes of other higher eucaryotes. The tRNAs also resembled the mitochondrial tRNAs found in mammals, amphibians, and insects in sequence and secondary structure. The general organization of the hake mitochondrial genome resembled the organization of the mtDNA, reported for other higher eucaryotes. The hake mtDNA had the most sequence homology to amphibian mtDNA.

CHAPTER 1

GENERAL INTRODUCTION

1.1 INTRODUCTION

1.1.1 General

Various species of hake (order Gadidae, family Merlucciidae, genus *Merluccius*) are found in the North and South Atlantic Oceans, the eastern Pacific Ocean and around New Zealand. With the exception of the hake population in the New Zealand waters, hake inhabit continental shelf waters (Kabata and Ho, 1981; Inada, 1981). Hake are an important resource in the world fisheries, and one of the problems associated with the industry has been the taxonomy of the species in the genus *Merluccius*. To date, twelve to fifteen different species of hake have been proposed (Inada, 1981). The reconstruction of the phylogenetic relationships among these species also remain a problem to be resolved.

One of the traditional methods of distinguishing species has been the examination of morphological characters such as number of scales in a lateral series, number of vertebrae, total number of gill-rakers, and relationship of standard length to head length (Inada, 1981). Often these morphological characters did not reflect the genetic variation between or within species. Several classical techniques exist today, which measure the relatedness in DNA sequences of individuals. Amongst these are comparison of amino acid sequences, immunological crossreactivity of proteins, protein electrophoresis, DNA hybridisation,

electron microscopy, and restriction endonuclease analysis (Upholt, 1977; Monnat and Loeb, 1985). The development of electrophoretic analysis of proteins made it possible to measure the amount of biochemical variation of protein encoding genes (nuclear). The measurement of genetic variation using protein electrophoresis, directly reflects differences at the DNA level. However, often nucleotide changes in the DNA go undetected as the change from one amino acid to another does not necessarily change the conformation or charge of the resulting protein, and thereby its electrophoretic mobility is not effected. Many of the nucleotide changes also occur in the third position of the codon, and these silent changes will not result in a change of an amino acid. This shows that protein electrophoresis cannot detect all genetic variations and it does not take into account changes which occur in noncoding DNA sequences.

About 1% of the DNA of an eucaryotic cell is located in the mitochondria. The mitochondrial DNA (mtDNA) has become a well characterised cytoplasmic gene system, as it is easy to purify and characterize. The discovery of restriction endonucleases resulted in the development of restriction endonuclease analysis of mtDNA. By digesting the mtDNA with restriction enzymes, it was possible to directly detect a change of one nucleotide. Mapping of restriction sites further allowed the detection of rearrangements on the mtDNA.

Restriction analysis of the mtDNA made it possible to detect changes at the DNA level, however, it may not detect randomly distributed differences in sequence changes that involve as much as 10% of all nucleotide positions in a population of DNA molecules (Monnart and Loeb, 1985). Nucleotide sequencing techniques, on the other hand, allow the detection of all nucleotide changes. The DNA sequence of a genetic system will also provide information about the overall organization of the genome and it gives insight into the nature of the evolutionary changes that the DNA undergoes.

This work was aimed at investigating the taxonomic status and the degree of genetic divergence between two nominal species of hake occurring along the coasts of southern Africa. The phylogenetic relationship between the southern African hake and the hake species from New Zealand, *M. australis*, was also studied. Four different techniques were employed: protein electrophoresis, restriction fragment analysis of mtDNA, restriction site analysis of mtDNA, hybridisation studies using cloned mtDNA, and sequencing of the mtDNA clones. In this chapter, the distribution of present-day *Merluccius* is discussed, the focus being on the two sympatric southern African hake species, *Merluccius capensis* and *Merluccius paradoxus*. In the second chapter, the genetic divergence between the two southern African hake species is determined using the distribution of electrophoretically-detectable protein variants.

Restriction fragment analysis of mtDNA, as a means of studying sequence divergence between and within species, is reviewed in chapter three; and in the fourth chapter the use of restriction-site mapping and cloning of mtDNA is discussed. In this chapter the amount of genetic divergence is also estimated between the southern African hakes and the New Zealand hake, *M. australis*, using restriction endonuclease analysis of mtDNA. The last chapter presents nucleotide sequences of parts of the cloned mtDNA from both southern African hake species.

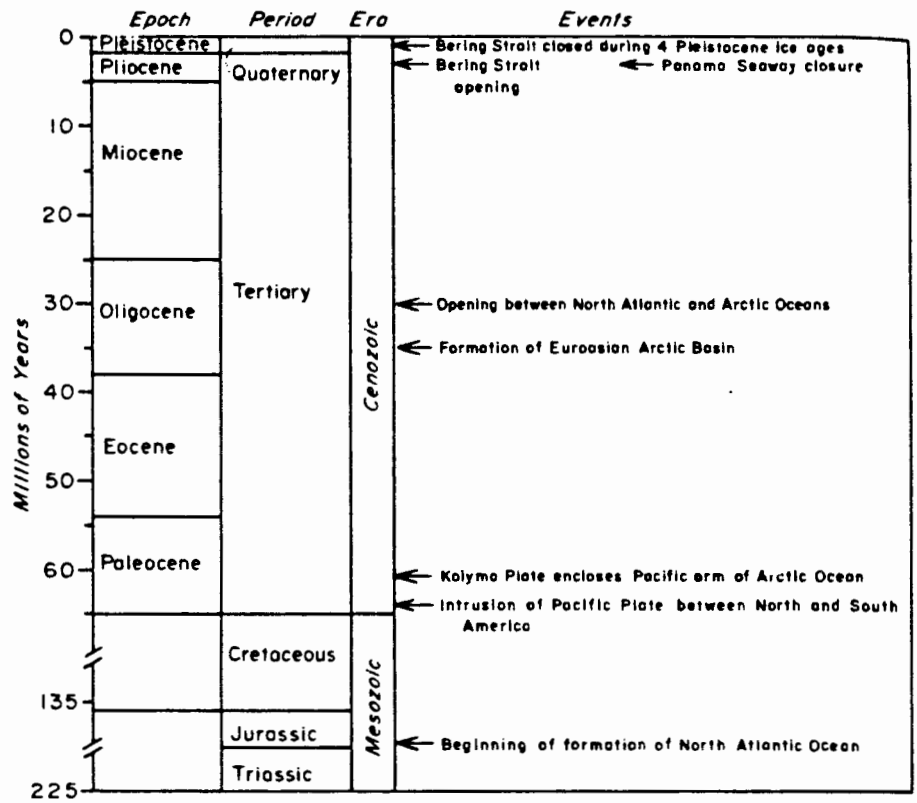
The methods chosen for this study, measure the amount of genetic variation of different parts of the genome, by examining the gene products of protein encoding loci on nuclear DNA, restriction analysis of mtDNA, and sequencing of mtDNA.

1.1.2 Evolutionary Origin of Hake

The course of dispersal of *Merluccius* has been investigated in order to explain the present day distribution of the 12-15 species of hake (Szidat, 1955; 1961; Kabata and Ho, 1981; Inada, 1981). The earliest records of Merlucciidae date back to fossils found in Oligocine and Miocene deposits in Europe (Inada 1981; Kabata and Ho, 1981) (Fig. 1.1.1 - Geological time scale). It has been proposed that the origin of Gadiformes is in the North Atlantic, and that the

Figure 1.1.1

Geological time scale (from Grant, 1987)



Merlucciidae also originated here during the middle Oligocene to the lower Miocene (Sventovidov, 1940; 1948; Inada, 1981). Szidat (1961), however, concluded from parasite studies of the Pacific hake, that the origin of Merlucciidae was in the North Pacific. Kabata and Ho (1981), and Inada (1981), in their papers, discuss Szidat's hypothesis and come to the conclusion that no proof exists to support the hypothesis. They, on the other hand, place the origin of hake near the southern tip of early Eocene Greenland or in the north-eastern Atlantic Ocean (Kabata and Ho, 1981; Inada, 1981). Hake must have existed in the early Atlantic prior to the break up of Laurasia into America and Eurasia, which happened during mid-Tertiary about, 40 million years ago (Kabata and Ho, 1981).

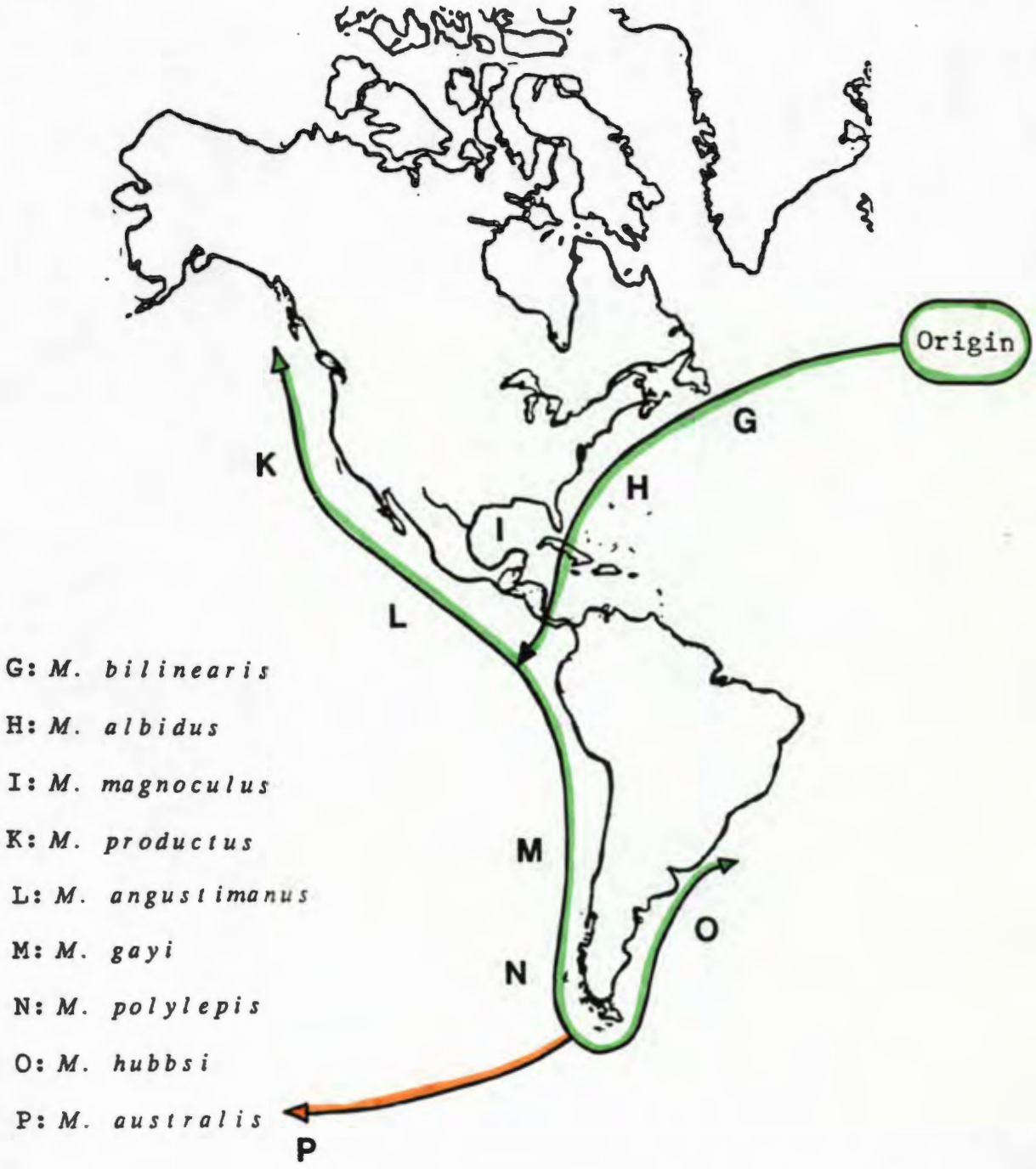
1.1.3 Dispersal of hake along the coasts of America

As America and Eurasia drifted apart, they presumably were accompanied by separate hake stocks. Kabata and Ho (1981) postulate that the dispersal of hake to other areas must have followed the margins of continental shelves and that the routes ran parallel to major currents. The hake then proceeded to disperse to the south along the Atlantic Coast of northern America. Northward migration was prohibited by cold temperatures, and each population moving south was not able to retain contact with populations further north

(Kabata and Ho, 1981). The southward migration of the hake continued, due to the cooling of the seas which occurred between the upper Eocene and mid-Oligocene. The hake occupied the continental shelf along the Atlantic coast of North America and evolved to form the present-day *Merluccius bilinearis*. Kabata and Ho (1981), and Inada (1981) further postulate that hake crossed from the Atlantic into the Pacific Ocean through Central America, when the isthmus of Panama was submerged (not later than the Pliocene). From there, hake dispersed northwards, giving rise to present-day stocks of the Pacific hake, *Merluccius productus*, and southwards forming the Chilean hake, *Merluccius gayi* (Kabata and Ho, 1981; Inada, 1981). One deep water species, *Merluccius angustimanus* (the Panama hake) evolved in the Bay of Panama, imprisoned below surface waters of unacceptably high temperature. The Offshore hake, *Merluccius albidus* developed along the Atlantic Coast of the United States, the Gulf of Mexico and the Caribbean Sea (Kabata and Ho, 1981; Inada, 1981). Kabata and Ho (1981) suggest that the Atlantic hake dispersed southward along the East Coast of South America and speciated into the Argentinian hake, *Merluccius hubbsi*, or alternatively, that *M. hubbsi* had a Pacific origin. Inada (1981) and Szidat (1955) also postulate that the hake invaded the eastern coasts of southern America around Cape Horn during the Interglacial period of Pleistocene (Fig. 1.1.2). The presence of the Amazon River, which already existed during the Pleistocene,

Figure 1.1.2

Dispersal of Merlucciidae along the coasts of the American continent



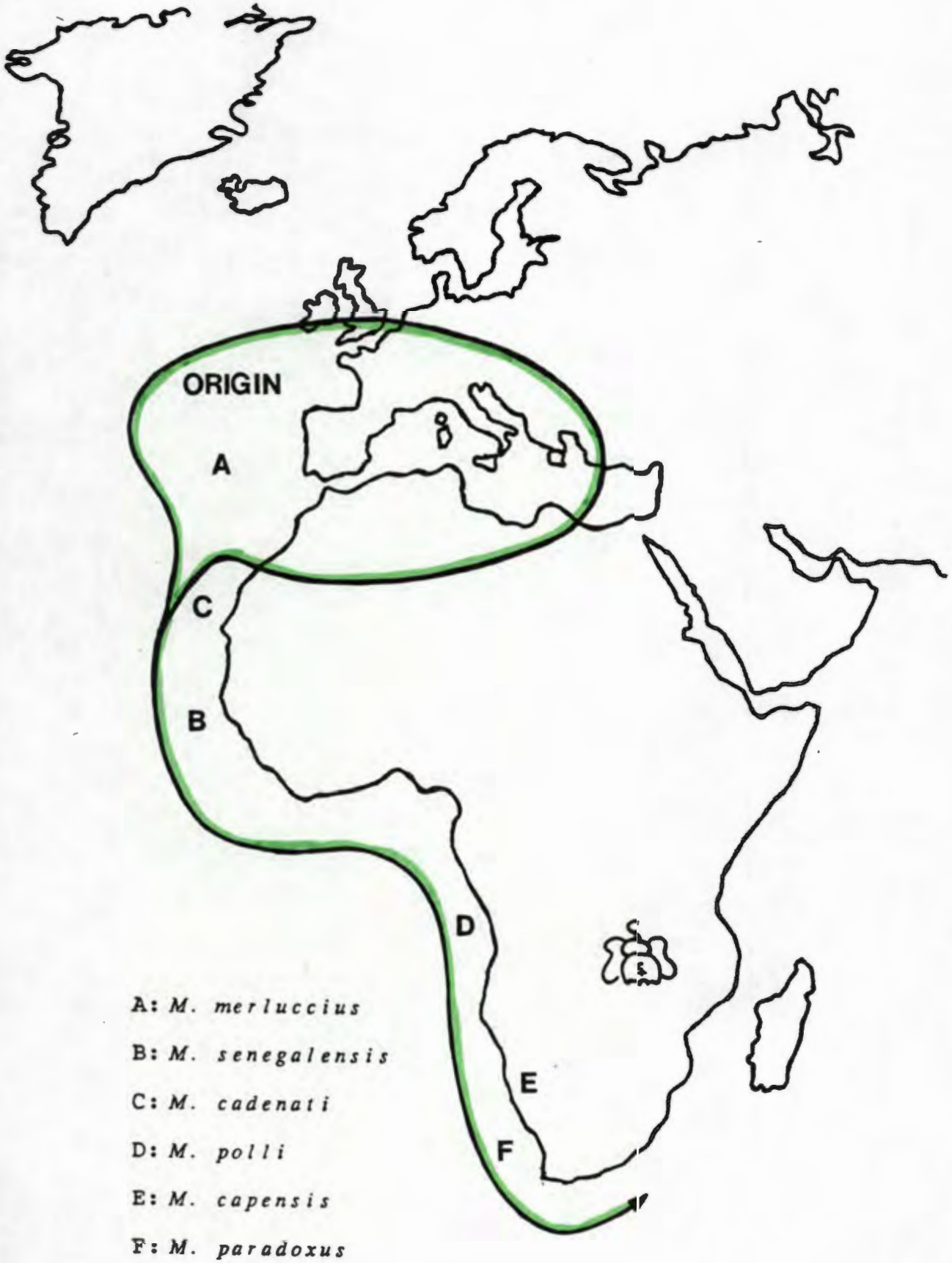
would have prohibited the southward migration of hake along the East Coast of South America, as it presented a barrier of low salinity water (Inada, 1981).

1.1.4 Dispersal of hake along the West Coast of Africa

Merlucciidae also migrated from the North Atlantic Ocean to southern warmer waters along the West Coast of the Eurasian landmass, and split into two groups when they reached the Mediterranean (Kabata and Ho, 1981) (Fig. 1.1.3). One group formed the European hake *Merluccius merluccius*, and the other group proceeded southward along the West Coast of Africa until it reached the southernmost point of Africa. Hake may have spread relatively rapidly along this coast and until recently, the populations of hake along the coast of Africa were regarded as subspecies of the European hake, *M. merluccius* (Kabata and Ho, 1981). Inada (1981) proposed that *M. merluccius* and *Merluccius senegalensis* (Senegalese hake) evolved to the north of western Africa and that the Benguela hake, *Merluccius polli*, which occurs along the Angolan coast, developed in deeper equatorial waters. Szidat (1961), on the other hand, suggests that *M. polli* is derived from the European *M. merluccius*, whereas *M. capensis*, the shallow-water Cape hake, was more similar to *M. senegalensis* from the West Coast of Africa. He postulated that *M. capensis* could have arrived in southern

Figure 1.1.3

Dispersal of Merlucciidae along the coast of the Eurasian landmass



Africa from the Pacific with the cold water currents, as it is morphologically very similar to *M. gayi*. He then further suggests that it is theoretically possible that the two species (*M. capensis* and *M. merluccius*) dispersed in opposite directions through the tropical zone, to speciate into *M. senegalensis* and *M. polli*, respectively.

None of these hypotheses include the fact that two sympatric species, *M. capensis* and *M. paradoxus*, occur along the coasts of southern Africa. It is still not clear whether these two species are derived from one common North Atlantic ancestor, or whether they are derived from different ancestral species. *Merluccius paradoxus* is reported to be closely related to the other two deep-water hake species, *Merluccius australis* and *M. polli*, in general appearance, and in that they have numerous vertebrae (Inada, 1981). *Merluccius capensis* is more closely related to *M. senegalensis* in general characteristics, and has some characteristics in common with *M. merluccius* and *M. hubbsi* (Inada, 1981).

1.1.5 Dispersal of hake to New Zealand

One of the controversies concerns the arrival of *M. australis* in New Zealand. Migration of the two southern African species to New Zealand via the Indian Ocean is prevented by the warm southward flowing Mozambique current

along the East Coast of southern Africa (Inada, 1981). Kabata and Ho suggest that either *M. gayi* (Chilean hake) or *M. productus* (Pacific hake) must have crossed from South America to Antartica and from there to Australia and New Zealand, about 30 million years ago. This migration was assisted by southward currents which existed prior to the Circumantarctic current (Kabata and Ho, 1981). Inada (1981, 1986) also found that the Patagonian hake, *Merluccius polylepis*, from southern South America, was the same as the New Zealand species. Other New Zealand fishes such as *Macruronus*, *Genypterus*, and *Callorhynchus* also show a close relationships with species from southern South America and southern Africa, thereby favouring the hypothesis of dispersal of hake from South America to New Zealand (Inada, 1981). Fossils of hake were found in Australia from the Miocene, showing that hake existed in Australia about 26 million years ago, and for some unknown reason, became extinct (Kabata and Ho, 1981; Inada, 1981).

1.1.6 Present-day distribution of Hake

The distribution of all the known species of hake is approximately limited by temperature. Two critical isotherms, 7°C and 23°C, form barriers to the expansion of hake populations (Kabata and Ho, 1981). Hake are confined in several areas by a combination of land and temperature barriers, which act as isolating mechanisms (Kabata and Ho,

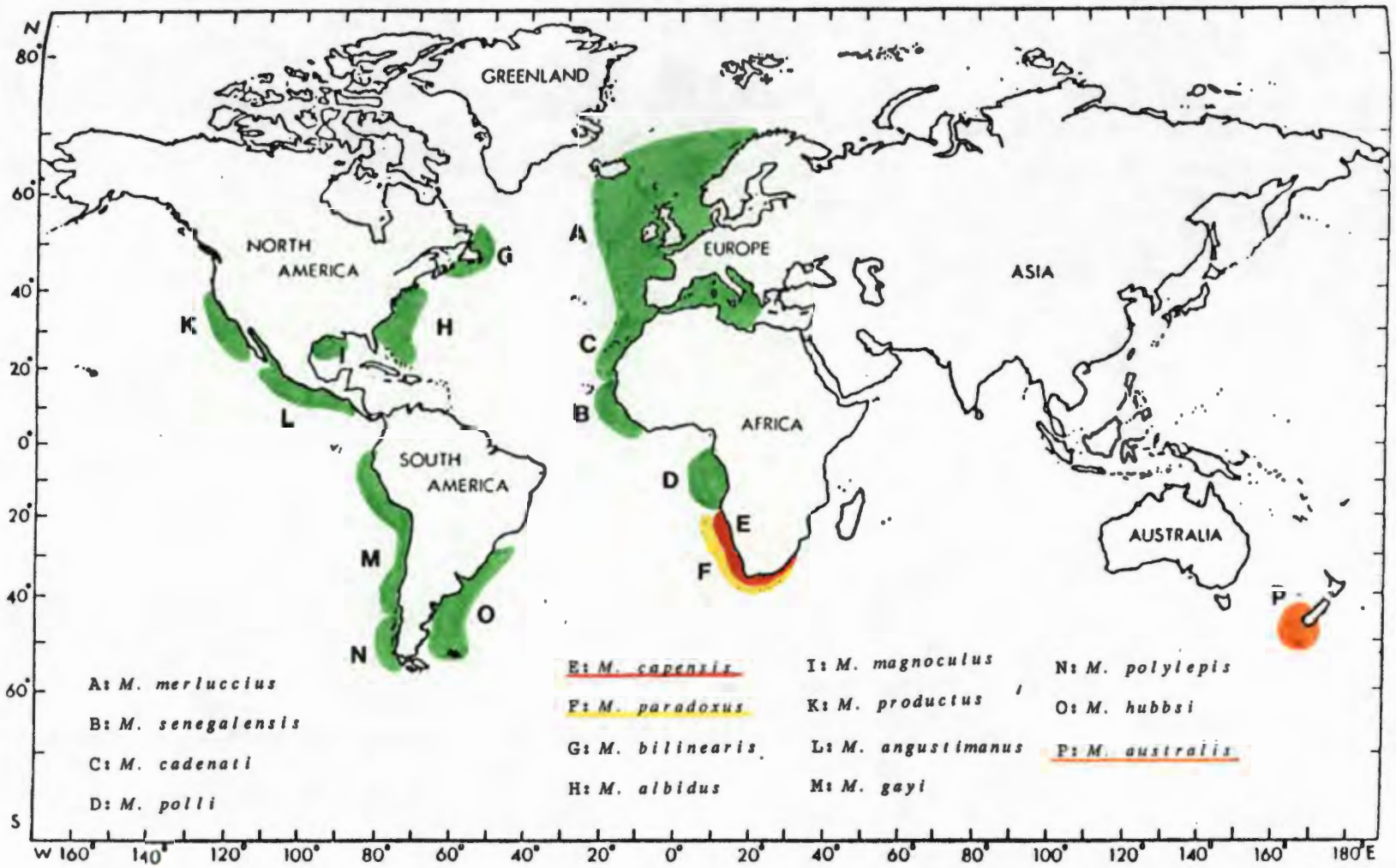


Figure 1.1.4

Present-day distribution of the hake species

1981). Present-day distributions of the fifteen hake species are presented in Fig. 1.1.4.

1.1.7 The southern African Hake Populations

1.1.7.1 Nomenclature of the Cape hake

The hake species exploited by the trawl fishery off Namibia (South West Africa) and South Africa are generally referred to as the Cape Hake or Stock fish. South African exploitation of the Cape Hake started as early as the end of the 19th century, and peaked in 1972 with landings of over 1 million tons, thereby establishing the Cape Hake fishery as the largest of its kind in the world (Botha, 1980; Inada, 1981).

Despite the increasing importance of the Cape Hake fisheries, very little research, other than early surveys, was done, and until 1954 it was assumed that the Cape Hake belonged to the species *M. capensis*. In 1954, Franca described two forms of *M. capensis* (Castelnau) which could be distinguished by the numbers of vertebrae (Franca, 1954). In a later paper, Franca described these two forms as subspecies of *M. capensis*, *Merluccius capensis capensis* (Cast.) Franca, and *Merluccius capensis paradoxus* Franca, but in 1962 he suggested that they were best considered subspecies of *M. merluccius* (Franca, 1960, 1962). In 1969, Jones and Mackie used zone electropherograms of muscle myogens to identify two forms of hake in the waters off

Namibia, and these results corresponded to the classification by vertebral counts (Jones and Mackie, 1970). Pchenitchny *et al*. (1969) reported that *M. m. capensis* had an average of 51 vertebrae, whereas *M. m. paradoxus* had 56 vertebrae. Jones and Mackie (1970) also suggested that the two forms of hake had different bathymetric distributions, with only a small degree of overlap, where the "*paradoxus*" form was found in deeper and cooler water, and the "*capensis*" form in shallower, warmer water. Van Eck (1969) reported similar results to those of Jones and Mackie using vertebral counts, and reported that the two subspecies could be distinguished with the naked eye by looking at the outer gill arch. The outer gill arch in *Merluccius* had outgrowths at both ends (low spinulose tubercles), which differ between species. *Merluccius m. capensis* had rounded tubercles of uniform white colour, whereas the tubercles of *M. m. paradoxus* were longer and thinner and had small black dots in the center (Fig. 1.1.5).

Although there is still some doubt as to whether these two forms are species or subspecies, they are generally referred to as the shallow-water Cape Hake, *M. capensis*, and the deep-water hake, *M. paradoxus*. They are morphologically very similar, as can be seen in Fig. 1.1.6.

Figure 1.1.5

Photo of the outer gill arch from *Merluccius capensis* (top) and *Merluccius paradoxus* (bottom). The spinulose tubercles from *M. capensis* are rounded and of uniform white colour. The tubercles from *M. paradoxus* are longer, thinner, and have small black dots in the center.

Figure 1.1.6

Photo of *Merluccius capensis* (top) and *Merluccius paradoxus* (bottom) showing the morphological similarity of the two species of hake.

1.1.7.2 Geographic Distribution

Merluccius capensis is distributed from Angola (about 12°S) on the West Coast, to 31°E, on the South East Coast of South Africa (Fig. 1.1.7). Its geographic range is limited by the cold, northward flowing Benguela Current along the West Coast, and by the warm, southward flowing Agulhas Current along the East Coast, and this hake is found most abundantly in the region between Cape Frio and East London. This species inhabits waters with temperatures between 4°C and 12°C, but is found chiefly in waters of 8°C (Chlapowski, 1975; Inada, 1981). It is mainly found in waters shallower than 440 m, and only large individuals, 50 cm or more, are found in deeper waters.

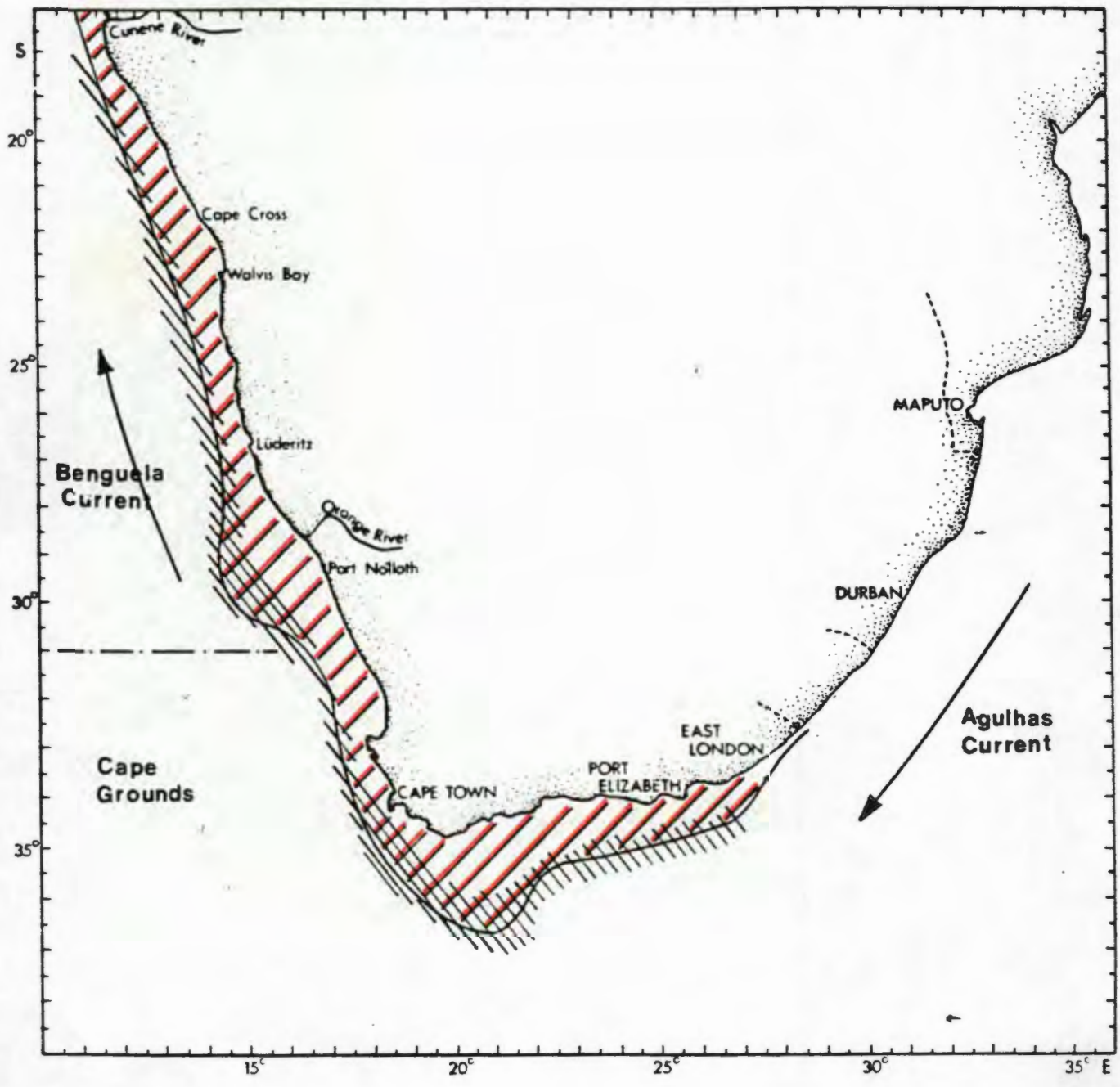
Merluccius paradoxus is distributed along the continental shelf off Namibia and South Africa, further offshore than *M. capensis* (Fig. 1.1.7). Its distribution is slightly more restricted than *M. capensis*, between 18°S on the West Coast to 27°E on the eastern edge of the Agulhas Bank. It is found in deeper waters, from 200 m to depths of 850 m. Generally, *M. paradoxus* is found in cooler waters, of 4°C - 8°C, and the two species coexist at depths ranging between 200 and 400 m (Jones and van Eck, 1967; Botha, 1973; Inada, 1981). *Merluccius capensis* females grow faster than males and this species grows faster than *M. paradoxus*. Since larger individuals inhabit deeper waters, the overlap

Figure 1.1.7

Geographic distributions of southern African hakes, *Merluccius capensis* and *Merluccius paradoxus*

M. capensis

M. paradoxus



between species exists between the large adults of *M. capensis*, and juvenile *M. paradoxus*.

1.1.8 Objectives

The two species of hake are of economic importance to the South African fishing industry. These two species are, however, still managed as one stock, due to the difficulty of splitting the catch into the two species, a process which requires a physical examination of each fish (Badenhorst, 1984).

There is a continuing dispute over the classification of the southern African hake, as to whether there are two species or two subspecies. With the heavy exploitation of the hake, an important consideration in stock assessment is whether each species represents a panmictic breeding unit, or whether subpopulations occur within each species. Results of this assessment will influence the conservation of the stock. It is not possible to determine different populations simply by physiological or morphological characteristics, and the previous discussion emphasises the difficulties experienced in distinguishing the two species.

The goal of this study was to determine the degree of genetic divergence between the two taxa of hake occurring off the coast of South Africa and Namibia. The intention was to employ two different methods to distinguish between the two taxa. One method was looking at the variation of

nuclear genes and the other was determining the amount of genetic variation in cytoplasmic genes. The first method used in this study was protein electrophoresis, and the second method was restriction endonuclease fragment- and restriction site-analysis of mtDNA. The results obtained by these methods were compared, and the data generated during this study were used to detect different stocks that might exist within each species.

Parts of the mtDNA from both species of hake were cloned and these clones were used as probes to determine the phylogenetic relationship of the two southern African hake species to the New Zealand hake, *M. australis*. Finally, functionally important mitochondrial genes were identified by sequencing parts of the mtDNA clones. At the same time, these data were used to elucidate the evolution of mtDNA between widely different organisms.

CHAPTER 2

PROTEIN ELECTROPHORETIC ANALYSIS OF *MERLUCCIUS CAPENSIS* AND *M. PARADOXUS*

2.1 INTRODUCTION

2.1.1 General

Species are defined as self-reproducing organisms which can interbreed among themselves but which are reproductively isolated from other such groups (Mayr, 1982). Often the limited mobility of each individual within a species, compared to the geographical range of the species, prevents genetic continuity. In fact, most species do not behave as simple panmictic groups (Shaklee, 1982). Non random reproduction within species does occur, because geographic and environmental barriers often reduce or prevent gene flow between individuals in different areas. This may give rise to subpopulations which are self-sustaining breeding units within the species. If the differentiation between the groups is great enough, they may be classified as discrete subspecies, or more frequently, they are referred to as stocks. Major factors influencing the amount of genetic differentiation among stocks are: (1) The amount of time since the populations were separated from one another; (2) Population size, as genetic drift and founder effects may be important mechanisms bringing about changes in allelic frequency; (3) The amount of gene flow between populations; and (4) the intensity of natural selection on individuals in different habitats. To formulate harvesting and conservation measures for fishery management, it is important to be aware of any subpopulations, as it is

crucial to maintain continued successful reproduction of each individual stock.

Normally species are distinguished according to their morphological traits, but often these traits are not very obvious, as, for instance, in the two "species" of hake found along the coast of southern Africa.

Over the past two decades, protein electophoresis was developed to measure the amount of inherited biochemical variation that existed in natural populations. Presently, electrophoretic analysis of proteins is used to determine the taxonomic status of species and the evolutionary interrelationships of populations and species (Allendorf and Utter, 1979). Electrophoretic protein analysis makes it possible to examine a large number of protein encoding genes without using breeding experiments, which are impractical. Other powerful uses of protein electrophoresis are: (1) identification of F_1 interspecific hybrids; (2) species-specific alleles can be used as markers to identify juvenile, larval and embryonic stages of fish; and (3) it can be used to identify isolated tissue samples such as fish fillets which lack diagnostic morphology.

Many fishes have been studied extensively using electrophoretic methods. They have been used to study rainbow trout (*Salmo gairdneri*), where a coastal and inland group were distinguished (Utter *et al.*, 1973). Reproductive mixing in spawning populations is small, as the adults

always return to their stream of birth. Another example of fish with similar behaviour is trout, and electrophoresis proved to be a valuable tool when runs of fish, from endangered spawning areas, had to be allowed to escape from the fishing grounds (Utter *et al.*, 1973).

Genetic stock identification was also applied to marine fishes such as Pacific herring (*Clupea pallasi*), showing that herring are not genetically subdivided, as tagging and morphological data suggested (Grant and Utter, 1984).

The amount of gene flow between populations of demersal fishes, appears to differ to a small degree from species to species. Oceanic populations of Pacific hake (*M. productus*) appear to act as a single panmictic unit, with low latitude spawning off southern California and Mexico, followed by annual feeding migrations to the north, up to Vancouver Island, Canada (Bailey *et al.*, 1982). In another study it was thought that the Benguela upwelling system of the West Coast of South Africa might act as a barrier between populations of the southern African pilchard (*Sardinops ocellata*). No genetic differences were found between the pilchards of South Africa, and Namibia (Grant, 1985). In general, marine fishes seldom have genetic subdivision between populations, because there are few physical or behavioural barriers to migration.

The aim of this study was to determine the taxonomic status of the two different hakes (*M. capensis* and *M. paradoxus*)

occurring along the coast of southern Africa, using a large number of biochemical markers. The gene products of 31 protein-coding loci were examined for Mendelian variation. The loci that were found to be polymorphic were used to describe the genetic population structure of both species of hake. This part of the work was done in close collaboration with W. S. Grant (see also Grant *et al.*, 1988a, 1988b).

2.2 MATERIALS AND METHODS

2.2.1 Sample collection and preparation

Samples of *M. paradoxus* and *M. capensis* were collected from ten and eleven locations, respectively, extending from the South Coast of South Africa to northern Namibia (Table 2.2.1 and 2.2.2; Fig 2.2.1.). Eyes, muscle, liver and heart tissues were collected from all individuals and kept frozen at -20°C until further analysis in the laboratory. Soluble proteins were extracted from muscle, liver or heart tissues by placing them in distilled water and homogenising with a glass rod by hand. Eye fluids were tested undiluted. The homogenate was centrifuged for five min at 1000*g* to remove cell debris.

2.2.2 Protein Electrophoresis

Horizontal starch-gel electrophoresis followed the procedure of May *et al.* (1979). All of the enzymes examined were resolved using three buffer systems (Table 2.2.3). Gels consisted of thirteen percent hydrolysed potato starch in 250 ml of the appropriate gel buffer. About one quarter of the buffer was used to suspend the starch in a 1000 ml Erlenmeyer flask. The remainder of the buffer solution was heated until boiling and then added to the starch-buffer suspension. The mixture was then heated to near boiling,

Table. 2.2.1

Locations, dates, depths and sample sizes of *Merluccius paradoxus* used for electrophoretic analysis.

Sample	Location		Date	Depth	Sample size		
number	S lat.	E.long	collected	(m)	total	male	female
Namibia							
1.	23°02′	13°05′	14.01.84	384	89		
2.	25°08′	13°38′	29.01.84	450	98	4	94
3.	27°03′	14°08′	28.01.84	420	76	3	73
4.	27°57′	14°58′	20.01.84	130	100		
West Coast of South Africa							
5.	29°01′	14°36′	22.01.84	299	99	58	41
6.	31°30′	16°00′	15.01.84	478	82	26	56
7.	32°05′	16°40′	14.01.84	346	94	38	56
8.	33°40′	17°30′	28.01.84	357	98	41	57
South Coast of South Africa							
9.	34°40′	18°25′	09.01.84	305	101	85	16
10.	34°20′	25°40′	27.11.84	205	46	42	4

Table. 2.2.2

Locations, dates, depths and sample sizes of *Merluccius capensis* used for electrophoretic analysis .

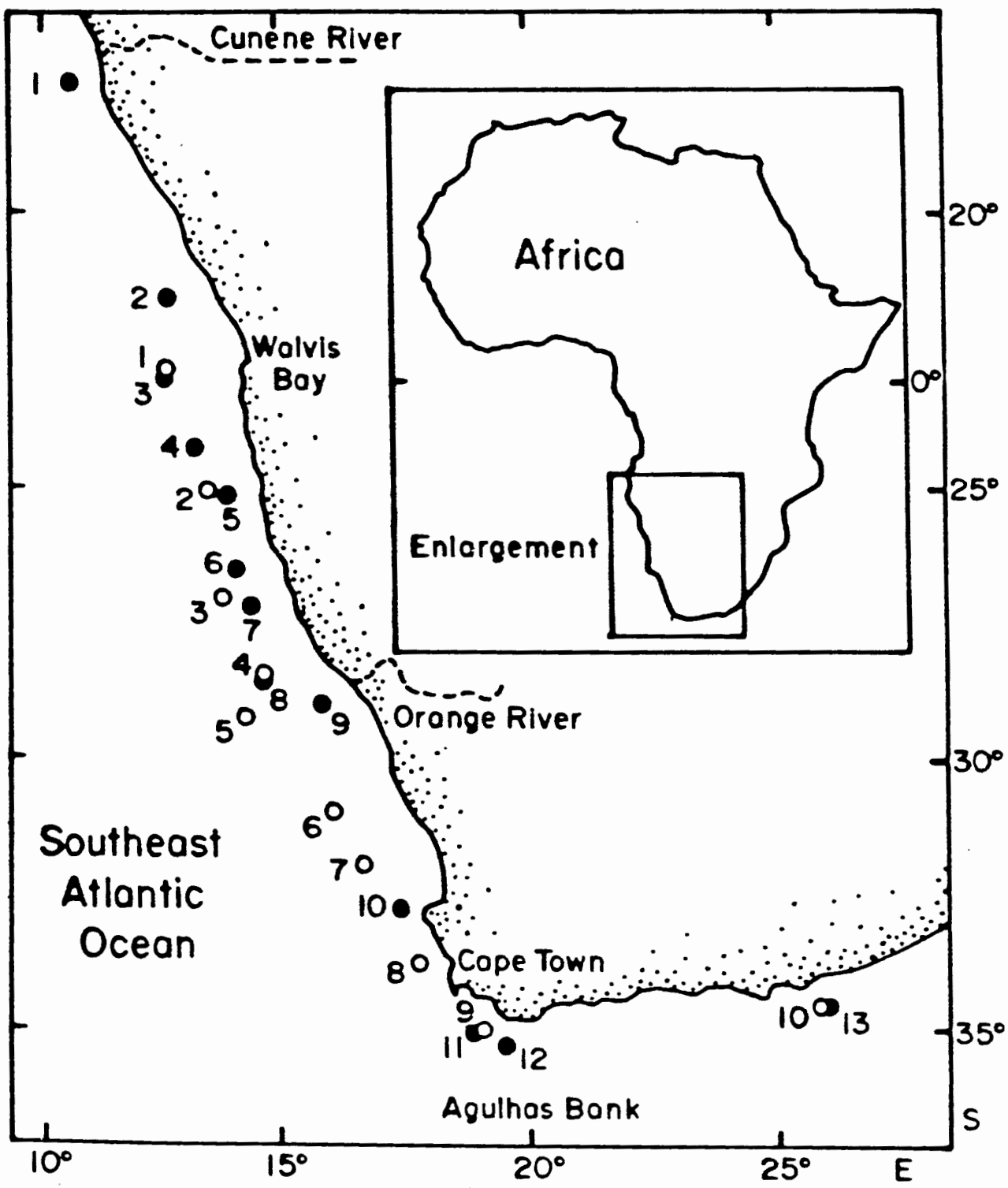
Sample	Location		Date	Depth	Sample size		
number	S lat.	E long	collected	(m)	total	male	female
Namibia							
1.	19°42′	11°48′	16.01.86	195	64	26	38
2.	22°03′	13°31′	28.01.86	136	100	50	50
3.	23°02′	13°05′	14.01.84	384	94		
4.	24°29′	13°36′	31.01.84	350	85	46	39
5.	25°01′	14°06′	17.01.84	179	101		
6.	26°07′	14°17′	18.01.84	216	100	46	54
7.	27°02′	14°35′	19.01.84	298	100	42	58
8.	27°57′	14°58′	20.01.84	193	91	36	55

Table 2.2.2 (continued)

Sample number	Location S lat.	Location E long	Date collected	Depth (m)	Sample size total	male	female
West Coast of South Africa							
9.	28°52′	16°28′	24.01.84	86	100	28	72
10.	32°55′	17°40′	12.01.84	301	101	51	50
11.	35°15′	19°25′	07.01.84	303	100	62	38
South Coast of South Africa							
12.	35°00′	14°05′	26.11.84	150	50	24	26
13.	34°20′	25°40′	29.11.84	205	50	5	45

Figure 2.2.1

Locations of samples of *Merluccius capensis* (●) and *M. paradoxus* (o) used in study of genetic variation.



swirling it vigorously all the time. The starch mixture was degassed using a suction pump, and was poured into a prepared form consisting of an 18 by 25 cm glass plate with four 7 by 18 mm plexiglass strips clamped around the perimeter. The hot starch mixture was allowed to cool, usually overnight, before use. The prepared gel was cut approximately three centimeters from the cathodal end. Supernatant from each sample prepared previously, was absorbed onto a filter paper wick. Forty to fifty of these wicks were placed vertically in a cut gel (origin). Disposable absorbent cloths were used to conduct the current from the electrode buffer to the appropriate ends of the gel. Electrophoresis was then started and the wicks were removed after five minutes of electrophoresis, the gel was pressed together firmly and an ice pack was placed on top of the gel to prevent denaturation of the proteins by heat. Electrophoresis was terminated after three to five hours, when an indicator dye marker had migrated 7-9 cm from the origin.

2.2.3 Staining Procedure

Following electrophoresis, each gel was sliced horizontally into four or five 1.5 mm thick slabs. This was done by sequentially placing pairs of 1.5 mm plastic strips on the sides of the gel and drawing thin fishing line through the gel. The locations of the enzymes in the gels were observed

using the histochemical staining procedures of Harris and Hopkins (1976). Twenty-five milliliters of the appropriate stain solution was mixed with molten agar to give a 1.0% agar concentration and this stain mixture was poured over each slab. The slabs were then incubated at 37°C in the dark. Genotypes for each individual were recorded as soon as the enzyme staining process was complete. The components of the staining solutions for the enzymes routinely examined in this study are listed in Table 2.2.3.

2.2.4 Nomenclature

Nomenclature of protein loci followed that of Allendorf and Utter (1979). Multiple loci, coding for functionally similar enzymes, were designated numerically, starting from the cathodal end of the gel. Variant alleles were designated according to their electrophoretic mobility relative to the most common allele, which was designated 100. Designations of allelic variants migrating cathodically were preceded by a minus sign. The abbreviations used to designate each enzyme are given in Table 2.2.3. When written in *italics*, these same abbreviations represent the genetic loci coding for these enzymes.

Table 2.2.3

Composition of staining solutions used for electrophoretic analysis of the enzymes routinely examined in this study. The Enzyme Commision numbers are given in parentheses.

Enzyme	Locus	Dye components	Electro-phoretic buffer	Best tissue
Creatine kinase (2.7.3.2.)	*Ck-A*	20 mg phosphocreatin	RW	M
	Ck-B	30 mg glucose	RW	M,E
		20 mg adenosine-diphosphate		
		1.5 ml glucose-6-P dehydrogenase		
		20 mg hexokinase		
		MTT, PMS		
		20 mg NADP		
Glucose-phosphate isomerase (5.3.1.9)	*Gpi-A*	10 mg fructose-6-P	RW	M
	Gpi-B	10 mg NADP	RW	M
		MTT, PMS		
Glyceraldehyde-phosphate dehydrogenase (1.2.1.12)	*Gap-1*	20 mg fructose-1,6-phospate	TC	M
	Gap-2	1 ml aldolase	TC	E
		30 mg arsenate		
		20 mg NAD		
		MTT, PMS		
Glycerol-3-phosphate dehydrogenase (1.1.1.8)	*Gpd-A*	30 mg 3-glycerol-phosphate	TC	M
	Gpd-B	20 mg NAD	TC	M
	Gpd-C	MTT, PMS	TC	M

Table 2.2.3 (continued)

Enzyme	Locus	Dye components	Electro-phoretic buffer	Best tissue
Guanine deaminase (3.5.4.3)	*Gda*	1.5 ml guanine stock solution 1 ml xanthine oxidase MTT, PMS	MF	H, L
Isocitrate dehydrogenase (1.1.1.42)	*Idh-B*	30 mg DL-isocitrate 10 mg NADP MTT, PMS	TC	M, H
Lactate dehydrogenase (1.1.1.27)	*Ldh-A* *Ldh-B* *Ldh-C*	2 ml lactate stock solution 20 mg NAD MTT, PMS	RW RW RW	M M L
Malate dehydrogenase (1.1.1.37)	*Mdh-A* *Mdh-B*	2 ml malate stock solution 20 mg NAD MTT, PMS	TC TC	M M
Malic enzyme (1.1.1.40)	*Me*	2 ml malate stock solution 10 mg MgCl 10 mg NADP MTT, PMS	TC	L, M
Mannose-phosphate isomerase (5.3.1.8)	*Mpi*	30 mg mannose-6-phosphate 10 mg NADP 1 ml glucose-P isomerase 1.5 ml glucose-6-P dehydrogenase MTT, PMS	MF	H, M

Table 2.2.3 (continued)

Enzyme	Locus	Dye components	Electro-phoretic buffer	Best tissue
General protein or Nonspecific protein	*Pt-1*	100 ml coomassie blue stain	RW	M
	Pt-2	soak for about 15 min.	RW	M
	Pt-3	then soak in destain	RW	M
	Pt-4		RW	M
	Pt-5		RW	M
Nucleoside phosphorilase (2.4.2.1)	*Np*	10 ml PO4 buffer,pH 7.5 10 mg inosine 1 ml xanthine oxidase MTT, PMS	MF	E
Peptidase (D) (3.4.13.9)		10 ml Tris-HCl pH 8, 0.5 M 20 mg of one of following substrates		
	Pep-A	leucyl-tyrosine	MF	M,H
	Pep-B	leucyl-glycyl-glycine	MF	M
	Pep-C	leucyl-tyrosine	MF	M
	Pep-D	phenylalanine-proline	MF	M
	Pep-X	appeared using leu-tyr and leu-gly-gly	MF	M
		10 mg snake venom 10 mg peroxidase 30 mg O-dianisidine-HCl		
Phosphogluco-mutase (2.7.5.1)	*Pgm-1*	20 mg glucose-1-phosphate	RW	M
	Pgm-2	with trace amounts of G-1,6-	RW	M,L
		10 mg NADP 1.5 ml glucose-6-P dehydrogenase MTT, PMS		

Table 2.2.3 (continued)

Enzyme	Locus	Dye components	Electro-phoretic buffer	Best tissue
6-Phospho-gluconate dehydrogenase (1.1.1.44)	*Pgd*	30 mg 6-phosphogluconate 10 mg NADP MTT, PMS	TC	M
Sorbitol dehydrogenase (1.1.1.14)	*Sdh*	40 mg sorbitol 20 mg NAD MTT, PMS	RW	L
Superoxide dismutase (1.15.1.1)	*Sod-1*	MTT, PMS	RW	H, L

All the stains were mixed in 10 ml 0.1M Tris-HCl pH 8.0, unless otherwise stated. 2 ml MTT (1g/100 ml) and 1 ml of PMS (1g/100 ml) were added and the solution was mixed. Then 10 ml of 2% agar was added to the mixture. The buffers indicated are those showing best results for a specific enzyme and the tissues that gave the best results are also shown, where E is eye fluids, L is liver, H is heart and M is muscle tissue. The buffer systems are listed in the Appendix.

2.2.5 Genetic Interpretations

Since no breeding experiments were performed to establish the genetic basis of the banding variants, several criteria were used to infer the mendelian nature of the phenotypes: (1) Gene expression in closely related organisms was considered, because gene expression tends to be similar in related groups of organisms,

(2) Initially, all of the different tissues were tested for expression of a locus, because the variant phenotypes should be parallel among tissues,

(3) The subunit construction of functionally similar isozymes is usually conserved across species and we compared our results to previously obtained results of related species,

(4) The phenotypic proportions in the sample fitted those expected with random mating, or Hardy-Weinberg proportions, which were obtained by binomial expansion of the sums of allelic frequencies,

(5) No unexpected phenotypes appeared on the gels. Unexpected phenotypes may indicate taxonomic heterogeneity among the samples, or it may mean that the particular genetic model used, must be revised.

2.2.6 Statistical Analysis

2.2.6.1 Hardy-Weinberg Proportions

In a randomly-mating population, mating occurs in the population without regard to genotype, and the expected proportions of genotypes in the generation following random mating are a binomial expansion of allelic frequencies. For a two allele system, Hardy-Weinberg proportions are:

$$(a + b)^2 = a^2 + 2ab + b^2$$

where a is the frequency of an allele, and $b = 1 - a$, a^2 and b^2 are the expected proportions of homozygotes, and $2ab$ is the expected proportion of heterozygotes. The expected numbers of individuals for each phenotypic class are obtained by multiplying each proportion by sample size N. The phenotypic proportions in a sample should fit those expected from random mating.

Deviations from Hardy-Weinberg proportions may be measured as the difference between the proportions of expected (H_E) and observed (H_O) relative to the proportion of expected heterozygotes

$$D = (H_E - H_O) / H_E .$$

A negative value of D indicates a deficit of heterozygotes and a positive value an excess of heterozygotes.

The test for goodness of fit (Sokal and Rohlf, 1981) was used to test the distributions of phenotypes at each locus, for fit of Hardy-Weinberg proportions. For loci having low

expected values of some phenotypes, phenotypic classes were pooled according to Swofford and Selander (1981).

2.2.6.2 Heterozygosity

Expected heterozygosity was the proportion of heterozygotes expected with the assumption of random mating (Hardy-Weinberg proportion), and was calculated by:

$$h = (1 - \Sigma p_i^2)$$

where p_i denotes the frequency of the ith allele. Average heterozygosity (H) was the average of h over all loci, including monomorphic loci.

2.2.6.3 Gene Diversity Analysis

Total gene diversity can be partitioned into within-and between-subpopulation components as a means of describing the dynamics of population structure (Nei, 1973). Nei (1973) gene diversity analysis with the computing algorithm of Chakraborty *et al.* (1982) was used to estimate the relative amounts of gene differentiation between groups at various levels of subdivision. The total population gene diversity (H_T; heterozygosity of pooled allele frequencies) was partitioned into two components

$$H_T = H_S + D_{ST}$$

where H_S is the average subpopulation heterozygosity and D_{ST} is the proportion of H_T that is due to differences between

subpopulations. The relative proportion of the total gene diversity that is contained within subpopulations is H_S/H_T and $G_{ST} = D_{ST}/H_T$ is the the relative proportion of the total diversity that is due to differences between subpopulations.

2.2.6.4 Genetic Distance

Nei's (1972) coefficients of identity (I) and distance (D) were used to measure genetic relationships between samples. I is defined as:

$$I = J_{xy} / (J_{xx} J_{yy})^{1/2}$$

where J_{xx} and J_{yy} are the averages over loci (including monomorphic loci) of the sum of the squares of allelic frequencies within populations X and Y, respectively. J_{xx} was calculated as follows:

$$J_{xx} = (\Sigma p_{xij}^2)/\text{loci}$$

where p_{xij} is the allelic frequency of the jth allele of the ith locus in population X. Similarly J_{yy} is calculated for population Y. J_{xy} is the average over loci of the cross products of allelic frequencies of the two populations

$$J_{xy} = (\Sigma p_{xij} p_{yij})/\text{loci}.$$

Genetic distance, D, is defined as the negative natural log of I

$$D = -\ln(I)$$

The standard error of D was computed according to Nei and Roychoudhury (1974). Values of I can range from 0.0, which represents total dissimilarity, to 1.0, representing identity, and the values of D can range from 0.0, representing no genetic difference, to infinity.

2.3 RESULTS

2.3.1 Electrophoretic variation

The gene products of 32 protein-coding loci were identified in *M. capensis*, and 31 loci were resolved in *M. paradoxus*, using the criteria of Allendorf and Utter (1979). The results were also compared to those reported for other species of hake (Utter and Hodgins, 1969, 1971; Utter *et al.*, 1970). In *M. capensis*, 19 loci were fixed for the 100 allele including, *Ck-B*, *Gap-1*, *Gap-2*, *Gpd-A*, *Gpd-B*, *Gpd-C*, *Ldh-C*, *Mdh-B*, *Mpi*, *Np*, *Pep-X*, *Pt-1*, *Pt-2*, *Pt-3*, *Pt-4*, *Pt-5*, *Sdh*, *Sod-1*, and *Pgd*. Seven loci (22.6%), *Ck-A*, *Gpi-A*, *Gpi-B*, *Pgm-1*, *Pgm-2*, *Pep-B*, and *Pep-D*, had common-allele frequencies less than 0.95 in at least one sample, and 6 loci, *Gda*, *Idh-B*, *Ldh-A*, *Ldh-B*, *Mdh-B*, and *Me*, had common-allele frequencies less than 1.00 but greater than 0.95.

In *M. paradoxus*, 13 loci, *Ck-B*, *Gap-2*, *Gda*, *Gpd-A*, *Gpd-B*, *Ldh-C*, *Mdh-B*, *Pep-X*, *Pt-1*, *Pt-2*, *Pt-4*, and *Sdh*, were fixed for the most common allele in *M. capensis*. Five loci, *Ck-A*, *Gap-1*, *Gdp-C*, *Pt-5*, and *Np*, were fixed for alleles other than the 100 allele in *M. capensis*; 8 loci (25.8%), *Gpi-A*, *Pgm-1*, *Pgm-2*, *Mdh-B*, *Mpi*, *Idh-B*, *Me*, and *Pep-D*, had common-allele frequencies less than 0.95; and *Gpi-B*, *Ldh-A*, *Ldh-B*, *Pgd*, and *Sod-1*, had common-allele frequencies less than 1.00 but greater than 0.95. Only the loci with common-allele

frequencies less than 0.95 were used for testing of geographic population structure.

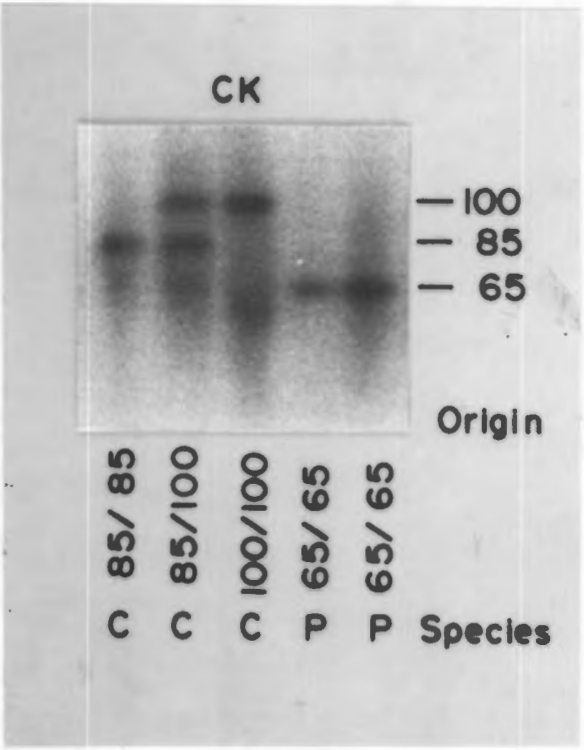
2.3.2 Description of Enzyme Systems

2.3.2.1 Creatine Kinase Ck

This enzyme showed two zones of activity, with *Ck-A* being expressed most strongly in skeletal muscle, while *Ck-B* appeared in eye fluids in both, *M. capensis* and *M. paradoxus*. Fisher *et al*. (1980) reported three CK loci in other Gadiformes. One isozyme was restricted to the eye and the brain, which agrees with the results found for these two species of hake. The most anodic zone *Ck-B* was very faint in *M. paradoxus*, but stronger in *M. capensis*. This locus was monomorphic in both species. The anodic zone *Ck-A* had double- and single-banded phenotypes in *M. capensis*, which were interpreted as heterozygotes and homozygotes respectively (Fig. 2.3.1). Ferris and Whitt (1978) have shown that although Ck is a dimer in fishes, the heterodimers of heterozygotes do not form because of temporal or spatial isolation of the Ck subunit synthesis and assembly. *Ck-A* of *M. paradoxus* was monomorphic.

Figure 2.3.1

Electrophoretic variability for creatine kinase (*Ck-A*) in muscle tissues of *Merluccius capensis* (C) and *Merluccius paradoxus* (P).



2.3.2.2 Glucosephosphate Isomerase - Gpi

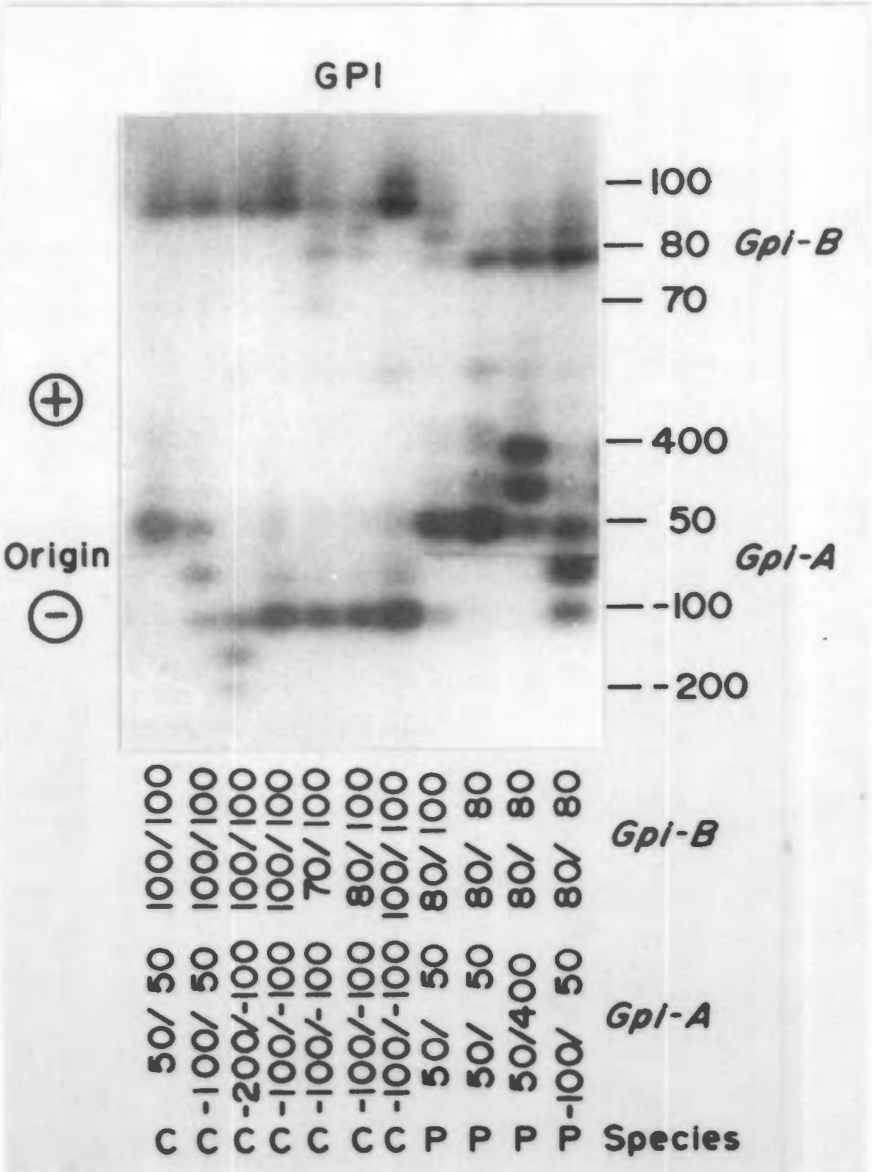
Three zones of activity appeared when gels were stained for this dimeric enzyme, and were interpreted to reflect the products of two loci, *Gpi-A* and *Gpi-B*, and a zone of intermediate mobility reflecting interlocus heterodimeric isozymes. This agreed with Dando (1974), who reported that all teleostean fishes have two distinct autosomal loci for Gpi. *Merluccius paradoxus* had two anodic zones of banding, whereas *M. capensis* had one cathodic zone of banding close to the origin and one anodic zone. In both species *Gpi* was most strongly expressed in skeletal muscle. Dando (1974) found that both loci were expressed equally well in heart muscle, but skeletal muscle also expressed both satisfactorily. In the two species of hake, both loci had single-banded and triple-banded phenotypes, which reflected homozygotes and heterozygotes, respectively (Fig. 2.3.2). The fast allele of *M. capensis*, which migrated anodally, had the same mobility as the most common allele of *M. paradoxus*.

2.3.2.3 Glyceraldehyde-phosphate dehydrogenase-Gap

Two loci, *Gap-1* and *Gap-2*, coding for this enzyme, were found in both species of hake. The protein encoded by *Gap-2* had the same electrophoretic mobility in both species, whereas the protein encoded by *Gap-1* of *M. paradoxus* had

Figure 2.3.2

Electrophoretic variability for glucosephosphate isomerase (*Gpi*) in muscle tissues of *Merluccius capensis* (C) and *Merluccius paradoxus* (P).



double the mobility of the one found in *M. capensis*. Both loci in both species of hake were monomorphic.

2.3.2.4 Glycerol-3-phosphate dehydrogenase - Gpd

The products of three loci were observed in both species; *Gpd-A* and *Gpd-B* were best expressed in muscle tissue extracts, and *Gpd-C* was best expressed in heart extracts. Fisher *et al*. (1980) found *Gpd-C* best expressed in liver, and three loci were reported. *Gpd-B* of *M. capensis* showed rare broad-banded heterozygotes. The other loci were monomorphic and had the same electrophoretic mobility in both species.

2.3.2.5 Guanine deaminase - Gda

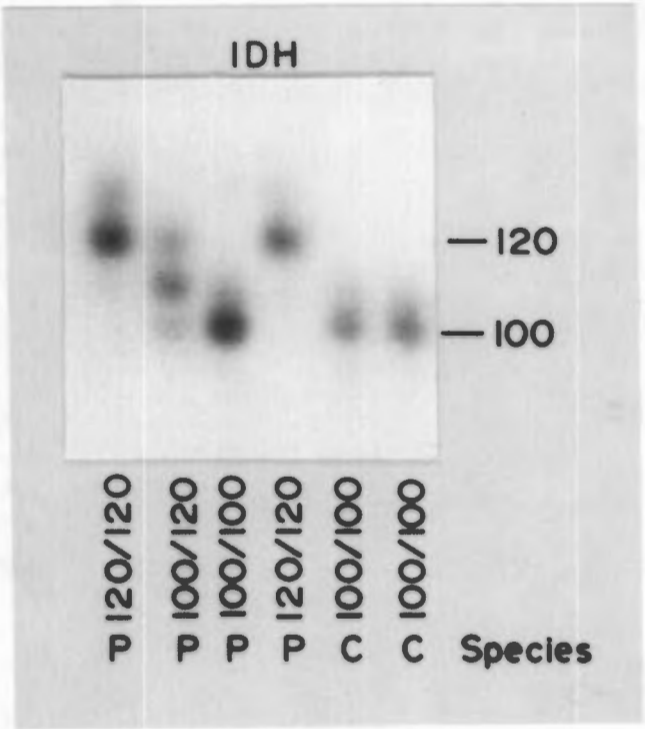
Only one zone of banding was observed for both species of hake when gels were stained for this enzyme. Both *M. capensis* and *M. paradoxus* were monomorphic for this enzyme, and it was expressed best in heart and liver extracts.

2.3.2.6 Isocitrate dehydrogenase - Idh

Staining for this dimeric enzyme revealed two zones of activity, with *Idh-B* being most prominent in extracts of skeletal muscle, and *Idh-A* most strongly expressed in liver

Figure 2.3.3

Electrophoretic variability for isocitrate dehydrogenase (*Idh-B*) in muscle tissues of *Merluccius capensis* (C) and *Merluccius paradoxus* (P).



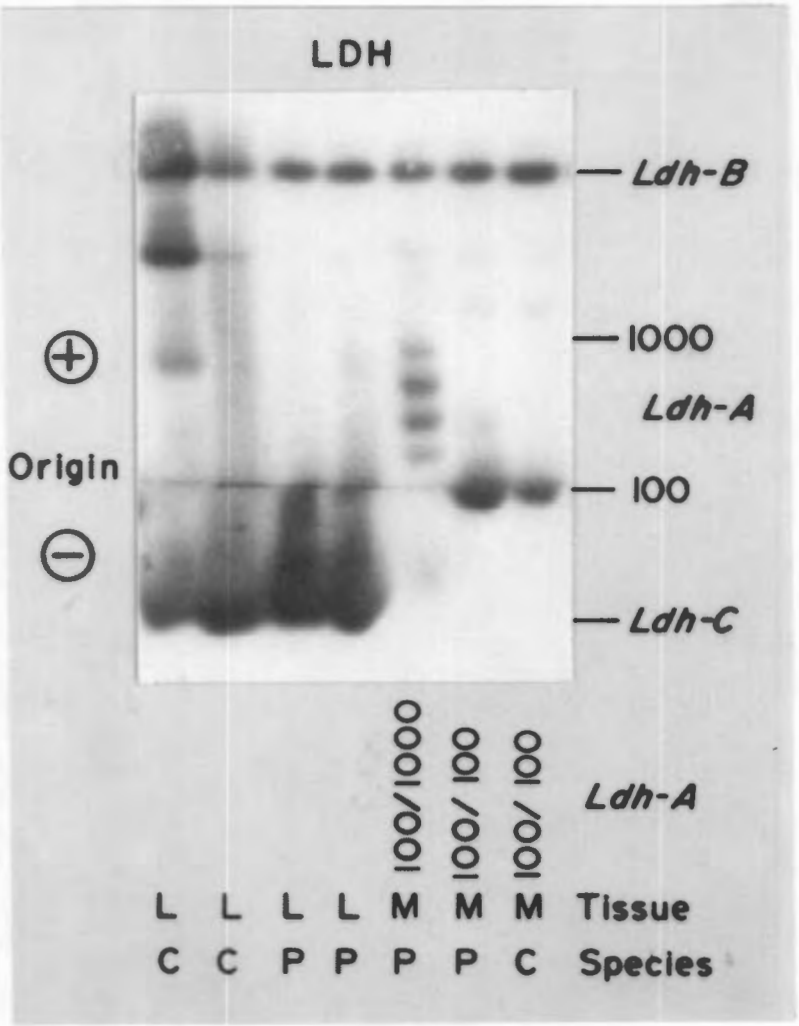
and heart tissue. In *M. paradoxus*, heterozygotes for *Idh-B* expressed symmetrical three-banded phenotypes, indicating that a single locus codes for this dimeric enzyme in *Merluccius* (Fig. 2.3.3). The *Idh-B* locus had rare triple-banded heterozygotes for *M. capensis*. The most common allele of *M. paradoxus* migrated with the same mobility as the allele most common in *M. capensis* (Fig. 2.3.3). The phenotypes of *Idh-A* were not well resolved on the gels and were not used for routine genotyping.

2.3.2.7 Lactate dehydrogenase - Ldh

The products of three loci were observed, where two zones of activity were cathodic and one was anodic. The most cathodic locus *Ldh-C* was only expressed in liver, and the other two loci, *Ldh-A* and *Ldh-B*, were expressed best in muscle tissue. Shaklee and Whitt (1981) analysed the Ldh isozymes of Gadiform fishes and reported that nearly all bony fishes had three distinct Ldh loci. In many Gadiform fishes the C4 isozyme is expressed as a relatively cathodal isozyme, predominating in the liver. Two species of Meriluccidae (*Merluccius bilinearis* and *Merluccius albidus*) showed that liver contained exclusively the C4 isozyme, and the other two isozymes were found in skeletal muscle (Shaklee and Whitt, 1981). Similarly, Utter and Hodgins (1971) reported that there were three loci coding for Ldh in the Pacific hake, *M. productus*. The loci detected in *M.*

Figure 2.3.4

Electrophoretic variability for lactate dehydrogenase (*Ldh*) in muscle and liver tissues of *Merluccius capensis* (C) and *Merluccius paradoxus* (P).



capensis and *M. paradoxus* agree with the results reported for other *Merluccius* species. In *M. capensis* and *M. paradoxus*, *Ldh-B* had rare heterozygotes, and *Ldh-A*, which migrated at the origin, had rare five-banded heterozygotes in *M. paradoxus* indicating that the enzyme is a tetramer (Fig. 2.3.4).

2.3.2.8 Malate dehydrogenase - Mdh

This dimeric enzyme was encoded by two loci in hake. Similarly, it was reported by Fisher *et al*. (1980) that two Mdh loci encoding cytosolic isozymes were present in fish. The more anodic zone, *Mdh-A*, was monomorphic for *M. paradoxus*, but rare broad-banded heterozygotes were found in *M. capensis*. Both loci were best expressed in muscle tissue extracts. Products of *Mdh-B*, which migrated more cathodally, showed no polymorphism for *M. capensis*, and only very rare broad-banded heterozygotes were detected for *M. paradoxus*. The most common allele of *Mdh-A*, of *M. capensis*, had the same electrophoretic mobility as the allele found in *M. paradoxus*.

2.3.2.9 Malic enzyme - Me

One anodal zone of activity was resolved for this tetrameric enzyme, which reflected a single locus. In *M. capensis* this

enzyme was rarely polymorphic. It was expressed well in liver and skeletal muscle of both species. In *M. paradoxus*, the homozygote gave rise to a narrow-banded phenotype, while the heterozygote gave rise to a broader-banded phenotype. Gene expression in these species of hake differs somewhat from other gadids. Grant and Stahl (1987), for instance, found expression of three loci in Atlantic cod (*Gadus morhua*) and of two loci in Pacific cod (*Gadus macrocephalus*).

2.3.2.10 Mannose-phosphate Isomerase - Mpi

One zone of banding was observed for both species. The zones were strongest in muscle tissue extracts, but the banding patterns were clearer in heart tissue extracts. *Mpi* was monomorphic for *M. capensis*, and *M. paradoxus* was highly polymorphic. Nine alleles could be distinguished for *M. paradoxus* in the form of double-banded heterozygotes and single-banded homozygotes.

2.3.2.11 General Protein - Pt

When staining for general proteins, five different zones of banding were observed (*Pt-1*, *Pt-2*, *Pt-3*, *Pt-4*, and *Pt-5*) which were all monomorphic for both species of hake.

2.3.2.12 Nucleoside phosphorylase - Np

Only single-banded homozygotes were detected in both *M. capensis* and *M. paradoxus* when staining for this enzyme. The expression was best in eye fluids and the enzyme of *M. paradoxus* migrated more anodally in comparison to the one of *M. capensis*.

2.3.2.13 Peptidases

Several different isozymes were detected using different di- and tripeptide substrates.

2.3.2.13.1 Pep-A (Leu-tyr)

Pep-A was expressed strongest in heart and skeletal muscle. *Pep-A* was highly polymorphic in both species of hake, but could not be adequately resolved for genotyping.

2.3.2.13.2 Pep-B (leu-gly-gly)

Pep-B in *M. paradoxus* had at least nine alleles which had mobilities close to one another, and interpretation of the products of this locus, for genotyping, was therefore not attempted. In *M. capensis*, only four alleles were found for *Pep-B*, of which the heterozygotes had double-banded

phenotypes with equal band intensity, whereas the homozygotes had single-banded phenotypes. These two-banded heterozygous phenotypes suggested that *Pep-B* encoded a monomeric enzyme.

2.3.2.13.3 Pep-C (leu-tyr)

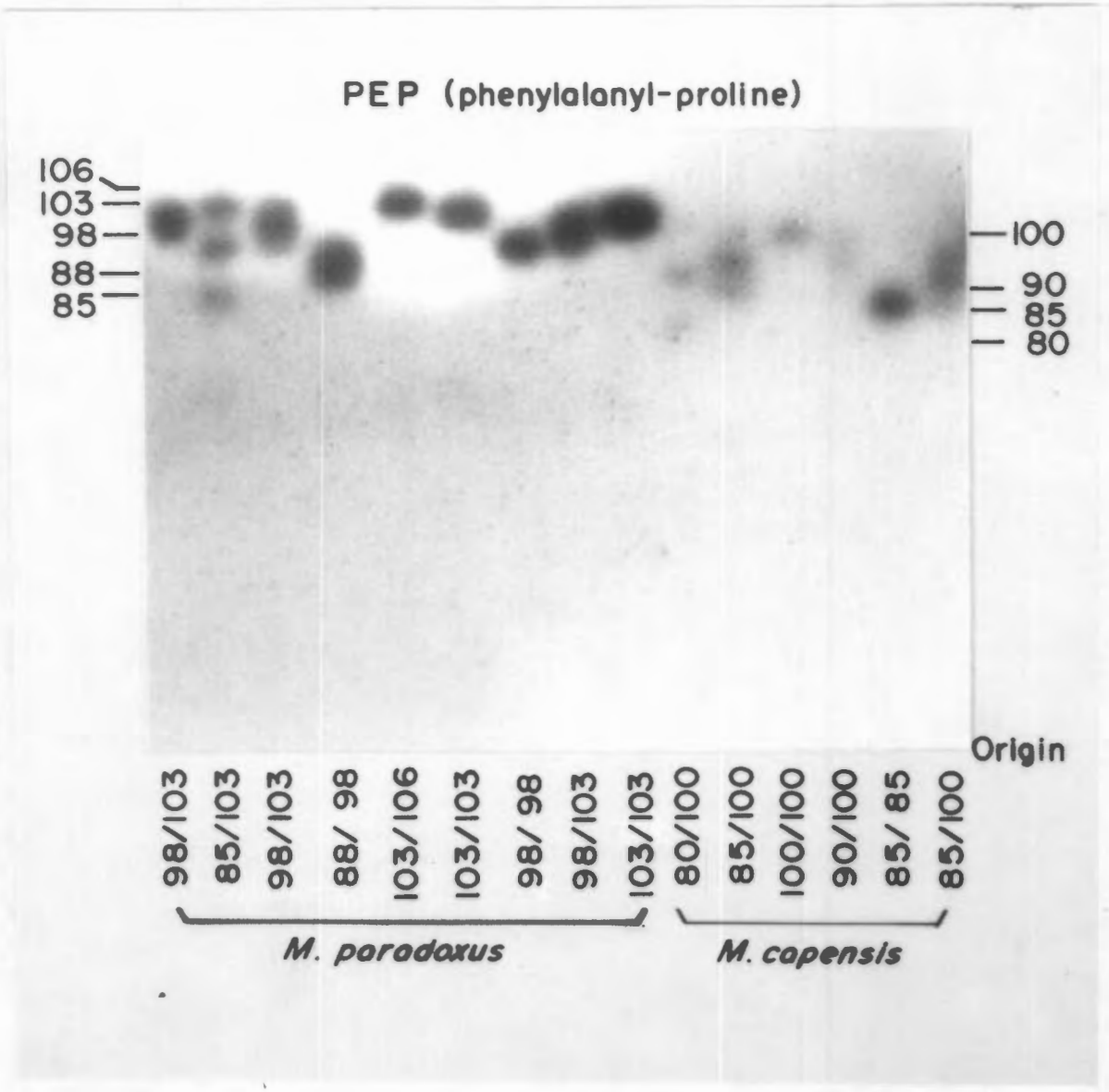
This peptidase appeared to be polymorphic in both species. Poor resolution with our buffers resulted in *Pep-A* and *Pep-C* migrating very close to each other. Genetic interpretation was not attempted for *Pep-C* in either species, because of poor resolution, complicated variable banding patterns, and partial overlap between *Pep-C* and *Pep-A*.

2.3.2.13.4 Pep-D (Phenylalanyl-proline)

Staining for this peptidase revealed one polymorphic locus and both species of *Merluccius* had at least six alleles. The heterozygote had symmetrical broad-banded phenotypes and also rare triple-banded phenotypes, suggesting a dimeric subunit construction. The homozygote had a narrower-banded phenotype (Fig. 2.3.5). In both species, *Pep-D* was expressed in the heart and skeletal muscle, and although banding was clearer in the skeletal muscle, it was much stronger in the heart extracts. This can be clearly seen in Fig. 2.3.5, where *M. paradoxus* samples were prepared from

Figure 2.3.5

Electrophoretic variability for the peptidase *Pep-D* (using substrate phenylalanyl-proline) in muscle tissues of *Merluccius paradoxus* (P) and in heart tissue of *Merluccius capensis* (C).



skeletal muscle, and *M. capensis*, from heart extracts. The most common allele of *M. paradoxus* was slightly more anodic than the most common allele of *M. capensis*.

2.3.2.13.5 Pep-X (leu-tyr and leu-gly-gly)

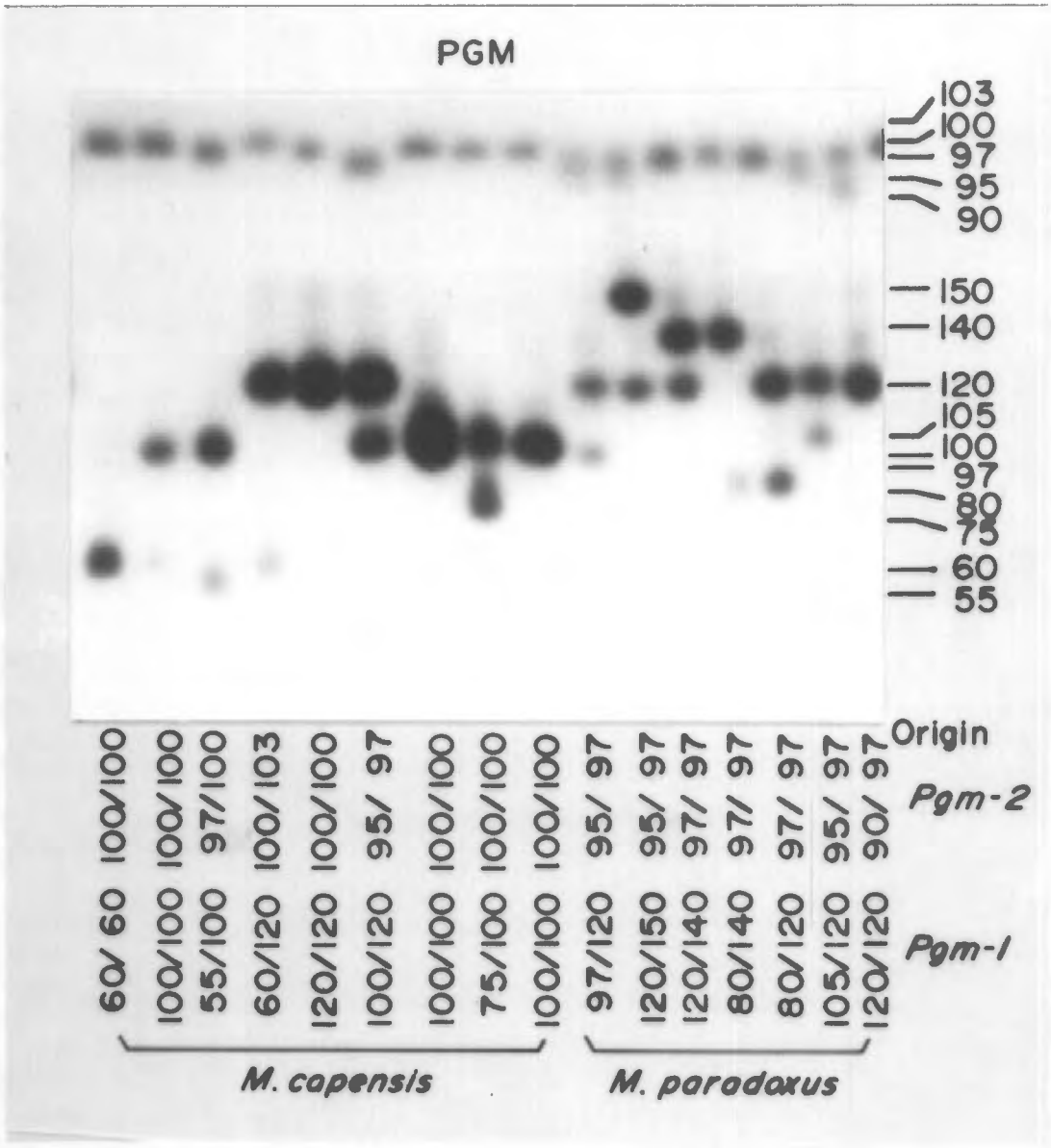
A more anodal zone of banding appeared when the two substrates leu-tyr and leu-gly-gly were used. This peptidase was designated *Pep-X* after the nomenclature of Frick (1983). This locus was invariant in both species.

2.3.2.14 Phosphoglucomutase - Pgm

Two loci coding for this monomorphic enzyme were observed; *Pgm-1* was best expressed in skeletal muscle and *Pgm-2* most strongly in liver extracts. The most anodic zone, *Pgm-2* showed narrow single-banded, and double-banded or broad-banded phenotypes. The former phenotype was interpreted as representing homozygotes, and the latter two phenotypes represented heterozygotes. This zone was also much lighter than the more anodic zone, *Pgm-1*. *Pgm-1* had single- and double-banded phenotypes, which reflected homozygotes and heterozygotes, respectively. The slow allele in *M. paradoxus* corresponded to the most common allele (100) in *M. capensis*, as can be seen in Fig. 2.3.6. Similarly, in *Pgm-*

Figure 2.3.6

Electrophoretic variability for phosphoglucomutase (*Pgm*) in muscle tissues of *Merluccius capensis* (C) and *Merluccius paradoxus* (P).



2, the most common allele of *M. paradoxus* had the same mobility as the most common allele in *M. capensis*.

2.3.2.15 6-Phosphogluconate dehydrogenase - Pgd

One zone of activity was detected in both species when staining for this enzyme. No heterozygotes were observed for *M. capensis*, and *M. paradoxus* had single-banded homozygotes and broad-banded heterozygotes. The allele found in *M. capensis* had the same mobility as the slow allele (60) in *M. paradoxus*, and they were expressed strongest in muscle tissue extracts.

2.3.2.16 Sorbitol dehydrogenase - Sdh

Sdh coded for one monomorphic enzyme, which was best expressed in heart tissue in both species.

2.3.2.17 Superoxide dismutase - Sod

Sod migrated cathodaly and had rare triple-banded heterozygotes. For the majority of species, only single-banded phenotypes were detected for *Sod*, in most tissues (Fisher *et al.*, 1980). No gene duplication was noted for this locus. Mangaly and Jamieson (1978), on the other hand,

reported two zones of activity for the Sod variants found in the European hake (*Merluccius merluccius*).

Allelic frequencies were pooled by species, and are presented in Table 2.3.1.

2.3.3 Statistical analysis

In *M. capensis*, average sample heterozygosities ranged from 0.046 to 0.055, and averaged 0.051 ($\pm$0.023) over 13 samples. The relative proportion of the total gene diversity that is contained within subpopulations was 0.983, and the relative proportion of the total gene diversity that is due to differences between subpopulations was 0.017. The average D between pairs of samples of *M. capensis* was 0.0007.

For *M. paradoxus*, sample heterozygosities ranged between 0.059 to 0.070 and averaged 0.063 ($\pm$0.026) over 10 samples. The difference between the average value of H for *M. capensis* and *M. paradoxus* was not significant (t-test on arcsin transforms of locus heterozygosities; Archie, 1985). The relative proportion of the total gene diversity that is contained within subpopulations was 0.987 for *M. paradoxus*. The relative proportion of the total gene diversity that is due to differences between subpopulations was 0.013. The average D between pairs of samples of *M. paradoxus* was 0.0006.

Table 2.3.1

Summary of alleles frequencies for southern African hakes *Merluccius capensis* and *M. paradoxus*. Alleles are arranged according to their electrophoretic mobility beginning from the cathodal end of the gel.

		Allele														
Locus	Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14	*h*
Ck-A	C	.001	.198	.004	.797	.001										.327
	P	-	-	-	-	1.0										.000
Gap-1	C	1.0	-													.000
	P	.002	.998													.004
Gda	C	.999	.001													.002
	P	1.0	-													.000
Gpd-C	C	-	1.0													.000
	P	.998	.002													.004
Gpi-A	C	-	-	.052	.753	.003	.190	.002	-	-	-	-				.394
	P	.001	.001	-	.005	-	.967	-	.005	.020	.001	.001				.059

Gpi-B	C	.001	.015	.001	.027	.001	.956	.001		.081
	P	.002	-	-	.993	.005	-	-		.016
Idh-B	C	-	-	.001	.999					.001
	P	.021	.816	-	.163					.308
Ldh-A	C	-	.001	.999	-					.002
	P	.001	.002	.985	.012					.030
Ldh-B	C	.005	-	.995						.009
	P	.002	.003	.995						.059
Mdh-A	C	.002	.997	.001						.000
	P	-	1.0	-						.000
Mdh-B	C	-	-	1.0						.000
	P	.002	.995	.003						.010
Me	C	-	-	.998	.002	-	-			.040
	P	.005	.018	.931	-	.043	.003			.131
Mpi	C	-	-	-	-	1.0	-	-	-	.000
	P	.011	.173	.010	.656	.001	.143	.003	.003	.520
Np	C	1.0	-							.000
	P	-	1.0							.000
Pep-D	C	.005	.358	.010	.057	-	.569	.001	- -	.551
	P	-	.001	-	-	.232	.005	.548	.209 .005	.602

Pgd	C	-	-	-	-	1.0										.000
	P	.001	.972	.002	.023	.001										.054
Pgm-1	C	.004	.009	.026	-	.003	-	.893	.004	.054	-	.007	-	-	-	.204
	P	-	-	-	.006	-	.038	.002	-	-	.906	-	.003	.042	.003	0.169
Pgm-2	C	.002	-	.006	-	.984	-	.008								.028
	P	-	.002	.065	.921	-	.012	-								.147
Pt-5	C	-	1.0													.000
	P	1.0	-													.000
Sod-1	C	-	-	-	1.0											.000
	P	.001	.008	.991	-											.019

The following loci were fixed for the same allele in both species *Ck-B*, *Gap-2*, *Gpd-A*, *Gpd-B*, *Ldh-C*, *Pep-X*, *Pt-1*, *Pt-2*, *Pt-3*, *Pt-4*, and *Sdh*.

Genetic identity, I, between *M. capensis* and *M. paradoxus* was 0.558 ($\pm$0.160) and genetic distance, D, between the two species was 0.583 ($\pm$0.160).

2.4 DISCUSSION

2.4.1 Taxonomy of the southern African hake

The first question addressed in this study was whether the two nominal species of hake occurring along the coasts of southern Africa are in fact two species or whether they should be considered subspecies. These two hakes are morphologically very similar and the amount of genetic divergence separating them has never been determined. Generally, genetic distance can be used as a guide to taxonomy, especially in cases where morphological differentiation is insignificant. The genetic distance, D, between the two hakes was found to be 0.583. Thorpe (1982) summarised genetic distance data for 2664 pairs of taxa of invertebrates and vertebrates, and found that an average D between congeneric species was 0.62. Shaklee *et al.* (1982) reviewed the genetic distances between fish taxa, and found that pairs of taxa, classified as related species, had D values ranging from 0.025 to 0.65 and averaged 0.30. These values show that the genetic distance between the two South African hake is such that they should be classified as full species, *M. capensis* and *M. paradoxus*.

Another significant observation was that several loci were fixed for different alleles which were not shared between these two species of hake. From this it appears highly likely that two species do not share a common gene pool.

No individuals with hybrid genotypes were detected during this study. Initially it was thought that some hybrids did exist along the coast where the Orange River opens into the Atlantic Ocean. These hake were thought to be hybrids because of their morphology. Their anal fins seemed to have hybrid colouring. North of the Orange River the anal fin of *M. capensis* is entirely white and that of *M. paradoxus* is black (Grant *et al.*, 1988a). These individuals were tested to determine if they had hybrid genotypes, but each individual showed either the *M. capensis* or *M. paradoxus* genotype. Jones and Mackie (1970) also showed, with the aid of zone electropherograms of muscle myogens, that only two forms of hake existed off the coast of Namibia. The absence of hybrid genotypes indicates that no interbreeding occurs between the two species of hake.

2.4.2 Evolution of the two sympatric species of hake

This then leads to the next question - whether these two sympatric species of hake are derived from the same ancestor, or whether they have evolved from two separate ancestors. In theory it is possible that sympatric speciation can occur, but as Mayr (1982) proposed, that it is most likely that isolation in allopatry and polyploidy are the primary modes of speciation. It is unlikely that sympatric speciation occurs. Normally, geographical

barriers lead to the formation of subpopulations and eventual speciation. If the two species of hake evolved from one common ancestor, they could have arrived in southern Africa as one population which then speciated into two species. This is unlikely, as no geographical barriers exist that could separate them into two populations, and they are still inhabiting the same waters today.

One model of sympatric speciation proposes that through assortive mating, a barrier eventually forms between two populations. This is often observed in birds, where behavioural rituals are very important (Futuyma, 1979). The two sympatric species of southern African hake rise off the bottom during spawning season and move into midwater where they spawn. They spawn at the same time of the year, but at different depth (Botha, 1973) and the pelagic eggs and larvae drift in nearshore currents at the surface. This indicates that there is some extent of selective mating between the two hake (e.g. depth), but the eggs mix readily in surface currents. It is not likely that selective mating led to sympatric speciation in these two hakes.

Other models of sympatric speciation assume that different genotypes adapt to different niches and if the selection for one genotype in the new niche is strong enough to prevent migration between the niches, these populations will form separate species. One possibility, in the case of the two South African hake, is that one population adapted to deeper

water, and eventually the two became reproductively isolated. However, the two species still spawn in the same waters, which makes it impossible for speciation to have occurred in such a manner.

Evidence exists that some reproductive isolation can evolve sympatrically, but it has never been proven that complete speciation ever took place in that way (Futuyma, 1979). It is assumed that closely related species, living in sympatry, must have made contact after the formation of reproductive barriers to gene flow, which prevents reassimilation into a single gene pool. Sympatric species do not necessarily share one common ancestor, but they may have descended through intermediate taxa. This suggests that the two South African species of hake, *M. capensis* and *M. paradoxus*, did not have the same ancestral species. They must have started to inhabit the same waters only after the formation of a barrier to gene flow.

Genetic distance can also give an estimation of the divergence time between two taxa, where $D = 1.0$ is about 19 million years of separation between two taxa (Carlson *et al.*, 1978; Vawter *et al.*, 1980). The value of $D = 0.583$ (± 0.160) for the two hakes gives an indication that they must have separated 7.6 to 13.6 million years ago.

Interpretation of the genetic stock structure of the southern African hake and further statistical analysis of the results is included in a separate study (see Grant *et al.*, 1988a, 1988b).

CHAPTER 3

RESTRICTION FRAGMENT ANALYSIS OF THE MITOCHONDRIAL DNA FROM *MERLUCCIUS CAPENSIS* AND *M. PARADOXUS*

3.1 INTRODUCTION

3.1.1 General

Eucaryotic cells contain many different organelles in their cytoplasm. The mitochondrion is one of these organelles, and it is the seat of the main energy-yielding reactions such as the citric acid cycle, and oxidative phosphorylation. Mitochondria vary in size and shape, depending on organism and tissue, and the abundance of mitochondria per cell also fluctuates. The generally bacterium-like shape of mitochondria, and the fact that they grow in length and then divide, suggested to early-twentieth-century cytologists that mitochondria might have their own genetic material, and that they might have evolved from a procaryotic ancestor, to live in symbiosis with their eucaryotic host (Watson *et al.*, 1987; Stent and Calendar, 1971; Metzler, 1977).

Around the 1950s, a special class of yeast mutants ("petite") became known. These mutations did not segregate as would be expected if the mutation was of nuclear origin, and it was concluded that the mutation was, in fact, cytoplasmic (Roman, 1982; Watson *et al.*, 1987). "Petite" mutants contain mitochondria that are missing several vital enzymes, and are therefore not able to generate ATP. The yeast strains survive by means of anaerobic glycolysis, forming small ("petite") colonies as a result. To explain

these mutant mitochondria, it was postulated that they lack the DNA, coding for certain proteins, that is found in normal mitochondria. Then, around 1963, it was discovered that all eucaryotic mitochondria possess small circular DNA molecules (mtDNA).

3.1.2 Structure and size of mtDNA

The amount of mtDNA present per cell varies from 50 - 100 copies per diploid cell. The mtDNA is isolated as a covalently closed circular (ccc) molecule, with a size ranging from 14.5 to 19.5 kb in higher eucaryotes (Watson *et al.*, 1987). The size of human, mouse, bovine, amphibian (*Xenopus laevis*), and invertebrate (*Drosophila yakuba*) mtDNA are 16,569, 16,295, 16,338, 17,553, and 16,019 bp, respectively, whereas the mitochondrial "long" genome of laboratory wild-type yeast strains, is 85 kb in size (Anderson *et al.*, 1981; 1982; Bibb *et al.*, 1981; Roe *et al.*, 1985; Clary and Wolstenholme, 1985; de Zamaroczy and Bernardi, 1986).

The size of the mtDNA is about 1% that of *Escherichia coli* chromosomal DNA, therefore it cannot code for all the mitochondrial proteins, and it was determined that the proteins present in the mitochondria are only partly self-made, while most of them are encoded by nuclear DNA (Watson

et al., 1987). The organization, replication, transcription and translation of the mtDNA is further reviewed in Chapter 5.

3.1.3 Size variation of mtDNA

Many instances of size variation of the mtDNA have been reported within species and between species. Most of these length-polymorphisms could be located on the Displacement loop (D-loop) of the mtDNA. One example is the goat D-loop, which is 75% as long as the sheep D-loop (Upholt and David, 1977). Another example is the D-loop of bovine mtDNA, which is 212 nucleotides shorter than the human D-loop (Anderson *et al.*, 1981, 1982). In separate studies, length mutations were discovered in humans (Cann and Wilson, 1983b; Greenberg *et al.*, 1983), apes (Ferris *et al.*, 1981b), mice (Ferris *et al.*, 1983b), lizards (Moritz and Brown, 1986), *Drosophila* (Fauron and Wolstenholme, 1976; Goddard *et al.*, 1982), and crickets (Harrison *et al.*, 1985). In the human mtDNA some of the additions or deletions were not located in the D-loop area (Cann and Wilson, 1983), and in the lizards, the length variation was due to a duplication of parts of the D-loop region, the rRNA genes, and some tRNA genes (Moritz and Brown, 1986).

3.1.4 Mitochondrial DNA as a means to measure Relatedness in Natural Populations

3.1.4.1 Rate of evolution of the mtDNA

The evolutionary rate of the molecule used to determine the evolutionary relationships between populations of organisms, should match the evolutionary rate of the taxa being studied. The function, gene content, and gene organisation of the mtDNA is highly conserved in all higher eucaryotes. It was assumed that the mtDNA nucleotide sequence might also be highly conserved. Brown *et al.* (1979) were the first to show that mtDNA has been evolving much more rapidly (5-10 fold) than single copy nuclear DNA. They constructed restriction endonuclease cleavage maps for mtDNA of four higher primates (Guinea baboon, rhesus macaque, guenon, and human), and by alignment of the maps with respect to origin of replication, the nucleotide sequence divergence was calculated from the number of restriction sites shared. Then the mitochondrial sequence divergence was plotted against divergence times estimated from fossil or protein data, and the rate of base substitutions could be calculated from the initial slope of the curve. The evolutionary rate of mtDNA was estimated to be 0.02 substitutions per base pair (2%) per million years. From the plateau of the curve, it seemed likely that certain highly variable positions in mtDNA become saturated by substitutions in about 10-20 million years, after which the accumulation of further

sequence differences slows dramatically (Brown *et al.*, 1979; Aquadro *et al.*, 1984; Avise, 1986). Ferris *et al.* (1983b) studied the mtDNA diversity in eight species of mice, and confirmed that the rate of sequence divergence of mtDNA was 2-4% per million years. For sea urchins (echinoid taxa), it was found, however, that the nuclear and mtDNA evolve at approximately the same rate, which was attributed to fluctuations in the rate of evolution of the nuclear DNA rather than mtDNA (Vawter and Brown, 1986).

Several hypotheses exist to explain the high rate of evolution of mtDNA:

(1) A high rate of mutation, which can arise from the higher turnover rate of the mtDNA, relative to the nuclear DNA, providing more rounds of replication, which allow more errors to be generated (Brown *et al.*, 1979). Another reason for the high rate of mutation, might be the fact that the mitochondria lack the repair system, which excises thymine dimers, and it is also likely that the mitochondria are inefficient in repairing other types of DNA damage (Brown *et al.*, 1979).

There are very high incidences of length mutations and transitions in mtDNA. The mismatch repair system of bacteria is specifically designed to repair length mutations and transitions, and microbial mutants lacking the mismatch repair system have a similar mutational spectrum as the one observed during evolution of the mammalian mtDNA (Wilson *et*

al., 1985). This observation confirms that the high evolutionary rate of the mtDNA, is due to inefficient repair of DNA.

(2) Rapid fixation of the mutations can occur, because of relaxed selective constraints on mitochondrial gene products. Another reason is that the translational apparatus only has to process few a kinds of mRNAs (Brown *et al.*, 1979; Avise, 1986). The mutations could also be fixed more readily in the mtDNA, because an individual would still survive, even if one of the mitochondrial genes was inactivated, due to the polyploidy of mtDNA in a single cell.

(3) Another reason for the rapid evolution of the mtDNA is the "wobble" in the genetic code, which will allow variability to exist in the switch from one code to another (Wilson *et al.*, 1985). The rapid evolution of the mitochondrial tRNAs and rRNAs, compared to their nuclear counterparts, also explains the high rate of evolution of the mtDNA.

Hauswirth and Laipis (1982) postulated various reasons for the rapid mitochondrial genotype variation. They reported that two mitochondrial genotypes, which were distinguished by the absence or presence of one *Hae* III restriction site, existed within one Holstein cow maternal lineage. Simple segregation of variant molecules could not explain the rapid variation, because numerous mitochondria are present in each

single cell. Several models, which all assume a heterozygosity of mtDNA within germ-line cells, were proposed to explain this phenomenon (Hauswirth and Laipis, 1982):

(4) There are only a few mtDNA molecules in germ cells, which would lead to a genetic founder effect, and the preferential amplification of one mtDNA molecule could lead to a shift from one mitochondrial genotype to another, in a single generation.

(5) In another model, it was proposed that the mitochondria are segregating into different fetal tissues in the developing oocyte (Hauswirth and Laipis, 1982). This would give rise to different mitochondrial genotypes.

A few examples exist where mtDNA heteroplasmy was detected. *Drosophila mauritiana* was shown to have two types of mtDNA molecules, of different length, occurring in the same individuals, and some crickets had more than one size class of mtDNA (Solignac *et al.*, 1983; Harrison *et al.*, 1985).

In any event, it is now a widely accepted fact that the mtDNA evolves faster than the nuclear DNA. This rapid rate of evolution makes it a useful molecule with which to study the relationships between species and populations. In general, it is the method of choice to estimate relatedness between species which diverged recently, about 5-10 million years ago (Brown *et al.*, 1979).

3.1.4.2 Inheritance of the mtDNA

Another interesting aspect of mitochondria is their manner of inheritance. Progeny inherit most, if not all, their mtDNA from their female parent, and they are not altered by recombination or meitotic segregation (Lansman *et al.*, 1981). Explanations for the maternal inheritance of mtDNA are that mitochondria, carried in the sperm midpiece, do not penetrate the egg at fertilization, or that the sperm mitochondria change, so that they cannot be used within the egg (Lansman *et al.*, 1983b). This means that individuals within a species that are derived from different females, are genetically isolated from each other with respect to mtDNA, even if they are members of one interbreeding population. Thus, all offspring from one female with a certain mitochondrial genotype, would form a mtDNA clone, all with the same mitochondrial genotype (Upholt and Dawid, 1977). Many studies were performed to determine the maternal inheritance of mtDNA in mice (Ferris *et al.*, 1982; Lindahl, 1985; Gyllensten *et al.*, 1985) humans (Giles *et al.*, 1980), and insects (*Heliothis* Lansman *et al.*, 1983). Gyllensten *et al.* (1985) tested the paternal inheritance of mtDNA in mice by extensive backcrossing of female interspecific hybrids to males of the paternal species. From the results obtained, they concluded that the father

contributes no more than one mtDNA per thousand maternal mtDNAs to the zygote.

In other studies with mouse-human cell hybrids, it was discovered that the only mtDNA which was detectable originated from the parent whose chromosomes were retained in the nucleus (de Francesco *et al.*, 1980). This demonstrates that only a single type of mtDNA is retained within cell hybrids. Ferris *et al.* (1982) determined, by restriction site mapping, that old inbred mice showed no variation in the mtDNA genotypes, and this was due to the fact that only one female lineage contributed to the formation of all old inbred lines. The maternal inheritance of mtDNA was thereby confirmed.

Evolutionary history of female movement and survival in a species can therefore be studied by examining variation of the mtDNA. Often the gene flow among populations is due to the migration of males, whereas the females remain more sedentary, resulting in a highly structured population with regard to female transmitted traits, but less differentiated with respect to nuclear genes (Lansman *et al.*, 1981). In populations of organisms where it is suspected that there is an unequal flow of genetic information between sexes, the determination of sequence divergence of mtDNA becomes even more applicable to the study, because of this maternal inheritance.

In other studies it was discovered that mtDNA can be transferred between species. Spolsky and Uzzell (1984), reported that the water frog, *Rana ridibunda* had two types of mtDNA, and the explanation was that an interspecific transfer of mtDNA must have occurred from *R. lessonae* to *R. ridibunda*. In another study, it was revealed that two closely related types of mtDNA from *Mus domesticus* (mouse) were found in a different species, *M. musculus* (Ferris *et al.*, 1983a). Both these results are instances where the distribution of the mitochondrial genotype does not agree with the biological species boundaries. One question arising from this is as follows: If the mtDNA can cross species boundaries, will it not obscure evolutionary histories of diverging populations? On the other hand, by studying the mtDNA, otherwise inaccessible information about complex evolutionary processes of closely related species can be obtained. These reports are isolated incidents, and one possible explanation is that these species could actually stem from a single ancestral female.

3.1.5 Restriction enzyme analysis of mtDNA

The mtDNA is a genetic system, well characterised at the molecular level. Most of the detailed information about the structure, function, replication, transcription, and the genetic code of the mtDNA was, however, obtained through sequencing of the total mitochondrial genome. To obtain the

total mtDNA sequence is a tedious and expensive task, and from a population geneticist's point of view, not feasible, as the aim is to examine a large number of individuals for mtDNA sequence polymorphisms. This is made possible through the use of restriction endonucleases, which cleave the DNA within or near their specific recognition sequences by breaking two phosphodiester bonds, one in each strand of DNA. The number of cleavage sites will differ from one DNA molecule to another, depending on the amount of sequence divergence between the DNAs. The fragments of DNA, created by these restriction endonucleases, can then be separated by gel electrophoresis on the basis of their molecular weight. In a typical population study, mtDNA from one individual is digested with a range of different restriction enzymes. Sequence divergence is then estimated by analysis of the number of restriction fragments, or restriction sites, shared between two DNAs. The restriction sites must be mapped, or the fragment changes must be sufficiently simple, to be interpreted in terms of specific site changes. Frequently, it is impractical to map specific sites, and in that case, comparisons may be based on the number of fragments shared between the two DNAs.

One assumption made in mtDNA sequence divergence studies, is that restriction site changes arise by single base pair substitutions, and not by rearrangements (inversions or transpositions), deletions, additions, or gross organizational changes.

Another assumption generally made is that variation in restriction patterns of mtDNA is due to differences in the sequence rather than in the methylation of the mtDNA. By sequencing a small restriction fragment with a variant restriction site, Castora *et al.* (1980) showed that the variation was, in fact, due to change of a nucleotide rather than due to methylation of the DNA.

It must also be assumed that recognition sequences occur with a distribution and frequency close to that expected in a random sequence of the same base composition (Upholt, 1977). The assumption of randomness, which is required to construct phylogenetic trees from restriction data, was tested by Adams *et al.* (1982). They found that only a limited number of enzymes cleave human mtDNA randomly in all respects (*Acy* I, *Bbv* I, *Bcl* I, *Bgl* I, *Hae* II, *Hind* III and *Sph* I). Unequal rates of nucleotide substitution among different nucleotides, rather than the deviation of the GC content, is another important factor causing bias in an estimate (Tajima and Nei, 1982; Aoki *et al.*, 1981). The bias is generally larger for four-base enzymes than for six-base enzymes. The problem can be overcome, to a certain extent, by using a large number of restriction enzymes to study DNA divergence. It is concluded that estimates of phylogenetic relationships, using restriction data, are biased, depending on the genomes under study, the distribution of restriction sites, and the enzymes chosen.

Another source of error in restriction enzyme analysis, is that one cannot differentiate between two different fragments that comigrate. In addition, very small fragments are normally ignored, because they are not easily observed by gel electrophoresis.

Avise *et al.* (1979a) compared protein electrophoresis with mtDNA restriction analyses in the pocket gopher, *Geomys pinetis*, to evaluate the mtDNA approach. Their study demonstrated that mtDNA restriction analysis alone or together with protein electrophoresis is an excellent means to study related populations, and that mtDNA may give more information on variation within populations than can be obtained by protein electrophoresis alone.

3.1.6 Mitochondrial DNA as a Genetic system for Evolutionary studies

In a recent review, Avise (1986) discusses the use of mtDNA genotypes as markers of evolutionary phenomena such as genetic structure of populations, dispersal, and historical zoogeography. One of the applications of the mtDNA restriction analysis is to estimate phylogenetic relationships among very closely related species. One area of great interest is to resolve the exact branching point between humans and apes, and to investigate the evolution, and recent history of humans (Brown and Vinograd, 1974;

Potter *et al.*, 1975; Brown, 1980; Ferris *et al.*, 1981a, 1981b; Templeton, 1983; Johnson *et al.*, 1983; Avise *et al.*, 1984; Nei and Tajima, 1985; Avise, 1986). To date, many other studies using mtDNA restriction analysis have been performed on mammals such as the pocket gopher, *Geomys pinetis*, (Avise *et al.*, 1979a), the deer mouse, *Peromyscus maniculatus*, (Avise *et al.*, 1979b; Lansman *et al.*, 1983b), mice, *Mus musculus*, (Yonekawa, 1981), and rats, (Brown and Simpson, 1981; Kessler and Avise, 1985a). Some phylogenetic studies using mtDNA also involved *Drosophila* (de Salle *et al.*, 1986a, 1986b; Solinac *et al.*, 1986); and birds (Glaus, 1980; Kessler and Avise, 1985b; Mack *et al.*, 1986).

Only a few related species of fish have been examined. In one such study, the gene flow between skipjack tuna from the Atlantic and Pacific oceans, was determined (Graves *et al.*, 1984). Skipjack tuna from the two oceans did not show a high degree of differentiation in their mtDNA restriction patterns, suggesting that there must have been a continued genetic contact between the two groups since the uplift of the Panama land bridge, 3.1 million years ago.

Avise and Saunders (1984) used mtDNA markers to study hybridisation and introgression among five sympatric species of sunfish (*Lepomis*), which are known to produce hybrids. They found that the mtDNA is a good means by which to detect hybrids produced by rare instances of backcrossing, in which a hybrid was fertilized by either of its parental strains.

One mitochondrial phenotype may replace another by a small amount of gene flow between species, but only a small amount of selection against the immigrant mitochondrial phenotype is sufficient to prevent its establishment. No evidence of introgression between species of *Lepomis* could be detected by mtDNA analysis, but it was successfully used to determine the parentage of F1 hybrids (Avise and Saunders, 1984).

Bermingham and Avise (1986a) used restriction length polymorphisms in mtDNA to reconstruct evolutionary relationships in populations of four freshwater fish *Amia calva*, *Lepomis punctatus*, *L. gulosus*, and *L. microlophus* and they estimated the divergence times between these species. They found extensive mtDNA polymorphisms within each species and they also reported that mtDNA clones that were closely related were also geographically contiguous.

Gyllensten and Wilson (1987) used restriction enzyme analysis of mtDNA of Salmonids to detect intra- and interspecific variability, and Gyllensten *et al.* (1985) used mtDNA analysis and protein electrophoresis to study secondary contact of two species of cutthroat trout which occur in western North America. Their results showed that secondary contact resulted in a randomly mating hybrid swarm.

Berg and Ferris (1984) used restriction enzyme analysis to examine systematic relationships within the Salmonidae at three taxonomic levels: tribe, genus and subgenus. They

conclude that this method is ideal to identify stocks of commercially important fish.

Kornfield and Bogdanowicz (1987) investigated the relationship among spawning stocks of Atlantic herring, *Clupea harengus*, using restriction analysis of mtDNA. Based on tagging experiments, it has been accepted that herring stocks return to specific spawning sites (Kornfield and Bogdanowicz, 1987). Restriction enzyme analysis of the mtDNA, however, revealed that 65% of all the individuals shared the same composite genotype and this implied that no discrete herring stocks exist throughout the Gulf of Maine.

3.1.7 Objective of this study

At present, several approaches to the use of mtDNA restriction analysis are in general use. The previous review shows that mtDNA polymorphism is an excellent means to study evolutionary phenomena in populations, especially to estimate the phylogenetic relationships among closely related species and to determine geographic variation within species.

In this part of the study, restriction fragment analysis has been used to estimate the nucleotide sequence divergence between two species of hake, *Merluccius capensis* and *M. paradoxus*, occurring along the coasts of southern Africa. The data generated in this study were also used to search

for possible genetic differences between regional populations.

At the same time, the restriction fragment data obtained was compared to the allozyme data, to determine the correlation between these sets of data.

Another objective was to determine whether it would be feasible to use this method for future, routine stock assessments of fish.

3.2 MATERIALS AND METHODS

3.2.1 Sample collection

Initially, samples from both *M. capensis* and *M. paradoxus* were collected on board the research vessel *S. A. Africana*. Samples of both species were collected along the South Coast of South Africa. Mitochondria were isolated after collection from fresh livers kept on ice, following the method of Powell and Zungia (1983). Isolated mitochondria were stored at -20°C until further analysis in the laboratory. Later on in the study, mtDNA was extracted on board and then transported in 95% ethanol (EtOH) to the laboratory, where the extraction was completed. All the initial studies on the mtDNA were done using these extracts. The yield in these samples, however, was very low and only a few digestions could be made with mtDNA from a single individual.

Later in the study, the method for rapid isolation of mtDNA from fish by Chapman and Powers (1985) was used. Samples were obtained from commercial trawlers, where whole fish were stored on ice soon after they were caught; catch positions were also recorded by the trawler crew (Table 3.2.1; Fig. 3.2.1). Generally, only fish caught in the last trawl were sampled, so the fish were seldom on ice for more

Table 3.2.1

Table of locations, dates collected, depth at which caught, tissue used and number of samples for both *Merluccius capensis* and *Merluccius paradoxus*. L = Liver and O = Ovaries

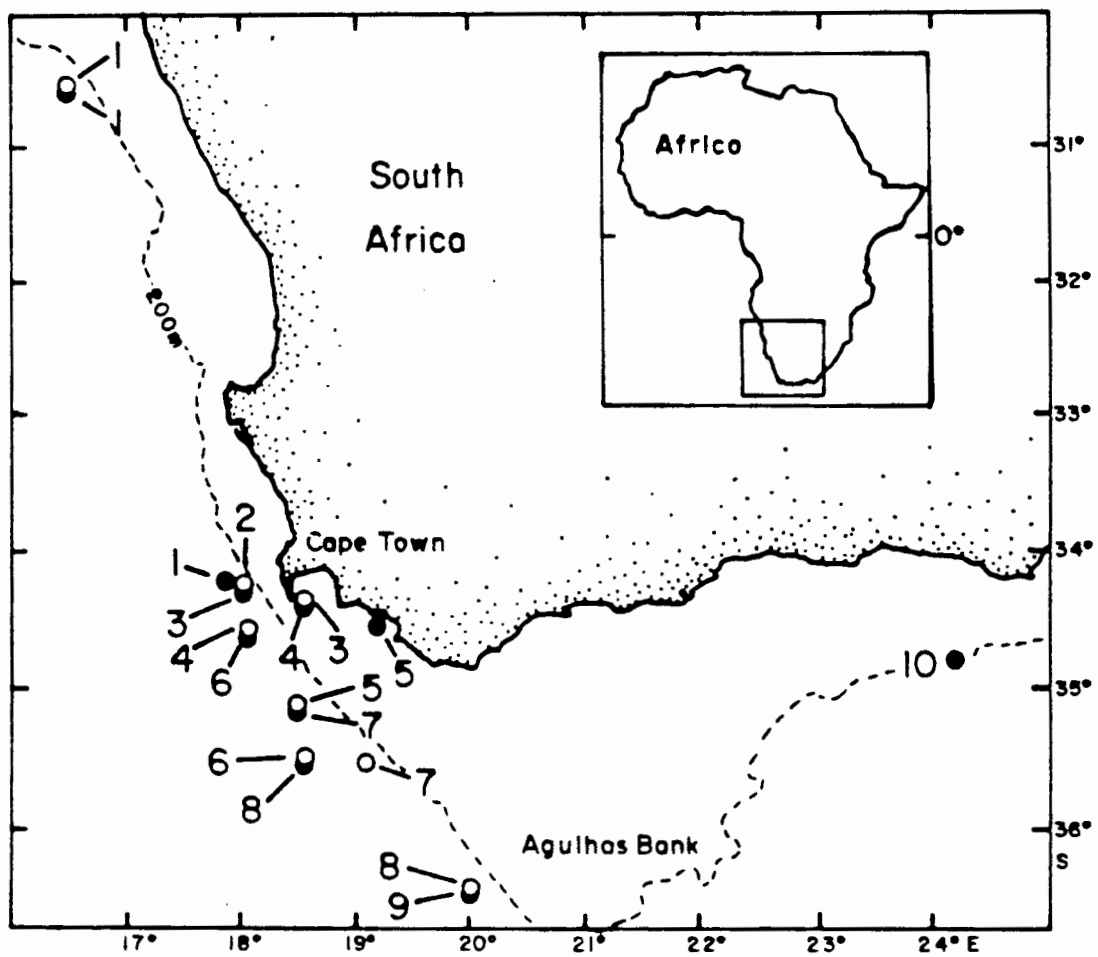
No	Sample No	Location		Date	Depth	Tissue	Sample size		
		S Lat	E Long		(m)		Tot	F	M
Merluccius capensis									
1	24,25,26	30°30	16°30	10.1.86	249	O	3	3	
2	7,8,9,10	34°10′	17°50′	3.12.85	200	L,O	4	2	2
3	11,12	34°10′	17°55′	6.12.85	180	O	2	2	
4	3,4	34°15′	18°35′	7.11.85		L	2		
5	1,2	34°20′	19°15′	26.9.85	200	L	2	2	
6	13	34°25′	18°02′	9.12.85	350	O	1	1	
7	14,15,16,17, 18	35°01′	18°35′	12.12.85	237	L,O	5	1	4
8	5	35°25′	18°45′	13.11.85		L	1		
9	6	36°21′	20°03′	21.11.85	284	L	1		
10	19,20,21 22,23	34°45′	24°10′	10.1.86	147	O	5	5	

Table 3.2.1 (continued)

No	Sample No	Location		Date	Depth	Tissue	Sample size		
		S Lat	E Long		(m)		Tot	F	M
Merluccius paradoxus									
1	47,48,49,50	30°30	16°30	10.1.86	249	O	4	4	
2	39,40	34°10′	17°55′	6.12.85	180	O	2	2	
3	27,28	34°15′	18°35′	7.11.85		L	2		
4	41,42,43,44	34°25′	18°02′	9.12.85	350	L,O	4	3	1
5	45,46	35°01′	18°35′	12.12.85	237	L	2	2	
6	29	35°25′	18°45′	13.11.85		L	1		
7	33,34,35, 36,37,38	35°30′	19°05′	27.11.85	425	O	6	6	
8	30,31,32 22,23	36°21′	20°03′	21.11.85	284	L	3		

Figure 3.2.1

Location of hake samples used for mtDNA analysis. Closed circles represent samples of *Merluccius capensis* and open circles represent samples of *Merluccius paradoxus*.



than 24 hours before mitochondria were extracted. This method proved to be simpler, much more rapid, and yielded more DNA.

3.2.2 Tissue Selection

It is important to select the best tissue for the isolation of mtDNA, as the aim is to obtain the highest yield of mtDNA with as little as possible contamination by nuclear DNA. For this reason, the ratio of mtDNA to nuclear DNA is important. In mature oocytes the ratio of mtDNA to nuclear DNA is 300:1 w/w (Chapman and Powers, 1985).

Another important factor influencing mtDNA extraction, is the ease of homogenization of the tissue. By choosing tissue that can be homogenized easily, the chances of fracturing nuclei during the process are reduced (Chapman and Powers, 1985). Again, mature oocytes proved to be the best material.

The tissue selected must be available in large quantities; during the spawning season, ovaries frequently outweigh all other tissues except muscle. All samples were collected from November to January, during the spawning season of hake. Another advantage of ripe ovaries, is that they have a low glycogen content.

Outside of the spawning season, the best tissue to use is liver as it is easily homogenized, and the ratio of mtDNA to nuclear DNA is higher than in other tissues, except gills.

Liver has a high glycogen content which interferes with further isolation steps (Chapman and Powers, 1985).

Ripe ovaries can be kept on ice for up to two weeks before mitochondria are isolated, which is very useful, as it is often impossible to obtain samples from distant locations within a few days. Liver tissue, on the other hand, can only be kept on ice for up to two days (Chapman and Powers, 1985).

During the study, the majority of mtDNA samples were isolated from mature ovaries (Table 3.2.1). The ovaries of samples 19-26 and 47-50 were obtained from the *S. A. Africana*, after they had been kept on ice for one week.

3.2.3 Tissue Homogenization.

Tissue homogenization is an important step in the isolation of mtDNA. In the initial mtDNA isolations, a hand glass homogenizer with a glass plunger was used. This method was time consuming and it was difficult to keep the homogenate cool. The advantage was that the cells were broken efficiently without lysing the nuclei.

Later, an Ultra-Turrax Type TP 18/10 was used for homogenization. The optimal time and speed were determined by homogenizing liver tissues at varying speeds for different lengths of time and by viewing the homogenate under a microscope. Homogenizing liver tissue for 3 to 5 seconds at top speed showed the best results, where most cells were broken, but the nuclei were still intact. When volumes larger than 40 ml were used, this method of homogenization was not very efficient. This step proved to have the largest yield loss. Intact eggs could be seen, but further homogenization with the Ultra-Turrax resulted in breaking of the nuclei.

3.2.4 Mitochondrial DNA Isolation

This method was adapted from Chapman and Powers (1985). Ten to 50 g of mature oocytes or liver, were homogenized in 5 volumes of TEK buffer (Appendix) with the Ultraturrax for up to 20 seconds. Samples were homogenized in 200ml Sorvall GSA rotor centrifuge tubes kept on ice. When smaller volumes were used, the samples were homogenized in 40ml Sorvall SS34 centrifuge tubes. The homogenate was underlaid with a 15% sucrose-TEK solution. Care was taken to let a sharp miniscus form between the two solutions. The homogenate was then centrifuged at 1000*g* for 10 min at 4°C. The supernatant was drawn off, and the above step was repeated. After the second low speed spin, the supernatant

was drawn off and mitochondria were pelleted by centrifugation at 18000*g* for 1h. The supernatant was drawn off and discarded, leaving behind a dense, clear glycogen layer overlayed with a loose mitochondrial pellet. The mitochondrial pellet was drawn off, leaving behind the clear pellet, and was resuspended in 20 ml TEK-buffer and again centrifuged at 18000*g* for 1h at 4°C. If a clear glycogen pellet was still present, the mitochondria were pelleted a third time. After the last centrifugation, the mitochondria were resuspended in 0.9 ml TEK-buffer per 25g starting tissue. By adding 0.1 ml 10% non-idet-TEK solution, the suspension was made 1% non-idet. The suspension was left to stand for 5-10 min, to insure that all mitochondria were lysed. The suspension was then centrifuged for 10 min at 12000*g* to remove debris. Non-idet\P40 is an anionic detergent that lyses all membranes, except nuclear membranes, therefore, any persistent nuclei were removed during the last step (Chapman and Powers, 1985). The supernatant was drawn off and mixed with an equal volume of TE saturated phenol, to precipitate proteins. Tubes were shaken thoroughly, allowed to stand for 5 min, and centrifuged at 12000*g* for 10 min. The upper aqueous phase was drawn off and reextracted with phenol. The phenol extraction was repeated a third time, if the aqueous phase was opaque after the second extraction. After the last extraction, the residual phenol was removed by extracting the aqueous phase with a 24 : 1, chloroform : iso-amyl

alcohol solution, which also removes residual protein. The upper aqueous phase was drawn off and was mixed with two volumes of cold 95% EtOH. The DNA was then precipitated by placing the samples at -20°C for at least two hours. Often, the mtDNA was stored in this state for up to three weeks. The DNA was then pelleted by centrifugation at 12000*g* for 10 min at 4°C. The ethanol was decanted and the pellets were dried at 37°C. Care was taken not to overdry the samples as this makes re-dissolving very difficult. The samples were then re-dissolved in approximately 1ml of TE buffer (Appendix). At this stage, the mtDNA could be stored at 4°C for three months or longer, without apparent loss of material. These mtDNA samples were subsequently used for digestion with restriction enzymes.

3.2.5 Restriction endonuclease Digestion

Portions of the extracted mtDNA from each individual were digested with one 5-base recognition site restriction enzyme, *Ava* I, and ten 6-base restriction endonucleases, *Bam*HI, *Bgl* I, *Eco*RI, *Hin*dIII, *Kpn* I, *Pst* I, *Pvu* II, *Sac* I, *Xba* I, and *Xho* I. Only a few samples from each species were digested with *Bgl* II, *Cla* I, *Hin*fI, *Hpa* II, *Sal* I, *Sma* I, and *Stu* I. The conditions used for digestion were recommended by the manufactures of the enzymes (Amersham Ltd., Anglian Biotechnology Ltd., and Boehringer Mannheim Ltd.). Before digesting the DNA, samples were electrophoresed on an

agarose minigel to determine the concentration of the DNA. The mtDNA of each individual was then diluted accordingly. The final concentration of the mtDNA was such that only 3µℓ were used per digestion.

DNAse free RNAse (Appendix) was added to each digestion at a concentration of 4µg/ml together with the restriction enzyme. This step reduced the background of low molecular weight RNA's, allowing much better resolution of the low molecular weight mtDNA fragments.

Digests were incubated for a minimum of four hours at temperatures recommended by the manufacturer. Some digests were left to incubate overnight to ensure complete digestion. The total digestion volume was then endlabeled with $\alpha^{32}P$ or loaded directly onto an agarose gel.

3.2.6 Agarose Gel Electrophoresis

The mtDNA fragments were separated on 0.7 to 1.0% agarose gels, depending on expected fragment sizes, determined by pilot digestions. The gels were electrophoresed with Ethidium Bromide (EtBr, see Appendix) added to the agarose at a concentration of 0.5 µg/ml, which enables one to distinguish between circular covalently closed DNA (cccDNA), nicked circular DNA, and linear DNA molecules, with the cccDNA having the fastest mobility.

Lambda phage DNA, digested with *Hin*dIII or a mixture of *Hin*dIII and *Eco*RI, was run on each gel and used to estimate sizes of the mtDNA fragments. Fragments with similar mobilities between genotypes were run side by side. Tris-borate EDTA (TBE, see Appendix) buffer was used for electrophoresis and the gels were usually run at 35V overnight.

After electrophoresis, the DNA fragments in gels were visualised by putting the gels on a 254 nm uv-transilluminator. Photographs were taken with a Polaroid camera Cu 5 using polaroid film type 667 or type 55. The type 667 film was exposed from 1 to 8 seconds whereas the type-55 film was exposed for 4 to 10 min. The latter film provided a negative for later use. In addition, a Kodak 23A orange filter, which absorbed all wavelengths above 500nm was used. EtBr bound to DNA flouresces below 500 nm, therefore the filter reduced the flourescent background caused by unbound EtBr in the gel (Chapman and Powers, 1985).

3.2.7 Endlabelling Procedure

Ten nanograms of mtDNA was digested with a restriction enzyme. In a separate tube, 1µℓ (10 µCi) of α^{32}P dCTP, 2µℓ Polymerase I large fragment (or Klenow), and 1µℓ of 10x *Hin*dIII buffer (Appendix) were mixed. One µℓ of this

solution, and 1 μℓ of a mixture of dATP, dGTP, and dTTP, (0.5 mM), were added to each tube containing digested mtDNA. This solution was mixed by pipetting it up and down, and the mixture was incubated on ice for 30 min.

Two μℓ of tracking dye was added to each digestion, and they were then electrophoresed on agarose gels. After electrophoresis the anodal end of the gel, which contained most of the unincorporated dCTP, was cut off. The gel was washed in water, blotted dry, placed onto 2 to 4 pieces of 3MM Whatman filterpaper, and dried for 2 hours using a vacuum gel dryer at 40°C.

The dried gel was placed into an X-ray cassette fitted with an intensifying screen, and a sheet of X-ray film (Kodak XAR5 or Chronex 4) was placed on top. When cooled to -70°C, the film could be developed after 3 hours. When the cassette was left at room temperature, the film was developed after an overnight exposure.

3.2.8 Data Analysis and Interpretations

One approach to the data analysis, was to consider the entire mtDNA digestion profile produced by each restriction enzyme for each individual as a quantitative genotype. Each individual organism was characterised by a composite genotype, which consisted of eleven different genotypes produced by each of the eleven restriction enzymes used in

this study. Hake sharing a given composite genotype, belong to the same mtDNA clone.

For each pair of individuals with a different mtDNA genotype, relative genetic similarity was calculated from the proportions of fragments shared in their mtDNA digestion profiles. Total proportions of shared fragments across all digestions were calculated by:

$$F = 2N_{xy} / (Nx - Ny)$$

where Nx and Ny are the numbers of fragments in genotype X and Y, respectively, and Nxy is the number of fragments shared between genotypes (Nei and Li, 1979).

This F value was then used to estimate p, the number of base substitutions per nucleotide, or the nucleotide sequence divergence, using Upholt's (1977) formula:

$$p = 1 - \left[\frac{-F + (F^2 + 8F)^{1/2}}{2}\right]^{1/n}$$

where n is the number of base pairs recognised by each restriction enzyme. We used ten restriction enzymes with recognition sequences of six bp and one enzyme (*Ava* I) with a five bp recognition site.

Since the variance of p value is large when n is small, it is important to use many different restriction enzymes to increase the reliability of p (Nei and Li, 1979).

When enzymes with the same n value were used, N_x, N_y, and N_{xy} were the numbers of unshared and shared fragments over

all restriction enzymes, but when enzymes with different n values were used, p was estimated for each n group. An average p was calculated by weighting the various p's with the number of nucleotide sites detected by each group of restriction enzymes.

Restriction enzymes recognizing a four bp sequence produce 100 to 150 bp fragments of mtDNA and are difficult to detect by normal agarose electrophoresis. An increase in the total number of fragments per digest, also increases the probability of scoring two non-homologous fragments as identical (Lansman *et al.*, 1981). Therefore, only five- or six-cutters were used for this study.

3.2.8.1 Expected number of fragments

The expected number of fragments in a mtDNA digest was calculated according to Nei and Li (1979).

$$a = (g/2)^{r1} \times [(1-g)/2]^{r2}$$

If nucleotides are distributed randomly in the mtDNA, then the above formula holds true. Here, r is the number of nucleotide pairs in the restriction site, $r1$, the number of guanines (G) plus cytosines (C), and $r2$, the number of thymines (T) plus adenines (A), so that $r = r1 + r2$. The GC content is shown in the g value.

The expected frequency of a restriction site is then converted to expected total number of restriction sites by $m^r a$. Where m^r is the size of the mtDNA in base pairs.

3.2.9 GC (Guanine + Cytosine) content determination

The GC content for both species of hake mtDNA was determined by R. Kirby. The mtDNA was run on an isopycnic gradient for 24 hours at 36,000*g* using the model E Beckman analytical ultracentrifuge, with CsCl starting density of 1.710 g/cm^3. Two DNA molecules with known GC contents were run at the same time, as standards. In this study DNA from *Clostridium acetobutylicum*, with a GC content of 28%, and *Streptomyces cattleya* DNA, with a GC content of 73%, were used to calibrate the gradient. The density of the mtDNA was then estimated from the position of the mtDNA in the gradient, relative to the other two DNA's.

3.3 RESULTS

3.3.1 Gel electrophoresis

Agarose gels of 0.8% concentration were used, which gave adequate resolution of fragment sizes up to 400 bp, using normal EtBr staining procedures.

The mtDNA of both species of hake, was digested with eleven restriction endonucleases. Endlabelling with $\alpha^{32}P$ was used for four of the restriction enzymes: *Ava* I, *Bam*HI, *Xba* I, and *Xho* I, but this only improved resolution significantly for *Ava* I digests.

3.3.2 Size and GC content of mtDNA

The length of the hake mtDNAs was determined by digesting the mtDNA with eleven different restriction enzymes, which each cut the mtDNA at least once. The sizes obtained across all digestions were averaged, and it was found that the mtDNA of *M. capensis* was 16.954 $\pm$512 bp and the size of the *M. paradoxus* mtDNA was 16.693 $\pm$405 bp. Fragment sizes are presented in Table 3.3.1.

The GC content of the mtDNA of both *M. capensis* and *M. paradoxus* was determined using an isopycnic gradient. The estimated GC content of *M. capensis* mtDNA was 47.0% ($\pm$2%) and that of *M. paradoxus* was 45.5% ($\pm$2%).

Table 3.3.1 (continued)

Pvu II		*Sac* I		*Xba* I		*Xho* I	
A:10000	(1)	A:12900	(1)	A:10000	(2)	A:11400	(1)
B:9600	(2)	B:8700	(1)	B:7500	(1)	B:11000	(2)
C:3000	(1,2)	C:8630	(2)	C:7100	(1)	C:5100	(1)
D:2100	(2)	D:3820	(2)	D:6900	(1)	D:5000	(2)
E:2000	(1,2)	E:3780	(1)	E:6000	(2)	E:4700	(1)
F:1900	(1)	F:3750	(1,2)	F:5800	(1)	F:4500	(1)
		G:3700	(1)	G:2800	(1)		
		H:3600	(2)	H:600	(1,2)		
		I:3450	(2)				
		J:700	(2)				
		K:600	(1)				
		L:550	(2)				
		M:500	(1)				
		N:400	(1)				
		O:150	(2)				

3.3.3 Expected and Observed number of Fragments

The expected and observed number of restriction sites for all the enzymes used in this study are presented in Table 3.3.2. The expected number of fragments, when using a six-bp restriction enzyme, varied from 3-4 for *M. capensis* and 3-5 for *M. paradoxus*. The expected number of fragments was constant for the six-bp cutters, as all the enzymes used in this study had recognition sites with either 2 or 4 GC molecules in the site, and the GC content of each restriction site is the only variable in the formula used. The observed number of restriction sites for these 6-bp restriction enzymes, was normally within the range of the expected number of restriction sites. Similarly, it was found that the expected number of restriction sites for the 5-bp restriction enzyme, *Ava* I, corresponded to the number of restriction sites observed. These results show that the interpretation of the digestion profiles was correct, as no unexpected number of restriction sites was reported.

3.3.4 Description of the Restriction Profiles

Fragment sizes are presented in Table 3.3.1, and the digestion patterns found for each individual are presented in Table 3.3.3, where an A in Table 3.3.3 refers to the fragment size A listed in Table 3.3.1.

Table 3.3.2

Expected and observed number of fragments produced by each enzyme when mtDNA of *M. capensis* and *M. paradoxus* was cut.

Enzymes and their recognition sequence	Number of Fragments			
	M. capensis		*M. paradoxus*	
	Expected	Observed	Expected	Observed
Ava I GPyCGPuG	13	10	12	11
*Bam*HI GGATCC	3-4	1	3	2
Bgl I GCCNNNNNGGC	3	4	2	5
*Eco*RI AATT	4	2	4-5	2
*Hin*dIII AAGCTT	4	2	4-5	2
Kpn I GGTACC	3-4	1,2	3	1,2
Pst I CTCGAG	3-4	1	3	2
Pvu II CAGCTG	3-4	4	3	4
Sac I GAGCTC	3-4	2,3,4	3	3,4,5
Xba I TCTAGA	4	4,5	4-5	3
Xho I CTCGAG	3-4	2	3	2

The observed restriction endonuclease patterns of the mtDNAs from the two species are also presented in Fig. 3.3.1. The following conventions were used to label the various fragment genotypes. The most common genotype in *M. capensis* was designated by the letter "A", and variant genotypes were "B", "C", "D", etc., including the genotypes in *M. paradoxus*. Identification of mutational transitions from one genotype to another was attempted, by examining the sizes of the restriction fragments. Genotypes that differed by one restriction endonuclease recognition site were designated by an arrow and the direction of the arrow indicated a loss of that restriction site. This direction does not imply evolutionary direction. Transitions involving a minimum of two steps, were designated by a line, with two hatch marks, connecting two genotypes. Unconnected genotypes represented transitions involving 3 or more mutation events. An insertion was represented by a line with a plus sign and a deletion was represented by a line with a minus sign.

The following section contains a brief description of the restriction fragment patterns, of all enzymes tested in this study.

Table 3.3.1

Sizes (bp) of fragments produced by digestion of the mtDNA of *Merluccius capensis* and *Merluccius paradoxus*.

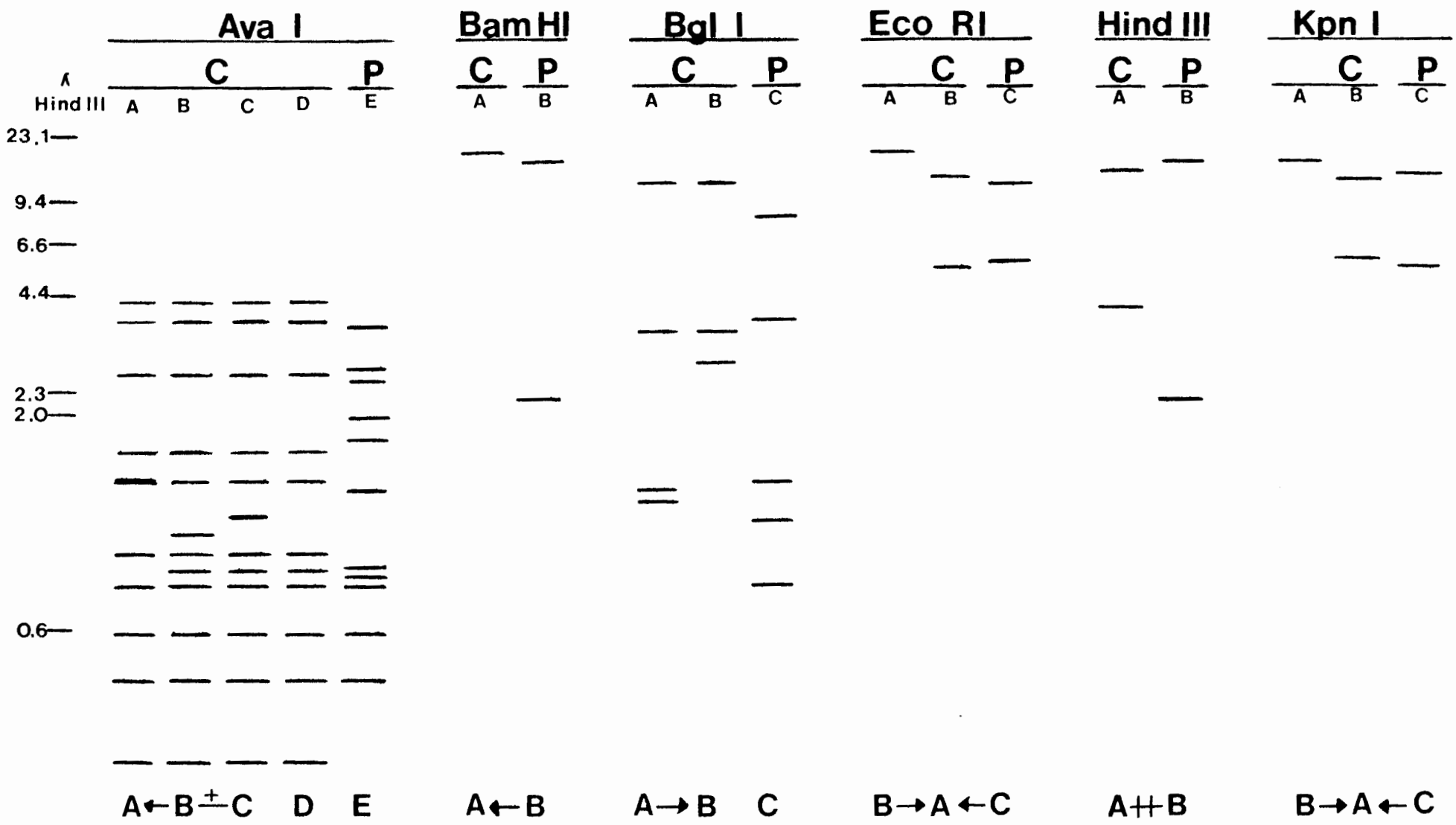
The size of the mtDNA of *M. capensis* was 16,954 ±512 bp, and *M. paradoxus* mtDNA was 16,693 ±405 bp.

The number in parenthesis indicates the species in which the fragment occurs, where (1) is *M. capensis* and (2) is *M. paradoxus*.

Ava I	*Bam*HI	*Bgl* I	*Eco*RI
A:4200 (1)	A:16900 (1)	A:10500 (1)	A:16900 (1,2)
B:3800 (1)	B:15000 (2)	B:9100 (2)	B:11800 (1)
C:3790 (2)	C:2100 (2)	C:4100 (2)	C:11000 (2)
D:2790 (2)		D:3800 (1)	D:5200 (2)
E:2700 (1)		E:2700 (1)	E:5100 (1)
F:2650 (2)		F:1850 (2)	
G:1900 (2)		G:1400 (1)	
H:1800 (2)		H:1200 (1)	
I:1700 (1)		I:1150 (2)	
J:1600 (1)		J:700 (2)	
K:1200 (2)			
L:1180 (1)			
M:950 (1)			
N:900 (1)	*Hin*dIII	*Kpn* I	*Pst* I
O:850 (1)	A:14000 (2)	A:16900 (1,2)	A:16900 (1)
P:800 (2)	B:12700 (1)	B:11150 (2)	B:9000 (2)
Q:750 (2)	C:4200 (1)	C:11100 (1)	C:8000 (2)
R:700 (1,2)	D:2200 (2)	D:5800 (1)	
S:500 (1,2)		E:5750 (2)	
T:400 (1,2)			
U:200 (1)			

Figure 3.3.1

Diagrammatic representation of restriction endonuclease digests of hake mtDNA. C stands for *Merluccius capensis* and P for *Merluccius paradoxus*. Mutational transitions were identified and indicated by an arrow. The direction of the arrow indicated a loss of a restriction site, but does not imply the direction of evolution. A double hash mark indicates that a minimum of two mutational steps connect the genotypes. Unconnected letters differed by 3 or more mutational events. Additions of base pairs are indicated by + and deletions by -.


Ava I
C
P
λ Hind III
A B C D E
23.1
9.4
6.6
4.4
2.3
2.0
0.6
A←B±C D E
BamHI
C P
A B
A←B
BgI I
C P
A B C
A→B C
Eco RI
C P
A B C
B→A←C
Hind III
C P
A B
A++B
Kpn I
C P
A B C
B→A←C

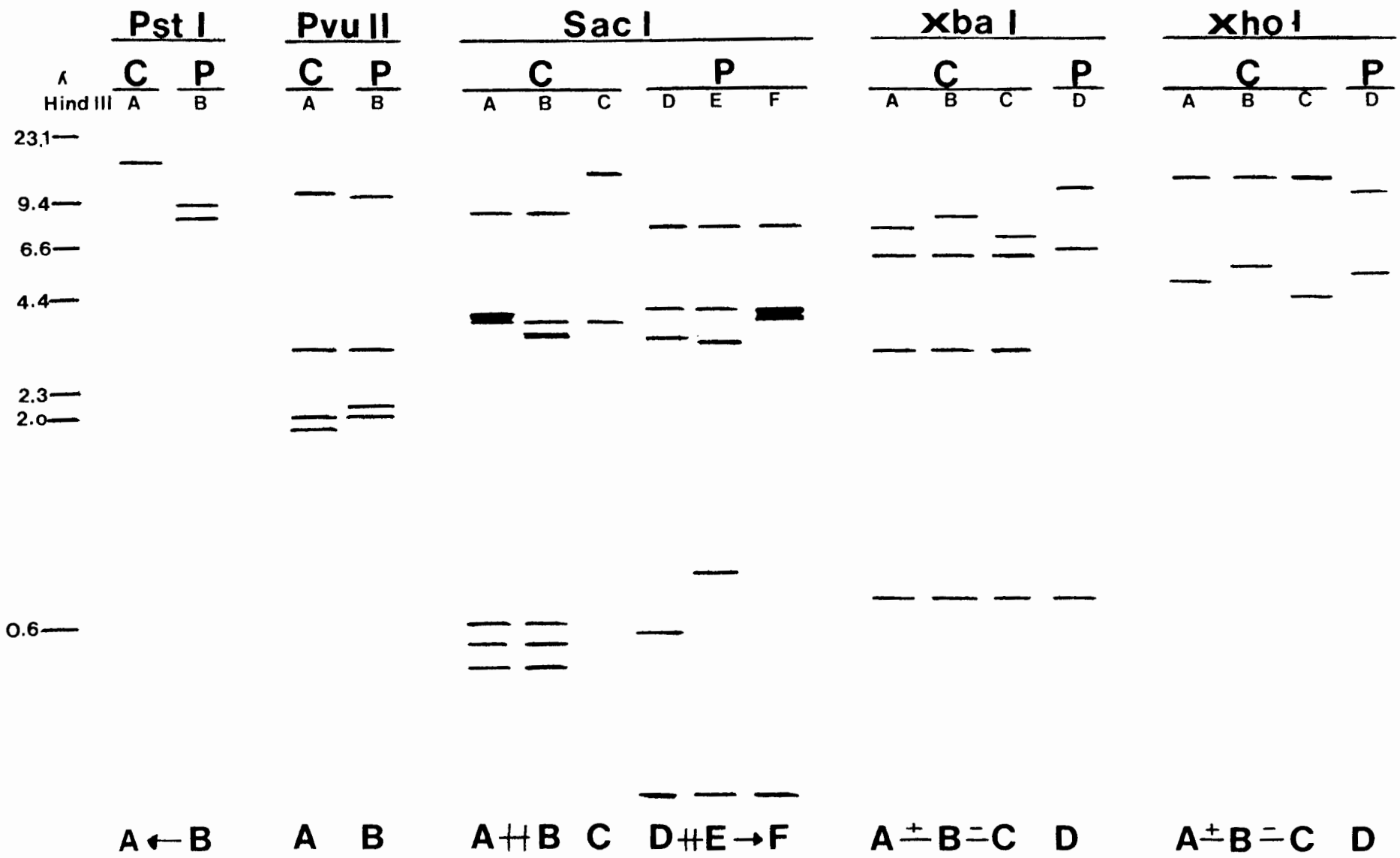

Pst I
Pvu II
Sac I
Xba I
Xho I
λ
Hind III
C
P
A
B
23.1
9.4
6.6
4.4
2.3
2.0
0.6
A←B
A B
A⫽B C D⫽E→F
A±B=C D
A±B=C D

Table 3.3.3

All the fragments detected for each individual from *Merluccius capensis* and *M. paradoxus*, using eleven different restriction enzymes. All the letters typed in **bold** indicate that the fragments were not observed on the gels, but it was assumed that they have the same pattern as the most common genotype. All the letters <u>underlined</u> indicate genotypes that are different to the most common genotype.

Table 3.3.3 (continued)

	Merluccius capensis					
No	*Kpn* I	*Pst* I	*Pvu* II	*Sac* I	*Xba* I	*Xho* I
1	A	A	ACEF	BEFKMN	BFGH	AC
2	A	A	ACEF	BEFKMN	CFGH	AE
3	CD	A	ACEF	BFGKMN	CFGH	AE
4	A	A	ACEF	BFGKMN	BFGH	AC
5	A	A	ACEF	BFGKMN	DFGH	AF
6	A	A	ACEF	BEFKMN	BFGH	AC
7	A	A	ACEF	BEFKMN	CFGH	AE
8	A	A	ACEF	BEFKMN	CFGH	AE
9	A	A	ACEF	BEFKMN	CFGH	AE
10	A	A	ACEF	BEFKMN	CFGH	AE
11	A	A	ACEF	BEFKMN	CFGH	AE
12	CD	A	ACEF	BEFKMN	DFGH	AF
13	A	A	ACEF	BEFKMN	BFGH	AC
14	A	A	ACEF	BEFKMN	CFGH	AE
15	CD	A	ACEF	BFGKMN	BFGH	AC
16	A	A	ACEF	BEFKMN	CFGH	AE
17	CD	A	ACEF	BEFKMN	CFGH	AE
18	A	A	ACEF	BEFKMN	DFGH	AF
19	A	A	ACEF	BEFKMN	CFGH	AE
20	A	A	ACEF	BEFKMN	CFGH	AE
21	A	A	ACEF	BEFKMN	CFGH	AE
22	A	A	ACEF	BEFKMN	CFGH	AE
23	A	A	ACEF	BEFKMN	BFGH	AC
24	A	A	ACEF	BEFKMN	CFGH	AE
25	A	A	ACEF	BEFKMN	CFGH	AE
26	A	A	ACEF	AD	CFGH	AE

Table 3.3.3 (continued)

	Merluccius paradoxus				
No	*Ava* I	*Bam*HI	*Bgl* I	*Eco*RI	*Hin*dIII
27	CDFGHKPQRST	BC	BCFIJ	A	AD
28	CDFGHKPQRST	BC	BCFIJ	CD	AD
29	CDFGHKPQRST	BC	BCFIJ	CD	AD
30	CDFGHKPQRST	BC	BCFIJ	A	AD
31	CDFGHKPQRST	BC	BCFIJ	A	AD
32	CDFGHKPQRST	BC	BCFIJ	A	AD
33	CDFGHKPQRST	BC	BCFIJ	CD	AD
34	CDFGHKPQRST	BC	BCFIJ	A	AD
35	CDFGHKPQRST	BC	BCFIJ	CD	AD
36	CDFGHKPQRST	BC	BCFIJ	CD	AD
37	CDFGHKPQRST	BC	BCFIJ	CD	AD
38	CDFGHKPQRST	BC	BCFIJ	CD	AD
39	CDFGHKPQRST	BC	BCFIJ	A	AD
40	CDFGHKPQRST	BC	BCFIJ	A	AD
41	CDFGHKPQRST	BC	BCFIJ	A	AD
42	CDFGHKPQRST	BC	BCFIJ	A	AD
43	CDFGHKPQRST	BC	BCFIJ	A	AD
44	CDFGHKPQRST	BC	BCFIJ	A	AD
45	CDFGHKPQRST	BC	BCFIJ	A	AD
46	CDFGHKPQRST	BC	BCFIJ	A	AD
47	CDFGHKPQRST	BC	BCFIJ	A	AD
48	CDFGHKPQRST	BC	BCFIJ	A	AD
49	CDFGHKPQRST	BC	BCFIJ	A	AD
50	CDFGHKPQRST	BC	BCFIJ	A	AD

Table 3.3.3 (continued)

	Merluccius paradoxus					
No	*Kpn* I	*Pst* I	*Pvu* II	*Sac* I	*Xba* I	*Xho* I
27	BE	BC	BCDE	CDFO	AEH	BD
28	BE	BC	BCDE	CDHLO	AEH	BD
29	A	BC	BCDE	CDHLO	AEH	BD
30	A	BC	BCDE	CDHLO	AEH	BD
31	A	BC	BCDE	CDHLO	AEH	BD
32	A	BC	BCDE	CDHLO	AEH	BD
33	A	BC	BCDE	CDHLO	AEH	BD
34	A	BC	BCDE	CDHLO	AEH	BD
35	A	BC	BCDE	CDHLO	AEH	BD
36	BE	BC	BCDE	CDHLO	AEH	BD
37	BE	BC	BCDE	CDHLO	AEH	BD
38	A	BC	BCDE	CDHLO	AEH	BD
39	A	BC	BCDE	CDHLO	AEH	BD
40	A	BC	BCDE	CDHLO	AEH	BD
41	A	BC	BCDE	CDHLO	AEH	BD
42	A	BC	BCDE	CDHLO	AEH	BD
43	A	BC	BCDE	CDHLO	AEH	BD
44	A	BC	BCDE	CDHLO	AEH	BD
45	A	BC	BCDE	CDHLO	AEH	BD
46	A	BC	BCDE	CDHLO	AEH	BD
47	A	BC	BCDE	CDFO	AEH	BD
48	A	BC	BCDE	CDFO	AEH	BD
49	A	BC	BCDE	CDHLO	AEH	BD
50	A	BC	BCDE	CDIJO	AEH	BD

3.3.4.1 Digestion profiles

A few mtDNAs from both *M. capensis* and *M. paradoxus* were digested with the following enzymes: *Bcl* I, *Bgl* I, *Cla* I, *Hinf* I, *Hpa* II, *Sac* I, and *Sma* I. These enzymes were not used for routine analysis. The enzymes *Bgl* II, and *Cla* I had no restriction sites on both *M. capensis* and *M. paradoxus* mtDNA, whereas *Sal* I had one restriction site. At least three restriction sites were observed for *Sma* I, but this enzyme was not used for routine analysis because of its high cost.

The two four-base pair restriction enzymes, *Hpa* II and *Hinf* I, cut the mtDNA from both species of hake into fragments smaller than 2.0 kb, making it not feasible to use these enzymes in routine analysis.

The mtDNA from both species of hake was cut by *Bcl* I into a few fragments, but mostly only partial digestions were obtained. Description of the restriction profiles, of the enzymes used routinely in this study, follow. Determination of the mutational events leading from one genotype to another, was also attempted. This, however, does not imply the evolutionary direction of the mutational events.

3.3.4.2 AvaI

This enzyme cut the mtDNA of *M. capensis* into ten fragments. Endlabelling with $\alpha^{32}P$ was used to visualise fragments in the gels, as only the first six bands (up to 900bp) could be seen when normal EtBr staining techniques were used.

Four different fragment genotypes were observed for *M. capensis* when cut with *Ava* I. Genotype "B" (sample 3) had two fragments, of 950 and 850 bp, which did not occur in genotype "A". This could be explained by assuming that the 1,600 bp fragment, present in the most common pattern, represented two bands superimposed upon one another. The 1,600 bp fragment of the most common genotype was a bright, broad band whereas in the variant genotype it was a band of average intensity, indicating that only one of the 1,600 bp bands was present in the "B" genotype. Two additional fragments of 850 bp and the 950 bp in the "B" genotype, could be explained by an addition of a restriction site in one of the 1,600 bp fragments. The size variation observed is within the error range. This indicates that one of the 1,600 bp fragments in the most common restriction pattern has lost one *Ava* I site to give rise to genotype "B".

The other variant genotype ("C") was found in sample 4. This individual had additional 1,180 bp and 850 bp fragments. The fragment of 1,180 bp must have arisen from an additional restriction site in one of the 1,600 bp fragments, creating a 1,180 and a 420 bp fragment. No 420

bp fragment was detected in the variant pattern, but a 850 bp fragment was observed. The "C" genotype could have arisen by an insertion of ±200 bp into the 950 bp fragment of the "B" genotype. This would give rise to the 1,180 and 850 bp fragments. This insertion is confirmed by looking at the restriction pattern of the enzymes *Xba* I and *Xho* I, of the same individual. The insertion is further discussed when describing the composite genotypes and the fragment patterns of the enzymes *Xba* I and *Xho* I.

A third variant genotype, "D", was found in the same area on the mtDNA in samples 10 and 11. These samples had an additional 850 bp band. The most common genotype must have gained one or more restriction sites to give rise to the 850 bp fragment and smaller fragments, which were not detected in the "D" fragment pattern as they will migrate with the other small fragments.

The results show that the area around the 900 bp fragment in *M. capensis* mtDNA is more variable than the rest of the mtDNA. Usually non-conserved areas do not code for any functional proteins or RNA species. Mutations giving rise to new restriction sites take place frequently in these areas. This indicates that the variable region is most likely situated in the D-loop area.

The mtDNA of *M. paradoxus* had eleven *Ava* I restriction sites, giving rise to eleven fragments. No variant genotypes were detected within *M. paradoxus* .

Three fragments, 700, 500, and 400 bp in length, were shared between the two species. Since these three fragments were also present in the variant genotypes found in *M. capensis*, it was assumed that these restriction sites fall into conserved areas.

3.3.4.3 BamHI

The enzyme *Bam*HI cut the mtDNA of *M. capensis* only once. All samples were initially endlabelled after digesting with *Bam*HI, and bands could be seen on X-ray films, showing that the mtDNA must have one restriction site, as only cut DNA can be endlabelled. All the individuals showed only one restriction site.

There were two *Bam*HI restriction sites in the mtDNA of *M. paradoxus*. All individuals screened showed the same digestion pattern (Fig. 3.3.2).

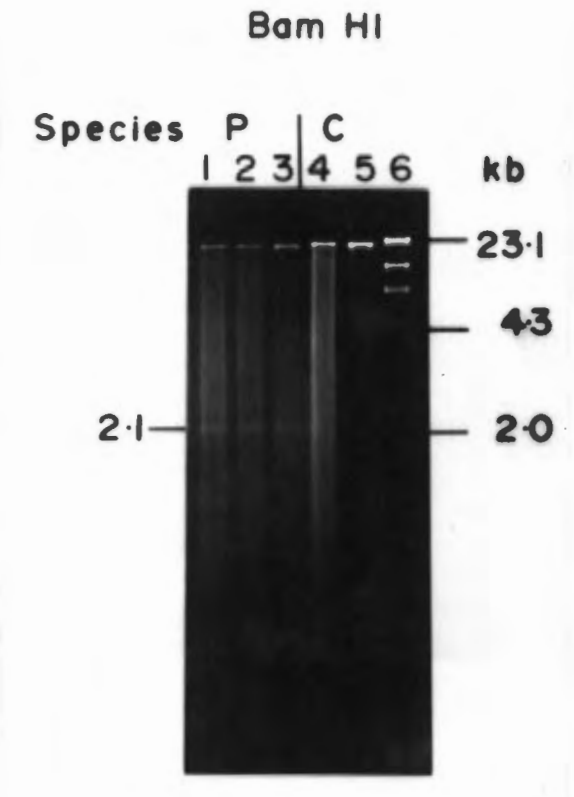
The two species do not share any restriction fragments for *Bam*HI. The transition from the *M. capensis* genotype to the *M. paradoxus* genotype could be explained by a loss of a *Bam*HI restriction site in the *M. capensis* mtDNA.

Table 3.3.3 (continued)

	Merluccius capensis				
No	*Ava* I	*Bam*HI	*Bgl* I	*Eco*RI	*Hin*dIII
1	ABEIJJNRSTU	A	ADGH	A	BC
2	ABEIJJNRSTU	A	ADGH	A	BC
3	ABEIJMNORSTU	A	ADGH	BE	BC
4	ABEIJLNORSTU	A	ADGH	A	BC
5	ABEIJJNRSTU	A	ADGH	BE	BC
6	ABEIJJNRSTU	A	ADGH	BE	BC
7	ABEIJJNRSTU	A	ADGH	BE	BC
8	ABEIJJNRSTU	A	ADGH	BE	BC
9	ABEIJJNRSTU	A	ADGH	A	BC
10	ABEIJNORSTU	A	ADGH	BE	BC
11	ABEIJNORSTU	A	ADGH	BE	BC
12	ABEIJJNRSTU	A	ADGH	BE	BC
13	ABEIJJNRSTU	A	ADE	BE	BC
14	ABEIJJNRSTU	A	ADGH	A	BC
15	ABEIJJNRSTU	A	ADGH	BE	BC
16	ABEIJJNRSTU	A	ADGH	A	BC
17	ABEIJJNRSTU	A	ADGH	A	BC
18	ABEIJJNRSTU	A	ADGH	A	BC
19	ABEIJJNRSTU	A	ADGH	A	BC
20	ABEIJJNRSTU	A	ADGH	BE	BC
21	ABEIJJNRSTU	A	ADGH	BE	BC
22	ABEIJJNRSTU	A	ADGH	BE	BC
23	ABEIJJNRSTU	A	ADGH	BE	BC
24	ABEIJJNRSTU	A	ADGH	A	BC
25	ABEIJJNRSTU	A	ADGH	BE	BC
26	ABEIJJNRSTU	A	ADGH	BE	BC

Figure 3.3.2

Mitochondrial DNA of *Merluccius paradoxus* (P, lanes 1-3) and *Merluccius capensis* (C, lanes 4 and 5) digested with the restriction enzyme *Bam*HI. Lane 6 contains phage lambda DNA digested with *Hin*dIII as a size marker.



3.3.4.4 Bgl I

The mtDNA of *M. capensis* gave rise to four fragments when cut with *Bgl* I. The bands were easily measured and had sizes which could be resolved with little error by using 0,8% agarose gels. Only one individual with a different fragment pattern was found (genotype "B"). Sample 13 had a variant band, which was 2,600 bp in length and the two bands of 1,400- and 1,200 bp were not present (Fig. 3.3.3). This can be explained by a loss of one restriction site between the 1,400 and 1,200 bp fragments, creating the 2,600 bp fragment.

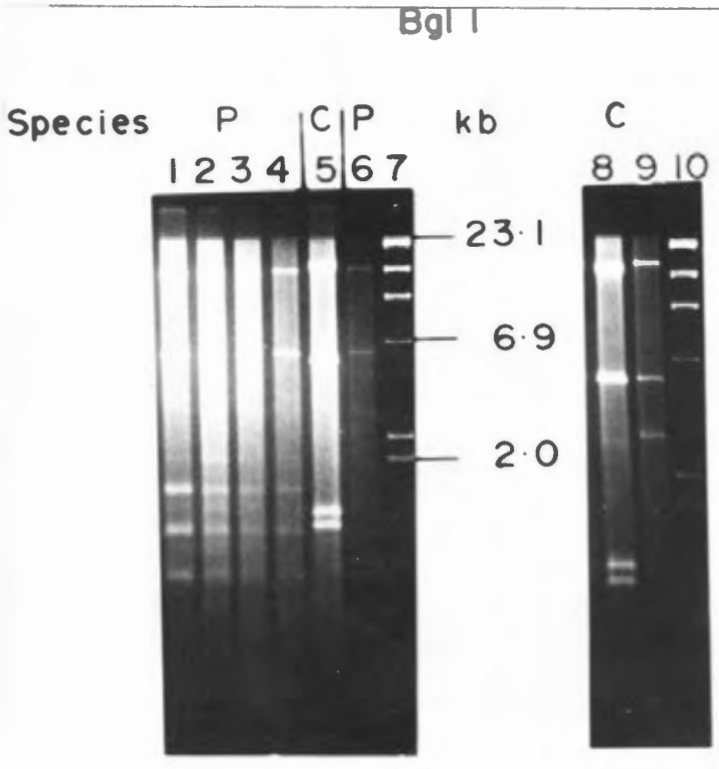
Bgl I cut the mtDNA of *M. paradoxus* into five fragments, and all individuals showed the same restriction pattern. *Merluccius capensis* and *M. paradoxus* do not share any fragments for this enzyme. Three or more mutational events must have taken place, to make it possible for the most common fragment genotype to evolve into the *M. paradoxus* genotype.

3.3.4.5 EcoRI

Some difficulty was encountered when digesting mtDNA of hake with *Eco*RI. The enzyme cut the DNA poorly even after adding an excess of enzyme and incubating overnight.

Figure 3.3.3

Merluccius paradoxus (P, lanes 3-4 and 6) and *Merluccius capensis* (C, lanes 5, 8 and 9) mtDNA digested with *Bgl* I. Phage lambda DNA cut with *Hin*dIII is run in lane 7 and 10. The mtDNA of *Merluccius capensis* digested with *Bgl* I in lane 9, represents the "B" genotype which has lost one *Bgl* I restriction site.



About half of all the *M. capensis* mtDNA molecules were cut into two fragments by this enzyme, while the other half had only one restriction site. All the individuals were digested twice with the enzyme and the same results were obtained each time.

Only seven individuals of *M. paradoxus* had two *Eco*RI sites. All individuals were also digested twice with this enzyme and the results were the same each time.

The two species of hake shared the fragment of genotype "A". An addition of one restriction site must have occurred within each species to give rise to the fragment genotypes "B" and "C". It can be assumed that the *Eco*RI restriction sites are not highly conserved in either species of hake.

3.3.4.6 HindIII

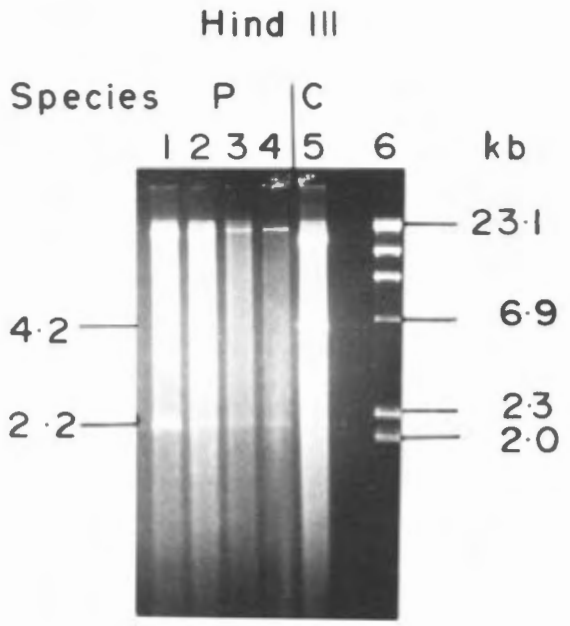
*Hin*dIII cut the mtDNA of *M. capensis* at two sites and all individuals that were digested showed the same restriction fragment pattern.

Mitochondrial DNA from *M. paradoxus* also had two restriction sites. The small fragment was about half the length of the one found in *M. capensis*, and no variant genotype was found in this species (Fig. 3.3.4).

Merluccius capensis and *M. paradoxus* do not share any of the

Figure 3.3.4

Restriction fragment pattern of mtDNA of *Merluccius capensis* (C) and *Merluccius paradoxus* (P) digested with *Hin*dIII. Lambda DNA digested with *Hin*dIII is in lane 6.



*Hin*dIII restriction sites. The transition of the "A" genotype to the "B" genotype would require a loss, followed by a gain, of *Hin*dIII restriction site.

3.3.4.7 Kpn I

Only four samples of *M. capensis* had two *Kpn* I restriction sites and the rest had one restriction site.

Four individuals of *M. paradoxus* had two *Kpn* I restriction sites while the remaining individuals had only one cutting site. This can be explained similarly to *M. capensis*, by the gain of one restriction site.

The genotype "A", which had only one restriction site, was common to both species of hake.

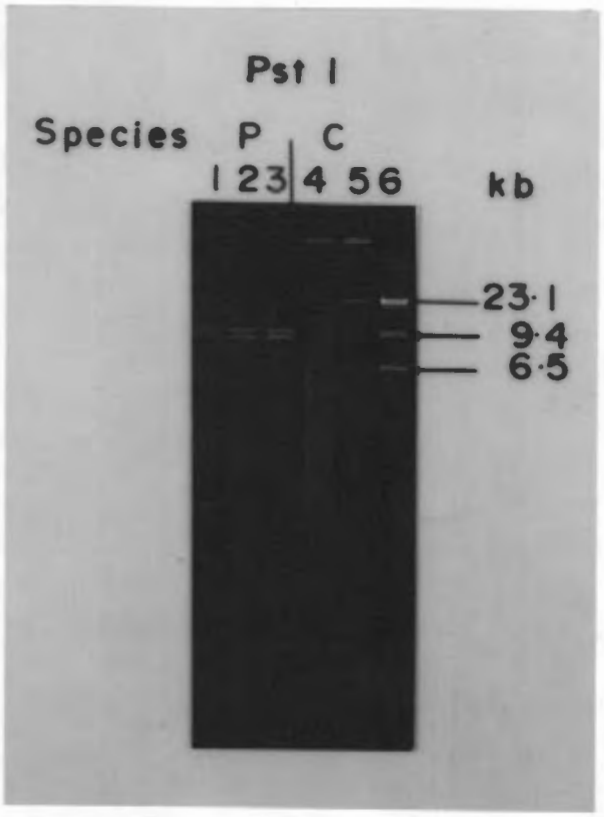
3.3.4.8 Pst I

All the samples of *M. capensis* had only one *Pst* I site on the mtDNA. No additional restriction sites were detected throughout this study.

Merluccius paradoxus had two *Pst* I restriction sites, which cut the mtDNA into two pieces of similar length (Fig. 3.3.5), and no smaller fragments could be detected by normal EtBr staining. The fragment pattern stayed constant throughout the study.

Figure 3.3.5

Mitochondrial DNA of *Merluccius paradoxus* (P) and *Merluccius capensis* (C) digested with *Pst* I. Lambda DNA digested with *Hin*dIII is in lane 6.



The two species had no *Pst* I fragment in common. The *M. paradoxus*, "B", genotype must have lost one *Pst* I restriction site to give rise to the *M. capensis*, "A", genotype.

3.3.4.9 PvuII

The enzyme *Pvu* II has four recognition sites on the mtDNA of *M. capensis* giving rise to one larger fragment of 10.000 kb, one medium sized fragment of 3.000 kb, and two small fragments of 2.000 and 1.900 kb, which form a doublet. No variant genotype was observed in this species.

The restriction pattern for *M. paradoxus* was almost identical to the one found in *M. capensis*. *Pvu* II also has four restriction sites in this species. The fragment sizes were: 9.6, 3.0, 2.1 and 2.0. Only by running both species side by side on the same gel, could the difference in the sizes of the fragments from each species be determined.

Only two of the four fragments are shared between the two species, and a loss of one restriction site followed by an addition of a new site was necessary to convert the "A" genotype into a "B" genotype (Fig. 3.3.1).

3.3.4.10 SacI

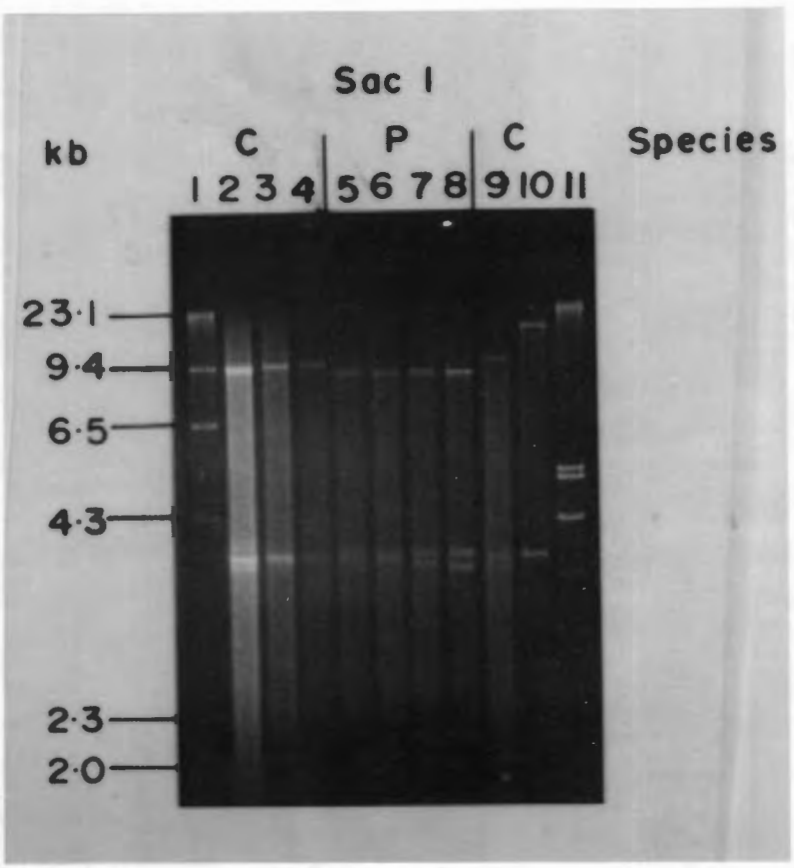
The mtDNA of *M. capensis* had six *Sac* I restriction sites. Most individuals tested had a fragment pattern of two larger fragments of sizes around 10.0 kb, and the other ones, around 4.0 kb. In addition to these fragments, some very small fragments ranging from 600 - 400 bp were detected after overexposing the gels. The "A" or the "B" genotype of *M. capensis* (Fig. 3.3.1) must have lost four restriction sites to create the "C" genotype. The conversion of the "A" genotype to the "B" genotype could be explained by the addition of one restriction site in the 3.78 kb fragment to give rise to the 3.7 kb and 80 bp fragments. The sensitivity of our method was not enough to detect the 80bp fragment.

Sac I cut the *M. paradoxus* mtDNA five times. The "E" genotype had a 3.6 kb fragment replacing the 3.45 kb fragment and the 550 bp piece was replaced by a 700 bp piece. This means that these individuals lost one site and then acquired another site at a different position on the mtDNA (Fig. 3.3.6). In a third genotype, "F", two 3.82 kb fragments were present and no small bands were detected, indicating that a loss of two restriction sites, giving rise to one large band, must have taken place (Fig. 3.3.1).

The *Sac* I sites are not highly conserved in both species, and three different genotypes were found within each species. Only one fragment was shared between the "A" and the "F" genotypes of the two species.

Figure 3.3.6

Merluccius capensis (C) mtDNA digested with *Sac* I is in lanes 2, 3, 4, 9, and 10. The sample run in lane 10 displays the "C" genotype, and the other digests have the most common genotype, "A". *Merluccius paradoxus* (P) mtDNA cut with *Sac* I is in lanes 5-8. Samples in lanes 5 and 6 have the "F" genotype, sample in lane 7 has the "D" genotype and the sample in lane 8 has the "E" genotype. The small mtDNA fragments from both *M. capensis* and *M. paradoxus* are run off the gel and lanes 1 and 11 contain lambda DNA cut with *Hin*dIII and *Hin*dIII plus *Eco*RI, respectively.



3.3.4.11 StuI

Cutting the mtDNA of *M. capensis* with *Stu* I, which has the recognition sequence of AGGCCT, resulted in eight fragments ranging from 4.3 to 0.4 kb. *M. paradoxus* was also cut into eight fragments of similar sizes. All the mtDNAs were digested with *Stu* I and all the digestions were partial. Due to the difficulty in obtaining samples of this enzyme with high activity, repeated digestions were not attempted. From the partial digests, it looked as though the five smaller fragments might be shared between the two species, but nothing conclusive could be established from these gels.

3.3.4.12 XbaI

The mtDNA of *M. capensis* was cut into four fragments, and all the samples were endlabelled, after digesting with *Xba* I. This made the evaluation of the low molecular weight DNA molecules much easier. Three genotypes were detected within *M. capensis*. In the "B" genotype, the top fragment of 7.1 kb was replaced by a bigger fragment of 7.5 kb. The other bands stayed unchanged. This means an insertion into that fragment must have taken place. In another genotype ("C"), the 7.1 kb fragment was replaced by a smaller fragment of 6.9 kb (Fig. 3.3.1), which indicates that a deletion had taken place inside the restriction fragment.

The mtDNA from *M. paradoxus* had three restriction sites, which were observed for all the individuals tested.

Only the 600 bp fragment was common to both *M. capensis* and *M. paradoxus*. The *M. paradoxus* genotype must have lost two *Xba* I restriction sites to give rise to the "A" genotype of *M. capensis*.

3.3.4.13 XhoI

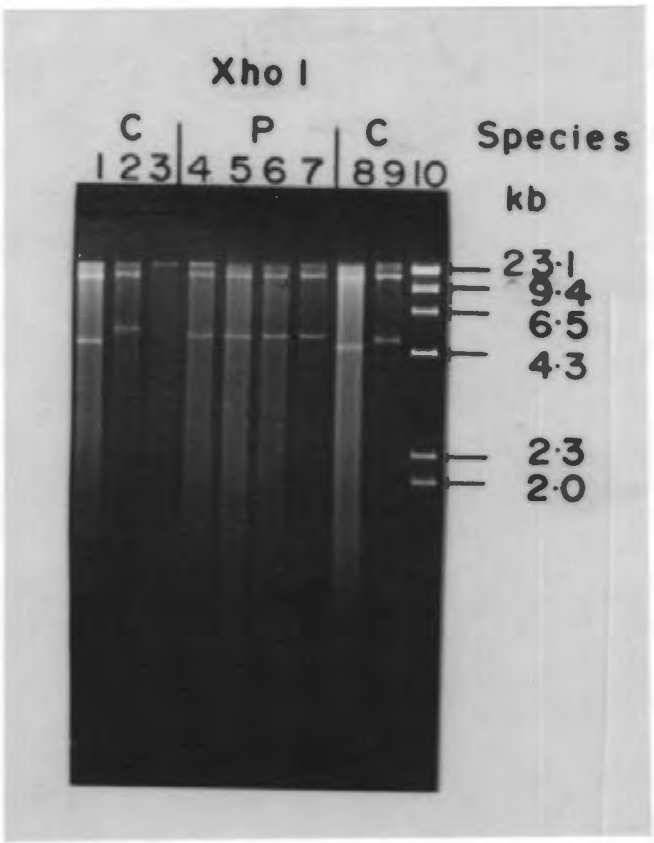
Two *Xho* I sites, which gave rise to fragments of 4.7 and 11.4 kb, were found in *M. capensis*. All of the digests were endlabelled, but the bands could also be seen clearly by EtBr staining Fig. (3.3.7). *Merluccius capensis* displayed three different genotypes when digested with *Xho* I. The one band was observed in all of the genotypes, whereas the other band varied from 4.5 to 5.1 kb. Again the "A" genotype lost a fragment of 4.7 kb and a 5.1 kb fragment appeared, showing that an insertion of 400 bp took place. Similarly, the "C" genotype had a 4.5 kb fragment indicating a deletion of 200 bp.

The mtDNA of *M. paradoxus* also had two *Xho* I sites which always remained the same (Fig 3.3.7).

No fragments were shared between the two species. The second fragment of *M. capensis* migrated at a very similar speed to the one found in *M. paradoxus*, but by running

Figure 3.3.7

Xho I restriction digest of mtDNA of *Merluccius capensis* (C) and *Merluccius paradoxus* (P). The most common fragment genotype of *M. capensis*, "A", is displayed in lane 9, the fragment genotype "C", which has a deletion, is seen lane 8, and the fragment genotype "B" is shown in lane 2. All *M. paradoxus* samples had the same fragment genotype "D". Lambda DNA digested with *Hin*dIII is in lane 10.



samples from each species side by side on the same gel, it was evident that they were of different size.

3.3.5 Data analysis

From these restriction patterns, the number of shared fragments between the two species were calculated. By taking restriction patterns of all eleven enzymes of each individual into account, each individual was allocated a composite genotype. It was found that *M. capensis* had 14 composite genotypes whereas *M. paradoxus* had 6 composite genotypes (Table 3.3.4). One most parsimonious network connecting the *M. capensis* genotypes with a minimum of 22 mutation steps, an insertion, and a deletion was found (Fig. 3.3.8). In this parsimonious network, genotype 6 was also equally connected to genotype 2 by three mutational steps, and genotypes 8 and 11 were connected by two mutational steps. The *M. paradoxus* genotypes were connected with a minimum of 7 mutational steps. In the *M. paradoxus* parsimonious network, genotype 15 and 20 were also connected by one mutational step, and genotype 16 and 18 were connected by two mutational steps. In Fig. 3.3.8, a link between two genotypes indicates a mutation for a single restriction site and a continuous hatch mark indicates a minimum of two mutational events. An insertion is indicated by a link between two genotypes with a plus sign and a deletion is indicated by a minus sign. For example, the

composite genotype AAAAAAAAAAA is related to AAABAAAAAAA by the appearance of an *Eco*RI restriction site, and it is related to the genotype AAAAABAAAAA by the addition of a *Kpn* I site. Similarly, composite genotype 8 (AAABAAAAABB) and 9 (AABBAAAAABB) differ only in the *Bgl* I genotype where the loss of one *Bgl* I restriction site in genotype 8 gave rise to genotype 9. Composite genotypes 2 and 5, also only differ in one *Ava* I restriction site, where at least two mutational steps are required to explain their differences. Insertions and deletions are apparent (a plus or minus sign in Fig. 3.3.8). Genotypes 7 to 11 contain a 400 bp insertion and genotypes 12 to 14 have a 200 bp deletion.

3.3.6 Sequence divergence

The number of shared fragments were also calculated between different composite genotypes, and that value was used to estimate the nucleotide sequence divergence, p. The F and p values estimated between the species, are presented in Table 3.3.5 and 3.3.6, respectively. It was assumed that the *Xba* I and *Xho* I restriction sites were shared between the various genotypes of *M. capensis*, as the change in the fragment size was not due to a change in the restriction site, but due to an insertion or deletion.

For *M. capensis*, the average number of base pairs examined per pairwise comparison was 235.6 (about 1.4% of the mtDNA

Table 3.3.4

Composite genotypes of *M. capensis* and *M. paradoxus*, fish identification numbers, and geographic locations of genotypes. Each letter in the composite genotype presents the restriction enzyme fragment genotypes of *Ava* I, *Bam*HI, *Bgl* I, *Eco*RI, *Hin*dIII, *Kpn* I, *Pst* I, *Pvu* II, *Sac* I, *Xba* I, *Xho* I, respectively.

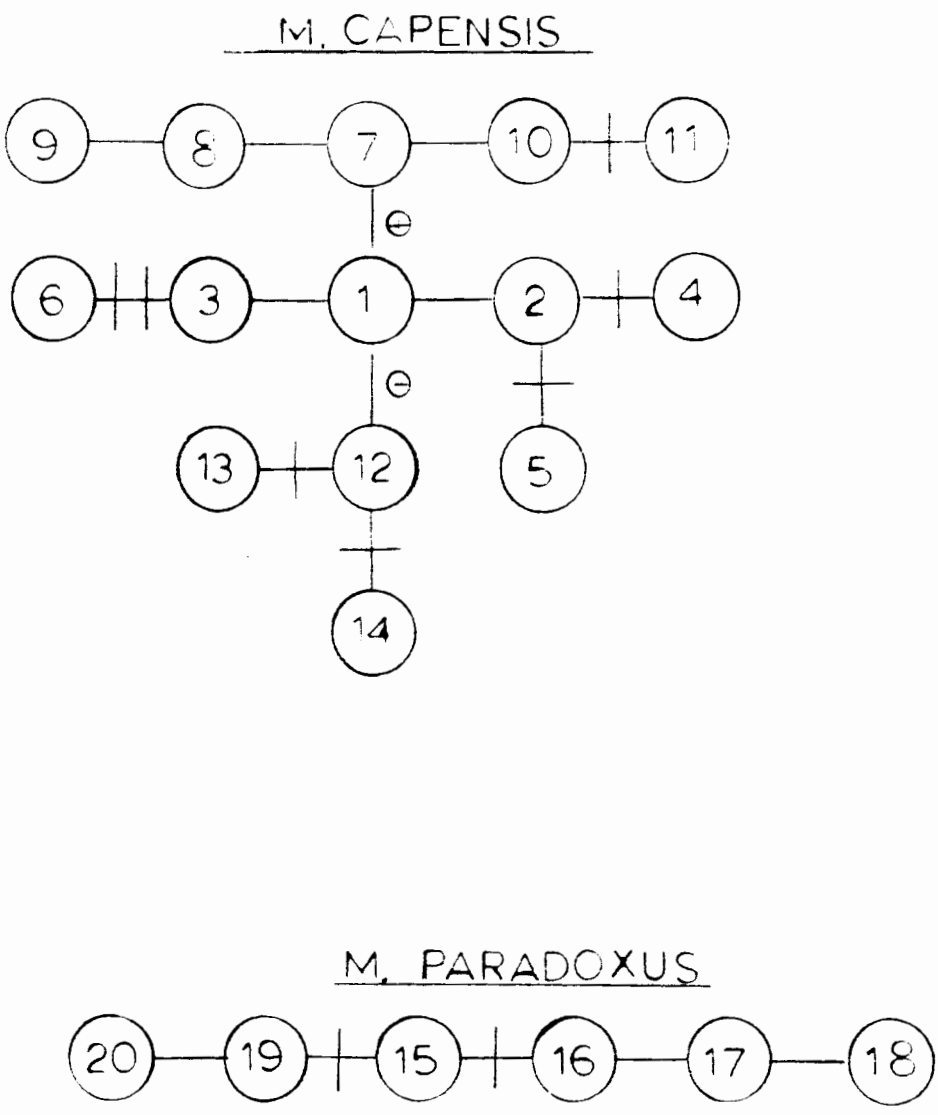
Clonal designation	Composite genotype	Fish number	Location
Merluccius capensis			
1	AAAAAAAAAAA	2,9,14,16,19,24	1,2,5,7,10
2	AAABAAAAAAA	7,8,20,21,22,25	1,2,10
3	AAAAABAAAAA	17	7
4	AAABAAAACAA	26	1
5	DAABAAAAAAA	10,11	2,3
6	BAABABAABAA	3	4
7	AAAAAAAAABB	1	5
8	AAABAAAAABB	6,23	9,10
9	AABBAAAAABB	13	6
10	CAAAAAAABBB	4	4
11	AAABABAABBB	15	7
12	AAAAAAAAACC	18	7
13	AAABABAAACC	12	3
14	AAABAAAABCC	5	8

Table 3.3.4 (continued)

Clonal designation	Composite phenotype	Fish number	Location
Merluccius paradoxus			
15	EBCABABBDDD	30,31,32,34,39,40, 41,42,43,44,45,46, 49	1,2,4,5,7, 8
16	EBCABABBEDD	50	1
17	EBCABABBFDD	47,48	1
18	EBCABCBBFDD	27	3
19	EBCCBCBBDDD	28,36,37	3,7
20	EBCCBABBDDD	29,33,35,38	6,7

Figure 3.3.8

A parsimony network of composite genotypes. Numbers correspond to composite genotypes in Table 3.3.4. Hatch marks indicate two or more mutational steps, a + indicates an insertion of 400 bp and a - indicates a deletion of 200 bp.



genome). Values of p ranged from 0.0000 to 0.0135 and averaged 0.0057. For *M. paradoxus*, the average number of bases per comparison was 224.4 (about 1.3% of the mtDNA). Values of p ranged from 0.0010 to 0.0093 and averaged 0.0055. Values of p between genotypes of the two species of hake varied from 0.0886 to 0.1352 and averaged 0.1160 (table 3.3.6).

3.3.7 Location of genotypes

It was not possible to match specific genotypes to certain areas. Two of the most common genotypes of *M. capensis* were, for instance, found in six individuals each, and the locations of those samples stretched over the entire sampling range (Table 3.3.4). The remaining genotypes appeared on individuals collected at various locations over the sampling area. The most common genotype found in *M. paradoxus* appeared in 13 individuals and the distribution of those samples stretched over most of the sampling range. Five less frequent genotypes occurred at various locations.

Table 3.3.5

The *F* value estimated between genotypes of *M. capensis* and *M. paradoxus*.

Genotype	1	2	3	4	5	6	7	8	9	10	11	12	13	14
15	.189	.148	.148	.160	.148	.109	.187	.148	.151	.187	.109	.187	.109	.148
16	.189	.148	.148	.160	.148	.109	.187	.148	.151	.187	.109	.187	.109	.148
17	.231	.187	.187	.163	.187	.111	.231	.187	.192	.192	.111	.231	.148	.151
18	.189	.148	.185	.123	.148	.109	.187	.148	.150	.187	.109	.187	.185	.148
19	.109	.107	.107	.115	.107	.105	.109	.107	.109	.109	.105	.109	.107	.107
20	.185	.145	.145	.159	.145	.107	.185	.145	.148	.185	.107	.185	.107	.145

Table 3.3.6

The *p* value estimated between the genotypes of *M. capensis* and *M. paradoxus*.

Genotype	1	2	3	4	5	6	7	8	9	10	11	12	13	14
15	.1010	.1149	.1149	.1102	.1149	.1330	.1006	.1149	.1136	.1014	.1328	.1011	.1327	.1149
16	.1010	.1149	.1149	.1102	.1149	.1330	.1006	.1149	.1136	.1014	.1328	.1011	.1327	.1149
17	.0888	.1010	.1011	.1091	.1011	.1320	.0886	.1006	.0995	.1000	.1137	.0887	.1149	.1317
18	.1010	.1149	.1017	.1257	.1149	.1330	.1011	.1149	.1140	.1013	.1328	.1011	.1160	.1149
19	.1327	.1339	.1339	.1296	.1339	.1352	.1327	.1339	.1328	.1330	.1350	.1328	.1350	.1339
20	.1017	.1160	.1160	.1106	.1160	.1341	.1017	.1160	.1149	.1020	.1339	.1017	.1339	.1020

3.4 DISCUSSION

3.4.1 Structure and size of mtDNA

The mtDNA of both species of hake was found to be a circular molecule. The size of the mtDNA of both *M. capensis* and *M. paradoxus* was estimated by summing the fragment sizes, which were determined by running a size standard on the gels (phage lambda DNA). The mean size was calculated by summing the fragment sizes obtained by digesting the mtDNA with 11 different restriction enzymes. *Merluccius capensis* mtDNA was found to be 16,954 $\pm$512 bp and *M. paradoxus* mtDNA was 16,693 $\pm$403 bp. Table 3.4.1 represents the sizes of mtDNA reported for other organisms. In general, the size of the mtDNA of higher eucaryotes ranges from 15,700 to 19,500 bp, whereas lower eucaryotes (yeast, fungus) and plants have much larger mtDNAs, which contain large non-coding regions.

The mtDNA of marine fishes seems to be slightly larger than the mtDNA of mammals. Berg and Ferris (1984) reported that the mtDNA of salmonids was 16,670 bp and Graves *et al.* (1984) showed that the skipjack tuna mtDNA is 16,900 bp in size. Mammalian mtDNA, on the other hand, ranges from 16,200 to 16,700 bp, and amphibian mtDNA seems to be larger, with a range of 17,000 - 23,000 bp. The size of the mtDNA of *M. capensis* and *M. paradoxus* estimated in this study agrees with the sizes reported for other marine fish, in that it is also about 400 bp larger than mammalian mtDNA.

Table 3.4.1

Size of the mitochondrial DNA reported for various organisms

Organism	Size of mtDNA bp	Reference
Mammals		
Human	16,569	Anderson *et al.*, 1981
Mouse	16,295	Bibb *et al.*, 1981
Bovine (cow)	16,338	Anderson *et al.*, 1982
Rodents		
Mus musculus	16,000 ±380	Yonekawa *et al.*, 1981
Deer mouse (*Peromyscus*)	15,600 ±2,000	Avise *et al.*, 1979b
Cotton rat *Sigmodon hispidus*	16,500 ±200	Kessler and Avise, 1985a
Birds		
Parus	16,700 ±800	Mack *et al.*, 1986
Waterfowl (*Anas*)	16,500	Kessler and Avise, 1985b
Warblers, sparrows	17,300	Kessler and Avise, 1985b
Amphibians		
Xenopus laevis	17,550	Roe *et al.*, 1985
Tree frog (*Hyla*)	23,000	Kessler and Avise, 1985b

Table 3.4.1 (continued)

Organism	Size of mtDNA bp	Reference
Marine fishes		
M. capensis	16,954 ±512	This study
M. paradoxus	16,693 ±405	This study
Skipjack tuna	16,900	Graves *et al.*, 1984
Clupea harengus	16,990 ±620	Kornfield and Bogdanowicz, 1987
Freshwater fishes		
Salmoid	16.670	Berg and Ferris, 1984
Salmo clarki lewisi	16,850	Gyllensten *et al.*, 1985
S. c. bouvieri	16,450	Gyllensten *et al.*, 1985
Fundulus heteroclitus	17,200 ±300	Gonzalez-Villasenor *et al.*, 1986
Insects		
Drosophila yakuba	16,019	Clary and Wolstenholme, 1985
Lower eucaryotes		
Yeast	±85,000	de Zamaroczy and Bernardi, 1986

3.4.2 GC content of mtDNA

The GC content of the mtDNA was estimated for both hake by sedimentation equilibrium analysis. It was found that *M. capensis* had a GC content of 47% ±2% and *M. paradoxus* had a GC content of 45.5% ±2%. Birds tend to have the greatest GC content (46-50%), followed by mammals (37-44%), and lower vertebrates (38-41%) (Brown, 1983). Freshwater trout, *Salmo gardeneri*, had a content of 41%. The estimates of the GC content of the hake are somewhat unusual in that they are higher than expected. Brown (1983) postulated that the GC content will influence the secondary structure of the mtDNA and in that way it may influence the efficiency of RNA processing. In organisms with large energy demands, where a high degree of mitochondrial efficiency is required, there may be a selective pressure for a strong GC bias. It is not clear how this mechanism can explain the difference in GC content between hake and trout. One explanation could be that the saline environment of the hake imposes greater metabolic demands on hake than does a low saline freshwater environment on trout.

3.4.3 Methylation of the mtDNA

Mammalian DNA containing 5-methylcytosine residues is not cleaved with the following restriction enzymes: *Hha* I, *Hpa* II,

Msp I, *Sal* I,and *Xho* I (Maniatis, 1982). The mtDNA from *M. capensis* and *M. paradoxus* was digested with *Sal* I, *Xho* I and *Hpa* II, and these enzymes cut the mtDNA from both species into 1, 2 and ±15 fragments, respectively. No variation was observed in the restriction pattern of *Sal* I, and *Hpa* II and it was concluded that the mtDNA of hake is not methylated. Castora *et al.* (1980) tested the assumption that variant restriction sites vary with respect to the sequence and not methylation. They cloned mtDNA fragments and the same fragment pattern was observed for the cloned fragments as for the original fragment, and the nucleotide sequence of the fragments revealed that the difference in restriction pattern was due to nucleotide replacements and not methylation of the mtDNA. The variation of the fragment pattern observed for *Xho* I, is also not due to methylation, but rather due to deletion or insertions into the mtDNA of *M. capensis*.

3.4.4 Size variation and heteroplasmy

Generally, it is assumed that nucleotide substitutions are responsible for variation in the restriction patterns of mtDNA. In this study, a length polymorphism was observed in the mtDNA of *M. capensis*. When the mtDNA was digested with the enzyme *Xba* I all but one, of the fragment sizes stayed constant. Upholt (1977) postulated that if an insertion into a fragment takes place, it will result in the loss of

that fragment and the appearance of a fragment larger than the one lost. It was found that the varying fragment was replaced by a larger fragment, indicating that a ±400 bp insertion took place. In other individuals the same fragment was replaced by a smaller fragment indicating that a ±200 bp deletion took place in that fragment. Similar results were obtained when cutting these mtDNAs with the enzyme *Xho* I. By looking at the composite genotypes, it was observed that if, in one mtDNA, an insertion was detected in the *Xho* I fragment pattern, an insertion was also detected in the *Xba* I fragment pattern. When cutting the mtDNA of *M. capensis* with *Ava* I, one individual (4) also showed an insertion. The conversion of genotype "A" or "B" to genotype "C" could only be explained by an insertion. The 950 bp fragment in genotype "B" is replaced by a 1,180 bp fragment. This indicates that an insertion of ±200 bp must have taken place. This insertion was confirmed when digesting this mtDNA (4) with the enzymes *Xba* I and *Xho* I.

The insertions and deletions observed in the restriction pattern of *Xba* I and *Xho* I were, however, not detected when cutting the mtDNAs with the other restriction enzymes used in this study. One explanation for this is that the insertions must have been into the large fragments, and the sensitivity of the analysis was not great enough to detect them.

Length-polymorphism in the mtDNA has been reported previously between species and also within species. Intraspecies length variation were discovered in humans (Cann and Wilson, 1983; Greenberg *et al.*, 1983), mice (Ferris *et al.*, 1983b), lizards (Moritz and Brown, 1986), tree frogs, *Hyla*, (Kessler and Avise, 1985b, Bermingham *et al.*, 1986b), fruit flies, *Drosophila*, (Goddard *et al.*, 1982, Fauron and Wolsten-Holme, 1976), and crickets (Harrison *et al.*, 1985). Bermingham *et al.* (1986b) noted that size polymorphisms tended to be more prevalent in lower vertebrates such as fish and reptiles than in mammals and birds. This shows that it is not unusual that insertions and deletions were observed for *M. capensis*.

Most of the additions or deletions occur in the D-loop region of the mitochondrial genome. In this part of the study it could not be determined whether the insertions and deletion observed, were in fact in the D-loop region of the hake mtDNA. No size variation could be detected within *M. paradoxus* mtDNA.

Size variation within one individual (or heteroplasmy) have also been reported for some species of lower vertebrates and a freshwater fish (Bermingham, 1986), *D. mauritiana* (Solignac *et al.*, 1983), and some crickets (Harrison *et al.*, 1985). In this part of the study no evidence was found to suggest that length variation does exist within individuals from either species of hake.

Another example where two types of mtDNA are present in one individual would be two mtDNAs with different restriction patterns for one restriction enzyme. In this part of the study we experienced some difficulties when digesting mtDNAs with *Eco*RI. One possible explanation for this is that the individuals were heteroplasmic. It was found that the smaller *Eco*RI fragment of both species of hake could only be visualised with some difficulty. If the one type of mtDNA has only one *Eco*RI restriction site, whereas the other one has two restriction sites, and both types are present in one individual, then one would expect to find a heavy top band and a much weaker small fragment. The heavy top band represents mtDNA linearized with *Eco*RI and mtDNA cut twice with *Eco*RI, as fragments above 10kb are not efficiently separated on 0.8% agarose gels. This indicates that the mtDNA from the two hake might be heteroplasmic for the *Eco*RI restriction sites. No evidence of heteroplasmy was, however, observed when digesting the mtDNA from both species of hake with other restriction enzymes.

3.4.5 Intraspecific variation

The levels of nucleotide site polymorphism differed between the two species of hake. This was reflected in the measures of polymorphism, which were significantly different between species at probability level of $p < 0.01$. The different level of nucleotide site polymorphism was also reflected in

the number of composite genotypes within each species and the difference in the complexity of the parsimonious trees connecting the composite genotypes with each other. *Merluccius capensis* had 14 composite genotypes, which were connected by a total of at least 22 mutational steps, whereas *M. paradoxus* had 6 composite genotypes connected by a minimum of 12 mutational steps. It is unlikely that the different levels of nucleotide site polymorphism are due to sample errors because the same number of fish were examined for each species of hake. It is most likely that both species have the same rate of mutation for mtDNA, and one possible explanation for the difference in polymorphism is that the two species experienced different population events. The first possibility is that the two species have different population sizes or different levels of gene flow between populations, which produce different levels of within-species polymorphism. In a previous study by Grant *et al.* (1988b) it was however determined that both species have large subpopulations and little genetic variation was detected between locations. The other possibility is that *M. paradoxus* experienced an ancient bottleneck in population size, whereas *M. capensis* did not. The variability of the nuclear DNA is less effected by bottlenecks than mtDNA, because of the fact that mtDNA is inherited maternally, and so the effective population size of mtDNA is immediately reduced by half (Ferris and Berg, 1987). By comparing mtDNA- and nuclear DNA heterogeneity, historical bottlenecks

can be detected (Wilson *et al.*, 1985; Ferris and Berg, 1987). In the first part of this study, it was shown that the genetic variability of the nuclear genes was high for both species of hake (Grant *et al.*, 1988b). Wilson *et al.* (1985) show that a low genetic variability of the mtDNA coupled with a high genetic variability of the nuclear genes, indicate that *M. paradoxus* may have experienced an ancient bottleneck in population size.

The amount of sequence divergence between the different mtDNA genotypes, found within each species, can also be used to estimate the relative ages of the species (Ferris and Berg, 1987). One assumption made, is that the species with many divergent mtDNA genotypes, is older than the species which has fewer more closely related mtDNA genotypes (Ferris and Berg, 1987). This would lead to the suggestion that *M. capensis* must be the older of the two species, because of the many different mtDNA genotypes found in that species.

The amount of intraspecific sequence divergence reported for other organisms, is summarised in Table 3.4.2. Generally, the sequence divergence found between mtDNA genotypes averages at 1%, but values as high as 8.7% have been reported between populations of freshwater fish (Bermingham and Avise, 1986). There appears to be little or no geographic ordering of mtDNA genotypes in marine fishes, most likely because of the general lack of barriers to migration in the sea. Although freshwater fishes and

Table 3.4.2

Mitochondrial DNA sequence divergence between intraspecific genotypes of various species estimated by the proportion of shared restriction fragments.

Species	% sequence divergence	Number of enzymes	Reference
Primates			
Homo sapien	2.3*	12	Cann *et al.*,1987
Rodents			
Geomys pinetis			
within regions	0.3	6	Avise *et al.*, 1979a
between regions	3.4	6	Avise *et al.*, 1979a
Peromyscus popionotus	1.1*	8	Avise *et al.*, 1979b
P. maniculatus	0.4*	8	Lansman *et al.*, 1983b
Rattus norvegicus	0.7(0.4-1.8)*	7	Brown and Simpson, 1981
R. rattus	4,2(0.4-9.6)*	7	Brown and Simpson, 1981
Birds			
Parus atricapillus	0.4	14	Mack *et al.*, 1986

* Estimated by restriction site maps.

Table 3.4.2 (continued)

Species	% sequence divergence	Number of enzymes	Reference
Marine fishes			
M. capensis	0.57(0.0-1.3)	11	This study
M. paradoxus	0.55(0.1-1.0)	11	This study
Katswonus pelamis	1.0	9	Graves *et al.*, 1984
Clupea harengus	1.7(0.2-4.4)	7	Kornfield and Bogdanowicz, in press
Freshwater fishes			
Lepomis cyanullus	1.4	13	Avise and Saunders, 1984
L. machrochirus	0.7	13	Avise and Saunders, 1984
L. puntatus	6.2	17	Bermingham and Avise, 1986
L. microlophus	8.7	17	Bermingham and Avise, 1986
Marine invertebrates			
Limulus polyphemus	2.0	12	Saunders *et al.*, 1986
Mytilus edulis	0.7	7	Skibinski, 1985
M. galloprovincialis	1.3	7	Skibinski, 1985

rodents also show a similar degree of differentiation among mtDNA genotypes, these genotypes usually correspond to geographic populations because physical barriers to migration between lakes or between areas prevent the homogenizing effects of gene flow (Bermingham and Avise, 1986).

No geographic variation was detected for the mitochondrial genotypes from both species of hake. No evidence of genetically discrete stocks within South African waters were detected using protein electrophoresis (Grant *et al.*, 1988a). Northern Namibian stocks of *M. capensis* did, however, show some genetic variation from the South African stocks, but no differences were detected for *M. paradoxus*. In this part of the study, no samples were collected from Namibian waters, therefore it was not possible to determine if the mtDNA genotypes in the Namibian waters are different to the one found in southern African waters.

3.4.6 Interspecific variation

Sequence divergence was estimated between *M. capensis* and *M. paradoxus* by the proportion of restriction fragments shared between the two species. It was found that the estimated sequence divergence was 11.6%. Sequence divergence reported between other related species is summarized in Table 3.4.3. By comparing the value of sequence divergence obtained for

different organisms, it can be determined if two species are closely related or not. Generally, the amount of sequence divergence estimated between conspecific mammals and fish are only a few percent. Lansman *et al.* (1981) reported that conspecific organisms exhibit sequence divergence of less than 3% and organisms belonging to closely related species show *p* values of 8 to 25%. The results reported in table 3.4.3 show that closely related sibling species have p values of about 5% (e.g. *Drosophila*; Solignac *et al.*, 1986), and more distantly related species have *p* values of 10% or greater (e.g. *Lepomis*; Avise and Saunders, 1984). The sequence divergence of 11.6%, estimated between the two hake species, appears to be typical of well differentiated congeneric species.

Ferris and Berg (1987) postulate that if a large population of fish is split into half by a geological event, which will prevent gene flow between the two resulting populations, both daughter populations will have initially the same mtDNA genotypes. The two populations will evolve separately, both acquiring new mtDNA genotypes, but some of the original mtDNA will be retained within each population. Therefore sibling species are expected to have some mtDNA genotypes in common. In this study, no common genotype was found between the two species of hake. From these results it was concluded that *M. capensis* and *M. paradoxus* are not sibling species.

Table 3.4.3

Mean (range) mtDNA sequence divergence between congeneric species estimated by the proportion of shared restriction fragments

Genus	Number of species	Sequence divergence (%)	Number of enzymes	Reference
Primates				
Pan	2	3.7*	15	Ferris *et al.*, 1981a
Rodents				
Mus	2	5.0	3	Ferris *et al.*, 1983b
Peromyscus	2	14.1	6	Lansman *et al.*, 1983b
Rattus	2	16.2 (13.5-18.4)*	7	Brown and Simpson, 1981
Birds				
Parus	3	7.0 (4.0-9.0)	14	Mack *et al.*, 1986
Marine fishes				
Merluccius	2	11.6 (8.9-13.52)	11	This study

Table 3.4.3 (continued)

Genus	Number of species	Sequence divergence (%)	Number of enzymes	Reference
Freshwater fishes				
Lepomis	9	20.1 (7.0-25)	13	Avise and Saunders, 1984
Salmo	4	9.3 (4.1-13.7)	13	Gyllensten and Wilson, 1987
Amphibians				
Rana	2	8.1	19	Spolsky and Uzzell, 1984
Insects				
Drosophila	10	5.7 (1.5-9.7)	9	Solignac *et al.*, 1986

*Estimated from restriction site maps

Sequence divergence of the mtDNA is also a measure of the time since two organisms diverged. Brown *et al.* (1979) showed that the mtDNA evolved 5-10 fold more rapidly than nuclear DNA and by plotting sequence divergence against divergence time, the evolutionary rate of the mtDNA was estimated to be 2% per million years. The times of divergence were estimated from fossil and protein data. Brown *et al.* (1979) showed that the *p* value was most accurate for mtDNA comparisons between species that diverged in the last 5 million years. Species that separated more than 10 million years ago accumulated differences at a slower rate, which produced an curvilinear relationship between mtDNA sequence divergence and time. Values of sequence divergence have to fall into the linear part of the graph, to give a meaningful estimate of the time of divergence (*p* values less than 0.15). Estimates of the evolutionary rate of mtDNA was based on the data obtained from higher primates. To date no time calibrations have been made for lower vertebrates or for fishes in particular, but time estimates using Brown's calibration curve appear to be at least approximately correct for freshwater fishes (Bermingham and Avise, 1986). The time of divergence between *M. capensis* and *M. paradoxus* was 5.8 million years ago as determined from the calibration curve.

3.4.7 Comparison of restriction fragment analysis and Protein electrophoresis

Restriction endonuclease analysis of mtDNA may be a more sensitive measure of nucleotide substitutions than protein electrophoresis, because of the degeneracy of the genetic code. A change in a nucleotide does not necessarily mean a change in an amino acid. Often an amino acid is replaced by a different amino acid with a side group having the same charge and polarity as the first amino acid. The resulting protein will migrate at the same speed in an electric field as the initial protein and hence the nucleotide substitution would go undetected. In endonuclease analysis of the mtDNA any nucleotide substitution in a endonuclease recognition site, will be detected. Ferris and Berg (1987) review studies which compare the two methods, and they conclude that greater resolution of genetic difference can be obtained by restriction analysis of mtDNA. One of the reasons being the higher rate of evolution of the mtDNA compared to the nuclear DNA. In this study, the time of divergence of *M. capensis* and *M. paradoxus*, was estimated at 7.6 to 13.6 million years by protein electrophoresis, compared with 5.8 million years estimated by mtDNA restriction analysis. This difference could be due to the fact that the calibration curve used to estimate mtDNA sequence divergence was based on primate data, and that this is not appropriate for lower vertebrates. The result most

likely also reflect the higher resolution power of mtDNA restriction analysis.

To reconstruct evolutionary events, it is often necessary to compare the levels of genetic variability of nuclear DNA to mtDNA. In this study it was detected that *M. paradoxus* had a lower level of nucleotide site polymorphism than *M. capensis*, but an average genetic variability of the nuclear DNA was detected for both species of hake. It was concluded that *M. paradoxus* must have experienced an ancient bottleneck in population size. This demonstrates that restriction endonuclease analysis of mtDNA can be used, alone or in conjunction with protein electrophoresis, to reconstruct evolutionary events of species, as mtDNA is often more sensitive than nuclear DNA to evolutionary events such as bottlenecks.

The results presented in this chapter also show that restriction endonuclease analysis is a fairly quick and reasonably expensive method, which can be readily used in routine analysis.

CHAPTER 4

CLONING AND CHARACTERISATION OF THE MITOCHONDRIAL DNA OF *MERLUCCIUS CAPENSIS* AND *M. PARADOXUS* AND MEASUREMENT OF GENETIC DIVERGENCE BETWEEN THESE SPECIES AND *M. AUSTRALIS*, FROM NEW ZEALAND.

4.1 INTRODUCTION

4.1.1 Mapping of the Mitochondrial DNA

Previously, immunological crossreactivity and protein electophoresis were used for the analysis of phylogenetic relationships of animals (Upholt, 1977). These methods did not take into account changes in DNA sequences that occur in non-coding or control areas. Several approaches are at present in use to determine mtDNA sequence divergence. In the previous chapter, the use of restriction fragment analysis of the mtDNA as a means to determine relatedness between and within species was discussed. By physical mapping of the restriction sites in the mitochondrial genome, a more precise determination of the number of base substitutions separating genotypes may be obtained. When the mtDNA is mapped, the position of each restriction site is determined relative to the restriction sites adjacent to it and eventually all restriction sites can be positioned on the mtDNA. Several strategies (double digests, partial digestion) are currently used for mapping (Parker *et al.*, 1977; Maniatis *et al.*, 1982). The advantage of mapping the mtDNA is that it is possible to determine exactly how many sites must be altered to change one restriction pattern to a different one. Restriction fragment analysis does not determine whether a specific site is shared by two genotypes or not, since two fragments of the same length may be

produced by the same enzyme, cutting different mtDNAs at different positions. By using mapping of the mtDNA the conservation of restriction sites, rather than restriction fragments is determined (Lansman *et al.*, 1981). Yet, often it is not practical to analyze changes of specific sites, especially if the restriction pattern is complicated or when extensive changes have occurred (Upholt, 1977a).

Brown *et al.* (1979) discovered, with the aid of restriction site mapping of the mtDNA, that the rate of evolution of the mtDNA is higher than nuclear DNA. In a similar study by Ferris *et al.* (1981a), human mtDNA was compared to mtDNA from five species of ape by mapping 50 restriction sites. The restriction maps were aligned with respect to 11 invariant positions and a genealogical tree was constructed. These restriction maps revealed that rearrangements took place during evolution of mtDNA of higher primates. Their results also showed that mapping of the mtDNA provided 6 fold more resolution than conventional protein electrophoresis. Lansman *et al.* (1981), however, compared the information obtained from fragment and site analysis of sequence variation within the species *Peromyscus maniculatus*, and their results showed that in analysing population structure, restriction site mapping was not a great improvement over fragment analysis.

Restriction site mapping also proved to be a valuable tool in reconstructing evolutionary histories of certain species.

In one such study, Johnson *et al.* (1983) mapped the mtDNA from five human populations using five restriction enzymes and they were able to demonstrate that different human mtDNA types are related by a series of mutations. By reconstructing the chronology of these mutations, an insight into the biological history of man was obtained.

Mapping studies dictate that it is essential to have an adequate supply of DNA. This can be obtained by cloning the mtDNA which makes the amplification of the DNA of one specific individual possible, and unlimited amounts of mtDNA can be produced as required. At the same time, this overcomes cumbersome mtDNA extractions, which yield only limited amounts of DNA.

4.1.2 Cloning of mtDNA

When cloning DNA, a suitable cloning vector and bacterial host strain, which allows replication and expression of the cloned genes, have to be selected. Very often plasmids, which are extrachromosomal circular genetic elements, are chosen as cloning vectors.

A suitable cloning vector must possess certain properties: (1) It should be relatively small and allow the insertion of large pieces of foreign DNA, (2) it should have copy numbers of 10-200, and (3) should carry one or more selectable markers (e.g. antibiotic resistance, or production of

restriction enzymes) for the identification of transformed cells, and to maintain the plasmid in the bacterial population. (4) The plasmid should contain some unique restriction sites that are located in regions not essential for replication and preferably they should be located in the areas coding for selectable markers (Maniatis *et al.*, 1982). By insertion of foreign DNA into these restriction sites the gene coding for the selectable marker will be inactivated.

Once the vector has been chosen, the mtDNA, or parts thereof, can be inserted into the vector, and by transferring this recombinant plasmid into a suitable bacterial host, a sufficient amount of mtDNA is made available for mapping.

In addition to mapping, cloned mtDNA may be used as a probe in various evolutionary studies. As an example, cloned mtDNA can be used as a hybridisation probe to differentiate between two species and it can be used to determine the phylogenetic relationship among species. A major advantage of using cloned DNA as a probe, instead of highly purified mtDNA, is that it eliminates the possibility of artifacts resulting from hybridisation to highly repeated nuclear DNA sequences, as no nuclear DNA contamination is present in the probe (Lansman *et al.*, 1981). Lansman *et al.* (1981) also describe the isolation and digestion of total cytoplasmic DNA from samples and the use of cloned mtDNA as a $\alpha^{32}P$ radioactively labeled probe to detect the mtDNA restriction

patterns. This technique made it possible to detect as little as 10 picograms of DNA, reducing the amount of mtDNA required by a factor of 1000, compared to the amount needed for direct visualisation on gels by EtBr staining. Hybridisation with a probe also eliminates the necessity of purifying the DNA, which is essential for direct labeling procedures, and makes it possible to use frozen or preserved tissue (Lansman *et al.*, 1981).

4.1.3 Objective of this study

Restriction fragment analysis can provide evidence of rearrangements, but no conclusive statement can be made about their nature (e. g. inversions, transitions). Cleavage map comparisons can, however, detect the nature of DNA rearrangements. Previously in this study, evidence that an insertion and a deletion had taken place in the mtDNA of *M. capensis*, was discovered through restriction fragment analysis of the mtDNA. This finding resulted in the mapping of five restriction enzyme sites with easily interpretable fragment patterns, on the mtDNA of both *M. capensis* and *M. paradoxus*.

At the same time, the additional information about the restriction sites shared between the two species of hake was used to evaluate the results previously obtained with restriction fragment analysis.

In this study cloned mtDNA was used to confirm restriction maps prepared from the mtDNA, because large quantities of relatively pure mtDNA was available. The clones were also used as probes to estimate sequence divergence between various species of hake. This phylogenetic study was started using the hake species, *M. australis*, which occurs along the coasts of New Zealand. This hake is the only species that does not inhabit continental shelf waters and it is reported that *M. australis* is closely related to the other two deep-water hake species, *M. paradoxus* and *M. polli* (which occur along the coast of Angola) (Inada, 1981). The mtDNA from *M. australis* was digested with the same battery of enzymes used in the fragment analysis of the two southern African hake. The cloned mtDNA from both species of hake was then used as a probe to determine the sequence divergence between these two species and the New Zealand hake. These probes also made it possible to map the restriction sites on the mtDNA of New Zealand hake. Cloning parts of the mtDNA of the two South African hake in turn permitted new methods to be used for estimating the sequence divergence between these two species of hake and other hake species. This can lead to reconstruction of the evolutionary history of hake.

4.2 MATERIALS AND METHODS

Unless otherwise stated, all techniques used were from Maniatis *et al.* (1982). All components of solutions are recorded in the Appendix.

4.2.1 Cloning of mtDNA

The mtDNA of *M. paradoxus* was cut into a number of fragments with the restriction enzymes, *Hin*dIII, *Pst* I and *Bcl* I, and *M. capensis* mtDNA was cut with *Hin*dIII. The mtDNA from both species of hake was cloned into the corresponding sites of the vector pEcoR 252 obtained from Dr. M. Zabeau, Plant Genetic Systems, Gent, Belgium. The plasmid pEcoR 252 is a derivative of the plasmid pBR 322, retaining the parent plasmid ampicillin resistance gene, and origin of replication. In addition, the plasmid contains the gene for the *E. coli Eco*RI restriction enzyme, under the control of the lambda rightward promotor/operator region (Pr). The plasmid is maintained in a strain containing lysogenic lambda, and therefore, expression of the EcoR1 gene is repressed. If this plasmid is transformed into a recipient strain that is not a lambda lysogen, the EcoRI gene is derepressed, and overproduction of *Eco*RI results in destruction of genomic DNA and of the cell. Insertion of foreign DNA into the EcoR1 gene causes insertional

inactivation of that gene, allowing the transformed cells to survive. Three separate fragments, using *Hin*dIII, *Pst* I and *Bcl* I of *M. paradoxus* mtDNA were cloned into pEcoR 252 and one *Hin*dIII fragment of *M. capensis* mtDNA was cloned into the same vector.

Vector- and mtDNA was digested with the appropriate enzymes for one hour at 37°C. After digestion, small aliquots were run on a minigel to determine whether or not the DNA was cut to completion. Both the mtDNA and the vector DNA were treated with an equal volume of TE saturated phenol (Appendix) to remove residual enzyme. Residual phenol was removed from the aqueous phase by ether extraction; this step was repeated several times. The DNA was then precipitated using 10% 4M LiCl and 2 volumes of 96% EtOH. The DNA was resuspended in TE buffer, and equal ratios of pEcoR 252 and mtDNA were ligated overnight at 15°C. As a control, pEcoR 252 was religated and run on an agarose gel with EtBr, which showed that the vector was religated. The ligation mixture was repurified and competent cells were transformed. The recipient *E. coli* strain for *M. paradoxus* mtDNA, was HB101, and for *M. capensis* it was LK111.

4.2.2 DNA Ligation reaction

The method of King and Blakesley (1986), for optimizing DNA ligations for transformation, by was adapted as follows:

1 pM/ml of vector and insert DNA was used where

1 pM = (mw expressed in Megadaltons) µg or

1 pM = (0.662 kb) µg

For example, vector pEcoR 252 was 3360 bp, therefore 1 pM = 0.662 x 3.36 = 2.22 µg of DNA was added to the ligation. Ligase (1.0 units) was added per 1.0 pM of DNA, per total volume of the ligation of 50 µℓ (H. Zappe, personal communication). Components and concentrations of the ligation buffer are listed in the Appendix.

Vector and insert DNA were added at a 1:1 ratio. The solution was mixed briefly, centrifuged for 5 seconds, and incubated at 15°C overnight. The DNA was cleaned and used to transform competent cells.

4.2.3 Preparation of plasmid DNA

Plasmid DNA was prepared according to Ish-Horowicz and Burke (1981). Plasmid pEcoR 252 was prepared by growth in a Lambda lysogen, i.e. in the *E. coli* strain K514. A 200 ml overnight culture was harvested at 5000 rpm for 10 min. The cells were resuspended in 3 ml of solution I (Appendix). The suspension was left at 22 °C for 5 min. and then 6 ml of solution II (Appendix) was added. It was mixed and left on ice for 5 min. Precooled solution III (4,5 ml) was added, it was mixed gently and left on ice for 5 min. Denatured proteins were removed by centrifugation at 12000 rpm for 5

min. The supernatant was retained, to which 2 volumes of 95% EtOH was added. The tubes were incubated for 30 min at -20°C and the DNA was pelleted at 12000 rpm for 30 min. The pellet was resuspended in 4 ml of TE buffer, then 5 g of CsCl and 0.2 ml of EtBr (10mg/ml). At this stage a clearing spin at 10000 rpm for 10 min was done to precipitate all the residual protein. The refractive index was adjusted to correspond to a density of 1.396 g/ml, by either the addition of CsCl or TE buffer. The solution was transferred to 5ml Beckman Quick Seal centrifuge tubes, which were then centrifuged overnight at 55000 rpm in a Beckman vertical rotor (VTi 65,2). The plasmid band was collected and EtBr was removed according to Maniatis (1982).

4.2.4 Preparation of Competent Cells

A single colony of *E. coli* (HB101 or LKIII) was inoculated into 10 ml Luria Broth (LB). This was grown overnight with shaking at 37°C. The overnight culture was diluted 1/100 in a 50 ml LB-preculture, and incubated at 37°C with shaking until the cells had grown to an OD_{600} = 0.2 (mid log phase, 2-4 h). The cells were diluted 1/50 in LB-400 ml main culture, and allowed to grow to an OD_{600} = 0.2 at 37°C, with shaking. The LB used for dilution, was prewarmed and glucose was added to a final concentration of 0.1%. The culture was cooled on ice for 5 min and then the cells were harvested in a GSA rotor at 5000 rpm for 5 min. The cells

were resuspended in 200 ml ice-cold 0.1 M $CaCl_2$, and left on ice for 20 min. The suspension was spun down in a GSA rotor at 4000 rpm for 5 min, and the cells were resuspended in 4 ml of ice-cold 0.1 M $CaCl_2$. These cells were then stored on ice and aged overnight at 4°C. The aging step was found to enhance the transformation frequency of the cells. The cells were then used, or stored in 0.1 M $CaCl_2$ containing 15% glycerol at -70°C for use at a later stage.

4.2.5 Transformation of Competent Cells

Competent cells (0.3 ml) were placed in 15 ml sterile culture tubes on ice. The DNA was added to the tubes, 5 µℓ of cleaned ligation mix was used per tube. As a control of transformation, competent cells were transformed with pBR-322 DNA. The tubes were kept on ice for 40 min, then the cells were heat shocked at 42°C for 3 min, and were returned to the ice bath. Two milliliters LB was added to the cells, and they were allowed to grow, with shaking, at 37°C for 30 - 60 min, which allows the cells to recover and start expressing ampicillin resistance. The cells were then plated out on Luria agar plates containing 100 µg/ml ampicillin, which were then incubated at 37°C overnight. Only cells transformed with the vector containing an insert, were expected to grow.

4.2.6 Colony Hybridisation

Colony hybridisation was accomplished by first transferring bacteria from a master plate containing 50 colonies in grid formation, to a nitrocellulose filter (millipore, filter type HA, 0.45 um), which was then transferred to an agar plate containing 100 µg/ml chloramphenicol, and was again incubated at 37°C for another 36 hours. This step amplified the number of recombinant plasmids, and was only necessary when low copy numbers of recombinant plasmids were expected. The colonies were then lysed, and the liberated DNA was fixed to the filter. The filter was placed in the lid of a petridish containing 0.75 ml of 0.5 M NaOH, making sure that the filter was evenly wetted, and that no solution reached the upper surface of the filter. It was left for 2-3 min, then the filter was blotted dry on paper towel, and the above step was repeated using a new petridish lid and fresh 0.5 M NaOH. The filter was blotted dry, transferred to a petridish with 0.75 ml of 1 M Tris-HCl (pH 7.4), for 5 min, then the filter was blotted dry and the step was repeated. Then the filter was transferred to a petridish with 0.75 ml of 1.5 M NaCl and 0.5 M Tris-HCl (pH 7.4). After 5 min the filter was blotted and transferred to a piece of Whatman No 3 filterpaper and allowed to air dry at room temperature for 30-60 min. The DNA was bound to the filter by baking for 2 hours at 80°C in a vacuum oven. The filter was prehybridised, and then hybridised to a α^{32}P-labelled probe, consisting of mtDNA from both *M. paradoxus* and *M. capensis*,

nick-translated to high specific activity with (α^{32}P)dCTP (10μCi/μℓ). The filter was then monitored by autoradiography and the colonies that gave a positive autoradiographic result, were recovered from the master plate.

4.2.7 Nick Translation of DNA

The nicktranslation kit and manual from Amersham (Amersham Int., UK, PB.5025) was used in this study.

The reagents were added into a microcentrifuge tube on ice, then they were mixed and placed at 15°C for 1-2 hours. Incorporation of (α^{32}P) dCTP into the DNA peaked at 2 hours. The reaction was terminated by making up the volume to 0.1 ml and loading the mixture directly onto a 10 cm column of Sephadex G-50 (Maniatis *et al.*, 1982).

The radioactive probe was denatured by boiling for 10 min and then immediately placing them on ice for a few min to prevent reannealing of the two strands. These probes were added to the filters, which had been allowed to prehybridize.

4.2.8 Hybridisation

A number of different hybridisation conditions were experimented with, which involved varying the stringency of hybridisation and washing procedures. Most hybridisation reactions and washes were done at high stringency (65°C). Initially, salmon sperm DNA, Denhardts solution, 20x SSC and BSA was used to prehybridize the blots, but the mtDNA present in sperm DNA resulted in a decrease in hybridisation of the probe with mtDNA bound to the filters. Subsequently, the method from Church and Gilbert (1984) was used. In that method, the blots are blocked by bovine serum albumin and phosphate buffer, which resulted in a ten fold increase of the signal. The blot was prehybridised for 5-10 min in CHB (Church Hybridisation Buffer, see Appendix) at 65°C, then the probe was added and allowed to hybridize overnight. The blot was washed 2x in WBA (Wash Buffer A), 10 min per cycle and then 4x WBB (Wash Buffer B), also 10 min per cycle. The blot was dipped in distilled water to remove SDS, and was then sealed in a plastic bag and subjected to autoradiography.

The colonies to which the mtDNA hybridised, were selected from the master plates and DNA was isolated as was described for preparation of plasmid DNA. This DNA was in turn $\alpha^{32}P$ labelled and then used as a probe to hybridize to mtDNA run on an agarose gel.

4.2.9 Southern blotting

Capillary blotting was done according to the recommendations of the manufacturer of Hybond-N (Amersham) and according to the method described by Reed (1987). After electrophoresis, using Tris-borate gels, the gel was placed in a pan containing two volumes of 0.25 M HCl, to nick the DNA. It was shaken for 15 min at room temperature, the HCl was poured off, and the step was repeated. Then the gel was rinsed in H_2O and the DNA was denatured with two volumes of 0.5 M NaOH, 1.5 M NaCl for 15 min, which was repeated twice. Following that, the gel was neutralized with 0.001 M Na_2EDTA, 1.5 M NaCl, 0.5 M Tris-HCl, pH 7.2 for 30 min, and this step was repeated twice. The agarose gel was placed on a glass plate, and Hybond-N, saturated in water, was placed on top of the agarose gel. Three sheets of filterpaper saturated with the same buffer were placed on top of the Hybond-N. Then a 2 inch layer of dry paper towels was placed on top of the filterpapers, and a weight of 1-2 kg was placed on top to ensure even contact. Transfer of DNA was allowed to proceed for 2 hours, to overnight. Then the DNA was bound to the filter, by placing the filter, DNA side down, on a standard short wavelength uv transilluminator for 2-5 min.

4.2.10 Isolation of cytoplasmic DNA

The method for rapid extraction of cytoplasmic nucleic acids from Lansman *et al.* (1981) was used to isolate mtDNA from *M. australis*. Ovaries, frozen in liquid nitrogen, were provided by Fisheries Research Division of New Zealand. The ovaries were subsequently stored at -70°C until they were used.

Two grams of liver were minced and then homogenized with 5 ml of 2x TSM (Appendix) using an Ultraturox homogenizer. The homogenate was made 0.2% nonidet P40 by addition of 0.1 ml of a 10% solution. Intact nuclei were removed by centrifugation for 5 min at 15000*g*. The supernatant was kept, an equal volume of TE-saturated phenol added to it, and this solution was held at 65°C for 5min. The phases were separated by centrifugation at 6000*g* for 10 min and the aqueous layer was reextracted with an equal volume of phenol. The residual phenol was removed by extracting the aqueous phase with ether. The DNA was precipitated by the addition of 95% EtOH and incubation at -20°C for 2h. The precipitate was collected by centrifugation at 6000*g* for 20 min. The pellet was washed with 70% EtOH, dried and dissolved in 1 ml TE buffer. The DNA prepared in this way was digested with restriction enzymes and run on agarose gels.

4.2.11 Mapping of mtDNA

Mitochondrial DNA was prepared using the method for rapid isolation of mtDNA from fish, by Chapman and Powers (1985). Mapping was achieved by digesting the mtDNA with two enzymes at the same time, then running it on agarose gels. By measuring the sizes of the fragments, a map of the restriction sites in relation to each other was obtained. One *M. paradoxus* mtDNA was mapped with six enzymes: *Bam*HI, *Bcl* I, *Hin*dIII, *Pst* I, *Sac* II, and *Xho* I. Three *M. capensis* mtDNAs were mapped with *Bam*HI, *Eco* RI, *Hin*dIII, *Pst* I, *Sac* II, and *Xho* I. In this species we chose three individuals, each one with a different genotype for the restriction fragment pattern of *Xho* I. The mtDNA clones from both species were mapped extensively using up to 27 different enzymes. The mtDNA of *M. australis* was digested, electrophoresed, and the gel was hybridised to a mixture of the clones of *M. paradoxus* and *M. capensis*, labelled with α^{32}P.

4.2.12 Data analysis

Sequence divergence, p, between *M. australis* and the two South African hakes was calculated, using the number of restriction fragments shared, as was described in Chapter 3. The number of shared cleavage sites, S, between two mtDNAs was calculated according to Nei and Li (1979).

$$S = 2N_{xy} / (N_x + N_y)$$

where N_{xy} is the number of sites in common to the two mtDNAs and N_x and N_y represent the total number of restriction

sites in each of the two mtDNAs. The amount of nucleotide sequence divergence of the two mtDNAs can be estimated from S by the following relationship:

$$p = - (\ln S)/n$$

where n is the number of nucleotides in the recognition sites of the restriction enzymes. Sequence divergence was estimated for 5-bp and 6-bp recognition sites separately, and was averaged by weighting each p value with the number of nucleotide sites detected by each group of restriction enzymes.

4.3 RESULTS

4.3.1 Cloning of mtDNA

Bacterial colonies containing recombinant plasmids, which had inserts of mtDNA, were selected by means of colony hybridisation (Fig. 4.3.1). Three different fragments of *M. paradoxus* mtDNA were cloned into the vector pEcoR 252 and one *Hin*dIII fragment of *M. capensis* was cloned into the same vector. The recombinant plasmids were isolated and the inserts were cut out by restriction enzymes and run on a gel in the lane adjacent to total mtDNA cut with the same enzyme. Clones contained bands which corresponded to a band in the restriction pattern of the mtDNA. The following fragments were cloned: (1) A *Pst* I fragment from *M. paradoxus*, of 500 bp, which was designated pIIB 101 (Fig. 4.3.2). This small fragment was not detected previously.
(2) pIIB 102 represents a 2,2 kb *Hin*dIII fragment of the same species (Fig. 4.3.2).
(3) An 8 kb *Bcl* I fragment of *M. paradoxus* was cloned into the *Bgl* II site of pEcoR 252 and was named pIIB 104 (Fig. 4.3.2).
(4) The *Hin*dIII fragment of *M. capensis* was 4 kb and the plasmid was designated pIIB 201 (Fig. 4.3.3).
Gels with digested total mtDNA were run, blotted and hybridised with each of the clones, ensuring that the insert was mtDNA.

Figure 4.3.1

Autoradiograph of colonies of *Escherichia coli* containing the plasmid pIIB 102 hybridised with mtDNA from *Merluccius paradoxus* labelled with $\alpha^{32}P$.

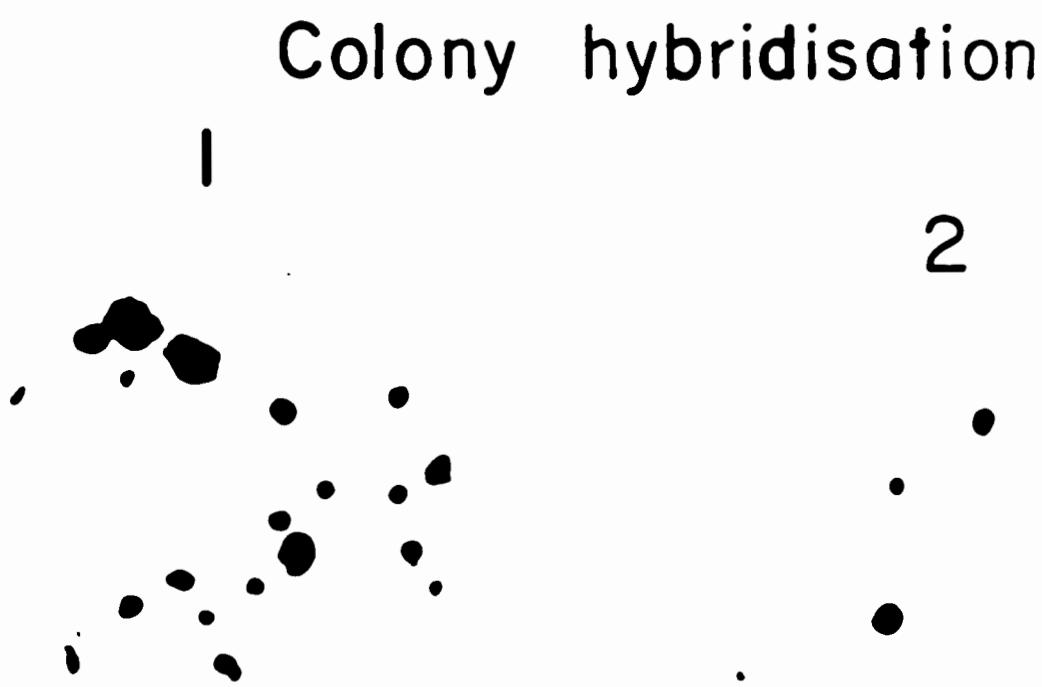


Figure 4.3.2

Restriction map of the mtDNA from *Merluccius paradoxus*. The outside circle represents the restriction maps of fractions of the mtDNA cloned into the cloning vector pEcoR 252.

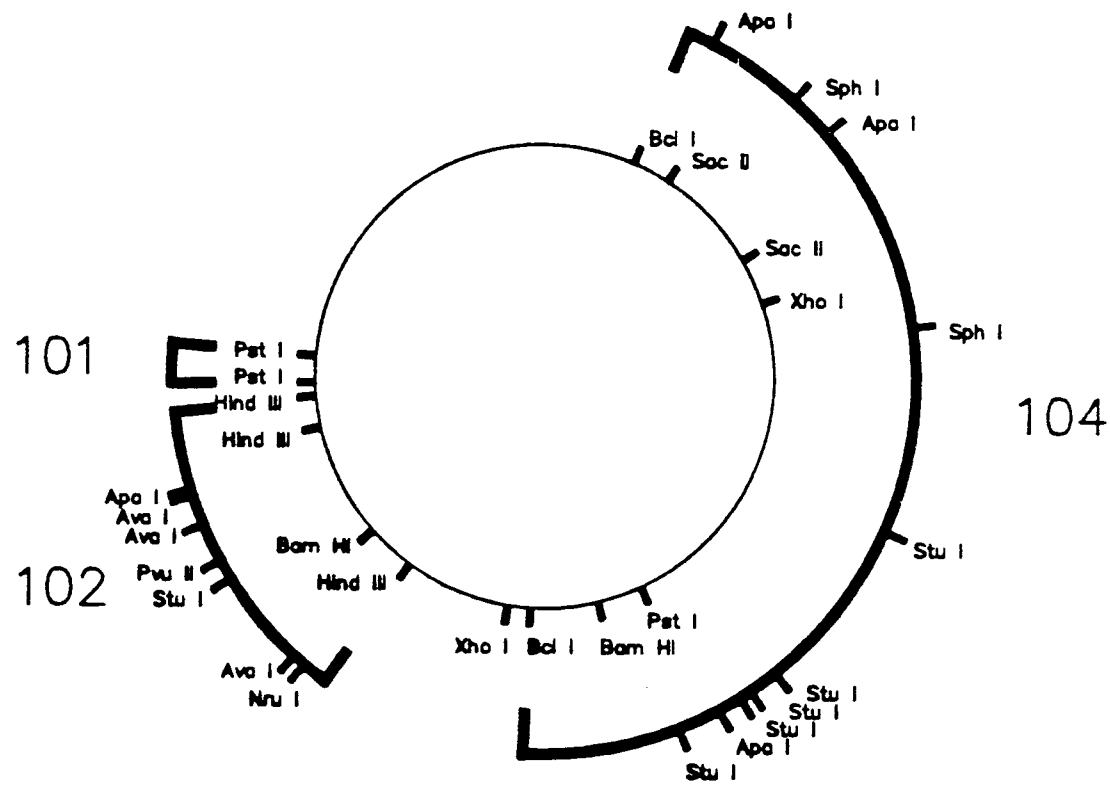
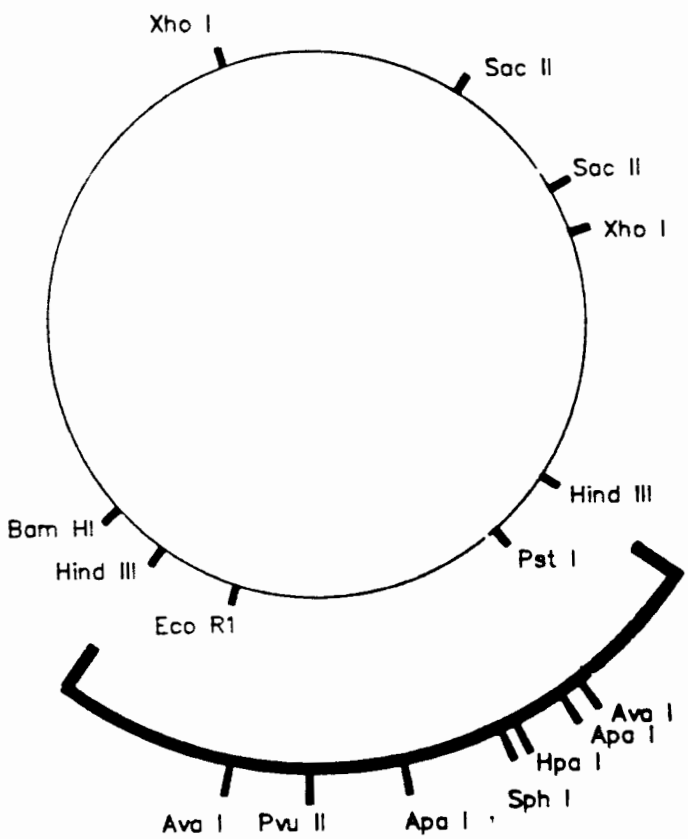


Figure 4.3.3

Restriction map of mtDNA from *Merluccius capensis*, and the restriction map of the cloned *Hin*dIII fragment, pIIB 201.



4.3.2 Cross hybridisation studies

Various cross hybridisation studies were performed between the four clones to establish the amount of sequence homology between the clones from the two species of hake. The *M. paradoxus* clone, pIIB 104, hybridised to *M. capensis* mtDNA with a slight decrease in the intensity of the signal, compared to binding to *M. paradoxus* mtDNA itself. This indicated that considerable sequence homology exists between the mtDNAs of these two species. The cloned *Hin*dIII fragment of *M. paradoxus*, pIIB 102, hybridised only to the small *Hin*dIII fragment of the mtDNA from both species and not to the large fragment, indicating that there are no extensive repeated sequences in the mtDNA.

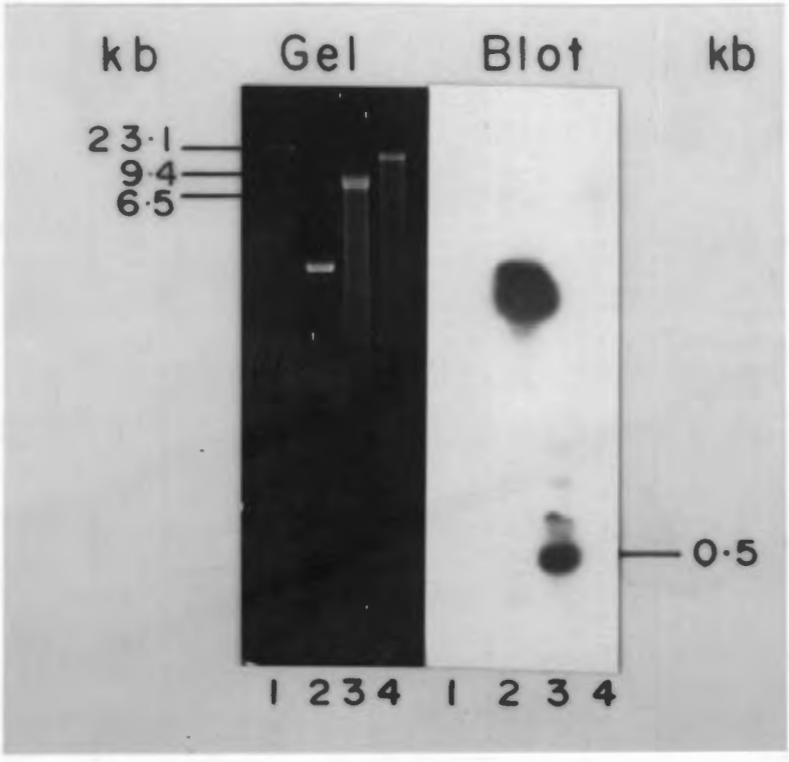
The small *Pst* I fragment of *M. paradoxus* was found to hybridize exclusively to *M. paradoxus* (Fig. 4.3.4).

The clone, pIIB 201, hybridised to the small *Hin*dIII fragment of *M. capensis*, and hybridize only to a small extent to the large *Hin*dIII fragment. This clone also hybridizes well to both *Hin*dIII fragments of *M. paradoxus*, again confirming the mtDNA sequence homology between the two species of hake.

The mtDNA clones from both species of southern african hake were used as probes to determine the restriction fragment pattern of the mtDNA from *M. australis*.

Figure 4.3.4

Picture of gel stained with Ethidium Bromide on the left and the same gel blotted and hybridised with pIIB 101 radioactively labelled with $\alpha^{32}P$ is on the right. Lane 1 contains phage lambda DNA digested with *Hin*dIII as a size marker. Lane 2 contains the vector pEcoR 252 DNA cut with *Hin*dIII as a control for blotting. Lane 3 contains *Merluccius paradoxus* mtDNA cut with *Pst* I and *Merluccius capensis* mtDNA cut with *Pst* I was run in lane 4. The blot shows that pIIB 101 only hybridizes to a small 500 bp fragment of *M. paradoxus* (lane 3) and not to the large fragments of *M. paradoxus* or the mtDNA from *M. capensis* (lane 4).



Other crosshybridization studies were performed between the total mtDNA of *Xenopus laevis*, cloned into pBR 322 (obtained from I.B. Dawid, National Institutes of Health, Maryland), and the cloned mtDNA of both species of hake. Using high stringency (65°C) hybridisation conditions, it was found that the *X. laevis* did not hybridize to the clones or the total mtDNA of both species of hake.

4.3.3 Restriction fragment analysis of M. australis mtDNA.

Mitochondrial DNA from seven individuals of *M. australis* was digested with the following enzymes: *Ava* I, *Bam*HI, *Bgl* I, *Eco* RI, *Hin* dIII, *Kpn* I, *Pst* I *Pvu* II, *Sac* I, *Xba* I and *Xho* I. Fragment sizes are presented in Table 4.3.1. The size of the *M. australis* mtDNA was determined to be 16,975 bp ±995 bp. The digests were run on gels and these were blotted. Blots were hybridised with a pool of pIIB 101, 102, and 104, or with pIIB 201 radioactively labelled with α^{32}P. The mtDNA clones from both species of southern African hake hybridize extensively to the mtDNA from New Zealand hake. The *M. paradoxus* clones hybridised to more *M. australis* mtDNA fragments than the *M. capensis* clone, which was expected, as a larger proportion of the *M. paradoxus* mtDNA was cloned. Seven *M. australis* mtDNAs were digested with the eleven restriction enzymes, and three composite fragment genotypes were identified (Table 4.3.2). The fragment

Table 4.3.1

Fragment sizes (bp) of *M. australis* mtDNA digested with eleven different restriction enzymes.

Ava I	*Bam*HI	*Bgl* I	*Eco*RI	*Kpn* I	*Hin*dIII
5500	16500+	10000	No sites	No sites	11500
4300		7200	17000*+	17000*+	2200
3500		1150*			1500
2900					700
2000					400*
500*+					

Pst I	*Pvu* II	*Sac* I	*Xba* I	*Xho* I
16000	8000	8000	12000	10500
200	2700	8000	4000	5000*
	2250			1900
	1800			
	1000			
	800			

* represents fragments shared between *M. australis* and *M. paradoxus*.

+ represents fragments shared between *M. australis* and *M. capensis*.

Table 4.3.2

Composite genotypes of *Merluccius australis*. Each letter in the composite genotype refers to the restriction enzyme fragment pattern of *Ava* I, *Bam*HI, *Bgl* I, *Eco*RI, *Kpn* I, *Pst* I, *Pvu* II, *Sac* II, *Xba* I, and *Xho* I respectively.

Clonal designation	Composite genotype	Number of individuals
21	FADECECCGEE	4
22	FADACACCGEE	1
23	FADACECCGEE	2

pattern of *M. australis* mtDNA is presented in Fig. 4.3.5, and the nucleotide sequence divergence between *M. australis* and the two South African species was calculated as described in chapter 3. Nucleotide sequence divergence was estimated separately for 5- and 6-bp restriction enzymes and was averaged by weighting each p value with the number of nucleotide sites detected by each group of enzymes. The F and p values are presented in table 4.3.3 and 4.3.4 respectively.

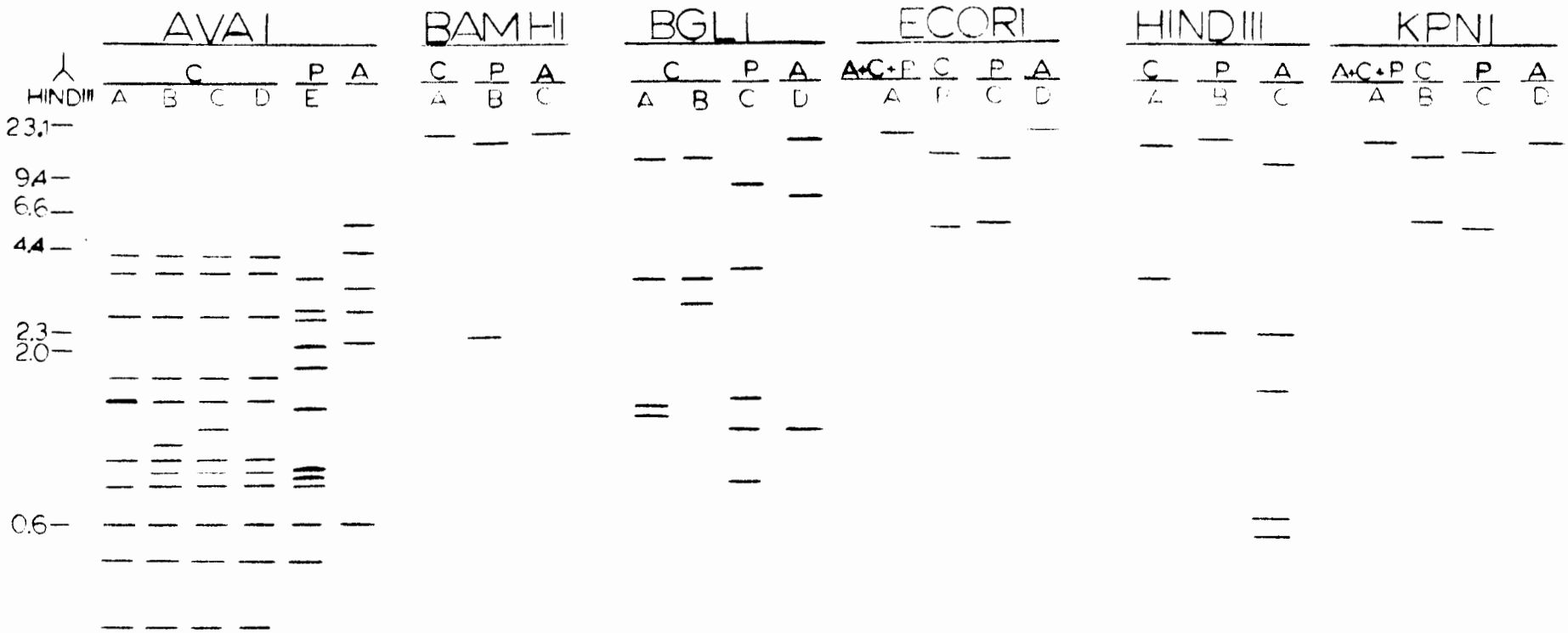
The amount of sequence divergence, p, between genotypes of *M. australis* ranged from 0.001 to .0021 and averaged 0.0014. Sequence divergence between the composite genotypes of *M. australis* and *M. capensis* ranged from 13,43 to 20,13% and averaged at 17,94% ($\pm$2.27%). Sequence divergence between *M. australis* and *M. paradoxus* ranged from 9,69 to 12,27% and averaged at 11,48% ($\pm$1.51%).

4.3.4 Mapping of mtDNA

The whole mtDNA from both southern African hake species was mapped with *Bam*HI, *Hin*dIII, *Pst* I, *Sac* II and *Xho* I. Three different genotypes were reported for the enzyme *Xho* I and three samples with different *Xho* I genotypes were mapped. The results confirmed that an insertion of 400 bp and a deletion of 200 bp had occurred, increasing or decreasing

Figure 4.3.5

Diagrammatic representation of restriction endonuclease digests of hake mtDNA. A stands for *Merluccius australis*, C stands for *M. capensis*, and P stands for *M. paradoxus*.



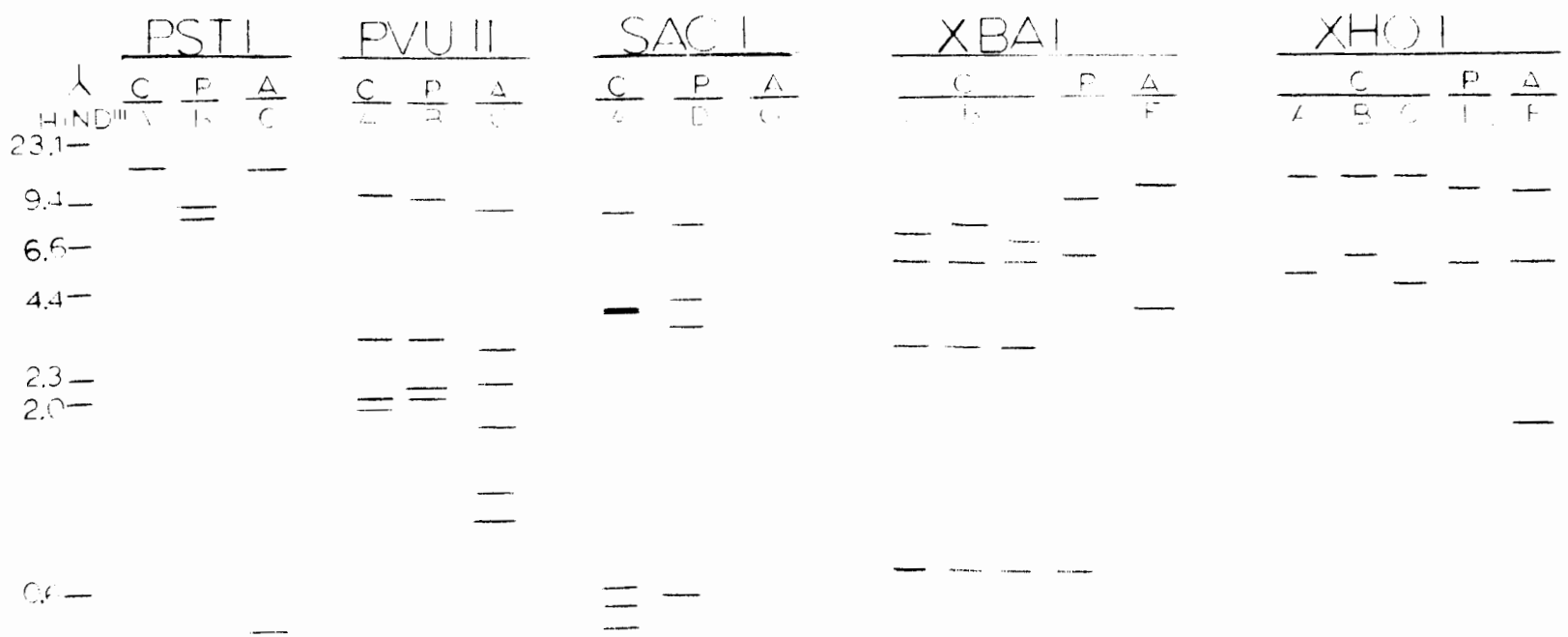

PST I
PVU II
SAC I
XBA I
XHO I
C
P
A
HIND III
23.1
9.4
6.6
4.4
2.3
2.0

Table 4.3.3

The *F* value estimated between the composite genotypes of *Merluccius australis* and the two South African species of hake, using the proportion of restriction fragments shared between the species.

Clonal designation	*Merluccius australis* 21	22	23
Merluccius capensis			
1	.0400	.1154	.0784
2	.0392	.0755	.0385
3	.0392	.0755	.0769
4	.0425	.0816	.0417
5	.0392	.0755	.0385
6	.0377	.0364	.0370
7	.0400	.0115	.0784
8	.0392	.0755	.0385
9	.0400	.0769	.0392
10	.0392	.0113	.0769
11	.0377	.0364	.0370
12	.0400	.1154	.0784
13	.0385	.0370	.0377
14	.0385	.0741	.0377

Table 4.3.3 (continued)

Clonal designation	*Merluccius australis* 21	22	23
Merluccius paradoxus			
15	.1501	.2182	.1852
16	.1501	.2182	.1852
17	.1538	.2222	.1887
18	.1501	.1818	.1852
19	.1455	.1404	.1429
20	.1501	.1818	.1481

Table 4.3.4

Sequence divergence, p, between *Merluccius australis* and the two South African hake, *M. capensis* and *M. paradoxus*. Sequence divergence was estimated for 5 bp- and 6 bp-specific restriction enzymes separately, and was averaged by weighting each p value with the number of nucleotide sites detected by each group of restriction enzymes.

Clonal designation	*Merluccius australis* 21	22	23
Merluccius capensis			
1	.1959	.1343	.1570
2	.1970	.1592	.1981
3	.1970	.1592	1581
4	.1741	.1546	.1935
5	.1979	.1601	.1991
6	.1826	.2016	.2007
7	.1993	.1343	.1570
8	.1970	.1592	.1981
9	.1959	.1581	.1970
10	.1974	.1357	.1584
11	.1823	.2013	.2003
12	.1959	.1343	.1570
13	.1981	.2003	.1993
14	.1981	.1601	.1993

Table 4.3.4 (continued)

Clonal designation	*Merluccius australis* 21	22	23
Merluccius paradoxus			
15	.1188	.0969	.1065
16	.1188	.0969	.1065
17	.1174	.0958	.1054
18	.1188	.1076	.1065
19	.1206	.1227	.1217
20	.1188	.1076	.1196

the size of the smaller of the two *Xho* I fragments, as was determined in the fragment analysis.

The cloned fragments of mtDNA were mapped extensively with up to twenty different restriction enzymes. The cloned *Hin*dIII fragment of *M. capensis* (pIIB 201) had no restriction sites for the following enzymes: *Bcl* I, *Cla* I, *Dra* I, *Eco* RV, *Nun* II, *Nru* I, *Sac* I, *Sac* II, *Sal* GI, *Sma* I, *Xba* I, *Xho* I. All three *M. paradoxus* clones had no restriction sites for *Alu* I, *Bcl* I, *Cla* I, *Dra* I, *Hpa* I, *Kpn* I, *Sac* I, *Sfi* I, and *Xba* I. The restriction maps of pIIB 101, 102 and 104 were aligned with the restriction map of *M. paradoxus* (Fig. 4.3.2), and pIIB 201 was aligned with the restriction map of *M. capensis* (Fig. 4.3.3).

The mtDNA from *M. australis* was mapped with *Bam*HI, *Hin*dIII, *Pst* I, *Sac* II, and *Xho* I using all four clones as probes (Fig. 4.3.6).

4.3.5 Restriction site analysis

All three species of hake were mapped with the five restriction enzymes, *Bam*HI, *Hin*dIII, *Pst* I *Sac* II, and *Xho* I (Fig. 4.3.7). This provided more information concerning sites that were shared among the species. The following sites were shared between *M. capensis* and *M. paradoxus*: Two *Sac* II sites, one *Bam*HI, one *Hin*dIII, one *Pst* I, and one *Xho* I site. *Merluccius australis* and *M. capensis* had two *Sac* II

Figure 4.3.6

Restriction map of the mtDNA from *Merluccius australis*.

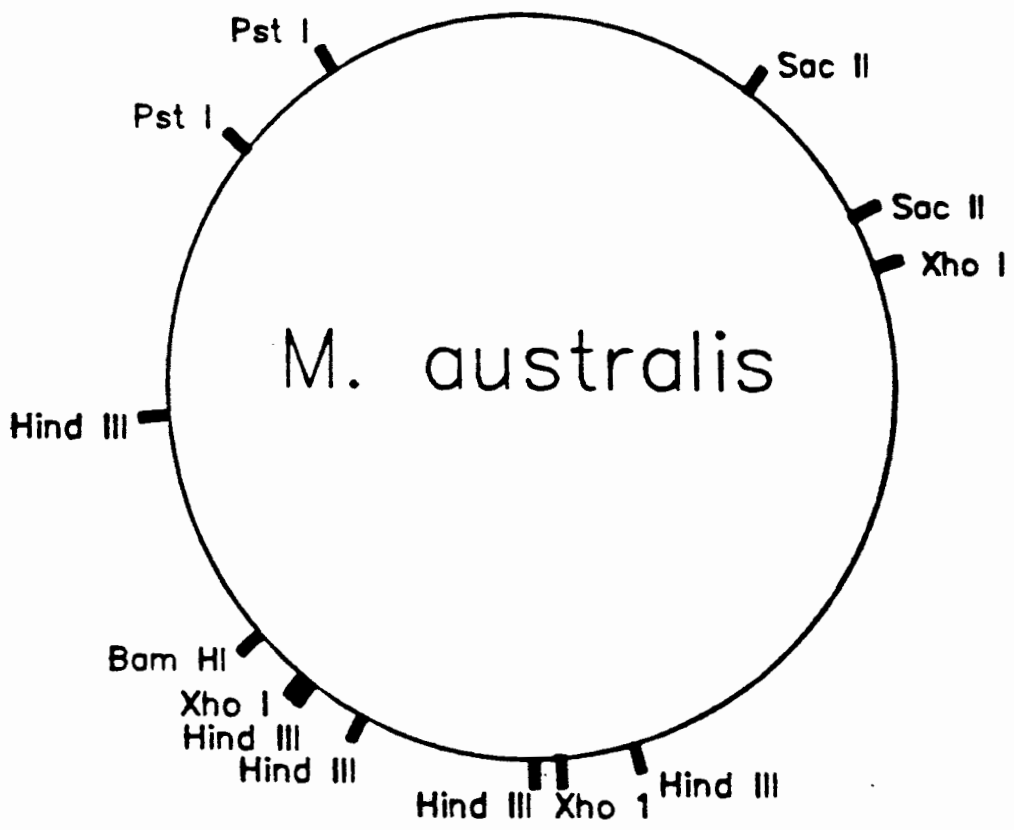
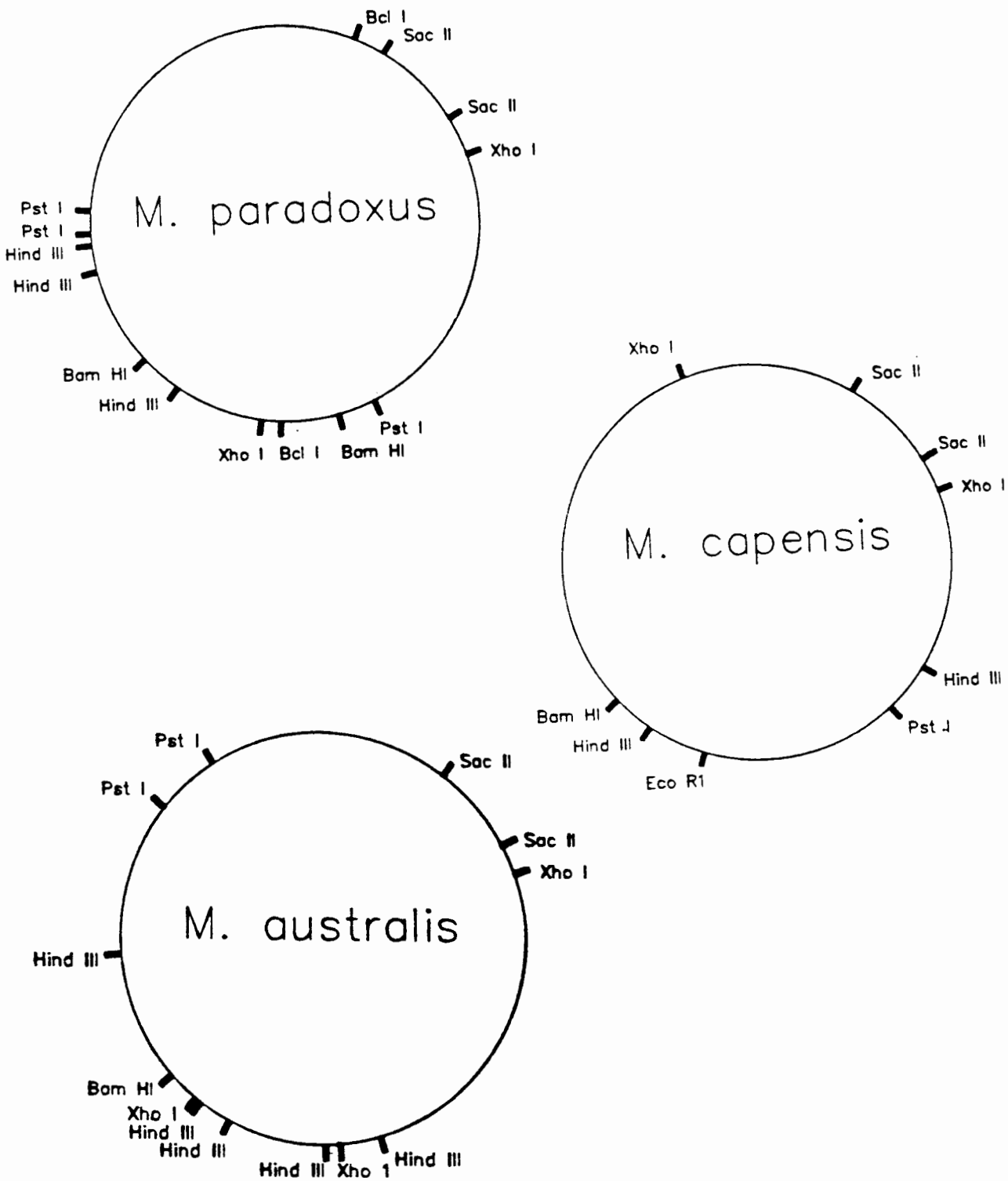


Figure 4.3.7

Restriction maps of the mtDNA from *Merluccius australis*, *M. capensis*, and *M. paradoxus* aligned with respect to the sites shared between these species of hake.



sites, one *Bam*HI, and one *Xho* I site in common. The restriction sites held in common between *M. australis* and *M. paradoxus* were two *Sac* II sites, two *Xho* I sites, one *Bam*HI site, and one *Hin*dIII site. Sequence divergence, *p*, between *M. capensis* and *M. paradoxus*, using the number of restriction sites held in common, was found to be 8.5%. Similarly, the sequence divergence between *M. australis* and *M. capensis* was 16.09%, and the *p* value calculated between *M. australis* and *M. paradoxus* was 12.23%.

4.4 DISCUSSION

4.4.1 Hybridisation methods

Initially hybridisation procedures used, all involved a prehybridisation step in which the nylon membrane is blocked with sonicated DNA, to prevent non-specific binding of the radioactive probe to the nylon membrane. Usually salmon- or herring sperm DNA is used for blocking of the membranes. In this study it was observed that the mtDNA probes of *M. capensis* and *M. paradoxus* did not bind very efficiently to the mtDNA from *M. australis*. This result was somewhat surprising as the mtDNA sequence is known to be highly conserved across taxa as far separated as amphibians and mammals (Wong *et al.*, 1983). The hake used in this study all belong to the same genus and more sequence homology (above 75%) was expected between the mtDNA from the various species.

Subsequently, a different hybridisation method was used (Church and Gilbert, 1984), which did not include any DNA in the prehybridisation solution. The result showed that the hybridisation was enhanced by a factor of 10. The method also proved to be quicker (only 10 min prehybridisation), easier (fewer solutions to be made up), and very efficient as no non-specific binding of the probe to the membrane was detected (and no Denhardts spots). These results show that it is an excellent method to be used in studies involving

hybridisation of a mitochondrial clone to mtDNA. It is likely to be the method of choice especially when studying fish mtDNA. The midpiece of mature sperm contains a small number (approximately 100) of mitochondrial DNA molecules (Avise and Lansman, 1983). It can be assumed that this mtDNA will also be present in the sperm DNA, which is used in the prehybridisation step. This sperm mtDNA will hybridize to the hake mtDNA, which is bound to a nylon membrane, and it presumably decreases the hybridisation of the radioactive probe to the hake mtDNA and will increase the background.

Usually herring- or salmon- sperm DNA is used for prehybridisation, and this sperm mtDNA would bind especially efficiently to the mtDNA of other fish. However, it has not been previously reported that the mtDNA of sperm DNA will interfere with the hybridisation of the probe. Lansman *et al.* (1981), describes the hybridisation techniques using *Peromyscus maniculatus* samples and cloned mtDNA as a probe. Salmon sperm DNA was used in the prehybridisation and hybridisation step. In other studies of fish, mostly sperm DNA is used in the prehybridisation step. Graves *et al.* (1984), for instance, used rainbow trout mtDNA as a probe to hybridize to skipjack tuna DNA, and Gonzàlez-Villaseñor *et al.* (1986) used cloned mtDNA from *Fundulus heteroclitus* as a hybridisation probe for other fish species. In both cases salmon sperm DNA was used in the prehybridisation step. The hybridisation was not carried out at high stringency

conditions (e. g. low temperature and low formamide concentration), and no decrease of hybridisation due to the sperm mtDNA was reported.

4.4.2 Cloned hake mtDNA as hybridisation probes

The fragments of mtDNA from *M. capensis* and *M. paradoxus*, cloned into pEcoR 252 were used as probes to detect mtDNA fragments from other species of hake. Size often limits the amount of mtDNA that can be isolated from one individual. Hake are large fish, but it is often difficult to obtain large quantities of tissue from one individual and the condition of the tissue on arrival will only allow isolation of a limited amount of mtDNA. The hake mtDNA clones will make it possible to digest each individual with a reasonable number of restriction enzymes, as very small amounts of mtDNA are detectable when using a hybridisation probe.

A pool of the *M. paradoxus* clones represented about 70% of the total mtDNA, and it appeared that they hybridised to most of the mtDNA fragments of *M. australis*. These clones represented an efficient probe, which facilitated the restriction endonuclease analysis of *M. australis*.

The clones can also be used in future studies to determine the evolutionary process and relationships among the other species of hake, which inhabit continental shelf waters around the world.

A small, 500 bp, *Pst* I fragment of *M. paradoxus* was only detected by cloning of the mtDNA. This clone, when used as a hybridisation probe, was found to bind exclusively to the mtDNA of *M. paradoxus*, and can be used in studies where it is necessary to distinguish between *M. paradoxus* and other species of hake. It could also be used in diagnostic studies to differentiate between the proportions of *M. capensis* and *M. paradoxus* present in a catch.

The *M. capensis* mtDNA clone represented only 25% of the mtDNA, but it proved to be useful as a hybridisation probe in the fragment analysis of *M. australis*, where it only hybridised to the complementary mtDNA fragments.

The *M. capensis* mtDNA clone also hybridised specifically to its complementary mtDNA fragment of *M. paradoxus* or *M. capensis*. This confirmed that there was no duplication of of the region in other parts of the mitochondrial genome, and that no extensive repeated sequences are present in the mtDNA of *M. capensis* and *M. paradoxus*. Similar results were obtained when hybridising the *M. paradoxus* mtDNA clones with mtDNA from *M. capensis* and *M. paradoxus*.

The total cloned mtDNA from *X. laevis* did not hybridize with the mtDNA clones of *M. capensis* and *M. paradoxus* or the total mtDNA from both species of hake. The hybridisation conditions used in this study were stringent (65°C) and the sequence homology of the mtDNAs was not enough for them to hybridize to each other (less than 70% homology, see DNA

sequence studies, chapter 5). To use the cloned mtDNA from *X. laevis* as a hybridisation probe for hake, the conditions of hybridisation and washing would have to be of lower stringency. The clones isolated in this study are species-specific probes, which can be used in further studies to elucidate the evolution of hake.

4.4.3 Restriction fragment analysis of M. australis mtDNA

The size of the mtDNA of *M. australis* was calculated by averaging the sum of the fragments produced by eight different restriction enzymes. The mtDNA was found to be 16,975 bp ±996, which is approximately the same size as the size of the mtDNA of the two southern African hakes. This agrees with the sizes of mtDNA reported for other fish (see chapter 3, Table 3.4.1).

Only three different composite genotypes were detected for *M. australis*. Two mutations were necessary to produce these three composite genotypes. This showed that *M. australis* displayed a low level of nucleotide site polymorphism compared to the levels reported for *M. capensis* (minimum of 22 mutational steps) and *M. paradoxus* (minimum of 12 mutational steps). This low level of nucleotide site polymorphism could be ascribed to the small number of fish examined in this study. Only by examining more *M. australis*

individuals, can differences in the level of nucleotide site polymorphism between the three species be quantified. As discussed in chapter 3, various factors could have attributed to the different levels of site polymorphisms. One possible explanation is that *M. australis* experienced different population events to the two southern African species. To reconstruct these events, additional information about the variability of the nuclear genes is needed. This could be obtained by protein electrophoresis. This again emphasizes the power of a dual approach, utilising restriction analysis of mtDNA in conjunction with protein electrophoresis. If it is assumed that the low level of polymorphism found in the mtDNA of *M. australis*, is not due to population size, sample error or lower rate of mutation, it can be assumed that *M. australis* experienced a more recent bottleneck in population size than *M. paradoxus*.

4.4.4 Sequence divergence

Sequence divergence was estimated between *M. australis* and the two South African hake using the number of restriction fragments held in common. Sequence divergence between the genotypes of *M. australis* and the composite genotypes of *M. capensis* averaged 17,94% and the estimated sequence divergence between *M. australis* and *M. paradoxus* averaged 11,48%. In the previous chapter, the p values for pairs of species from various taxa, were summarised in table 3.4.3.

In general, sequence divergence within species ranges from 0 to 5%, and distantly related species have *p* values of 10% and above. Closely related sibling species will have *p* values of approximately 5% to 10%. Gyllensten and Wilson (1987), on the other hand, report that the divergence between the two genera, *Salmo* and *Salvelinus*, was approximately 13%. This shows that no definite scale exists for the amount of sequence divergence, but sequence divergence can indicate whether two populations are conspecific or congeneric. Sibling species are also expected to share some of the mtDNA genotypes, because these genotypes existed prior to the split into two species (Ferris and Berg, 1987). No genotypes were found to be shared between any of the three species of hake. The amount of sequence divergence estimated between *M. australis* and the two southern African species of hake also shows, as expected, that all three are well separated congeneric species.

The amount of sequence divergence can be used to estimate the branching times between two species. As was discussed in chapter 3, mtDNA accumulates mutations at a rate of 2% per million years, and this value is slowed down for species that were separated more than 10 million years ago. Generally, values of *p* below 0.15 are within the linear portion of the curve. Using this time scale, it was determined that *M. australis* and *M. paradoxus* separated 5.7 million years ago. The *p* value of 0.18, estimated between

M. capensis and *M. australis*, lies in the non-linear range of the calibration curve. Assuming that a similar rate of mutation existed, this *p* value indicates that the time of divergence between the two species was approximately 7 to 9 million years ago.

4.4.5 Mapping of mtDNA

The mtDNA from all three species of hake, *M. australis*, *M. capensis*, and *M. paradoxus*, were mapped with five different restriction endonucleases (Fig. 4.3.7). Mapping of restriction sites allows more rigorous explanation of the mutations that produce the various restriction fragment patterns, and it also makes it possible to detect deletions, insertions, and other rearrangements (Ferris and Berg, 1987). By mapping the three different *Xho* I restriction fragment genotypes of *M. capensis*, it was confirmed that an insertion of 400 bp, or deletion of 200 bp took place into the small *Xho* I fragment.

By mapping all three hake species with the same restriction enzymes, the amount of sequence divergence could also be calculated using the number of restriction sites held in common by a pair of species. The amount of sequence divergence, *p*, estimated between *M. capensis* and *M. paradoxus*, averaged 0.085 (0.116), the *p* value between *M. capensis* and *M. australis* was 0.160 (0.179), and the

sequence divergence between *M. paradoxus* and *M. australis* averaged 0.122 (0.115). The value of *p* estimated using the number of restriction fragments held in common between each pair of species, is shown in parenthesis. There is variation of approximately 2% between the two methods used to determine the amount of sequence divergence between pairs of species. The difference in the values of sequence divergence reported here, might be due to the different number of restriction enzymes used for both methods. Ferris and Berg (1987) compared sequencing, mapping and fragment analysis of mtDNA between species and within species. Between species, the difference in the percentage of sequence divergence, comparing mapping and fragment analysis of mtDNA, ranged from 1% to 7%. The difference of 2% reported in this study seems to be well within the expected range. This also shows that the amount of sequence divergence is a rough estimate of the time elapsed since two species separated.

Cloned fragments of mtDNA from *M. capensis* and *M. paradoxus* were mapped in detail to acquire more information about the clones, and to facilitate subcloning into sequencing vectors. Subcloning and sequencing of the clones is described in the following chapter.

4.4.6 Biogeographic scenarios

Data obtained by means of restriction analysis of mtDNA can be used to reconstruct phylogenies. The hypotheses existing at present about the dispersal of hake, are discussed in chapter 1. In short, Kabata and Ho (1981), and Inada (1981) proposed that hake dispersed along two main routes from the North Atlantic. One group migrated from the North Atlantic along the coast of Africa, till they reached the southern most tip of Africa and the other group of hake dispersed southward along the coast of north America and then crossed from the Atlantic Ocean into the Pacific Ocean before the closure of the Panama Seaway. This group then migrated southward and around the Cape Horn. It was also postulated that the New Zealand hake must have been derived from southern South America (Inada, 1981).

At present no single hypothesis can explain the presence of two sympatric species of hake in the southern African waters. In this study, with the aid of protein electrophoresis and restriction endonuclease analysis, it was determined that the two southern African species of hake are not closely related and may not share an immediate common ancestor. This indicates that the two species of hake did not speciate sympatrically .

Another hypothesis is that the two hake species, *M. capensis* and *M. paradoxus*, arrived via different routes or by secondary dispersal of the North Atlantic taxa, to southern

Africa. If both hake species arrived in southern Africa by two different dispersals of the North Atlantic hake along the West Coast of Africa, the species which dispersed more recently would be expected to be genetically more closely related to the North Atlantic species. In this study, no hake mtDNA from the European hake, *M. merluccius*, was available to test this hypothesis. Another assumption that can be made, is that older species will have many divergent mtDNA genotypes, whereas the younger species will have fewer, more closely related mtDNA genotypes (Ferris and Berg, 1987). This leads to the suggestion that *M. capensis* is an older species compared to *M. paradoxus*, and that *M. capensis* most probably represents the species that originated from the first dispersal of the North Atlantic hake to southern Africa.

Inada (1981) showed, by morphological studies, that *M. australis* was closely related to the other two deep water hakes, *M. paradoxus* and *M. polli*. He also reported that *M. australis* is different from *M. hubbsi* and other Pacific hake species, and that in general appearance it is more closely related to eastern Atlantic species such as *M. merluccius*, *M. senegalensis*, and *M. capensis*. Other demersal fish from southern South America, New Zealand and South Africa also showed close relationships, and Inada (1981) reported that there is a possibility that dispersal from east to west occurred in the southern Atlantic. From the results obtained by endonuclease analysis of mtDNA, it could be

concluded that *M. paradoxus* was more closely related to *M. australis* than *M. capensis* was to *M. australis*. These results fit in with the observations that *M. australis* is closely related to the other deep water hake, and that a close relationship exists between the fish from southern South America, New Zealand, and southern Africa. One possible explanation is that an ancient *M. paradoxus* or *M. polli* species migrated to southern South America and from there dispersed to New Zealand, speciating into *M. australis*. *Merluccius australis* and *M. paradoxus* do not have any mitochondrial genotypes in common, and the amount of sequence divergence estimated between the two species ($p = 0.114$) indicates that they are not sibling species. This implies that they must be connected through an intermediate species or that they have been isolated from one another for a very long time. Not enough data is yet available to reconstruct with certainty the historical biogeography of hake, but this study confirms hypotheses existing about the dispersal of *M. australis* to New Zealand. Further work will be needed to confirm the origin of *M. capensis* and *M. paradoxus*.

CHAPTER 5

SEQUENCING PARTS OF THE MITOCHONDRIAL DNA OF *MERLUCCIUS CAPENSIS* AND *M. PARADOXUS*

5.1 INTRODUCTION

5.1.1 Sequencing the mtDNA

Several classical techniques exist today, which measure the relatedness in DNA sequences of individuals. In the previous sections, the use of restriction endonuclease analysis to determine the sequence divergence between species was discussed. This method, however, does not detect all randomly distributed differences in DNA sequences. Nucleotide sequencing techniques allow the detection of single nucleotide changes and also provide quantitative information about the overall organization of the genome. At the same time, many questions about the evolution of species, how closely different species are related to one another, and the amount of variation within species can be answered through sequencing of the mtDNA. Sequencing data from mtDNA also allowed the presentation of a new archigenetic hypothesis on the evolution of mitochondria (Mikelsaar, 1987).

To date, complete mtDNA sequences are known for humans (Anderson *et al.*, 1981; Attardi *et al.*, 1982), cow (Anderson *et al.*, 1982), mouse (Bibb *et al.*, 1981), *Xenopus laevis* (Roe *et al.*, 1985), *Drosophila yakuba* (Clary and Wolstenholme, 1985), and yeast *Saccharomyces cerevisiae* (de Zamaroczy and Bernardi, 1986). All of the information about the coding capacity and gene arrangement of the mtDNA was

obtained through sequencing. Sequencing of the mtDNA also revealed the mode of replication, transcription and translation of the mtDNA. This in turn led to the discovery that the mitochondrial code differs in many ways from the universal genetic code, and that mitochondrial tRNAs and rRNAs differ from nuclear tRNAs and rRNAs. Discoveries made through sequencing mtDNA are reviewed in the following section.

5.1.2 Replication of the mtDNA

Mitochondria have a separate, autonomously replicating DNA genome, and replication of the mtDNA is unidirectional, in contrast to bacteria, which have bidirectional replication (Metzler *et al.*, 1977, Anderson *et al.*, 1981; Van Etten *et al.*, 1982; Attardi *et al.*, 1982). The origin of replication is situated within a Displacement (D) loop region (Watson *et al.*, 1987; Anderson *et al.*, 1981). These D-loops were first observed by visualizing the mtDNA with the aid of electron microscopy, and it was thought that these structures arose due to low GC content in that region. Later on, it was found that initiation of synthesis of the heavy (H) strand, results in the formation of the D-loop, with a newly synthesized H strand of $\pm$ 680 bases, known as the 7S DNA (14S in amphibians), contained within the D-loop. The D-loop region is maintained as a triple stranded structure by repeated synthesis and degradation of this 7S

DNA, and displacement of the parental H-strand as a single-stranded loop. It was also found that a large proportion of the mtDNA exists in the D-loop form, but only few of the nascent H-strands are synthesized to full length (Anderson *et al.*, 1981; Bibb *et al.*, 1981; Attardi *et al.*, 1982; van Etten *et al.*, 1982).

Replication of the mammalian mtDNA proceeds after initiation of the H-strand synthesis, at the origin of replication. Once the replication fork has moved about two-thirds of the way around the genome, and has exposed the origin of replication of the light (L) strand, replication of the L-strand commences (Watson *et al.*, 1987; Anderson *et al.*, 1981; Attardi *et al.*, 1982; van Etten *et al.*, 1982). The DNA sequence surrounding the L-strand origin is an inverted repeat that may form a hairpin structure once the origin is exposed (Bibb *et al.*, 1981). This secondary structure may play a role in initiation of replication of the L-strand (van Etten *et al.*, 1982).

Replication of the mitochondrial genome is catalyzed by the γDNA polymerase, which occurs only in the mitochondrion. In addition, topoisomerase activity is also located within the mitochondrion, catalyzing the nicking and closing of the mtDNA, which results in untwisting of the parental strands, making the separation of the daughter double helixes possible (Watson *et al.*, 1987). A type-II topoisomerase, which is unable to supercoil DNA, but which can catenate,

decatenate, knot, and relax DNA, has also been reported to be present in the mitochondrion, and may form part of a gyrase (Castora *et al.*, 1982).

No open reading frames of any significant size could be assigned to the D-loop region and it is generally accepted that this region has no coding function (Anderson *et al.*, 1981; 1982; Bibb *et al.*, 1981; van Etten *et al.*, 1982; Attardi *et al.*, 1982). Very little sequence homology of the D-loop region could be detected between rat, human, and bovine mtDNA, and the size of this region also seems to vary much more than other intergenic spaces on the mtDNA (Anderson *et al.*, 1981).

5.1.3 Organization of the mitochondrial genome

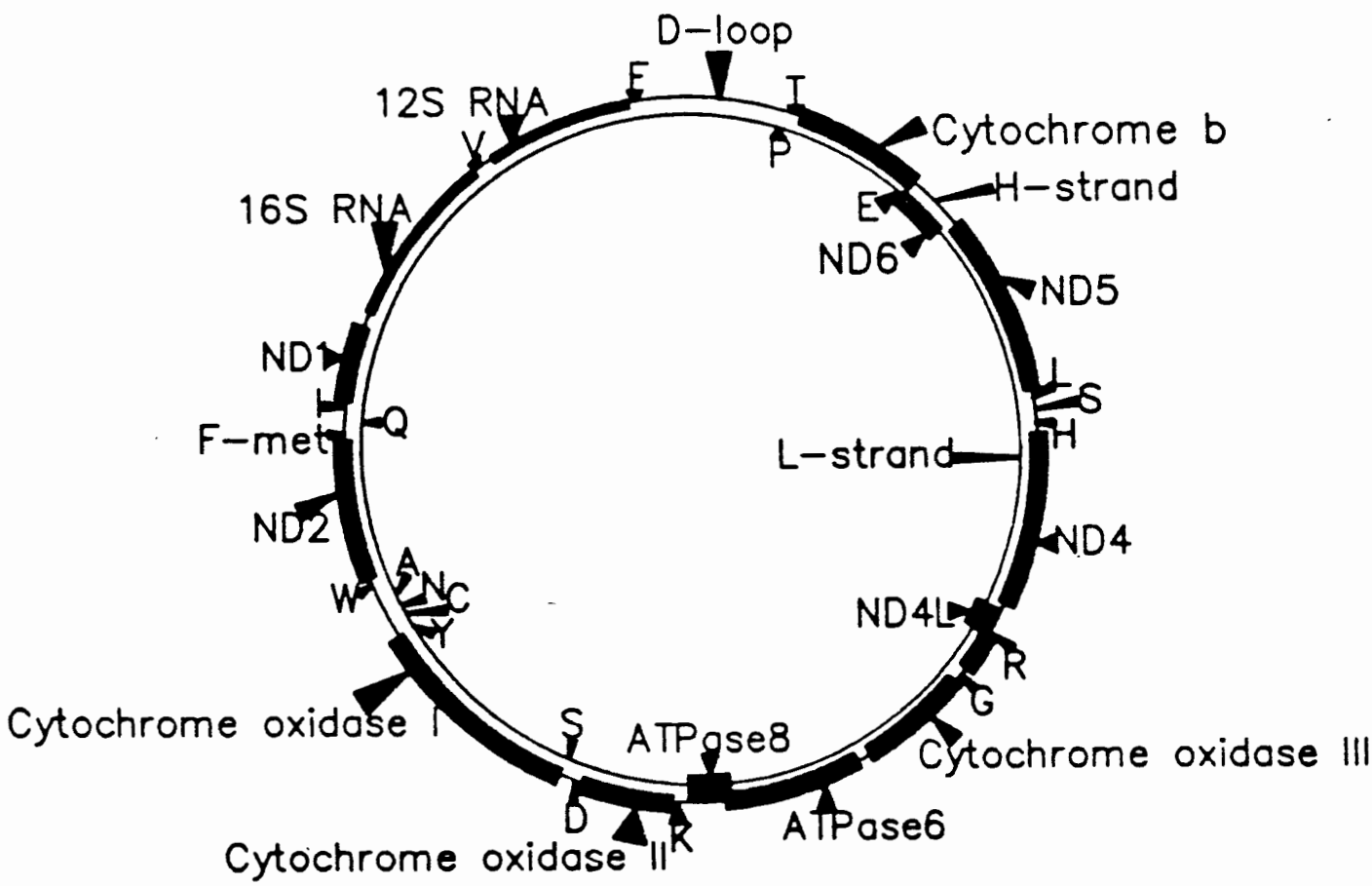
By means of the complete mtDNA sequences known to date, various genes have been located on the mitochondrial genome. The organization of the genes within the mtDNA shows extreme economy, as there are virtually no non-coding bases between adjacent genes and none of the mitochondrial genes contain introns. The D-loop region is the only stretch of DNA to which no coding function can be assigned (Anderson *et al.*, 1981). The remaining genome basically reflects a continuous coding function, with tRNAs interspersed between the rRNA genes and the protein encoding genes (Bibb *et al.*, 1981; Attardi *et al.*, 1982; van Etten *et al.*, 1982).

The following genes have been identified on the mitochondrial genome of higher eucaryotes: The 12S and 16S ribosomal RNAs (rRNA), 22 transfer RNAs (tRNA), and several proteins, which are components of the mitochondrial inner membrane (Watson *et al.*, 1987; Anderson *et al.*, 1981, 1982; Bibb *et al.*, 1981; Roe *et al.*, 1985; Clary and Wolstenholme 1986). The mitochondrial genomes of higher eucaryotes contain 13 long open reading frames, five of which have been identified to code for the cytochrome c oxidase subunit I, II, and III (COI, COII, COIII), cytochrome b, and ATPase subunit 6 (Anderson *et al.*, 1981, 1982; Wong *et al.*, 1983; Avise and Lansman, 1983). Eight unidentified reading frames (URFs) have been reported, seven of which are encoded by the H-strand, and one (URF6), which is encoded by the L-strand (Attardi *et al.*, 1982). Three pairs of these reading frames, URFA6L and ATPase6, URF4L and URF4, and URF5 and URF6, were found to exist as overlapping reading frames in the mouse mtDNA sequence (van Etten *et al.*, 1982). It is generally accepted that these unidentified reading frames represent genes for mitochondrial proteins. Poly-adenylated mRNAs, corresponding directly to the reading frames (and antisense to URF6), were identified (Anderson *et al.*, 1981, 1982; Attardi *et al.*, 1982). At the same time, an average 70% nucleotide and amino acid homology was found between the human and bovine URFs, which is unlikely to have occurred by chance (Anderson *et al.*, 1981, 1982). More recently, it has been shown that the URFA6L has sequence homology, and

similarity in transcriptional control, to the ATPase8 (Clary and Wolstenholme, 1986). Chomyn *et al.* (1985) then reported that the products of URF1-5 and URF4L were precipitated from a mitochondrial lysate by antibodies against highly purified bovine NADH-ubiquinone oxidoreductase (complex I) and they concluded that these six reading frames code for components of the respiratory chain NADH dehydrogenase and are now designated ND1-5 and ND4L. Even later Chomyn *et al.* (1986) found that URF6 also codes for an NADH dehydrogenase subunit. A map representing the organization of the mammalian mitochondrial genome is shown in Fig. 5.1.1 (from Watson *et al.*, 1987). The gene order is the same for human, bovine, mouse, amphibian and invertebrate mtDNA, and in general, the mtDNA of higher eucaryotes is highly conserved in size, function, and organization (Avise and Lansman, 1983; Brown, 1980). The mtDNA of lower eucaryotes (*Saccharomyces cerevisiae*, *Neurospora*, and *Paramecium*), especially yeast, has also been studied extensively, and the functions of their mtDNA are very similar to those reported for higher eucaryotes (Watson *et al.*, 1987; de Zamaroczy and Bernardi, 1986). However, the size of the mtDNA is not conserved. For example, it is reported that the size of the yeast mtDNA is five times larger than the human mtDNA, and that over two thirds of yeast mtDNA consists of non-coding sequences (Watson *et al.*, 1987; de Zamaroczy and Bernardi, 1986). Another example is

Figure 5.1.1

Map of the mammalian mitochondrial genome (Adapted from Watson *et al.*, 1987). The tRNAs are indicated by their one letter code. ATPase6 and 8 are components of the mitochondrial ATPase complex. ND1-6 code for components of the respiratory chain NADH dehydrogenase. L-strand is the light strand and H-strand is the heavy strand.



of the plant mtDNA, in which a range of sizes (1.5 - 30 kb) of mtDNA were found within single plants (Dale, 1982; Bendich, 1982).

5.1.4 The mitochondrial genetic code

The genetic code of mtDNA differs from the universal code. The fact that only 22 tRNAs exist in the mitochondria, means that more "wobble" is permitted than is the case with the 32 nuclear coded tRNAs. In the mitochondrion, a single amino acid is specified by four codons with the same first and second positions (e.g., the four valine codons GUU, GUC, GUA, GUG). A single tRNA is sufficient to read all four codons, because wobble permits pairing of any third codon base with the U that resides in the first position of the tRNA anticodon. The amino acids that are specified by pairs of codons with pyrimidines in their third positions are read by G in the tRNA anticodon, while pairs with purines in the third position are read by U in the anticodon of the tRNA (Watson *et al.*, 1987). The genetic code for mtDNA is presented in Table 5.1.1. The mitochondrial genetic code varies from the universal code in that UGA codes for tryptophan and not termination, AUA codes for methionine not isoleucine, and AGA and AGG are termination, rather than arginine codons. Thus, there are four stop codons (UAA, UAG, AGA, and AGG) in the mitochondrial code. Internal methionine is encoded by both AUG and AUA, and initiation

Table 5.1.1

Mitochondrial genetic code

Phe	UUU	Ser	UCU	Tyr	UAU	Cys	UGU
GAA	UUC	UGA	UCC	GUA	UAC	GCA	UGC
Leu	UUA		UCA	Stop	UAA	Trp	UGA
UAA	UUG		UCG		UAG	UCA	UGG
Leu	CUU	Pro	CCU	His	CAU	Arg	CGU
UAG	CUC	UGG	CCC	GUG	CAC	UCG	CGC
	CUA		CCA	Gln	CAA		CGA
	CUG		CCG	UUG	CAG		CGG
Ile	AUU	Thr	ACU	Asn	AAU	Ser	AGU
GAU	AUC	UGU	ACC	GUU	AAC	GCU	AGC
Met	AUA		ACA	Lys	AAA	Stop	AGA
CAU	AUG		ACG	UUU	AAG		AGG
Val	GUU	Ala	GCU	Asp	GAU	Gly	GGU
UAC	GUC	UGC	GCC	GUC	GAC	UCC	GGC
	GUA		GCA	Glu	GAA		GGA
	GUG		GCG	UUC	GAG		GGG

Differences between the mitochondrial and universal genetic code are shown by shading. The tRNA anticodon is written (5′-3′) in bold.

Boxes of four codons each are read by a single tRNA with U in the first position of the anticodon. Boxes of two codons with U/C (pyrimidines) in the third position are read by G in the first position of the anticodon of the tRNA, and codons with A/G (purines) in the third position are read by tRNAs with U in the first position of the anticodon. Only the C in the first position of the anticodon of the Methionine tRNA, engages in unusual pairing.

methionines are specified by AUG, AUA, AUU, and AUC. The mitochondrial genetic code seems to be the same for all the higher eucaryotes.

5.1.5 Transcription of the mtDNA

Both the H- and L-strands of the mtDNA are transcribed completely, starting from promotors situated in the D-loop region (Watson *et al.*, 1987; Anderson *et al.*, 1981). In the human mtDNA, transcription of the H-strand starts in the D-loop region, about 15 nucleotides upstream of the $tRNA^{Phe}$ gene (Fig. 5.1.2). The L-strand is defined as the main coding strand in the mtDNA, as it contains the sense sequence of the rRNAs and of most tRNAs and mRNAs. These RNAs are transcribed from the H-strand, and therefore the transcripts are of L-strand sequence and complementary to the H-strand (Anderson *et al.*, 1981, 1982; Bibb *et al.*, 1981). Only a few tRNAs and the URF6 are H-strand coded.

Transcription gives rise to very long RNA molecules, which are then processed to give the 12S and 16S rRNAs, tRNAs and a number of messenger RNAs (mRNA), which are not capped, but are polyadenylated. These mRNAs are then translated into the corresponding proteins. The open reading frames and the rRNA genes are butt-jointed at their 3' termini to a tRNA gene, and the large precursor mRNAs are processed by the precise excision of the tRNA sequences. Initiator

codons lie directly at, or very close to the 5′ end of these processed mRNAs. Montoya *et al*. (1981), for instance, showed that the 5′ end of the COII mRNA corresponded precisely with the first nucleotide of the coding sequence, thereby confirming the lack of leader sequences on the 5′ end of mitochondrial mRNAs. They also aligned all the poly-A containing RNA sequences with the mtDNA sequences, and found that all putative mitochondrial mRNAs have initiation codons within the first six nucleotides. All mRNAs (procaryote and eucaryote) have a stretch of sequence preceding the initiation codon and it was thought that this area was the ribosome binding site. Montoya *et al*. (1981) proposed that in the mitochondrial mRNAs the ribosomes attach to the 5′ end, with recognition of the initiation codon, and that internal AUG sites could be hidden by the secondary structure of the mRNA. In that way, the absence of 5′ leader sequences in the mitochondrial mRNAs, yet retention of efficient initiation, could be explained.

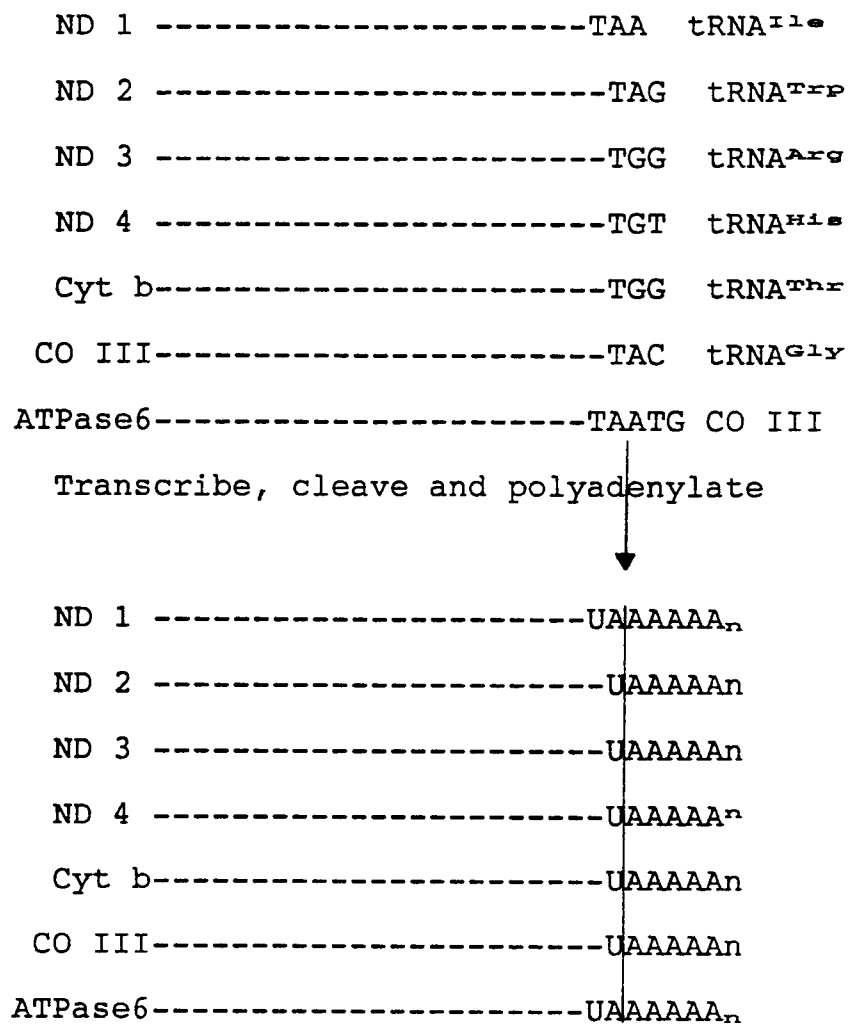
The coding regions run all the way to the 3′ end of the mRNA and, generally, no termination codons were found on the DNA. Anderson *et al*. (1981) proposed a model of post-transcriptional creation of UAA termination codons in the mRNAs. The model was proposed because of the observation that when tRNAs are cleaved precisely out of the primary transcript, by endonucleolytic processing (by RNase P type enzymes, see Watson *et al*., 1987 for review on RNA processing), the reading frames with no termination codons

are left with a U or UA at their 3′ end, in phase with the reading frame. By polyadenylation of these mRNAs a UAA stop codon is created in phase with the reading frame, as is demonstrated in Fig. 5.1.2 (Watson *et al.*, 1987; Anderson *et al.*, 1981; Ojala *et al.*, 1981).

This model also provides a termination codon for ATPase6, which overlaps with COIII, so that they have overlapping termination and initiation codons in the UAAUG sequence. The cleavage site is between the two A residues so the termination site is destroyed in the process of primary RNA cleavage, but the polyadenylation model explains the formation of a new termination codon. The overlap between ATPase6 and COIII shows that the primary RNA transcript is processed in the same manner as in the case where a tRNA abuts a coding region. The signal for processing of the primary transcript remains a puzzle, as it was assumed that the functional role of the tRNA sequences, relative to their position in the mtDNA, was to serve as recognition signals for RNA processing (Anderson *et al.*, 1981; Attardi *et al.*, 1982). One plausible model proposes that it is possible for the tRNA sequences to acquire the cloverleaf structure while they are still part of the nascent RNA transcript, and that this structure may be recognised by the RNaseP-like enzyme, resulting in the precise excision of the tRNAs (Ojala *et al.*, 1981; Attardi *et al.*, 1982).

Figure 5.1.2

Processing and polyadenylation model of primary RNA transcripts to generate translational termination codons at the ends of the coding regions for mitochondrial proteins (Anderson *et al.*, 1981). The top represents the DNA sequences, indicating how the coding regions abut tRNAs or other coding regions. The bottom shows the resulting RNA sequences that are created by site-specific cleavage of the primary transcripts, at the vertical line, and the addition of A residues.



5.1.6 The mitochondrial tRNAs

The mitochondrial tRNAs not only obey unusual wobble rules, but the sequences of the mitochondrial tRNAs are also unusual (Anderson *et al.*, 1981). Hence, it is expected that the mitochondrial tRNAs will differ from nuclear tRNAs, both in their three-dimensional structure and in their interaction with the mitochondrial ribosomes (Watson *et al.*, 1987). It also seems that the mitochondrial tRNAs are stabilized by fewer tertiary interactions. Not surprisingly, it was found that the mitochondrial tRNAs have bizarre structures which are different to the cloverleaf structures of non-mitochondrial tRNAs; an extreme case is one of the serine tRNAs, which completely lacks the D arm of the usual cloverleaf structure (Anderson *et al.*, 1981, 1982; Wong *et al.*, 1983).

5.1.7 The mitochondrial rRNA genes

Both the 12S rRNA and the 16S rRNA genes are flanked by tRNAs (Fig. 5.1.1). Thus, the rRNAs are released by the precise excision of the tRNAs from the primary transcript. Mitochondrial rRNAs are distinctive in that they are posttranscriptionally polyadenylated (Dubin *et al.*, 1982). It appears that all ribosomes are constructed in a similar way and that the conserved regions are necessary for the function of the ribosomes, whereas the remainder is free to

change (Watson *et al.*, 1987; Anderson *et al.*, 1982). The exact function of the rRNAs is, however, still unknown (Dubin *et al.*, 1982).

5.1.8 Rate of evolution of mitochondrial DNA

Information about the evolution of the mtDNA was obtained by direct comparisons of mtDNA sequences. Anderson *et al.* (1982) compared the human and bovine mitochondrial sequences of three genes (COII, ND5, URF6), for each of the three codon positions. They found that the average rate, over all three codon positions, was 0.003 substitutions per base pair per million years. This rate of evolution is lower than the one estimated by Brown *et al.* (1979), but it nevertheless confirms that the mtDNA has an extremely high rate of evolution. These results also demonstrated that the third codon position showed the highest rate of change, followed by the first and then the second position. This is expected, because of the degeneracy of the genetic code.

It is also expected that the rate of sequence divergence on the mitochondrial genome will vary from site to site. Bibb *et al.*, (1981) for instance, compared the mouse mtDNA sequence to the human sequence and found a 74% nucleotide conservation for protein-coding genes and an 82% amino acid conservation, showing that the rate of silent base substitutions is higher than the rate of base substitutions

leading to amino acid replacements. Brown and Simpson (1982), confirmed this observation. They determined the sequence of the cytochrome oxidase subunit II, of two closely related rat species, *Rattus norvegicus* and *R. rattus*, and found that 94.4% of the nucleotide substitutions were silent. As a result of this high proportion of silent substitutions, they then proposed that the rapid evolution of mtDNA, relative to nuclear DNA, was due to silent changes, whereas the rate of amino acid altering substitutions accumulate at comparable rates in the mitochondrial and nuclear DNA (Brown and Simpson, 1982).

The amino acid conservation between human and bovine mitochondrial proteins also showed that uncharged and non-polar amino acids were most frequently interchangeable, whereas the charged amino acids were more conserved (Anderson *et al.*, 1982).

In another study, it was reported that there is a strong bias in favor of transitions (purines to purines, A:G, or pyrimidines to pyrimidines, C:T) among individuals within a species, but individuals from divergent species (between mouse and cow) show a smaller bias towards transitions (Aquadro *et al.*, 1984).

Anderson *et al.* (1981, 1982) also determined that the average nucleotide conservation among tRNAs was 80%, and they were found to be more conserved than either the rRNA genes or the protein coding genes. However, this

conservation was much less than that found in non-mitochondrial tRNAs (Anderson *et al.*, 1981, 1982; Bibb *et al.*, 1981; Avise, 1986). This, again, reflects the rapid rate of evolution of the mtDNA, relative to nuclear DNA.

The D-loop region contains some areas that show variable, and in some species, extremely rapid rates of nucleotide divergence (Bibb *et al.*, 1981; Avise, 1986). There are regions within this highly divergent area, though, that are extremely conserved, and which might be involved in the initiation of the H-strand synthesis (Anderson *et al.*, 1982).

5.1.9 Objective of this study

In this section of the work, parts of the hake mitochondrial clones, obtained previously from both *Merluccius capensis* and *M. paradoxus*, were cloned into suitable sequencing vectors. Then both ends of each sub-clone were sequenced. The sequences obtained in this way were identified by aligning them with mtDNA sequences of bovine (Anderson *et al.*, 1982), *Xenopus laevis* (Roe *et al.*, 1985), and *Drosophila yakuba* (Clary and Wolstenholme, 1985). These sequences were used to identify specific genes in hake mtDNA. Once the genes were identified, the genetic code and structure of the molecules were determined. The Genbank database from August 1986 was used in this study and the

only other fish mitochondrial DNA sequence was that of carp, *Cyprinus carpio* (Araya, 1984). The mitochondrial sequences from carp were parts of the cytochrome oxidase I, glutamic acid tRNA, part of the cytochrome b, and URF4; none of these sequences overlapped with the hake mitochondrial sequences.

5.2 MATERIALS AND METHODS

5.2.1 Subcloning of mtDNA into sequencing vectors

Cloned mtDNA was amplified in the the pEcoR 252 vector and purified by CsCl/EtBr centrifugation. This DNA was digested with appropriate restriction endonucleases and ligated into puc 19 which had been digested to completion with the same endonucleases. The recombinant plasmids were transformed into *E. coli* JM109 cells Maniatis *et al.*, 1982). Transformed cells were selected for on Amp, Xgal (Appendix) plates. The plasmid was prepared as described (Section 4.2.3), DNA was digested, run on a gel, and southern blotted. By using mtDNA as a probe it was determined whether or not a plasmid contained mtDNA. The selected plasmids were purified by two successive CsCl/EtBr centrifugation steps. The DNA thus prepared was then used for double stranded sequencing.

5.2.2 Sequencing reactions

All DNA sequences were determined using the Sanger dideoxy chain termination procedure (Sanger *et al.*, 1977) as modified by Chen and Seeburg (1985).

5.2.2.1 Double stranded DNA priming

Two µg of DNA was mixed with 2 µℓ of 2M NaOH and incubated at room temperature for 5 min. To this, 3 µℓ 3 M NaAc pH 4.5, 2µℓ H_2O, 5 µℓ Primer (4.4 µg/µℓ), and 75 µℓ 96% EtOH was added. M13 forward sequencing primer and reverse sequencing primer (Amersham) were used. The mixture was placed on ice for 10 min, centrifuged for 30 min and then resuspended in TE buffer. The DNA was reprecipitated with 96% EtOH, washed with 70% EtOH, and dried. It was then resuspended in 12 µℓ of 1x sequencing buffer (Appendix) to which 1 µℓ Klenow polymerase (4U/µℓ), and 1µℓ (10 µCi) of ^{35}S dATP was added. The mixture was split into four aliquots which were allowed to react with appropriate deoxynucleotide/dideoxynucleotide mixes (Boehringer, Mannheim). This chain termination reaction continued for 20 min at 37°C. Each reaction was then saturated with 1.3 µℓ chase solution (Appendix) and allowed to react for a further 15 min, after which the reaction was stopped by the addition of 3 µℓ of sequencing stop buffer (Appendix). The mixtures were heated to 80°C for 3 min and run on denaturing polyacrylamide gels (Sanger *et al.*, 1977).

5.2.3 Polyacrylamide gel electrophoresis

A 6% (34 x 40 cm x 0.25 mm) polyacrylamide gel was cast in a BRL model SO apparatus using 0.2 mm spacers. A sharktooth comb was used (BRL 1045 SC, 24 teeth, 6mm point-to-point). Electrophoresis took place at 35 mA for 4-6 h. After electrophoresis the gels were dried on a slab gel dryer (Hoeffer, Model 1125B) and the gels were placed under Kodak XAR-5 X-ray film for between 20h and 4 days.

5.2.4 Analysis of sequences

The DNA sequences were read, then typed into a data file and analyzed using an IBM XT computer. The sequences were analyzed using two computer programs, Genepro (Riverside Scientific, Seattle) and Microgenie (Beckmann). Sequences were compared to all available mtDNA sequences on the genebank from August 1986, to determine open reading frames and coding areas (tRNAs).

5.3 RESULTS

5.3.1 Subcloning of mtDNA into puc 19

The following mtDNA fragments were cloned into puc 19:

1. A 1.4 kb *Hin*dIII- *Sph*I fragment from the *M. capensis* clone pIIB 201, and this subclone was named pIIB 1211.

2. The 2 kb *Hin*dIII- *Hin*dIII fragment of *M. paradoxus* (pIIB 102), which was then called pIIB 1102.

3. A small *Hin*dIII -*Hin*dIII fragment from *M. paradoxus*. This subclone was 400 bp in size and was named pIIB 1122.

4. The fourth subclone was named pIIB 1134, and it was a *Bam*HI - *Sph*I fragment of 1.8 kb from *M. paradoxus* (pIIB 104).

The subcloned fragments were identified by cutting them out of the sequencing vector and running the DNA on gels adjacent to the clones that they originated from, cut with the same enzymes.

5.3.2 Mitochondrial DNA sequences from M. capensis and M. paradoxus

DNA from these four subclones was prepared, purified, and used for double stranded DNA sequencing. Both ends of the

subclones were sequenced using the double stranded sequencing method. The subclone pIIB 1122 was sequenced completely. The mtDNA sequences obtained from both species of hake were aligned with known mtDNA sequences from other organisms. The following genes were identified on the mtDNA of *M. capensis* adjacent to the *Hin*dIII site of pIIB 1211: ND1, isoleucine tRNA, glutamine tRNA, and f-methionine tRNA (Fig 5.3.1). Genes identified on the other side of the fragment were ND2, tryptophan tRNA, and half of the alanine tRNA (Fig 5.3.2).

The following genes were identified on the mtDNA subclones of *M. paradoxus* : cytochrome oxidase subunit I, cytochrome oxidase subunit II, lysine tRNA, ATPase8, and ATPase6. The DNA sequence of the small *Hin*dIII fragment, pIIB 1122, contained sequences of cytochrome oxidase subunit II, lysine tRNA, ATPase8, and ATPase6 (Fig. 5.3.3). The DNA sequence of cytochrome oxidase subunit I was located on the other end of the large *Hin*dIII fragment, pIIB 1102 (Fig. 5.3.4). The 16S ribosomal sequence was located on the fragment pIIB 1134 (Fig. 5.3.5).

The sequences were identified by aligning them with mitochondrial sequences of mammals, (bovine, Anderson *et al.*, 1982); amphibians (*Xenopus laevis*, Roe *et al.*, 1985); and insects (*Drosophila yakuba*, Clary and Wolstenholme *et al.*, 1985). The sequences are presented in Fig. 5.3.1 - 5.3.5, where the nucleotides typed in bold are in common

with the nucleotide in the hake mtDNA. The amino acid sequence of the coding regions were inferred. Here, the amino acids in common with either *M. capensis* or *M. paradoxus*, were typed in **bold**, and conserved amino acids were underlined. The change from one conserved amino acid to another will influence the charge or conformation of the resulting protein only minimally.

The percentage of DNA and amino acid homology, between the taxa is presented in table 5.3.1, and the predicted secondary structures of the lysine tRNA, tryptophan tRNA, isoleucine tRNA, glutamine tRNA, and f-methionine tRNA, were prepared by R. Kirby and are presented in Fig. 5.4.1.

Fig. 5.3.1

Comparison of *Merluccius capensis* ND 1, Isoleucine tRNA, Glutamine tRNA, and f-Methionine tRNA genes with the same genes of *Xenopus laevis* (X. laev), bovine (Bov.), and *Drosophila yakuba*. The sequences are shown aligned, with spaces introduced to maximise the homology. Nucleotides and amino acids conserved between *M. capensis* and the other species, are typed in **bold**, amino acids underlined are considered to be conserved. Asterisks represent the termination codon.

```
          ND 1---------------------------------------------------->
            LeuValIleTrpHisLeuSerLeuPro   SerIleSerGlySerProProAsnPro*
M. cap.   GCCTGGTAATTTGGCACCTCTCACTACCA---TCCATTAGCGGGTCCCCTCCCAATCCCT
            IleThrLeuTrpHisIleSerLeuProIleSerMetLeuGlyLeuProSerGlnThr
X. laev.  CCATAACATTATGACATATCTCATTACCAATTTCTATGCTAGGCCTACCATCACAAACC-
            LeuCysMetTrpHisValSerLeuProIleLeuThrSerGlyIleProProGlnThr
Bov.      CCCTGTGCATGTGACACGTATCCCTACCCATCCTTACATCAGGCATCCCACCACAAACA-
            IleLeuLeuPheSerPheLeuLeuTrpIlePhePheSerLysLysLeuMetGluAsn*
D. yakuba AAATTTTATTATTTTCTTTTTTATTGTGAATTTTTTTTAGTAAAAAGTTAATAGAAAATT

          **    <Ile tRNA----------------------ACD------------------
M. cap    AGCAAGGGAG--TG-GCCCTGAATTAAGGGGCCACTTTGATAGAGTGTAATATAAGGGTT
               ***
X. laev   -----TAGGAAATGTGCCCGAAAGTCAGGGATCACTTTGATAGAGTGAAATATATGGGTT
               ***
Bov.      -----TAAGAAATATGTC-TGACAAAAGAGTT-ACTTTGATAGAGTAAATAATAGAGGTT
               **
D. yakuba AA----AATGAAT-TGC-CTGATAAAAAGGGTTACCTTGATAGGGTAAATTATGCAGTTT
```

Figure 5.3.1 (continued)

```
           ------------------<>-----------------------------------ACD----
M. cap.    AAATTCCCTTCCACTCCTTAGAAAAGGGTTAAG-GAACCAT--CTTAG-AGATCAATTCT

X. laev.   CAAACCCCATCATCTCCTTAGAAAGACAGGAATTGAACCTGCACCTGAGAGATCAAAACC

Bov.       CAAACCC--TCTTATTTCTAGAACTATAGGAATCGAACCTACTCCTAAGAATCCAAAACT

D. yakuba ---TCTGCATTCATTGTT-ATATTTAATAGAATTAAACTATTTCTAAAAGTATCAAAAAC


           ----------------------Gln tRNA>       <f-Met tRNA-----------
M. cap     CTATCGTTCACTAAC-TAT----CTTCCTAG--G--AGGTCAGCTAATTAAGCCTTTGGG

X. laev.   -CTCCGTACTCCCAC-TATACCACTTCCTA---GTAAAGTCAGCTAAAAAAGCTTTTGGG

Bov.       TCTTCGTGCTCCCAATTACACCAAATTCTATTAGTAAGGTCAGCTAATTAAGCTATCGGG

D. yakuba -TTTTGTGCAT-CA-T--CACCAAAATATATTTATTATAAAAAGA---TAAGCTACTGGG


           --ACD---------------------------------------->
M. cap.    CCCATACCCCG---ATGT-GGTAACTCCCCTTAGCTATGGTCTA

X. laev.   CCCATACCCCAAACATGTTGGTTAAACCCCTT--CCTTT-ACTA

Bov.       CCCATACCCCGAAAATGTTGGTTTATATCCTT--CCCGT-ACTA

D. yakuba TTCATACCCCATTTATAAAGGTTATAATCCTT--TTCTT-TTTA
```

Fig. 5.3.2

Comparison of ND2, Tryptophan tRNA, and Alanine tRNA of *M. capensis* and other mitochondrial genomes, where M. cap = *Merluccius capensis*, X. laev. = *Xenopus laevis*, Bov. = bovine, and *Drosophila yakuba*. The genes are shown aligned, with spaces introduced to maximise the homology. Nucleotides and amino acids in common with *M. capensis* are typed in **bold**, amino acids underlined are considered conserved, and asterisks represent termination codons.

```
          -ND2-------------------------------------------------------
          PheThrLysGlu---PheLeuLeuProLeuThrAlaSerThrMetIleThrThrMet
M. cap.   TTCACAAAAGAG---TTTCTTCTCCCCCTAACAGCCTCCACTATAATTACCACAATG
          HisSerLysGln---ProThrLeuLeuLeuSerIleAlaLeuIleLeuSerSerPhe
X. laev.  CACTCTAAACAA---CCAACACTTTTATTATCAATCGCATTAATCCTATCCTCATTT
          MetLysLysMetThrPheLeu---ProThrMet------ValValLeuSerThrMet
Bov.      ATGAAAAAAATAACTTTTCTA---CCAACAATA------GTCGTATTATCTACCATA
          AspSerAspAspThrAspLeuTyrLeuIleMet------ThrPhePheSerIlePhe
D. yakuba AATAGTAATAATACTAATTTATATTTAATTATA------ACTTTTTTTTCAATTTTC


          -ND2------------------------------------>
          LeuLeuProIleAlaProAlaThrValIleLeuMetSerGly*--<*Tryptophan
M. cap.   CTCCTCCCCATCGCTCCAGCAACTGTAATTTTAATGTCTGGCT--AGGGGCTTAGAT
          IleIleProIleSerPro------LeuThrLeu---Thr---***
X. laev.  ATTATTCCAATTTCACCA------TTAACTTTA---ACA---TAAAGAGATTTAAGT
          MetLeuProLeuThrProMetLeuSerValLeuGlu------*--**
Bov.      ATACTACCACTCACGCCAATACTATCAGTGTTAGAA------T--AGGAATTTAGGT
          GlyLeuPheLeuIleSer------LeuPhePhePheMetLeu***
D. yakuba GGATTATTTTTAATTTCT------TTATTTTTTTTTATACTTTAAAAGGCTTTAAGT
```

Fig. 5.3.2 (continued)

```
           tRNA--------------ACD-----------------------------------
M. cap     GAACCAGACCAAGAGCCTTCAAAGCTCTCAGC--GAGTGAGAAC-CCTCCAGCCCCT

X. laev.   TAACAAGACTAAGAGCCTTCAAAGCCCTAAGCAGGAGTTAGAATCTCCT-AATCTCT

Bov.       TAAACAGACCAAGAGCCTTCAAAGCCCTAAGCAAG---TACAATTTACTTAATTCCT

D. yakuba TAACTAAACTAATA-CCTTCAAAGCTGTAAATAA---AGGGTATTCCTTTAAGTCTT


           > <Alanine tRNA--------------------------------ACD---------
M. cap.    G-ATAAAA-TCTGCGGGA-TACTACCCA-CATCTTCTCAT--GCCTGCAAGTC

X. laev.   G-ATAAGG-CTTGCAGGA-TTTTATCCAACATCA----AT-TGAATGCAACTC

Bov.       G-ATAAGG-ATTGCAAGACTAC-ACCTTACATCA----AT-TGAATGCAAATC

D. yakuba AGGGTTGTAGTT--AA---TTATAACAT---TTG----ATTTGCATTCAAAAA
```

Fig. 5.3.3

Comparison of the Cytochrome oxidase subunit II, Lysine tRNA, ATPase8 and ATPase6 sequences of *Merluccius paradoxus* with sequences from *Xenopus laevis* (X. laev.), bovine (Bov.), and *Drosophila yakuba*. The sequences were aligned and spaces were introduced to maximise homology. Nucleotides and amino acids typed in bold were conserved between *M. paradoxus* and the relevant species. Asterisks represent termination codons.

```
          -Cytochrome oxidase subunit II------------------------------
          AlaSerProThrThrLeuAsnGlnThrAlaPheIleLeuProAspProAlaValTyrTyr
M. para.  GCAAGCCCGACGACTCTAAACCAAACAGCCTTCATTCTGCCCGACCCAGCAGTATACTAC
          AlaIleProGlyArgLeuHisGlnThrSerPheIleAlaThrArgProGlyValPheTyr
X. laev.  GCAATCCCAGGACGACTTCATCAAACATCATTTATTGCTACTCGTCCGGGAGTATTTTAC
          AlaIleProGlyArgLeuAsnGlnThrThrLeuMetSerSerArgProGlyLeuTyrTyr
Bov.      GCAATCCCAGGCCGTCTAAACCAAACAACCCTTATATCGTCCCGTCCAGGCTTATATTAC
          GlyThrProGlyArgLeuAsnGlnThrAsnPhePheIleAsnArgProGlyLeuPheTyr
D. yakuba GGAACTCCTGGACGATTAAATCAAACTAATTTTTTTATTAACCGACCAGGGTTATTTTAT

          ------------------------------------------------------------
          GlyGlnCysSerGluIleCysAspArgIleHisSerPheMetProIleValValGluAla
M. para.  GGACAATGTTCAGAAATTTGCGATCGAATCCACAGCTTCATACCCATCGTCGTCGAAGCC
          GlyGlnCysSerGluIleCysGluAlaAsnHisSerPheIleProIleValValGluAla
X. laev.  GGACAATGTTCAGAAATTTGCGAGGCAAACCACAGCTTTATACCAATTGTAGTTGAAGCA
          GlyGlnCysSerGluIleCysGlySerAsnHisSerPheMetProIleValLeuGluLeu
Bov.      GGTCAATGCTCAGAAATTTGCGGGTCAAACCACAGTTTCATACCCATTGTCCTTGAGTTA
          GlyGlnCysSerGluIleCysGlyAlaAsnHisSerPheMetProIleValIleGluSer
D. yakuba GGTCAATGTTCAGAAATTTGCGGGGCTAATCATAGTTTTATGCCAATTGTAATTGAAAGT
```

Fig. 5.3.3 (continued)

```
           ------------------------------------------------>>
           ValProLeuGlnHisPheGluSerTrpSerSerSerMetLeuGluAspAla*<-Lysine
M. para.   GTTCCCCTACAACACTTCGAATCTTGATCCTCTTCAATACTTGAAGACGCCTCATTAACA
           ValProLeuThrAspPheGluAsnTrpSerSerSerMetLeuGlu---Ala*
X. laev.   GTACCGCTAACCGACTTTGAAAACTGATCTTCATCAATACTTGAA---GCATCACTAAGA
           ValProLeuLysTyrPheGluLysTrpSerAlaSerMetLeu***
Bov.       GTCCCACTAAAGTACTTTGAAAAATGATCTGCGTCAATATTATAAAAT----CACTAAGA
           ValProValAspAspPheIleLysTrpIleSerSerAsnAsnSer*
D. yakuba  GTTCCTGTAAATAATTTTATTAAATGAATTTCTAGAAATAATTCTT------CATTA-GA


           tRNA-------------------------ACD----------------------------
M. para.   AGCTATTAGGGCTATAGGCACAAGCCTTTTAAGCTGGAAGTGGGTG-CTCCCAACCACCC

X. laev.   AGCTAATAGGGCATTAGCGACA-GCCTTTTAAGCTGTAGATTGGTGACTCCCAACCACCC

Bov.       AGCTATATAGCACTAAC--TAA-CC-TTTTAAGTTAGAGATTGAGAGCCATATACTCTCC

D. yakuba  TGACTGAAAGCAAGTACTGGT----CTCTTAAACCATTTTATAGTAAATTAGCACTTAC-


                     <-ATPase 8-----------------------------------------
           --------> MetProGlnLeuIleProAspProTrpPheAlaIlePheMetPheThrTr
M. para.   TT--AGTGACATGCCCCAATTAATCCCTGACCCCTGGTTCGCGATTTTCATGTTCACCTG
                     MetProGlnLeuAsnProGlyProTrpPheLeuIleLeuIlePheSerTr
X. laev.   TT--AATGATATGCCACAGTTAAACCCAGGCCCATGATTCCTAATCCTAATCTTTTCCTG
                     MetProGlnLeuAspThrSerThrTrpLeuThrMetIleLeuSerIlePh
Bov.       TT--GGTGACATGCCGCAACTAGACACGTCAACATGACTGACAATGATCTTATCAATATT
                     IleProGlnMetAlaProIleSerTrpLeuLeuLeuPheIleValPheSe
D. yakuba  TTCTAATGA-ATTCCACAAATAGCACCAATTAGATGATTATTACTATTTATTGTTTTTTC
```

Fig. 5.3.3 (continued)

```
            ------------------------------------------------------------
            pThrIleSerLeuThrIleLeuProGlnLysValIleAlaHisSerPheProAsnGluPr
M. para.    AACAATTTCTTTAACAATTCTCCCACAAAAAGTAATAGCACACTCTTTCCCTAATGAACC
            pLeuValLeuLeuThrPheIleProProLysValLeuLysHisLysAlaPheAsnGluPr
X. laev.    ACTTGTCCTTTTAACATTTATCCCACCAAAAGTTTTAAAACACAAAGCATTTAATGAACC
            eLeuThrLeuPheIleIlePheGlnLeuLysValSerLysHisAsnPheTyrHisAsnPr
Bov.        CTTGACCCTTTTTATCATCTTTCAACTAAAAGTTTCAAAACACAACTTTTATCACAATCC
            rIleThr---PheIleLeuPheCysSerIleAsnTyrTyrSerTyrMetProThrSerPr
D. yakuba   TATTACA---TTTATTTTATTTTGTTCTATTAATTATTATTCATATATACCAACTTCACC


            ------------------------------------------------------<---->>
            oThrProGlnThrValGlnThrLeuLysProPro---ProTrpAsnTrpProTrpThr**
                                                                  MetAsnLeu
M. para.    AACACCCCAAACCGTGCAAACCCTGAAACCGCCC---CCATGAAACTGACCATGAACCTA
            oThrThrGlnThrThrGluLysSerLysProAsn---ProTrpAsnTrpProTrpThr**
                                                                  MetAsnLeu
X. laev.    AACTACACAAACCACAGAAAAATCTAAACCTAAC---CCTTGAAACTGACCATGAACCTA
            oGluLeuThrProThrLysIleLeuLysGlnAsnThrProTrpGluThrLysTrpThrLy
                                                                  MetAsnGlu
Bov.        AGAACTGACACCAACAAAAATATTAAAACAAAACACCCCTTGAGAAACAAAATGAACGAA
            oLysSerAsnGluLeuLysAsnIleAsnLeuAsn---SerMetAsnTrpLysTrp***
                                                                  MetMetThr
D. yakuba   TAAATCTAATGAATTAAAAAATATTAATTTAAAT---TCTATAAACTGAAAATGATAACA
```

Fig. 5.3.3 (continued)

```
            *
            -ATPase 6----------------------------------------------------
            SerLeuPheAspGlnPheGlnValProProIleIleGlyIleProLeuIleIleValAla
M. para.    AGCCTCTTTGACCAATTTCAAGTCCCACCTATTATAGGAATTCCCCTCATTATAGTTGCC
            SerPhePheAspGlnPheMetSerProValIleLeuGlyIleProLeuIleAlaIleAla
X. laev.    AGCTTCTTCGACCAATTTATGAGCCCTGTAATTTTAGGTATTCCACTTATCGCAATCGCT
            sIleTyrLeuProLeuLeuLeuProLeu***
            AsnLeuPheThrSerPheIleThrProValIleLeuGlyLeuProLeuValThrLeuIle
Bov.        AATTTATTTACCTCTTTTATTACCCCTGTAATTTTAGGTCTCCCTCTCGTAACCCTTATC
            AsnLeuPheSerValPheAspProSerAlaIlePheAsnLeuSerLeuAsnTrpLeuSer
D. yakuba   AATTTATTTTCTGTATTTGACCCTTCTATATACCAACTTCACCTAAATCTAATGAATTAA


            ---------------
            LeuSer--
M. para.    CTAAGCTT
            IleLeu--
X. laev.    ATACTTGA
            ValLeu--
Bov.        GTACTATT
            ThrPhe--
D. yakuba   AAAATATT
```

Fig. 5.3.4

Nucleotide sequence of the *Merluccius paradoxus* cytochrome oxidase subunit I, aligned with the COI sequences from bovine (bov.), *Xenopus laevis* (X. laev.), and *Drosophila yakuba* (D. yakuba). Nucleotides or amino acids held in common with *M. paradoxus* are typed in bold. Amino acids underlined are considered conserved.

```
            LeuLeuProProSerPheLeuLeuLeuLeuAlaSerSerGlyValGluAlaGlyAlaG
M. para.    GCTTCTCCCTCCGTCTTTCCTGCTCCTCCTAGCATCTTCTGGGGTAGAAGCCGGGGCCG
            LeuLeuProProSerPheLeuLeuLeuLeuAlaSerSerMetValGluAlaGlyAlaG
Bov.        ACTCCTCCCTCCCTCATTCCTACTACTCCTCGCATCCTCTATAGTTGAAGCTGGGGCAG
            LeuLeuProProSerPheLeuLeuLeuLeuAlaSerSerGlyValGluAlaGlyAlaG
X. laev.    ACTTCTTCCCCCATCATTTCTTTTATTACTAGCATCATCTGGGGTTGAAGCAGGAGCCG
            LeuLeuProProAlaLeuSerLeuLeuLeuValSerSerMetValGluAsnGlyAlaG
D. yakuba   ATTACTACCTCCTGCTCTTTCTTTACTATTAGTAAGTAGAATAGTTGAAAATGGAGCTG

            lyThrGlyTrpThrValTyrProProLeuAlaSerAsnLeuAlaHisAlaGlyAlaSerV
M. para.    GGACAGGCTGAACAGTCTATCCCCCTCTTGCAAGCAATCTTGCCCACGCTGGCGCCTCCG
            lyThrGlyTrpThrValTyrProProLeuAlaGlyAsnLeuAlaHisAlaGlyAlaSerV
Bov.        GAACAGGCTGAACCGTGTACCCTCCCTTAGCAGGCAACCTAGCCCATGCAGGAGCTTCAG
            lyThrGlyTrpThrValTyrProProLeuAlaGlyAsnLeuAlaHisAlaGlyAlaSerV
X. laev.    GCACAGGTTGAACTGTGTACCCGCCTTTAGCTGGAAACCTAGCACATGCTGGAGCATCAG
            lyThrGlyTrpThrValTyrProProLeuSerSerGlyIleAlaHisGlyGlyAlaSerV
D. yakuba   GAACAGGATGAACTGTTTATCCACCTTTATCCGCTGGAATTGCTCATGGTGGAGCTTCAG

            alAspLeuThrIlePheSerLeuHisLeuAlaGlyValSerSerIleLeuGlyAlaIleA
M. para.    TTGACCTCACCATCTTCTCACTCCACCTAGCAGGTGTTTCATCAATTCTAGGGGCAATTA
            alAspLeuThrIlePheSerLeuHisLeuAlaGlyValSerSerIleLeuGlyAlaIleA
Bov.        TAGATCTAACCATTTTCTCTTTACACTTAGCAGGAGTTTCCTCAATTTTAGGAGCCATCA
            alAspLeuThrIlePheSerLeuHisLeuAlaGlyIleSerSerIleLeuGlyAlaIleA
X. laev.    TTGACCTAACAATTTTCTCCCTTCACTTAGCTGGTATTTCATCTATTTTAGGAGCAATTA
            alAspLeuAlaIlePheSerLeuHisLeuAlaGlyIleSerSerIleLeuGlyAlaValA
D. yakuba   TTGATTTAGCTATTTTTTCTCTACATTTAGCAGGGATTTCTTCAATTTTAGGAGCTGTAA
```

Fig. 5.3.4 (continued)

```
            snPheIleThrThrIleIleAsnValGlnLysProProAlaIleSerGlnTyrGlnThrP
M. para.    ATTTTATTACTACTATTATCAACGTACAAAAACCCCCTGCAATCTCACAGTACCAAACAC
            snPheIleThrThrIleIleAsnMet   LysProProAlaMetSerGlnTyrGlnThrP
Bov.        ACTTCATTACAACAATTATCAACATA    AAGCCCCCCGCAATGTCACAATACCAAACCC
            snPheIleThrThrThrIleAsnMet   LysProProAlaMetSerGlnTyrGlnThrP
X. laev.    ACTTCATCACAACAACAATTAACATA    AAACCACCAGCTATATCTCAATACCAAACCC
            snPheIleThrThrValIleAsnMet   ArgSerThrGlyIleThrLeuAspArgMetP
D. yakuba   ATTTTATTACAACTGTAATTAATATA    CGATCAACAGGAATTACATTAGATCGTATAC

            roLeuPheValTrpSerValLeuIle
M. para.    CCCTCTTTGTTTGGTCCGTCCTTATT
            roLeuPheValTrpSerValMetIle
Bov.        CTCTGTTCGTATGATCCGTAATAATT
            roLeuPheValTrpSerValLeuIle
X. laev.    CACTATTTGTTTGATCAGTATTAATC
            roLeuPheValTrpSerValValIle
D. yakuba   CTTTATTTGTTTGATCAGTAGTTATT
```

Fig. 5.3.5

Comparison of the *Merluccius paradoxus* 16S Mitochondrial Ribosomal RNA with mitochondrial sequences from *Xenopus laevis* (X. laev.), and bovine (Bov.). Sequences were aligned, with spaces introduced to maximise the amount of homology. Nucleotides in common with *M. paradoxus* were typed in **bold**.

M. para. ATGCCATAATAGGCCAAGGAAACTAAGGAGTTAATCAGAAGGGGTACAGCCCTTCTGATA

X. laev. **AA**CAAT**TAAAG**TAAAAA**G**TCT**ACTTAGGATTTATTCAA**TCAGGGTACAGCCTGAT**TGAAA**

Bov. TTCAAA**TAAA**CCCC**CA**CTGTAGCTTTAAAAG**TTAGTCTAAAAAAGGTACAGCCTTTTAGAAA**

M. para. AAAGAAACAACTTTAACAGGTGACCCAAGATCATATTAACCAAGGACTTTAAATTAAGTG

X. laev. CAGGAT**ACAAC**CTATAATACTGGGTA**AAGATTATAATCTTCAAGGAAAGTTGAGTCAGTG**

Bov. C-GG**ATACAAC**CTTGA**C**TAGAGAGTAAA-ATT-TAA-CACTAC----------CAT**AGTA**

M. para. GGCCTAAGAGCAGCCATCTGTACAGAAAGCGTTAAAGCTCAAATTAAACATCGAC

X. laev. GGCCTAAAAGCAGCCACCTGTAAAGACAGCGTCAAAGCTCAACTCAATCATTAAC

Bov. GGCCTAAAAGCAGCCATCAATTAAGAAAGCGTTAAAGCTCAACAACAAAAATTAA

Table 5.3.1

Percentage of homology of tRNAs and proteins. DNA = % Nucleotide conservation, AA = % Amino acid conservation, and silent changes = % Silent base changes. M. cap = *Merluccius capensis*, M. para = *Merluccius paradoxus*, X. laev = *Xenopus laevis*, bov = bovine, and D. yak = *Drosophila yakuba*.

		% Homology				
Coding region		M. cap	M. para	X. laev	Bov	D. yakuba
Proteins						
ND1	DNA	100		46	45	26
	AA	100		68	68	26
Silent changes		100		75	70	100
ND2	DNA	100		40	40	29
	AA	100		58	50	28
Silent changes		100		63	60	66
COI	DNA		100	75	73	61
	AA		100	98	97	83
Silent changes			100	61	70	68
COII	DNA		100	75	67	52
	AA		100	74	68	63
Silent changes			100	41	35	57
ATPase8	DNA		100	85	65	49
	AA		100	66	38	34
Silent changes			100	37	53	80

Fig. 5.3.1 (continued)

	% Homology				
Coding region	M. cap	M. para	X. laev	Bov	D. yakuba
ATPase6 DNA		100	64	52	34
AA		100	76	60	36
Silent changes		100	50	40	50
tRNAs					
Isoleucine	100		69	51	51
Glutamine	100		63	44	37
F-Methionine	100		72	65	50
Tryptophan	100		66	62	52
Lysine		100	75	52	39
16S rRNA		100	64	54	

The percentage of amino acid conservation was calculated assuming that amino acid changes which do not affect the charge or conformation of the proteins, are conserved. Amino acids within each of the following sets are considered conserved (underlined in Fig. 5.3.1 - 5.3.4): (Ala, Ser, Thr), (Asn, Gln), (Asp, Glu), (Ile, Leu, Met, Val), (Arg, His, Lys), and (Phe, Tyr, Trp). Silent nucleotide changes represent the nucleotides that change without affecting the amino acid that would be incorporated in that position.

5.4 DISCUSSION

5.4.1 Organization of the hake mitochondrial genome

By aligning the genes sequenced in this section, with the restriction maps of the mtDNA from both *M. capensis* and *M. paradoxus*, the location of the genes on the mitochondrial genome could be inferred. For both species of hake it was found that the basic organization of the genes on the mitochondrial genome resembled the organization of the mitochondrial genes reported for other higher eucaryotes, such as bovine, mouse, human, and *X. laevis*. Clary and Wolstenholme (1985) compared the gene arrangement on the mitochondrial genome of *D. yakuba* to the arrangement of the genes on mouse mtDNA. Many genes on the *D. yakuba* mtDNA were arranged differently to the mammalian mtDNA, relative to their adjacent genes. They postulated that most of the rearrangements could stem from a translocation and inversion of a single fragment of DNA on the ancestral mtDNA molecule of either *D. yakuba* or mouse. From the partial sequence of the mtDNA of *M. paradoxus*, it was determined that arrangement of these genes were more similar to the mammalian mtDNA than to insect mtDNA.

The organization of the hake mitochondrial genes also show extreme economy as was reported for the mitochondrial genes of other higher eucaryotes (Anderson *et al.*, 1981, 1982;

Bibb *et al.*, 1981; Clary and Wolstenholme, 1985; Roe *et al.*, 1985). It was reported that there are none or very few non-coding bases between adjacent genes (Anderson *et al.*, 1981). From the mtDNA sequences of both *M. capensis* and *M. paradoxus*, it was also determined that there were no or only a few non-coding bases between adjacent genes. One example in the *M. capensis* mtDNA sequence was the ND1 gene, which is adjacent to the isoleucine tRNA gene, and only four non-coding bases were found between these genes. Similar results were obtained for the *M. paradoxus* mtDNA sequence, and in that sequence, an overlap of ten bases was also found between the ATPase8 and ATPase6 genes. Van Etten *et al.* (1982) reported that there are three occurrences of reading frames in the mouse mtDNA sequence which overlap, including ATPase8 and ATPase6. The overlap between these two genes was also reported for the human (46 bp overlap, Anderson *et al.*, 1981), bovine (40 bp overlap, Anderson *et al.*, 1982), *X. laevis* (10 bp overlap, Roe *et al.*, 1985), and the *D. yakuba* (6 bp overlap, Clary and Wolstenholme, 1985) mtDNA. The number of base pairs found in the overlap seems to be quite variable across the species. *Merluccius paradoxus* had the same number of nucleotides found in the *X. laevis* mtDNA sequence.

The arrangement of the genes on the mtDNAs examined was highly conserved across species (e. g. bovine to *X. laevis*). From this it can be assumed that the gene arrangement on the mtDNA of *M. capensis* will most likely be the same as for *M.*

paradoxus. By aligning the two restriction maps of *M. capensis* and *M. paradoxus*, it was determined that one *Hin*dIII site was common to both species. If the gene arrangement is the same on the mtDNA from both species, it is expected that the sequence of COI from *M. paradoxus* would be continued on the *M. capensis* mtDNA. However, the sequence found at that *Hin*dIII site on the *M. capensis* mtDNA was the alanine tRNA, which is about 500 bp away from the sequence on the *M. paradoxus* mtDNA. It is unlikely that there will be such a major difference in the mtDNA organization of the two hake, and it was concluded that another undetected *Hin*dIII site must occur on the mtDNA from *M. capensis*. This *Hin*dIII site must be situated about 500 bp from the other *Hin*dIII site. This is a similar result to that obtained for *M. paradoxus*, where the small *Hin*dIII fragment was only detected through cloning of the fragment. Subsequently the mtDNA that was originally cloned, was again digested with *Hin*dIII and run on a gel adjacent to other mtDNA samples digested with *Hin*dIII. It was then observed that the large, 12 kb *Hin*dIII fragment of the cloned mtDNA did in fact migrate faster than the *Hin*dIII fragments from the other individuals, which indicated that this fragment was indeed 500 bp smaller. Not enough mtDNA was available, to redigest all individuals with *Hin*dIII, therefore it could not be determined whether the additional *Hin*dIII site was present in most of the mtDNA from *M. capensis* or not. It is

possible that this *Hin*dIII site represents a polymorphic site, which is found in some of the *M. capensis* mtDNAs.

One striking difference in the gene arrangement was detected between *M. paradoxus* and *X. laevis*. The distance from gene COI to the 16S RNA is approximately 4.2 kb on the mitochondrial genome of *X. laevis*, but with *M. paradoxus*, this distance was approximately 8 kb. One explanation is that an inversion took place around the D-loop region, resulting in a different gene arrangement of the genes in that area. More sequence information is required to determine the exact arrangement of the genes on the mtDNA of hake around the D-loop region.

The sequences of the mitochondrial genes from both *M. capensis* and *M. paradoxus*, reported in this study, represent one of the first reported mtDNA sequences of fish (except the sequences reported for carp). From these sequences it is clear that the organization of the genes on the mitochondrial genome of fish is similar to the arrangement of the genes published for other higher eucaryotes.

5.4.2 Size variation of the mtDNA

In chapter 2, insertions and deletions in the mtDNA of *M. capensis*, were detected by means of restriction fragment analysis. These rearrangements were apparent when cutting the mtDNA with the restriction enzymes *Xho*I, and *Xba*I.

Subsequently, in chapter 3, these deletions and insertions were confirmed, using restriction site analysis. At that stage it could not be determined whether the additions and deletions occurred in a coding or non coding region (D-loop). Through sequencing parts of the mtDNA, the location of the restriction sites on the mtDNA were determined relative to the position of the genes on the mtDNA. This then led to the discovery that the insertions and deletions most probably occurred in the D-loop region of the mtDNA from *M. capensis*. The D-loop region has the least sequence homology between different mammals, and generally length polymorphisms observed within and between species could be located in this region. In this study, mapping and sequencing parts of the mtDNA made it possible to locate the D-loop region on the mtDNA. The results confirmed that hake mtDNA has similar properties to that of other higher eucaryotes, in that the observed variation in size of the mtDNA was due to an insertion or deletion in the D-loop region. More sequence determination around and within the D-loop region is necessary to verify the results.

5.4.3 Fish mitochondrial genetic code

The mitochondrial genetic code varies from the universal genetic code as was summarised in section 5.1.4 and table 5.1.1. In this section it could be confirmed that the mitochondrial genetic code of the two hakes also differs

from the universal genetic code in the same manner as described for other higher eucaryotes (Anderson *et al.*, 1981. 1982; Bibb *et al.*, 1981; Roe *et al.*, 1985, Watson *et al.*, 1987). One of the changes from the universal code was that tryptophan is not coded for by UGG, but by UGA, which is a termination codon in the universal code. From the sequence of the tryptophan tRNA of *M. capensis*, it is clear that the anticodon of the tryptophan tRNA is UCA, thereby confirming that UGA codes for tryptophan in the fish mtDNA. Similarly, the methionine tRNA (anticodon CAU) is also unusual in that it can recognise all four AUN codons (to specify methionine) when they occur as initiation codons, but only AUA and AUG are recognised by the anticodon, when they occur internally (Anderson *et al.*, 1981, 1982). This means that the C in the first position of the f-methionine anticodon must engage in unusual pairing. From the sequence data in Fig. 5.3.1, it is clear that *M. capensis* f-methionine tRNA possesses the CAU anticodon, confirming that hake mitochondrial genomes conform to the mitochondrial genetic code.

The mitochondrial termination codons AGA and AGG, differ from the termination codons, UGA, UAA, and UAG, of the universal code. The codons UAA and UAG are also termination codons in the mitochondrial genetic code. In the mtDNA sequences of *M. capensis* (Figs. 5.3.1 and 5.3.2), it was observed that ND1 and ND2 were terminated with the codon TAG. The ATPase8 from *M. paradoxus* was terminated with the

stop codon TAA. In these partial sequences, no mitochondrial termination codons were observed, but in both hake mtDNAs, the codon TGA, specifies tryptophan and not termination.

In summary, it can be said that the mtDNAs from both species of hake, *M. capensis* and *M. paradoxus*, use the mitochondrial genetic code which seems to be well conserved for all higher eucaryotes.

5.4.4 Protein genes

Parts of ND1 and ND2, of the *M. capensis* mtDNA, were sequenced and ATPase8 and parts of ATPase6, COI, and COII of the *M. paradoxus* mtDNA, were sequenced. The percentage of nucleotide and amino acid conservation between hake and bovine, amphibian (*X. laevis*), and insect (*D. yakuba*) is presented in table 5.3.1. The *M. capensis* and *M. paradoxus* mtDNAs have the highest nucleotide homologies (63%) to the mtDNA sequence of the amphibian, *X. laevis*. A fairly high sequence homology (57%) was found between the mtDNA sequences of hake and the mtDNA sequences of mammals (bovine), and least sequence homology (42%) was found between the insect (*D. yakuba*) and the hake mtDNA. Anderson *et al*. (1982) reported that the COI was the most highly conserved protein which had 91.2% amino acid homology between human and bovine. In this study, it was also found that the COI had the highest amino acid homology, with 97%

homology between *M. paradoxus* and bovine, 98% homology between *M. paradoxus* and *X. laevis*, and 83% homology between *M. paradoxus* and *D. yakuba*. Anderson *et al*. (1982) also reported that the ATPase8 (URF A6L) was the least conserved protein. In this study the ATPase8, ND1 and ND2 sequences were the least conserved sequences. Bibb *et al*. (1981) pointed out that the C- and the N-terminal amino acids of the URFs (NDs) were less conserved than the rest of the amino acids in those proteins. Only the C-terminal parts of ND1 and ND2 were sequenced in this study, which may have resulted in this low level of sequence homology.

All of the protein sequences, except ATPase8, show a higher amino acid conservation than nucleotide conservation between the species. When aligning the reading frames of hake with homologous reading frames of bovine, *X. laevis*, and *D. yakuba*, a high proportion of the observed codon changes are silent ones, occurring in the third position of the codon. The percentage of silent base changes is presented in table 5.3.1, showing that over half (55%) of the conserved codons are different in the third position. This result reflects the degeneracy of the genetic code and agrees with previously reported results that the mtDNA evolves approximately 5-10 times faster than single copy nuclear DNA, yet the evolutionary rate is about the same for nuclear and mitochondrial proteins (Brown and Simpson, 1982).

5.4.5 Termination of translation

The mitochondrial genetic code has four different termination codons: TAA, TAG, AGA, and AGG. In previous studies it has been observed that many mitochondrial reading frames did not contain termination codons, but rather end in T (Anderson *et al.*, 1981, 1982; Montoya *et al.*, 1981; Watson *et al.*, 1987). Normally this T is found adjacent to the 5'-terminal of a tRNA gene, and by cleavage of the tRNA out of a primary transcript, the reading frame is left with a U or a UA at the 3' end. By polyadenylation of the transcript, a termination codon (UAA) is created (Ojala *et al.*, 1981; Anderson *et al.*, 1981) (see section 5.1.5 and Fig. 5.1.2, for post-transcriptional creation of termination codons). It was observed that the hake mtDNA sequences had the UAG termination codon at the end of the ND1 reading frame, whereas the UAA termination codon was found at the end of the ATPase8 reading frame. The ND2 and the COII reading frames were terminated with a T, and they were adjacent to the Trp and Lys tRNA, respectively. By precise excision of the tRNAs, a UAA termination codon is also created as described for other organisms. The various termination codons, or the presence of T as termination codon, are not strongly conserved across all species. For instance, only ND2 from human, cow, mouse, *D. yakuba*, and *M. capensis* ends with a T, whereas the *X. laevis* ND2 ends with a TAG.

5.4.6 Conserved amino acids

Anderson *et al.* (1982) calculated the percentage of conservation between human and bovine amino acids. They found that tryptophan was the most conserved amino acid followed by arginine and glycine, and the amino acids with neutral non-cyclic side groups were least conserved between the two species. This indicated that most probably the hydrophobicity of the side group determined the degree of conservation of the amino acid (Anderson *et al.*, 1982). The average percentage of conserved amino acids between the two hake and bovine was calculated and the results are presented in tables 5.4.1 and 5.4.2. It was determined that tryptophan was the most conserved amino acid found in the sequences of *M. paradoxus* and bovine. The next most conserved amino acids were tyrosine and histidine, followed by proline and glycine. In this study it was determined that the amino acids with neutral, non-cyclic side groups were also the ones least conserved between the two species. For *M. capensis* and bovine, it was determined that the amino acids seemed to be conserved in the same proportion as was reported between human and bovine, but not enough sequence information was available to express this quantitatively. It seems to be a general rule that amino acids with structurally similar side groups, possessing the same charge, are readily interchangeable.

Table 5.4.1

Degree of conservation for each amino acid determined for ND1 and ND2 of the *Merluccius capensis* mtDNA compared to the equivalent DNA sequence from bovine mtDNA. MC = *M. capensis*

Amino Acid	Total bovine	Total MC	Conserved	Average % conserved
Trp	1	1	1	100
Gly	1	2	1	76
Arg	0	0	0	-
Gln	1	0	0	-
Pro	6	7	6	92
Glu	1	1	0	-
Tyr	0	0	0	-
His	1	1	1	100
Phe	1	2	1	76
Asn	0	1	0	-
Asp	0	0	0	-
Lys	2	1	1	76
Leu	9	9	5	56
Cys	1	0	0	-
Ala	0	3	0	-
Ser	4	6	2	40
Ile	2	5	0	0
Met	7	3	1	20
Val	4	1	0	0
Thr	6	6	1	17

Average percentage conserved = 2 x (number conserved)/(total *M. capensis* + total bovine). The order of the amino acids is according to Anderson *et al*. (1982), where the top amino acid is the most conserved amino acid found between human and bovine.

Table 5.4.2

Conservation of amino acids (in COI, COII, ATPase8, and ATPase6) between *Merluccius paradoxus* (MP) and bovine mtDNA.

Amino Acid	Total bovine	Total MP	Conserved	Average % conserved
Trp	6	7	6	92
Gly	12	9	8	76
Arg	2	1	0	0
Gln	7	12	5	53
Pro	17	26	16	74
Glu	7	6	4	62
Tyr	6	4	4	80
His	5	5	4	80
Phe	11	13	9	75
Asn	10	7	5	59
Asp	2	6	1	25
Lys	9	3	3	50
Leu	34	26	19	63
Cys	2	2	2	100
Ala	12	18	11	73
Ser	20	21	13	63
Ile	18	19	11	59
Met	10	8	3	33
Val	12	15	9	67
Thr	20	15	8	46

Average percentage conserved = 2 x (number conserved)/(total *M. capensis* + total bovine). The order of the amino acids is according to Anderson *et al*. (1982), where the top amino acid is the most conserved amino acid found between human and bovine.

5.4.7 Codon usage

The total number of each kind of codon found in the open reading frames of *M. capensis* and *M. paradoxus* mtDNA are summarized in table 5.4.3. A preference for codons ending in C is evident. *Merluccius capensis* used codons ending in C with frequently (35.3%), codons ending with A and T were used less frequently (25.5% and 23.5%, respectively), and codons ending with G were seldom used (15.7%). *Merluccius paradoxus* frequently used codons ending with C and A (36.2% and 32.6%, respectively), codons ending with T were used less frequently (22.8%), and codons ending with G were used rarely (8.5%). Mouse mtDNA mainly uses codons ending in A (Bibb *et al.*, 1981); in human and bovine mtDNA, codons ending in A and C predominate (Anderson *et al.*, 1981. 1982); in *X. laevis* codons ending in T were used most frequently (Roe *et al.*, 1985); and in *D. yakuba*, mostly codons ending in A and T were used (Clary and Wolstenholme, 1985). The preference for using codons ending in a certain base seems to be determined by the increased presence of that base in the major coding strand of mtDNA. For instance, in the human mtDNA the major coding strand was composed up to 30% of T and C, whereas the amphibian mtDNA had an increased presence of T in the major coding strand of the mtDNA. The result reported here reflects a high level of C and A in the major coding strand of hake mtDNA, and also explains the high GC content reported in chapter 3.

Table 5.4.3

Codon usage in the mitochondrial protein genes of *Merluccius capensis* and *Merluccius paradoxus*, sequenced in this study. MC = *M. capensis*, and MP = *M. paradoxus*

		MC	MP			MC	MP			MC	MP			MC	MP
Phe	TTT	1	4	Ser	TCT	1	7	Tyr	TAT	0	1	Cys	TGT	1	1
GAA	TTC	1	9	**UGA**	TCC	3	3	**GUA**	TAC	0	3	**GCA**	TGC	0	1
Leu	TTA	1	2		TCA	1	6	Stop	TAA	0	1	Trp	TGA	0	5
UAA	TTG	0	0		TCG	0	0		TAG	2	0	**UCA**	TGG	1	2
Leu	CTT	1	5	Pro	CCT	1	6	His	CAT	0	0	Arg	CGT	0	0
UAG	CTC	4	9	**UGG**	CCC	4	11	**GUG**	CAC	1	5	**UCG**	CGC	0	0
	CTA	2	7		CCA	2	6	Gln	CAA	0	11		CGA	0	1
	CTG	1	3		CCG	0	3	**UUG**	CAG	0	1		CGG	0	0
Ile	ATT	4	13	Thr	ACT	2	3	Asn	AAT	1	3	Ser	AGT	0	0
GAU	ATC	1	6	**UGU**	ACC	1	5	**GTT**	AAC	0	4	**GCU**	AGC	1	5
Met	ATA	1	5		ACA	3	7	Lys	AAA	1	3	Stop	AGA		
CAU	ATG	2	3		ACG	0	1	**UUU**	AAG	0	0		AGG		
Val	GTT	0	5	Ala	GCT	1	1	Asp	GAT	0	1	Gly	GGT	0	1
UAC	GTC	0	5	**UGC**	GCC	1	8	**GUC**	GAC	0	5	**UCC**	GGC	1	2
	GTA	1	4		GCA	1	8	Glu	GAA	0	6		GGA	0	2
	GTG	0	1		GCG	0	1	**UUC**	GAG	1	0		GGG	1	4

5.4.8 The tRNA genes

Five different tRNA genes were identified on the mtDNA of *M. capensis*, and one tRNA gene, lysine tRNA, was identified on the *M. paradoxus* mtDNA. These tRNA were identified on the basis of the sequence homology to other known mitochondrial tRNAs. The sequences of the tRNAs, aligned with other tRNA sequences, are presented in Figs. 5.3.1 - 5.3.3 and the cloverleaf structures of the tRNAs are presented in Fig 5.4.1. The hake mitochondrial tRNAs were interspersed between protein coding genes, indicating that processing of a primary transcript by precise excision of the tRNAs most probably also exists in fish mitochondria.

Mitochondrial tRNAs have several unique features compared to other tRNAs: (1) in general they do not have the invariant GTψCRA sequence which is found in other tRNAs;

(2) nucleotides which are located on the D and T loops are invariant in other tRNAs, but they vary in the mitochondrial tRNAs;

(3) the TψC loop, which is always seven nucleotides long in non-mitochondrial tRNAs, can be three to nine residues long in the mitochondrial tRNAs.

The structure of the mitochondrial tRNAs, determined in this study, was also different to non-mitochondrial tRNAs. One

Fig. 5.4.1

Sequences of *Merluccius capensis* Ile tRNA, Gln tRNA, f-Met tRNA, and Trp tRNA, and of *Merluccius paradoxus* Lys tRNA genes represented in the cloverleaf structure. Normal base pairing in the stem structures is represented by a dash, and unusual base pairing is represented by an asterisk.

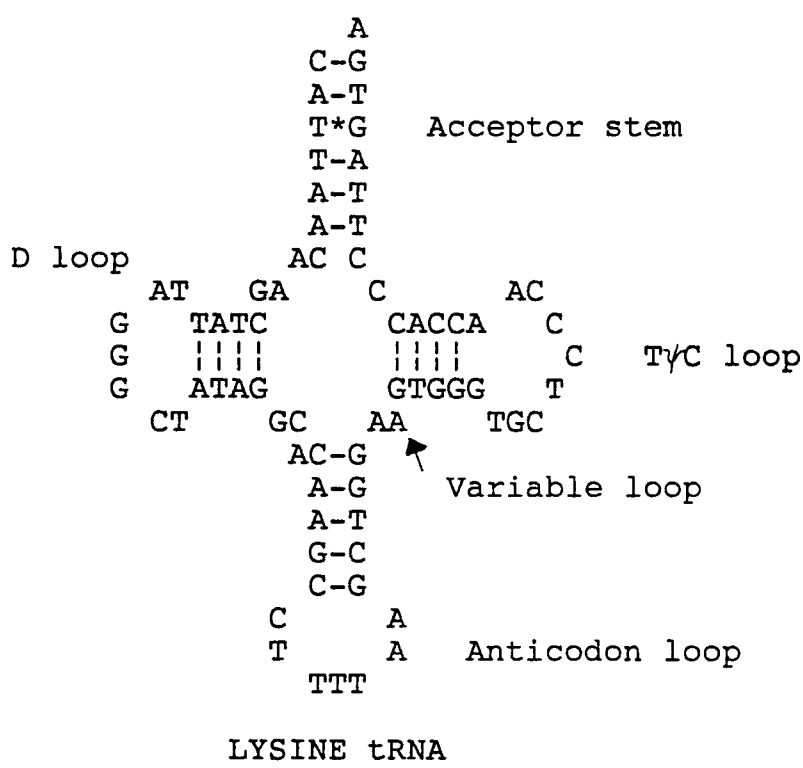
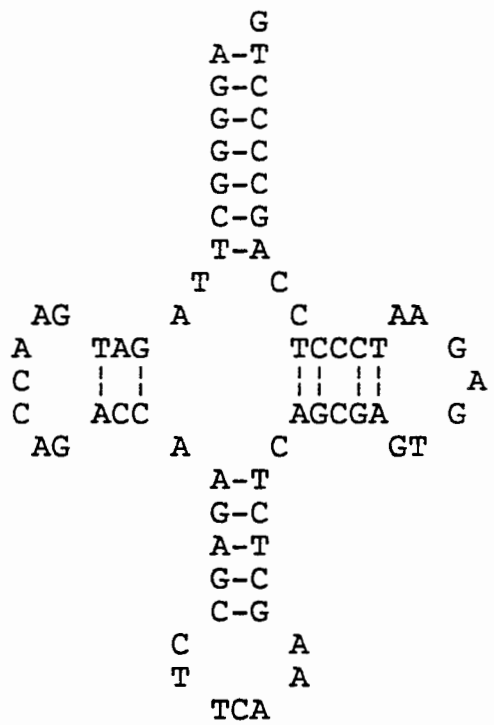
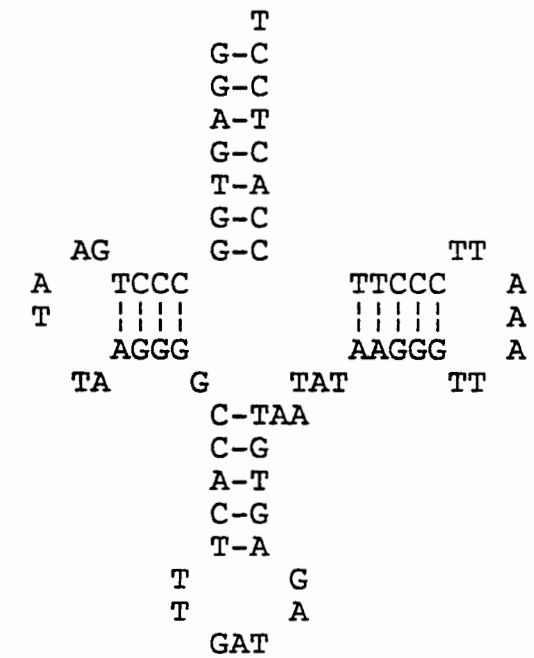


Fig. 5.4.1 (continued)

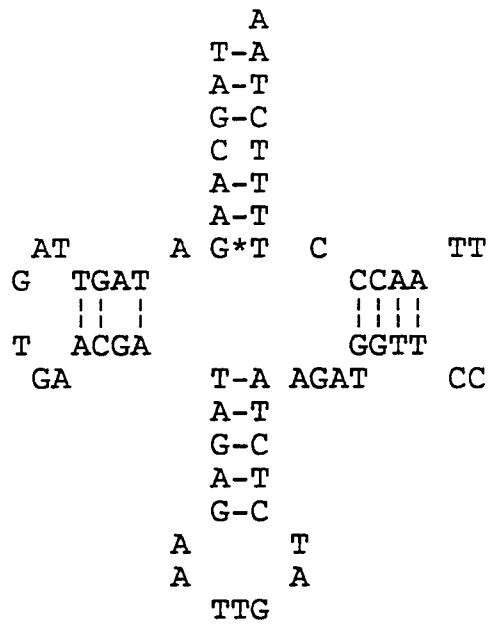


TRYPTOPHAN tRNA

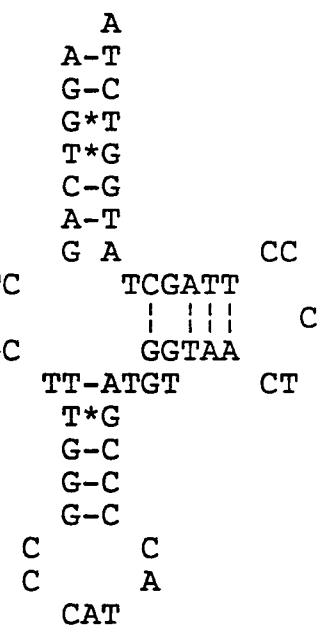


ISOLEUCINE tRNA

Fig. 5.4.1 (continued)



GLUTAMINE tRNA



f-METHIONINE tRNA

example being the T ψ C loop which, was 4, 5, 7, and 8 nucleotides long in the glutamine tRNA, f-methionine tRNA, tryptophan tRNA and the lysine tRNA, respectively. The TψC stem was four or five nucleotide pairs long and the invariant GT ψ CRA structure of non-mitochondrial tRNAs was also not observed in these hake mitochondrial tRNAs. It was reported that the bovine mitochondrial tRNAs are A-U rich and that they contain several instances of mismatched base pairing in their stem regions (Roe *et al.*, 1982). The hake mitochondrial tRNAs were found to be A and U rich and some mismatched base pairing was detected in the stem regions of these tRNAs, which is indicated by asterisks in Fig 5.4.1. The hake mitochondrial tRNAs had certain conserved features as was reported for other mitochondrial tRNAs. The anticodon loop was always seven nucleotides long, the anticodon stem was five nucleotide pairs in length, the acceptor stem was always seven nucleotide pairs long, and none of the tRNAs had the trinucleotide sequence CAA, which occurs at the 3′ end of non-mitochondrial tRNAs. In general, the structure of the hake mitochondrial tRNAs resemble the structure reported for the mitochondrial tRNAs of other higher eucaryotes.

The sequence homology between the hake tRNAs and the *X. laevis* tRNAs was about 69%, and the sequence homology between the tRNAs from hake and bovine was about 53%. This value of nucleotide conservation is lower than the reported 80% sequence homology between the tRNA genes of human and

bovine (Anderson *et al.*, 1981, 1982), but a similar value was reported for the conservation of tRNAs of six primates, mouse, and cow (Avise, 1986). The level of nucleotide conservation of the tRNA genes is most likely higher, when comparing hake tRNAs to other fish tRNAs.

5.4.9 Ribosomal RNA genes

The 16S rRNA sequence of *M. paradoxus* was identified by comparison with other known mitochondrial sequences. The 16S rRNA gene was then positioned on the mitochondrial genome by aligning the sequence with the restriction map of the mtDNA. The sequence homology between hake and *X. laevis* 16S rRNA was 64%, and between hake and bovine was 54%. It appears that the 16S-like RNA from the small ribosomal subunit of all organisms has the same four-domain structure. The most conserved sequences amongst those rRNAs are found in the single-stranded regions. Not enough information was available to predict any secondary structure of the 16S rRNA of hake, but the high percentage of homology observed between the hake 16S rRNA and those from amphibian and bovine, suggest that the 16S rRNA sequence is located in the single stranded domain of the rRNA.

5.4.10 Evolution of mitochondrial genomes

Several hypotheses have been proposed to account for mitochondrial evolution. The most widely accepted hypothesis is that mitochondria represent ancient procaryotic cells which invaded eucaryotic cells forming an endosymbiotic relationship. The fact that the mitochondrial rRNAs are similar to procaryote rRNAs, in that they are chloramphenicol sensitive, confirmed this hypothesis. It has also been demonstrated that mitochondrial rRNAs can be chloramphenicol resistant, and that this insensitivity to the drug is the result of a change of one base pair in the 16S rRNA gene, in sequences that are identical in *E. coli* and yeast (Kearsey and Craig, 1982; Wallace *et al.*, 1982). The most recent archigenetic hypothesis (Mikelsaar, 1987) predicts an ancient separate formation of animal, fungal and plant mitochondria and independent invasions of free-living protoeucaryotes by different free-living protomitochondria. This hypothesis is based on sequencing data which became available in recent years. More plant mtDNA sequences are required to test this archigenetic hypothesis, which proposes that animal mitochondria are the oldest living fossils preserved today. The mtDNA sequence presented in this study represents one of the building blocks required to prove that all animal mitochondria are relatively simple in structure and function, with no introns, a small number of tRNAs (22), a non-universal genetic code, and a different transcription-translation system.

CHAPTER 6

CONCLUSION

CONCLUSION

This study was aimed at determining the amount of genetic divergence between the two sympatric species of hake, *Merluccius capensis* and *M. paradoxus*, which occur along the coasts of southern Africa. Protein electrophoresis, restriction endonuclease fragment- and site-analysis of the mtDNA and sequencing of parts of the mtDNA, were the methods employed. Cloning parts of the mtDNA made the estimation of sequence divergence between the New Zealand hake, *M. australis*, and the two southern African hake, possible.

Through the use of protein electrophoresis, it was determined that the two southern African hake are genetically well differentiated and should be classified as two separate species which do not interbreed. The results of the restriction analysis of mtDNA also confirmed this. Generally, it is accepted that sympatric speciation is unlikely to occur and that isolation in allopatry is the primary mode of speciation (Mayr, 1982). It is most likely that the two southern African species of hake evolved in allopatry and that they only started to live in sympatry after the formation of a barrier to gene flow. One model of allopatric speciation is that an ancestral southern African hake population got divided and that this resulted in the formation of the two species. This is unlikely and the genetic data also suggest that the two species did not share

one common ancestor. Another model is that the presence of the two sympatric hake species might be the result of repeated dispersal of ancestral populations of other Atlantic Ocean hakes to southern Africa.

Restriction endonuclease analysis of the mtDNA made it possible to determine evolutionary events such as bottlenecks. A low level of nucleotide polymorphism, which was reflected in the number of composite genotypes and the complexity of the parsimonous trees connecting the genotypes with each other, was detected for *M. paradoxus* and *M. australis*. This low level of nucleotide site polymorphism showed that these species most likely experienced ancient bottlenecks in population size.

Results also showed that the deep water hake *M. paradoxus* was closely related to the hake species of New Zealand. This confirmed the hypothesis that there is some connection between the fish species of southern Africa, southern South America and New Zealand. The close relationship between *M. paradoxus* and *M. australis* may also indicate that an ancient *M. paradoxus* dispersed in the southern Atlantic Ocean from east to west and gave rise to the New Zealand species, *M. australis*.

The time of divergence was estimated between the New Zealand hake and the two southern African species, using sequence divergence of mtDNA. The time of divergence between *M. australis* and *M. paradoxus* was 5.75 to 6.1 million years and

between *M. australis* and *M. capensis*, 7 to 8.95 million years. This again confirmed the close relationship between *M. paradoxus* and *M. australis*.

Restriction analysis of the mtDNA also showed that all three species are well separated congeneric species and not siblings. The species of hake analyzed in this study, however, do not present enough data to reconstruct all the evolutionary events. More hake species need to be analyzed to determine the origin of the two species of hake inhabiting the southern African waters.

The time of divergence was calculated using genetic distance and sequence divergence of the mtDNA, and it was found that *M. capensis* and *M. paradoxus* separated 4.25 to 13.6 million years ago. The large range in the time of divergence is due to the different methods used, and the errors inherent in each method. The discrepancy in the time of divergence between protein electrophoretic results (7.6 to 13.6 million years) and estimates calculated using mtDNA sequence divergence (4.25 to 5.8 million years) may be due to the incorrect estimation of the mutation rate of mtDNA. In recent reviews on evolution of man, it is clear that the mutation rate of mtDNA is not as yet accepted, with Wolpoff claiming: "They have calibrated the mutation rate incorrectly," (Lewin, 1987a), and Cann saying: "I see big variation in rate in different parts of the molecule" (Lewin, 1987b). The regularity of the mitochondrial clock

can only be proved by detailed comparisons of sequences of significant sections of the mtDNA (Lewin, 1987b). The data reported in this study form a building block towards the construction of a mtDNA clock for marine fish.

Restriction fragment analysis of the mtDNA further revealed that the *M. capensis* mitochondrial genome was 16.9 kb in length, *M. paradoxus* mtDNA was 16.7 kb long, and *M. australis* mtDNA was 16.9 kb in length. The size of the mtDNA of hake proved to be of similar length, as has been reported for other fish species. Size variations were detected within *M. capensis* and, through restriction site analysis of the mtDNA, these size variations could be recognised as insertions and deletions which are most likely to be situated on, or close to, the D-loop structure.

The mtDNA sequence data from *M. capensis* and *M. paradoxus* allowed the prediction of much of the general organization of the mitochondrial genome of hake. As expected, the fish mtDNA has strong sequence homology to amphibian mtDNA. It was also found that the fish mitochondrial protein genes were similar in conservation, genetic code, codon usage, overlapping of reading frames, and termination of coding regions, to the equivalent protein genes of other higher eucaryotes. The genes on the mitochondrial genome of fish showed extreme economy, with some protein coding genes having overlapping reading frames, and none or only a few non-coding bases were situated between adjacent genes. It

was also observed that the mitochondrial genetic code seems to be conserved across all species. The mitochondrial genetic code of hake was found to be different to the universal genetic code in that the codon UGA, which is a termination codon in the universal code, specified tryptophan, and the C in the first position of the methionine tRNA must engage in unusual pairing. The protein genes reflected the degeneracy of the genetic code in that over 55% of the nucleotide substitutions did not result in an amino acid substitution. It was also determined that some of the protein coding genes had no termination codons and that these coding regions must have formed postranscriptional termination codons by the precise excision of a tRNA from the primary transcript, followed by polyadenylation, resulting in a termination codon.

The mitochondrial tRNAs also resembled (in sequence and secondary structure) the mitochondrial tRNAs found in mammals, amphibians and insects. It is apparent that future detailed analysis of the hake mtDNA sequences may provide important information about the evolution of hake and will also provide information about the evolution and function of mtDNA in general.

LITERATURE CITED

LITERATURE CITED

Adams J., Rothman E.D. 1982. Estimation of phylogenetic relationships from DNA restriction patterns and selection of endonuclease cleavage sites. Proc. Natl. Acad. Sci. USA 79: 3560-3564.

Allendorf W.F., Utter F.M. 1979. Population genetics. In Fish Physiology edited by W.S. Hoar, D.J. Randall, J.R. Brett. Academic Press, New York, NY. p. 407-454.

Allendorf F.W., Phelps S.R. 1981. Use of allelic frequencies to describe population structure. Can. J. Fish. Aquat. Sci. 38(12): 1507-1514.

Anderson S., Bankier A.T., Barrell B.G., de Bruijn M.H.L., Coulson A.R., Drouin J., Eperon I.C., Nierlich D.P., Roe B.A., Sanger F., Schreier P.H., Smith A.J.H., Staden R., Young I.G. 1981. Sequence and organisation of the human mitochondrial genome. Nature 290: 456-465

Anderson S., de Bruijn M.H.L., Copulson A.R., Eperon I.C., Sanger F., Young I.G. 1982. Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. J. Mol. Biol. 156: 683-717.

Araya A., Amthauer R., Leon G., Krauskopf M. 1984. Cloning, physical mapping and genome organization of mitochondrial DNA from *Cyprinus carpio* oocytes. Mol. Gen. Genet. 196: 43-52.

Aoki K., Tateno Y., Takahata N. 1981. Estimating evolutionary distance from restriction maps of mitochondrial DNA with arbitrary G+C content. J. Mol. Evol. 18: 1-8.

Aquadro C.F, Kaplan N., Risko K.J. 1984. An Analysis of the Dynamics of Mammalian Mitochondrial DNA Sequence Evolution. Mol. Biol. + Evol. 1(5): 423-434.

Archie J.W. 1985. Statistical analysis of heterozygosity data: independent sample comparisons. Evolution 39: 623-637.

Attardi G., Cantatore P., Chomyn A., Crews S., Gelfand R., Merkel C., Montoya J., Ojala D. 1982. A comprehensive view of mitochondrial gene expression in Human cells. In Mitochondrial genes edited by Slonimski, Borst, Attardi, CSH 51-72.

Avise J.C. 1974. Systematic value of electrophoretic data. Syst. Zool. 23(4): 465-481.

Avise J. C., Giblin-Davidson C., Laerm J., Patton J.C., Lansman R.A. 1979a. Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, *Geomys pinetis*. Proc. Natl. Acad. Sci. USA 76(12): 6694-6698.

Avise. J.C., Lansman A., Shade R.O. 1979b. The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. I. Population structure and evolution in the genus Peromyscus. Genetics 92: 279-295.

Avise J.C., Lansman R.A. 1983. Polymorphism of mitochondrial DNA in populations of higher animals. In Evolution of genes and proteins edited by M. Nei and R.K. Koehn Sinauer Assoc., Sunderland Mass. USA. Chapter 8

Avise J.C., Saunders N.C. 1984. Hybridization and Introgression among species of Sunfish (Lepomis): Analysis by Mitochondrial DNA and allozyme markers. Genetics 108: 237-255.

Avise J.C. Neigel J.E., Arnold J. 1984. Demographic influences on mitochondrial DNA lineage survivorship in animal populations. J. Mol. Evol. 20: 99-105.

Avise J.C. 1986. Mitochondrial DNA and the evolutionary genetics of higher animals. Phil. Trans. R. Soc. Lond. B 312: 325-342.

Badenhorst 1984. Development of the chondrocranium of *Merluccius capensis* Cast. and its possible contribution to the distributional ecology of the Cape Hakes. PhD Thesis, University of Stellenbosch, South Africa.

Bailey K.M., Francis R.C., Stevens P.R. 1982. The life history and fishery of Pacific whiting, *Merluccius productus*. Calif. Coop. Oceanic Fish. Invest. Rep. 23: 81-98.

Bendich A.J. 1982. Plant mitochondrial DNA: The last frontier. In Mitochondrial genes edited by Slonimski, Borst, Attardi, CSH 477-482.

Berg W.J., Ferris S.D. 1984. Restriction endonuclease analysis of Salmonid Mitochondrial DNA. Can. J. Fish. Aquat. Sci. 41: 1041-1047.

Bermingham E. Avise J.C. 1986. Molecular zoogeography of freshwater fishes in the southeastern United States. Genetics 113: 939-965.

Bermingham E., Lamb T., Avise J.C. 1986. Size polymorphism and heteroplasmy in the mitochondrial DNA of lower vertebrates. Heredity 77: 249-252.

Bibb M.J., van Etten R.A., Wright C.T., Walberg M. W., Clayton D.A. 1981. Sequence and Gene organization of mouse mitochondrial DNA. Cell 26: 167-180.

Botha L. 1973. Migrations and spawning behaviour of the Cape hakes. The S. Afr. Shipp. News Fishg Ind. Rev., 28(4): 62-67

Botha L. 1980. The biology of the Cape Hakes *M. capensis* Cast. and *M. paradoxus* Franca in the Cape of Good Hope area. PhD Thesis, University of Stellenbosch, South Africa. 182 pp.

Brown W.M., Vinograd J. 1974. Restriction endonuclease cleavage maps of animal mitochondrial DNAs. Proc. Natl. Acad. Sci. USA 71(11): 4617-4621.

Brown W.M., George M.Jr., Willson A.C. 1979. Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. USA 76(4): 1967-1971.

Brown W.M. 1980. Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. Proc. Natl. Acad. Sci. USA 77(6): 3605-3609.

Brown G.C., Simpson M.V. 1981. Intra- and interspecific variation of the mitochondrial genome in *Rattus norvegicus* and *Rattus rattus*: Restriction enzyme analysis of variant mitochondrial DNA molecules and their evolutionary relationships. Genetics 97:125-143.

Brown G.G., Simpson M.V. 1982. Novel features of animal mtDNA evolution as shown by sequences of two rat cytochrome oxidase subunit II genes. Proc. Natl. Acad. Sci. USA 79: 3246-3250.

Brown W.M. 1983. Evolution of animal mitochondrial DNA. In Evolution of Genes and Proteins edited by M. Nei and R.K. Koehn. Sinauer Assoc., Sunderland Mass. USA Chapter 4.

Cann R.L., Wilson A.C. 1983. Length mutations on human mitochondrial DNA. Genetics 104: 699-711.

Cann R.L., Stoneking M., Wilson A.C. 1987. Mitochondrial DNA and human evolution. Nature 325: 31-36.

Carlson S.S., Wilson A.C., Maxson R.L. 1978. Do albumin clocks run on time? A reply. Science 200: 1183-1185.

Castora F.J., Arnheim N., Simpson M.V. 1980. Mitochondrial DNA polymorphism: Evidence that variants detected by restriction enzymes differ in nucleotide sequence rather than in methylation. Proc. Natl. Acad. Sci. USA 77(11): 6415-6419.

Castora F.J., Sternglanz R., Simpson M.V. 1982. A new mitochondrial topoisomerase from rat liver that catenates DNA. In Mitochondrial genes edited by Slonimski, Borst, Attardi, CSH 143-154.

Chakraborty R., Haag M., Ryman N., and Stahl G. 1982. Hierarchical gene diversity analysis and its application to brown trout population data. Hereditas 97: 17-22.

Chapman R.W., Powers D.A. 1984. A method for the rapid isolation of mitochondrial DNA from fishes. Tech. Rep. Maryland Sea Grant Prog. No. UM-SG-TS-84-05, 11p.

Chen E.Y., Seeburg P.H. 1985. Supercoil sequencing: A fast and simple method for sequencing plasmid DNA. DNA 4(2): 165-172

Chlapowski K. 1975. Hake (*Merluccius capensis* and *Merluccius paradoxus*) in Polish catches in Southwest African Area. Colln scient. Pap, int. Comm SE Atl, Fish, 2: 39-42.

Chomyn A., Mariottini P., Cleeter M.W., Ragan C.I., Matsuno-Yagi A., Hatefi Y., Doolittle R.F., Attardi G. 1985. Six unidentified reading frames of human mitochondrial DNA encode components of the respiratory-chain NADH dehydrogenase. Nature 314: 592-597.

Chomyn A., Cleeter M.W., Ragan C.I., Riley M., Doolittle R.F., Attardi G. 1986. URF6, last unidentified reading frame of human mtDNA, codes for an NADH dehydrogenase subunit. Science 234: 614-618.

Church G., Gilbert W. 1984. Proc. Natl. Acad. Sci. USA 81: 1991-1995.

Clary D.O., Wolsten-holme D.R. 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: Nucleotide sequence, Gene organization, and Genetic Code. J. Mol. Evol. 22: 252-271.

Cooper D.W. 1968. The significance level in multiple tests made simultaneously. Heredity 23: 614-617.

Dale R.M.K. 1982. Structure of plant mitochondrial DNAs. In Mitochondrial genes edited by Slonimski, Borst, Attardi, CSH 471-476.

Dando P.R. 1974. Distribution of multiple glucosephosphate isomerases in teleostean fishes. Comp. Biochem. Physiol. 47B: 663-679.

Denaro M., Blanc H., Johnson M.J., Chen K.H., Wilmsen E., Cavalli-Sforza L.L., Wallace D.C. 1981. Ethnic variation on *Hpa* I endonuclease cleavage patterns of human mitochondrial DNA. Proc. Natl. Acad. Sci. USA 78(9): 5768-5772.

De Francesco L., Attardi G., Croce C.M 1980. Uniparental propagation of mitochondrial DNA in mouse-human cell hybrids. Proc. Natl. Acad. Sci. USA 77(7): 4079-4083.

De Salle R., Giddings L.V., Templeton A.R. 1986a. Mitochondrial DNA variability in natural populations of Hawaiian *Drosophila*. I. Methods and levels of variability in *D. silverestris* and *D. heteroneura* populations. Heredity 56: 75-85.

De Salle R., Giddings L.V., Kaneshiro K.Y. 1986b. Mitochondrial DNA variability in natural populations of Hawaiian Drosophila. II. Genetic and phylogenetic relationships of natural populations of *D. silverestris* and *D. heteroneura*. Heredity 56: 87-96.

De Zamaroczy M., Bernardi G. 1986. The primary structure of the mitochondrial genome of *Sachromyces cerevisiae* - a review. Gene 47: 155-177.

Dubin D.T., HsuChen C.C., Timko K.D., Azzolina T.M., Prince D.L., Ranzini J.L. 1982. 3' Termini of mammalian and Insect mitochondrial rRNAs. In Mitochondrial genes edited by Slonimski, Borst, Attardi, CSH 89-98.

Fauron C.M.R., Wolsten-holme D.R. 1976. Structural heterogeneity of mitochondrial DNA molecules within the genus *Drosophila*. Proc. Natl. Acad. Sci. USA 73: 3623-3627.

Ferris S.D., and Whitt G.S. 1978. Genetic and molecular analysis of nonrandom dimer assembly of the creatine kinase isozyme of fishes. Biochem. Genet. 16: 811-829.

Ferris S.D., Wilson A.C., Brown W.M. 1981a. Evolutionary tree for apes and humans based on cleavage maps of mitochondrial DNA. Proc. Natl. Acad. Sci. USA 78(4): 2432-2436.

Ferris S.D., Brown W.M., Davidson W.S., Wilson A. C. 1981b. Extensive polymorphism in the mitochondrial DNA of apes. Proc. Natl. Acad. Sci. USA 78(10): 6319-6323.

Ferris S.D., Sage R.D., Wilson A.C. 1982. Evidence from mtDNA sequences that common laboratory strains of inbred mice are descended from a single female. Nature 295: 163-165.

Ferris S.D., Sage R.D., Huang C-M., Nielsen J.T., Ritte U., Wilson A.C. 1983a. Flow of mitochondrial DNA across a species boundary. Proc. Natl. Acad. Sci. USA 80: 2290-2294.

Ferris S.D., Sage R.D., Prager E.M., Ritte U., Wilson A.C. 1983b. Mitochondrial DNA evolution in mice. Genetics 105: 681-721.

Ferris S.D., Berg W.J. 1987. The utility of mitochondrial DNA in fish genetics and fishery management. In Population genetics and Fishery Management edited by N. Ryman and F. Utter. Washinton Sea Grant Prog., Univ. Washington Press, Seattle, Washington. Chapter 11: 277-299.

Fisher S.E., Shaklee J.B., Ferris S.D., Whitt G.S. 1980. Evolution of five multilocus isozyme systems in the chordates. Genetica (52)53: 73-85.

Franca P. 1954. Contribuicao para o conhecimento do genero *Merluccius* no Atlantico oriental ao sul do equador. Trabhs Miss Biol. marit. 8: 46-98.

Franca P. 1960. Nova contribuicao para conhecimento do genero *Merluccius* no Atlantico oriental ao sul do equador. Mems. Junta Invest. Ultramar, (2) 18: 57-101.

Franca P. 1962. Considerations sur la taxonomie des *Merluccius* de l'Atlantique oriental. Mems. Junta Invest. Ultramar. (2) 36: 7-48.

Frick L. 1983. An electrophoretic investigation of the cytosolic di- and tripeptidases of fish: molecular weights, substrate specificities, and tissue and phylogenetic distributions. Biochem. Genet. 21: 309-322.

Futuyma D.J. 1980. Variation in Evolutionary Biology. Chapter 9: 189-413.

Giles R.E., Blanc H., Cann H.M., Wallace D.C. 1908. Maternal inheritance of human mitochondrial DNA. Proc. Natl. Acad. Sci. USA 77 (11): 6715-6719.

Glaus K.R. 1980. The structure, organization and evolution of avian mitochondrial DNA. PhD Thesis, Graduate School of the Ohio State University.

Goddard J.M., Fauron C.M.-R., Wolsten-holme D.R. 1982. Nucleotide sequences within the A + T-rich region and the large-rRNA gene of mitochondrial DNA molecules of *Drosophila yakuba*. In Mitochondrial genes edited by Slonimski, Borst, Attardi, CSH 99-104.

Gonzàlez-Villaseñor L.I., Burkhoff A.M., Corces V., Powers D.A. 1986. Characterization of cloned mitochondrial DNA from teleost *Fundulus heteroclitus* and its usefulness as an interspecies hybridization probe. Can. J. Fish. Aquat. Sci., 43: 1866-1872.

Grant W.S., Utter F.M. 1984. Biochemical population genetics of the Pacific herring (*Clupea pallasi*). Can. J. Fish. Aquat. Sci. 41: 856-864.

Grant W.S. 1985. Population genetics of the southern African Pilchard, *Sardinops ocellata*, in the benguela upwelling system. In Simposio Internacional sobre las areas de afloramiento mas Importantes del oeste Africano edited by C. Bas, R. Margalef, P. Rubies. Invest. Pesq, Barcelona. 551-562.

Grant W.S., Stahl G. 1987. Description of electrophoretic loci in Atlantic cod, *Gadus mornua*, and comparison with Pacific cod, *G. macrocephalus*. Hereditas. In press.

Grant W.S. 1987. Genetic divergence between congeneric Atlantic and Pacific Ocean fishes. In Population genetics and Fishery Management edited by N. Ryman and F. Utter. Washinton Sea Grant Prog., Univ. Washington Press, Seattle, Washington. Chapter 9: 225-246.

Grant W.S., Leslie R.W., Becker I.I. 1988a. Genetic stock structure of the southern African hakes, *Merluccius capensis* and *M. paradoxus*. Mar. Ecol. Prog. Ser. In press.

Grant W.S., Becker I.I., Leslie R.W. 1988b. Evolutionary divergence between sympatric species of southern African hakes, *Merluccius capensis* and *M. paradoxus*. I. Electrophoretic analysis of proteins. Heredity. In press.

Graves J.E., Ferris S.D., Dizon A.E. 1984. Close genetic similarity of Atlantic and Pacific skipjack tuna (*Katsuwonus pelamis*) demonstrated with restriction endonuclease analysis of mitochondrial DNA. Mar. Biol. 79: 315-319.

Greenberg B.D., Newbold J.E., Sugino A. 1983. Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. Gene 21: 33-49.

Gyllensten U., Leary R.F., Allendorf F.W., Wilson A.C. 1985. Introgression between two cutthroat trout subspecies with substantial karyotypic, nuclear and mitochondrial genomic divergence. Genetics 111: 905-915.

Gyllensten U., Wilson A.C. 1987. Mitochondrial DNA of salmonids Inter- and Intraspecific variability detected with restriction enzymes. In Population genetics and Fishery Management. edited by N. Ryman and F. Utter Washinton Sea Grant Prog., Univ. Washington Press, Seattle, Washington. Chapter 12: 301-317.

Harris H., Hopkinson D.A. 1976. Handbook of Enzyme Electrophoresis in Human Genetics. Elsevier, Amsterdam.

Harrison R.G., Rand D.M., Wheeler W.C. 1985. Mitochondrial DNA Size variation within individual Crickets. Science 228: 1446-1447.

Hauswirth W.W., Laipis P.J. 1982. Mitochondrial DNA polymorphism in a maternal lineage of Holstein cows. Proc. Natl. Acad. Sci. USA 79: 4686-4690.

Inada T. 1981. Studies on the Merlucciid Fishes. Bull. Far Seas Fish. Res. Lab. 18: 172 pp.

Inada T., Takeda M., Hatanaka H. 1986. Important Fishes Trawled off Patagonia. Japan Marine Fisheries Resource Research Center.

Ish-Horowicz D., Burke J.F. 1981. Rapid and efficient cosmid vector cloning. Nucleic Acids Res. 9: 2989.

Johnson M.J., Wallace D.C., Ferris S.D., Rattazzi M.C., Cavalli-Sforza L.L 1983. Radiation of human mitochondrial DNA types analyzed by restriction endonuclease cleavage patterns. J. Mol. Evol. 19: 255-271.

Jones B.W., Van Eck T.H. 1967. The cape hake: its biology and the fishery. S. Afr. Shipp. News Fishg Ind. Rev. 22(11): 90-97.

Jones B.W., Mackie I.M. 1970. An application of electrophoretic analysis of muscle myogens to taxonomic studies in the genus *Merluccius*. Comp. Biochem. Physiol. 32: 267-273.

Kabata Z., Ho Ju-Shey 1981. The Origin and Dispersal of Hake (Genus Merluccius: Pisces: Teleostei) as indicated by its Copepod Parasites. Oceanogr. Mar. Biol. Ann. Rev., 19, 381-404.

Kearsey S.E., Craig I.W. 1982. Genetic basis of chloramphenicol resistance on mouse and human cell lines. In Mitochondrial genes edited by Slonimski, Borst, Attardi, CSH 117-120.

Kessler L.G., Avise J.C. 1985a. Microgeographic lineage analysis by mitochondrial genotype: variation in the cotton rat (*Sigmodon hispidus*). Evolution 39(4): 831-837.

Kessler L.G., Avise J.C. 1985b. A comparative description of mitochondrial DNA differentiation on selected avian and other vertebrate genera. Mol. Biol. + Evol. 2(2): 109-125.

King P.V. and Blakesley R.W. 1986. Optimizing DNA Ligations for Transformation. Focus 8(1): 1-3.

Kornfield I., Bogdanowicz S.M. 1987. Differentiation of mitochondrial DNA in Atlantic herring. Fish. Bull. USA (in press).

Lansman R.A., Shade R.O., Shapira J.F., Avise J.C. 1981. The use of restriction endonucleases to measure Mitochondrial DNA sequence relatedness in natural populations. J. Mol. Evol. 17:214-226.

Lansman R.A., Avise J.C., Huettel M.D. 1983a. Critical experimental test of the possibility of "paternal leakage" of mitochondrial DNA. Proc. Natl. Acad. Sci. USA 80: 1969-1971.

Lansman R.A., Avise J.C., Aquadro C.F., Shapira J.F., Daniel S.W. 1983b. Extensive genetic variation in mitochondrial DNA's among geographic populations of the deer mouse, *Peromyscus maniculatus*. Evolution 37(1): 1-16.

Lewin R. 1987a. Africa: Cradle of modern humans. Science 237: 1292-1295.

Lewin R. 1987b. The unmasking of Mitochondrial Eve. Science 238: 24-26.

Lindahl K.F. 1985. Mitochondrial inheritance in mice. TIG May 135-139.

Mack A.L., Gill F.B., Colburn R., Spolsky C. 1986. Mitochondrial DNA: A source of genetic markers for studies of similar passerine bird species. The Auk 103: 676-681.

Mangaly G., Jamieson A. 1978. Genetic tags applied to European hake, Merluccius merluccius(L.)[1]. Anim. Blood Groups Biochem. Genet 9: 39-48.

Maniatis T., Fritsch E.F., Sambrook J. 1982. Molecular Cloning. Cold Spring Harbor Lab., New York.

Markert C.L., Faulhaber I. 1965. Lactate dehydrogenase isozyme patterns of fish. J. Exp. Zool. 159: 319-332.

May B., Wright J.E., Stoneking M. 1979. Joint segregation of biochemical loci in Salmonidae: results from experiments with *Salvelinus* and review of the literature on other species. J. Fish. Aquat. Sci. 36: 114-1128.

Mayr E. 1982. Processes of Speciation in Animals. In Mechanisms of Speciation edited by C. Barigozzi. Alan R. Liss, Inc., New York.

Metzler D.E. 1977. Biochemistry The chemical reactions of living cells. Academic Press.

Mikelsaar R. 1987. A view of early cellular evolution. J. Mol. Evol. 25: 168-183.

Monnat R.J., Loeb L.A. 1985. Nucleotide sequence preservation of human mitochondrial DNA. Proc. Natl. Acad. Sci. USA 82: 2895-2899.

Montoya J., Ojala D., Attardi G. 1981. Distinctive features of the 5'-terminal sequences of the human mitochondrial mRNAs. Nature 290: 465-470.

Mork J., Ryman N., Stahl G., Utter F., Sundes G. 1985. Genetic variation in Atlantic cod (*Gadus morhua*) throughout its range. Can. J. Fish Aquat. Sci. 42: 1580-1587.

Moritz C., Brown W.m. 1986. Tandem duplication of D-loop and Ribosomal RNA Sequences in Lizard mitochondrial DNA. Science 233:1425-1427.

Nei M. 1972. Genetic distance between populations. Am. Nat. 106: 283-292.

Nei M. 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. (USA) 70: 3321-3323.

Nei M., Roychoudhury A.K. 1974. Sampling variances of heterozygosity and genetic distance. Genetics 76: 379-390.

Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583-590.

Nei M., Li W-H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 76(10): 5269-5273.

Nei M., Tajima F. 1985. Evolutionary change of restriction cleavage sites and phylogenetic inference for man and apes. Mol. Biol. + Evol. 2(3): 189-205.

Ojala D., Montoya J., Attardi G. 1981. TRNA punctuation model of RNA processing in human mitochondria. Nature 290: 470-474.

Parker R.C., Watson R.M., Vinograd J. 1977. Mapping of closed circular DNAs by cleavage with restriction endonucleases and calibration by agarose gel electrophoresis. Proc. Natl. Acad. Sci. USA 74(3):851-855.

Pchenitchny B.P., Assorov V.V. 1969. Some biological characteristics of hake (*Merluccius*) of the Atlantic Ocean near the south west coast of Africa. Vopr. Ikhtiol. 9(3): 423-430.

Potter S.S., Newbold J.E., Hutchison III C.A., Edgell M.H. 1975. Specific cleavage analysis of mammalian mitochondrial DNA. Proc. Natl. Acad. Sci. USA 72(11): 4496-4500.

Powell J.R., and Zuniga M.C. 1983. A Simplified Procedure for studying mtDNA Polymorphisms. Biochem. Genet. 21: 1051-1055.

Reed K.C. 1987. Alkaline southern blotting to zeta-probe membrane. Bio-Rad Laboratories, 1414 Habour Way South, Richmond.

Ridgway G.J., Sherburne S.W., Lewis R.D. 1970. Polymorphism in the esterase of Atlantic herring. Trans. Am. Fish. Soc. 99: 147-151.

Rigby P.W. J., Dieckmann M., Rhodes G., Berg P. 1977. Labelling DNA to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. 113: 327-251.

Roe B.A., Wong J.F.H., Chen E.Y., Armstrong P.W., Stankiewicz A., Ma D.P., McDonough J. 1982. Mammalian mitochondrial tRNAs: A modified nucleotide 3' to the anticodon may modulate their codon response. In Mitochondrial genes edited by Slonimski, Borst, Attardi, CSH 45-51.

Roe B.A., Ma D.P., Wilson R.K., Wong J.F.H. 1985. The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome. J. Biol. Chem. 260(17): 9759-9774.

Roman H. 1982. Boris Ephrussi and the early days of cytoplasmic inheritance in *Sachromyces*, 1. In Mitochondrial genes edited by Slonimski, Borst, Attardi, CSH 1-4.

Sanger F., Nicklen S., Coulson A.R. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467.

Saunders N.C., Kessler L.G., Avise J.C. 1986. Genetic variation and geographic differentiation in mitochondrial DNA of the horseshoe crab, *Limulus polyphemus*. Genetics 112: 613-627.

Shaklee J.B., Whitt G.S. 1981. Lactate dehydrogenase isozymes of Gadiform fishes: divergent patterns of gene expression indicate a heterogeneous taxon. Copeia 3: 563-578.

Shaklee J.B., Tamaru C.S., Waples R.S. 1982. Speciation and evolution of marine fishes studied by the electrophoretic analysis of proteins. Pac. Sci 36: 141-157.

Shaklee J.B. 1983. The utilization of isozymes as gene markers in fisheries management and Conservation. Isozymes: Current topics in biological and medical research. Alan R. Liss Inc., New York. 11: 213-247.

Skibinski D.O.F. 1985. Mitochondrial variation in *Mytilus edulis* and the Padstow mussel. J. Mar. Biol. Ecol. 92: 251-258.

Sneath P.A.H., and Sokal R.R. 1973. Numerical taxonomy. Freeman and Co., San Francisco, CA.

Sokal R.R., and Rohlf R.J. 1981. Biometry. 2nd. Ed., Freeman and Co., San Francisco, CA. 859 p.

Solignac M., Monnerot M., Mounolou J-C. 1983. Mitochondrial DNA heteroplasmy in *Drosophila mauritiana*. Proc. Natl. Acad. Sci. USA 80: 6942-6946.

Solignac M., Monnerot M., Mounolou J.C. 1986. Mitochondrial DNA evolution in the *melanogaster* species subgroup of *Drosophila*. J. Mol. Evol. 23: 31-40.

Spolsky C., Uzzell T. 1984. Natural interspecies transfer of mitochondrial DNA in amphibians. Proc. Natl. Acad. Sci. USA 81: 5802-5805.

Stent G.S., Calendar R. 1978. Molecular Genetics an introductory narrative. W.H. Freeman and Company, San Francisco.

Sventovidov A.N. 1940. The problem of distribution of Gadiae and of the other families of Gadiformes. Bull. Soc. Nat. Moscow, Sec. Biol. 94(1): 50-60.

Sventovidov A.N. 1948. Fauna of the U.S.S.R. Fishes. Gadiformes. Zool. Inst. Akad. Nauk. (English translation) 9(4): 221p.

Swofford D.L., and Selander R.B. 1981. Biosys-1: a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. Heredity 72: 281-283.

Szidat L. 1955. La fauna de parasito de *Merluccius hubbsi* como caracter auxiliar para la so-lucion de problemas sistematicos y zoogeograficos del genero Merluccius. Commun. Inst. Nac. Inv. Cs. Nat., Buenos Aires 3: (1), 1-54.

Szidat L. 1961. Versuch einer Zoogeographie des Süd-Atlantik mit Hilfe von Leitparasiten der Meeresfische. Parasitol. Schriftenr. H. 13: 1-97.

Tajima F., Nei M. 1982. Biases of the estimates of DNA divergence obtained by the restriction enzyme technique. J. Mol. Evol. 18: 115-120.

Templeton A.R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37(2): 221-244.

Thorpe J.P. 1982. The molecular clock hypothesis: biochemical evolution genetic differentiation and systematics. Ann. Rev. Ecol. Syst. 13: 139-168.

Upholt W.B. 1977. Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. Nucleic Acids Res. 4(5): 1257-1265.

Upholt W.B., Dawid I.B. 1977. Mapping of Mitochondrial DNA of Individual Sheep and Goats: Rapid Evolution in the D loop Region. Cell 11: 571-583.

Utter F.M., Hodgins H.O. 1969. Lactate dehydrogenase isozymes of Pacific Hake (*Merluccius productus*). J. Exp. Zool. 172: 59-68.

Utter F.M., Stormont C.J., Hodgins H.O. 1970. Esterase polymorphism in vitreous fluid of Pacific hake, *Merluccius productus*. Anim. Blood Groups. Biochem. Genet. 1: 69-82.

Utter F.M., Hodgins H.O. 1971. Biochemical polymorphisms in the Pacific hake (*Merluccius productus*). Internat. Council Explor. Sea. 161: 87-89.

Utter F.M., Hodgins H.O., Allendorf F.W., Johnson A.G., Mighell J.L. 1973. Biochemical variants in Pacific salmon and rainbow trout: their inheritance and application in population studies. Genetics and Mutagenesis of Fish. Springer-Verlag, Berlin.

Van Eck T.H. 1969. The South African Hake: "*Merluccius capensis*" or "*Merluccius paradoxus*"? S. Afr. Shipp. News Fishg Ind. Rev. 24(5): 5-6.

Van Etten R.A., Michael N.L., Bibb M.J., Brennike A., Clayton D.A. 1982. Expression of the mouse mitochondrial DNA genome. In Mitochondrial genes edited by Slonimski, Borst, Attardi, CSH 73-88.

Vawter A.T., Rosenblatt, Gorman G.C. 1980. Genetic divergence among fishes of the Eastern Pacific and the Caribbean: Support for the molecular clock. Evolution 34: 705-711.

Vawter L., Brown W.M. 1986. Nuclear and mitochondrial DNA comparisons reveal extreme rate Variation in the molecular clock. Science 234: 194-196.

Wallace D.C. Oliver N.A., Blanc H., Adams C.W. 1982. A system to study human mitochondrial genes: application to chloramphenicol resistance. In Mitochondrial genes edited by Slonimski, Borst, Attardi, CSH 105-117.

Watson J.D., Hopkins N.H., Roberts J.W., Steitz J.A., Weiner A.M. 1987. Molecular Biology of the gene. Fourth edition The Benjamin/Cummings Publishing Company.

Whitt G.S. 1970. Developmental genetics of the lactate dehydrogenase isozymes of fish. J. Exp. Zool. 175: 1-35.

Wilson A.C., Cann R.L., Carr S.M., George M., Gyllensten U.B., Helm-Bychowski K.M., Higuchi G.R., Palumbi S.R., Prager E.M., Sage R.D., Stoneking M. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linn. Soc. 26: 375-400.

Wolsten-holme D.R., Fauron C.M-R., Goddard J.M. 1982. Nucleotide sequence of *Rattus norvegicus* mitochondrial DNA that includes the genes for $tRNA^{ile}$, $tRNA^{gln}$ and $tRNA^{f-met}$. Gene 20: 63-96.

Wong J.F.H., Ma D.P., Wilson R.K., Roe B.A. 1983. DNA sequence of the *Xenopus laevis* mitochondrial heavy and light strand replication origins and flanking tRNA genes. Nucleic Acids Res. 11(14): 4977-4995.

Wright S. 1943. Isolation by distance. Genetics 28: 114-138.

Yonekawa H., Moriwaki K., Gotoh O., Hayashi J.I., Watanabe J., Miyashita N., Petras M., Tagashira Y. 1981. Evolutionary relationships among five subspecies of *Mus musculus* based on restriction enzyme cleavage patterns of mitochondrial DNA. Genetics 98: 801-816.

APPENDIX

GENERAL MEDIA, BUFFERS, AND SOLUTIONS

GENERAL MEDIA, BUFFERS AND SOLUTIONS

Ampicillin
Make up a stock solution of 25 mg/ml of the sodium salt of ampicillin in water. Filter sterilize and store at -20°C.

ATP (0.1 M)
Dissolve 60 mg of ATP in 0.8 ml of H_2O. Adjust th pH to 7.0 with 0.1 M NaOH and the volume to 1.0 ml with H_2O. Store at -70°C.

10x HindIII buffer
600 mM NaCl
67 mM $MgSO_4$
100 mM Tris-HCl pH 7.9
5 mM β-mercaptoethanol
Store at -20°C.

Hybridisation buffers

Church Hybridisation buffer (CHB)
1% Bovine Serum Albumin (Sigma)
7% SDS
0.5 M Na_2PO_4 pH 7.0 (PB)
1 mM EDTA

Wash Buffer A (WBA)
0.5% Bovine Serum Albumin
5% SDS
40 mM PB
1 mM EDTA

Wash Buffer B (WBB)
1% SDS
40 mM PB
1 mM EDTA
Make up stock solutions (10% BSA, 20% SDS, 1M NaH_2PO_4, pH 7.0, and 0.5 M EDTA) and store separately.

DNase free RNase
Dissolve pancreatic RNase (Sigma) in 0.1 M Sodium acetate pH 4.5. Boil for 10 min and store in small aliquots at 70°C.

Denhardt's solution (1x)

0.02 % ficoll (Sigma)
0.02 % polyvinylpyrolidine
0.02 % bovine serum albumin
Make up a 50x stock solution and store in small aliquots at
-20°C.

DNA extraction solutions

Solution I
50 mM Glucose
25 mM Tris pH 8.0
10 mM EDTA
adjust pH with HCl, autoclave and store at 4°C.

Solution II
0.2 N NaOH
1 % SDS
Make up stock solutions separately.
10 N NaOH
25% SDS
Autoclave stock solutions and store at room temperature. Take 2 ml of NaOH stock solution and 4 ml of SDS stock solution and add 94 ml H_2O. Mix solution weekly and keep at room temperature.

Solution III
5 M Potassium acetate, pH 4.8
Made from 2 M acetic acid and 3 M potassium acetate. Make up potassium acetate and dissolve in H_2O, add acetic acid until pH 4.8. Autoclave and store at 4°C.

Dithiothreitol DTT

Dissolve DTT (Boehringer, Mannheim) in sodium acetate, pH 5.2 and sterilize by filtration. Store small aliquots at -20°C.

Dialysis Tubing

Cut the tubing into pieces of convenient length. Boil for 10 min in a large volume of 2% sodium bicarbonate and 1 mM EDTA. Rinse the tubing thoroughly in distilled water, boil for 10 min in distilled water or autoclave. Allow to cool and store at 4°C. Before use, wash the tubing inside and out with distilled water. Handle the tubing with gloves at all times.

Denatured salmon sperm DNA
Dissolve DNA in water at concentration of 10 mg/ml. Stir 2-4 hours to dissolve and then shear DNA by passing it several times through a 18-gauge hypodermic needle. Boil for 10 min and store at -20°C in small aliquots. Before use heat the DNA for 5 min in a boiling water bath and chill on ice.

Ethidium Bromide - EtBr
Make up a stock solution of 10 mg/ml Ethidium Bromide (Sigma) in H_2O. Stir for several hours to make sure that the dye has dissolved. Wrap the container in aluminium foil and store at 4°C.

Gel- loading buffer
0.25% bromophenol blue
0.25% Xylene cyanol
25% Ficoll (type 400) in H_2O
Store at room temperature

5x Ligation buffer
250 mM Tris-Cl, pH 7.6
50 mM $MgCl_2$
5 mM DTT
25% (w/v) polyethylene glycol 8000
5 mM ATP
Make up ATP separately and store both solutions at -20°C or -70°C.

Luria broth
10 g Bacto tryptone
5 g Bacto yeast extract
10 g NaCl
Make up to 1 ℓ with distilled water.

Luria agar
Luria broth containing 15 g Bacto nutrient agar.

Phenol preparation
Melt Phenol at 68°C and add 8-Hydroxyquinoline (Merck, Darmstadt) to a final concentration of 0.1%. 8-Hydroxyquinoline is a partial inhibitor of RNase, a weak chelator of metal ions and an antioxidant. The melted phenol is extracted several times with an equal volume of 1.0 M Tris-HCl, pH 8.0 followed by 0.1 M Tris-HCl, pH 8.0 and 0.2% β-mercaptoethanol until the pH of the aqueous phase is above 7.6. The phenol is stored at 4°C under equilibration buffer.

Restriction endonuclease digestion buffers

Low-salt buffer
10 mM Tris-HCl, pH 7.5
10 mM $MgCl_2$
1 mM dithiothreitol

Medium-salt buffer
50 mM NaCl
10 mM Tris-Cl, pH 7.5
10 mM $MgCl_2$
1 mM dithiothreitol

High-salt buffer
100 mM NaCl
50 mM Tris-HCl, pH 7.5
10 mM $MgCl_2$
1 mM dithiothreitol

Sequencing Solutions

10x Sequencing buffer
0.5 M NaCl
0.1 M Tris-HCl pH 7.5
0.1 M $MgCl_2$
0.001 Dithiothreitol

Chase solution
A 0.5 M uniform mix of all four dNTPs
Store at -20°C

Stop buffer
100 ml formamide
5 g Amberlite ™ MB1 (Mixed Bed resin)
Stir gently for 30 min and then filter to remove resin.
Add 0.1 g xylene cyanol FF
0.1 g Phenol red
4 ml 500 mM Na_2EDTA
Store at room temperature. Keep for 1 month maximum

Acrylamide stock solution

40% acrylamide stock
38 g acrylamide
2 g bis-acrylamide (NN'-methylenebisacrylamide)
Make up to 100 ml with distilled water.
Add 5 g Amberlite MB1 and stir gently for 30 min.
Filter through a sintered glass filter to remove resin.
Remove all fine particles with millipore ™filter (0.45 μm) and store in the dark at 4°C.

SSC
0.15 M NaCl
0.015 M Trisodium Citrate
Make up a 20 x stock and autoclave

STE buffer
10 mM Tris-Cl, pH 8.0
100 mM NaCl
1 mM EDTA

Starch gel buffers

MF buffer
Electrode: 0.18 M Tris, pH 8.7
0.1 M Boric acid
0.004 M NaEDTA (Markert and Faulhaber, 1965)
gel: 1:4 dilution of electrode buffer

RW buffer
Electrode: 0.06 M lithium hydroxide
0.3 M boric acid, pH 8.1
Gel: 0.03 M Tris, pH 8.5
0.005 M citric acid
(Ridgway et al., 1970)

TC buffer
Electrode: 0.15 M Tris, pH6.9
0.05 M citric acid
Gel: 1:15 dilution of electrode solution
(Whitt, 1970)

Southern blotting buffers

Denaturing solution
1.5 M NaCl
0.5 M NaOH

Neutralisation solution
1.5 M NaCl
0.5 M Tris-HCl, pH 7.2
0.001 M Na_2EDTA

Transfer buffer for electro blotting

0.025 M Sodium phosphate, pH 6.5
Use 16 ml 0.5 M Na_2HPO_4 and 34 ml 0.5 M NaH_2PO_4 per 50 ml.

TEK buffer

50 mM	Tris pH 7.5
10 mM	EDTA
1.5 %	KCl

Tris-EDTA buffer - TE

10 mM Tris-Cl, pH 8.0
1 mM EDTA

TBE Gel buffer

8.9 mM Tris-borate, pH 8.0
8.9 mM Boric acid
0.2 mM EDTA

Tris-acetate gel buffer

40 mM Tris-Cl, pH 7.8
5 mM Sodium acetate
1 mM EDTA

2x TSM

0.06 M Tris-HCl, pH 7.4
0.3 M NaCl
3.0 mM $MgCl_2$

IPTG

Isopropyl-β-D-thio-galactopyranoside (Sigma, St. Louis)
Make up a 100 mM solution in H_2O.
Store small aliquots at -70°C.
40 µℓ is added to each petridish.

X-Gal

5-Bromo-4-chloro-3-indoyl-β-galactoside (BRL Research Laboratories)
Make up a 2% solution in dimethyl formide.
Store small aliquots at -70°C.
40 µℓ is added per petridish.