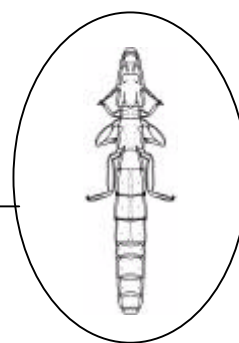


Mitochondrial DNA analysis of
intraspecific variation in the
Phthirapteran chewing louse,
Halipeurus pelagicus from
Madeiran storm petrels,
Oceanodroma castro.



Honours Project Report, March 2001
Katherine Sumner - 9705658

1. Introduction

1.1. Background

The Madeiran storm petrel, *Oceanodroma castro*, is a small Procellariiform seabird (6-8 inches long) belonging to the family Hydrobatidae. It has dark plumage with a narrow, rectangular, white rump patch (Fig 1.) giving them their alternative name, the band-rumped storm petrel. They are pelagic seabirds that spend the majority of their time at sea, only returning to land to breed. They feed during the day and night but only return to the breeding colony at night due to the threat of gulls and other predators during the day. Nests are made in lava tubes, cavities between rocks and burrows in the ground.



Fig 1. Picture of *Oceanodroma castro* showing its rectangular, white rump patch (Source, www.bham.net/aos/pelagicphotos).

O. castro is found in subtropical areas of the Pacific and Atlantic Oceans (Fig 2), with breeding colonies off Japan, Hawaii, Galapagos, off Portugal, Madeira, Salvages,

Canary Islands, Cape Verde Islands, Ascension and St Helena islands and the Azores (Monteiro & Furness, 1998).



Fig 1. Range of *Oceanodroma castro*. The shaded areas indicate the at sea distribution while breeding colonies are marked by black dots (Source, www.bird-stamps.org/species/12011)

It was noticed that at some of these breeding colonies there appeared to be two breeding seasons: one in the spring (hot-season) and one in the autumn (cool-season) with birds time-sharing the same nest sites. A study of mercury concentrations in the plumage of *O. castro* (Monteiro et al., 1995) was the first evidence that two discrete breeding populations existed. Storm petrels breeding in the hot-season were found to have significantly lower levels of mercury in their plumage than those breeding in the cool-season, indicating a difference in ecology. A subsequent study confirmed the presence of two breeding populations, one breeding in the spring, the other in the autumn, with an overlap during August and early September (Monteiro & Furness, 1998). Hot-season breeders were found to be smaller, lay smaller eggs, have longer wings and longer and deeper forked tails than cool-season breeders. Their chicks were also found to be 15% lighter than those of cool-season breeders. Monteiro & Furness (1998) believe the birds have undergone sympatric speciation. Sympatric speciation occurs when “reproductive isolation takes place within a single population without the help of geographical separation” (Page & Holmes, 1998). In the case of *O. castro*, the

mechanism of reproductive isolation is differences in the breeding season. There is little or no switching between the seasons, with only one case of movement between the colonies having been recorded (Monteiro & Furness, 1998). The favoured breeding season appears to be the cool-season but it is hypothesised that overcrowding of breeding sites led to unfavourable breeding conditions and caused the displacement of some of the birds to the hot-season (Monteiro & Furness, 1998). The presence of two breeding populations has been confirmed by the discovery of genetic differences in the mitochondrial (mt) control region between birds from the hot- and cool-season (Lodha *et al.*, unpublished data)

Results from these previous studies have been the basis for this project which involves DNA analysis of the phthirapteran feather louse, *Halipeurus pelagicus* (Fig 3), to investigate whether the genetic differences found between the two populations of host storm-petrels also occur in their parasitic lice. This would be the first stage in the process known as cospeciation. As early as 1896 Kellogg began examining the evolutionary relationships of avian hosts and their parasitic lice. Since then, scientific advances, particularly in the field of genetics, have enabled in depth studies of these fascinating associations. Much recent research has focused on cospeciation of pocket gophers and their parasitic lice (Hafner & Nadler, 1998; Page, 1996; Page and Hafner 1996). Lice provide the clearest evidence for host-parasite cospeciation (Hafner and Nadler, 1988). They are likely to coevolve with their hosts because they are nonpathogenic, complete their whole life cycle on a host, generally dying within 1 to 2 days if removed (Paterson *et al.*, 1993) therefore reducing the likelihood of host-switching, and they also tend to be very host specific (Furness & Palma, 1992). This host specificity suggests a long evolutionary history with birds (Smith, 2000). This

study looks for evidence that such a relationship is developing between Madeiran storm-petrels and one of their parasitic lice, *H. pelagicus*.

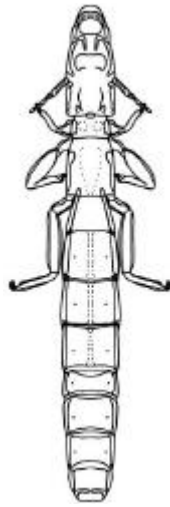


Fig 3. *Halipeurus pelagicus* (Ischnocera: Philopteridae). Actual adult size is around 3-4mm long. (Source- Smith, 2000)

If genetic variation between breeding seasons is not found there are two possible explanations. The first is simply that there is no genetic variation among lice from the two seasons and therefore there is no indication that cospeciation will occur. Alternatively, genetic variation may be present but not detectable by the methods used in this study.

An additional possibility to variation between breeding seasons is geographical variation between parasites from birds at different breeding colonies. Table 1. shows the breeding populations for genetic analysis involved in this study for which samples were available.

Island	Hot	Cold
Graciosa (Azores)	✓	✓
Selvages (Madeira)	✓	-
Raso (Cape Verde)	✓	-
Branco (Cape Verde)	✓	-
Santa Maria (Azores)	-	✓
Desertas (Madeira)	✓	✓

Table 1. Colonies and breeding seasons involved in the study.

Of these, the islands of the Azores are furthest north and the Cape Verde islands the furthest south (Fig 4). Studies of *Calanthes* beetles (Emerson *et al.*, 1999) from the Canary islands, and North American Rubber Boas (Rodrigues-Robles *et al.*, 2000) have shown genetic variation corresponding to geographical variation. However, these species are relatively immobile compared to avian lice (which could easily be transported between islands by their highly mobile hosts) and therefore represent examples of allopatric speciation. Allopatric speciation is defined as “reproductive isolation caused by geographical isolation of populations perhaps by natural barriers such as mountains and rivers, so that gene flow between them is prevented and genetic divergence occurs” (Page and Holmes, 1998). Therefore, geographical variation may be unlikely as the populations of lice may not be isolated.



Fig 4. Map showing the relative situations of the study islands

There is also evidence that although there is phenotypic and genetic variation between storm-petrels breeding in hot and cool seasons, there is little variation between same season breeders from different colonies (Monteiro & Furness, 1998). However,

storm-petrels and other pelagic seabirds tend to be highly faithful to their natal colony and rarely move to breed in another location, so that the behaviour of the hosts might reduce possibilities for gene flow among lice from different localities.

By using mtDNA analysis it is possible to compare homologous genes from dissimilar species such as birds and lice, which share few morphological features (Page *et al.*, 1998; Smith, 2000). Lice themselves are not particularly suitable for morphological analysis, as generally they are very similar (Clayton *et al.*, 1996) as shown by the similar appearance of *Philoceanus* and *Halipeurus*, which are separate genera. Therefore, it was predictable that there would be little or no morphological difference between hot- and cool-season lice (but see Ramli *et al.*, 2000). MtDNA is favourable for use in genetic studies as it shows extensive interspecific variation and is thought not to undergo genetic recombination (Awise, 2000). The mt genome consists of 37 functionally distinct regions (Fig 5.), with 12S, Cytochrome Oxidase 1 (CO1) and the control region being the focus of this study.

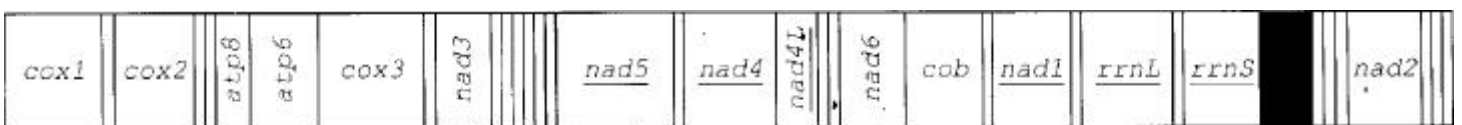


Fig 5. Gene arrangement of ancestral insect mtDNA. Cytochrome oxidase subunits 1-3 (cox1-3), ATP synthase subunits 6 and 8 (atp6, atp8), NADH dehydrogenase subunits 1-6 and 4L (nad1-6, nadh4L), 12S and 16S (rrnS, rrnL), and 22 tRNA genes, Control region represented by shaded black area. (Taken from Shao, unpublished data)

The 12S and CO1 regions are relatively slowly evolving genes but have been successfully sequenced for lice before (Hafner *et al.*, 1994). The control region, the region where genetic variation was detected among *O. castro* (Lodha, unpublished data), is particularly useful for detecting intraspecific variation because it has a fast

pace of nucleotide substitution and a high level of intraspecific polymorphism (Avice, 2000). It also has suggested clock calibrations 3-5 times greater than the rest of the mt genome (Avice, 2000), meaning it is a faster evolving region and would be more likely to reveal any genetic variation in closely related individuals.

1.2. Summary of Aims

The aim of the project is to use mitochondrial DNA analysis to study intraspecific variation of *Halipeurus pelagicus*, a feather louse colonising the Madeiran storm-petrel, *Oceanodroma castro*. The two possible types of variation to be investigated are;

1. Genetic variation between lice from hot- and cool-season breeders, corresponding to the genetic variation discovered in *O. castro*
2. Genetic variation between lice from different island populations.

2. Materials and Methods

2.1. Louse Sampling

The collection of the lice was carried out during a previous study of *Oceanodroma castro* involving ornithologists from Portugal and the UK, led by Dr. L.R. Monteiro and Professor R.W. Furness, who provided information on the sampling methods used. The birds were caught by use of fine nylon mesh mist nets suspended from vertical poles. Each net is approximately 10m long by 3m high, with four shelf strings, which create a pocket in which the birds hitting the net are collected. The nets are set up during the day at known nest site locations to catch the birds, which only return to land during darkness, as they visit or explore nest sites. Teams of 2-6 people removed the birds from the net pockets, placed them into cloth bags and took them to a central processing point. Here species identification was checked from plumage characteristics and birds were ringed with individually numbered metal rings, weighed and measured. Two different methods were used to collect lice from the birds; manual examination or the use of a 'delousing chamber'.

Manual examination consists of running a finger through the feathers, exposing the underside of each one in turn, concentrating firstly on the underwing covert feathers, an area where the greatest number of lice tend to be found. The tail feathers, body feathers, neck and head feathers were then examined. When a louse was located it was removed using insect forceps and placed directly into an Eppendorf tube containing 95% alcohol. Each tube contained all lice collected from a single bird. This method proved impractical and time consuming when large numbers of birds were being caught so for large numbers 'delousing chambers' were used.

The 'delousing chamber' was a small plastic bucket in which the birds were suspended by the neck in a soft rubber gasket (made from car inner tube). A few drops of chloroform were placed into the bucket. The chloroform vapour caused ectoparasites to fall off the bird and collect in the bucket. After around 5 minutes the birds were removed and released, and the ectoparasites collected and placed into Eppendorf tubes containing 95% alcohol. This method did not sample lice from the head. Several birds were sampled simultaneously in separate 'delousing chambers', allowing quicker collection. Both methods of collection were used to obtain a reasonable number of lice for study and are not quantitative.

The tubes were placed in zip-lock bags, labelled with the sampling date and colony (island) name. The sampling date was used to determine whether the birds were from hot- or cool-season breeding populations. The sample was labelled uncertain if it was taken from a time of year when both breeding populations may have been present. All samples were stored at ambient temperature and taken back to Glasgow University, where they were stored in a 4°C fridge.

2.2. DNA Extraction

Using a compound microscope a large, adult *H. pelagicus* specimen was selected from a collection tube which was then labelled with the site and sample number e.g. CSM1 for the first bird sampled from the Cold breeding population from the island of Santa Maria. Initially, three lice were sampled from three different birds in each population¹. A large, adult louse was used to avoid possible misidentification of juvenile instars and to maximise DNA yield. Initial identification was carried out

¹ Sampling carried out by myself, Rob Cruickshank and Gareth Hood (M Res).

with the assistance of Vince Smith, PhD. *Halipeurus* looks very similar to the louse *Philoceanus*, which is also found on *O. castro*. They can be distinguished by the presence of a continuous ventral carina in *Philoceanus* and discontinuous ventral carina in *Halipeurus* (Smith, 2000).

Latex gloves were worn throughout the remainder of the experimental work. This was to help prevent contamination of the samples. However, since the primers used were insect-specific, it was unlikely that contamination would affect the extraction. Negative controls were included where appropriate to monitor the possibility of contamination.

The selected louse was placed on a glass slide and a scalpel used to detach the head from the body. The extraction was carried out, with a few modifications, following the Qiagen DNeasy protocol for animal tissue (April, 2000). Both parts of the louse were added to a 1.5ml micro-centrifuge tube containing 180µl Buffer ATL and 20µl Proteinase K, vortexed, and placed in a 55°C water bath for two nights. These tubes were also labelled with the site and sample number. Proteinase K was added to break down the extracted proteins which, if not removed, would break down the DNA. An experimental negative was included that contained Buffer and Proteinase K, but lacked a louse.

After two nights the tubes were removed from the water bath. The liquid was siphoned off using a pipette into another labelled micro-centrifuge tube and vortexed. 50µl of water was added to the tubes containing the lice, which were then stored in the refrigerator. The louse parts were retained so identification could be checked if

unusual results were obtained for a particular specimen and for eventual slide mounting as voucher specimens. 200µl of Buffer AL was added to the extraction, mixed by vortexing and incubated in a 70°C waterbath for 10 minutes, or until the white precipitate formed with the addition of the buffer had dissolved. 200µl of 100% ethanol was added to precipitate the DNA and again mixed thoroughly by vortexing. The 600µl of solution was pipetted into the DNeasy mini column sitting in a 2-ml collection tube provided in the kit and centrifuged at 8000rpm for 1 minute. The flow-through tube was discarded, since the DNA was now bound to the filter of the column, and the mini column placed into a new collection tube. 500µl of two high salt concentration wash buffers, AW1 and AW2, were used in turn to remove any remaining proteins and other contaminants that could inhibit the polymerase chain reaction. 50µl of distilled water was then pipetted directly onto the centre of the DNeasy filter and incubated at room temperature for thirty minutes before centrifuging. This extended period of time helps maximise the yield of DNA. The addition of water causes the purified DNA that was bound to the filter to be released and upon centrifuging, the DNA is collected in the microcentrifuge tube. The final DNA extraction was then stored in a -70°C freezer.

2.3. Polymerase Chain Reaction (PCR)

A PCR is carried out to generate multiple copies of the desired sequence of the DNA from small amounts of extracted DNA. Primers specifically designed to bind to the target sequence anneal at low temperature to single DNA strands that have been separated at high temperature. Taq polymerase then synthesises the target DNA at an intermediate temperature and as newly generated sequences can act as templates, the yield of DNA doubles during each cycle. Taq polymerase is obtained from the hot

spring-dwelling bacterium *Thermophilus aquaticus* and is therefore able to withstand the high temperatures reached during a PCR that would denature most other enzymes.

PCR reactions contained 2.5µl of specimen DNA, 9.875µl of water, 2.5µl of Mg free Buffer, 3.5µl of MgCl₂, 2.5µl of free deoxyribonucleotides (dNTPs), 2.0µl of forward and reverse primers, and 0.125µl of Taq Polymerase. Six PCR reactions were carried out for each sample in order to obtain enough DNA for sequencing. A PCR negative was also included that contained 2.5µl of water instead of DNA. PCR amplification of the 12S mtDNA region was carried out using the insect primers 12Sai and 12Sbi (Simon *et al.*, 1994). The primers L6625 and H7005 were used for CO1 (Hafner *et al.*, 1994) and the primers used in the attempt to sequence the control region were 12Sair and ND2 (Simon *et al.*, 1994).

The PCR was programmed as follows: denaturation at 94°C for 1 minute which ensures the DNA is separated into single strands, followed by 41 cycles of denaturation at 92°C for 30s, annealing at 45°C for 40s, and a ramp stage with an increase of 0.33°Cs⁻¹ to extension at 65°C, followed by a final extension at 72°C for 10 minutes. The ramp stage is included to assist annealing. If the temperature rise is too rapid, any sequences not completely bound to the primer will detach. Therefore, this ramp stage promotes maximum amplification. The products were then held at 10°C in the machine until use.

2.4. Gel Electrophoresis

Gel electrophoresis is the process in which substances of different charge and molecular weight are separated using an electric field. These differences result in

substances travelling different distances through agarose gels.

200ml of 1X TAE (Tris, Acetic acid and EDTA) was added to 4g of Agarose, mixed together and boiled up in the microwave until the agarose was completely dissolved. Whilst in the microwave the gel was watched to ensure it didn't bubble over. If this looked as if it was going to happen, the microwave was stopped and then restarted once the bubbles had settled. The gel was stained with 8 μ l of Ethidium bromide. This process was carried out in a fume cupboard due to the harmful nature of this compound. The Ethidium bromide stains the DNA allowing bands to be viewed under ultra-violet (UV) light. The gel was then poured into a gel frame, covered and left to set for half-an-hour.

The gel was then placed in the electrophoresis tank and filled with 1X TAE Buffer until the gel was fully covered. Before the PCR products were added to the wells of the gel plate, 5 μ l of dye was added to them so the distance travelled on the plate could be seen. DNA from each individual was run in 6 consecutive wells so a sufficient volume of purified DNA could be obtained for sequencing. 5 μ l of marker was added to the first well of each row to indicate the base pair length. A charge of 150 volts was passed through the gel. This causes the negatively charged DNA to travel towards the positively charge electrode. It took about half-an-hour for the pigment to travel far enough (roughly three quarters of the way between the two rows of wells) for the bands to become separated. The gel was then removed, viewed under UV light and photographed.

2.5. Gel Extraction

Gel extraction was carried out, with a few alterations, following the Qiagen QIAquick

Gel Extraction protocol (April, 2000). The desired bands were cut out from the gel using a scalpel, trying to include as little of the gel as possible. The bands were placed into labelled tubes each containing 2 gel slices. Each tube was weighed and 3 volumes of Buffer QG to 1 volume of gel was added, i.e. 300µl of Buffer for every 100mg of gel. Tubes were then incubated in a 50°C water bath for 10 minutes to dissolve the gel, before addition of 1 gel volume of isopropanol, i.e. 100µl of isopropanol for every 100mg of gel. This solution was transferred to a QIAquick spin column and centrifuged at full speed (13000rpm) for 1 minute. As before, this leaves the DNA attached to the filter membrane and the flow-through can be discarded. Frequently, the volume of solution was greater than the maximum volume the spin column can hold (800µl). When this was the case the column was reloaded with the excess volume and centrifuged again. To remove all traces of agarose, 0.5ml of Buffer QG was added before centrifuging at full speed for 1 minute. 0.75ml of wash Buffer PE was then added. The column was left to stand for 2-5 minutes before being centrifuged at full speed for 1 minute. The flow-through was discarded and the column spun for an additional minute at full speed. The collection tube was replaced by a 1.5-ml microcentrifuge. 40µl of H₂O was added to the centre of the membrane, left to stand for 10-30 minutes and then centrifuged at full speed for 1 minute. Again, the addition of water causes the DNA attached to the filter membrane to be released and collected in the tube.

A small amount of the DNA collected for each individual (5µl) was run on a small test gel to check if the extraction had been successful. This small gel plate was mixed as before but contained 1g of agarose, 50ml of TAE and 2µl of Ethidium bromide and was run for 30 minutes using a charge of 50 volts. This was photographed and the

remaining DNA was sent to the molecular biology support unit for sequencing.

2.6. Sequencing

Sequencing determines the exact running order of bases in a gene. The method used by the molecular biology support unit is the Sanger dideoxy chain termination method. It uses sequencers that detect DNA molecules labelled with fluorescent dyes, and automatically analyses the data to determine nucleotide sequence. Each deoxyribonucleotide contains a specific fluorescent dye that can be excited by a laser. The signals are collected and analysed by specific software to produce an electropherogram, showing peaks of different colours for each base (T =Red, A =Green, C =Blue, G =Black). Samples were sequenced in both directions to reliably determine the correct base sequence. Any missing data on one sequence, or disagreements between the two sequences can be checked.

2.7. Random Amplified Polymorphic DNA (RAPDs)

RAPDs are used to study the amount of genetic variation present within a population. Random primers are used to amplify random DNA fragments. These are then separated out using gel electrophoresis and resulting bands can be compared. Bands are marked as to whether they are present (1) or absent (0). The advantage of RAPDs is that it supplies an immediate indication of variation. However, there are problems, such as reproducibility and accurate scoring that will be discussed further in the discussion.

The PCR was carried out as above using the primers OPAB 1 & 11, 2 & 12, 3 & 13, 4 & 14. Each is a random, 10 base long primer. As before, a PCR negative was included. The PCR was programmed as follows; 95°C for 90seconds, followed by 38 cycles of annealing at 37°C for 1 minute, a ramp step of 0.17°Cs⁻¹ to 72°C for 3

minutes, followed by extension at 94°C for 1 minute, followed by a final annealing at 37°C for 5 minutes and a ramp step of 0.17°Cs⁻¹ to a final extension at 72°C for 5 minutes. The products were then held at 10°C in the machine until use.

Gel electrophoresis was carried out as above but for a much longer period (1.5 to 2 hours). This allows the maximum number of bands to be obtained for comparison. However, running the gel for a longer period resulted in the slightly positively charged Ethidium bromide moving towards the negative electrode and leaving the latter half of the gel only faintly stained. As a consequence, the gels were scored whilst under the UV light as photographs were poor. Due to known problem of reproducibility (Perez et al., 1998), the experiments were then repeated in exactly the same way to see if the same banding patterns could be reproduced.

2.8. Data Analysis

2.8.1. 12S

The sequences received from the lab were aligned to outgroup sequences² using the program Clustal X. From the data set a distance matrix was created using the computer program PAUP* 4. This matrix compares every sequence to the others and indicates how many bases differ between the two. This distance matrix was then used to produce a neighbour joining phylogram and a bootstrap analysis, again using PAUP* 4. The neighbour joining phylogram presents the distance matrix in the form of a tree. Bootstrapping gives an indication of how strongly the data support each node on the tree. It produces a pseudoreplicate by resampling at random from within the data set provided. The pseudoreplicate contains only sites found in the original

² Outgroup sequences obtained from Rob Cruickshank, Molecular & Evolutionary Biology, University of Glasgow.

data set but it differs in the frequencies of different sites. In this case the process was repeated 1000 times to obtain bootstrap values.

2.8.2. RAPDs

The scores from the gels were entered into the computer program MacClade 4. This data set was then transferred to PAUP* 4 where a Permutation Tail Probability (PTP) test was carried out. The PTP test looks to see if there is a tree-like relationship in the data set. It does this by randomly reassigning the character states to the taxa and produces trees (100) from these new data sets. The tree length for each of these is calculated. The data set will pass the PTP test if the length of the tree constructed from the empirical data set is significantly shorter than those constructed from the permuted data sets.

3. Results

3.1. 12S

Gel electrophoresis for this region always produced strong, clear bands for all individuals (Appendix 1a). Gel extraction was also successful (Appendix 1b). The colour coding of bases makes it easy to see that there is little variation present between the majority of individuals sequenced in this study (Appendix 2). However, one individual, HCVB 1A, is noticeably different.

The neighbour joining phylogram produced for the data set (Fig 7.) placed all *H. pelagicus* individuals sequenced, regardless of host, into one large group which divides into two subgroups. All individuals from this study, apart from HCVB 1A form one subgroup, while HCVB 1A forms a separate subgroup with a *H. pelagicus* louse found on a white-faced storm-petrel, *Pelagodroma marina*, from the Cape Verde island of Branco. The phylogram shows that there are between 13 and 14 base changes between the ancestor of HCVB 1A and the other lice sequenced, compared with a maximum of 3 base changes between the ancestor of all other *H. pelagicus* individuals sequenced. Excluding HCVB 1A, all the other lice sequenced were found to be very similar to *H. pelagicus* previously sequenced from *P. marina* from New Zealand and one from a Bulwer's storm-petrel, *Bulweria bulwerii*.

Due to the considerable difference found for HCVB 1A, the DNA for this individual was resequenced. DNA from a second louse from the same bird (HCVB 1B) was also sequenced. HCVB 1B was found to be genetically similar to the rest of the lice studied.

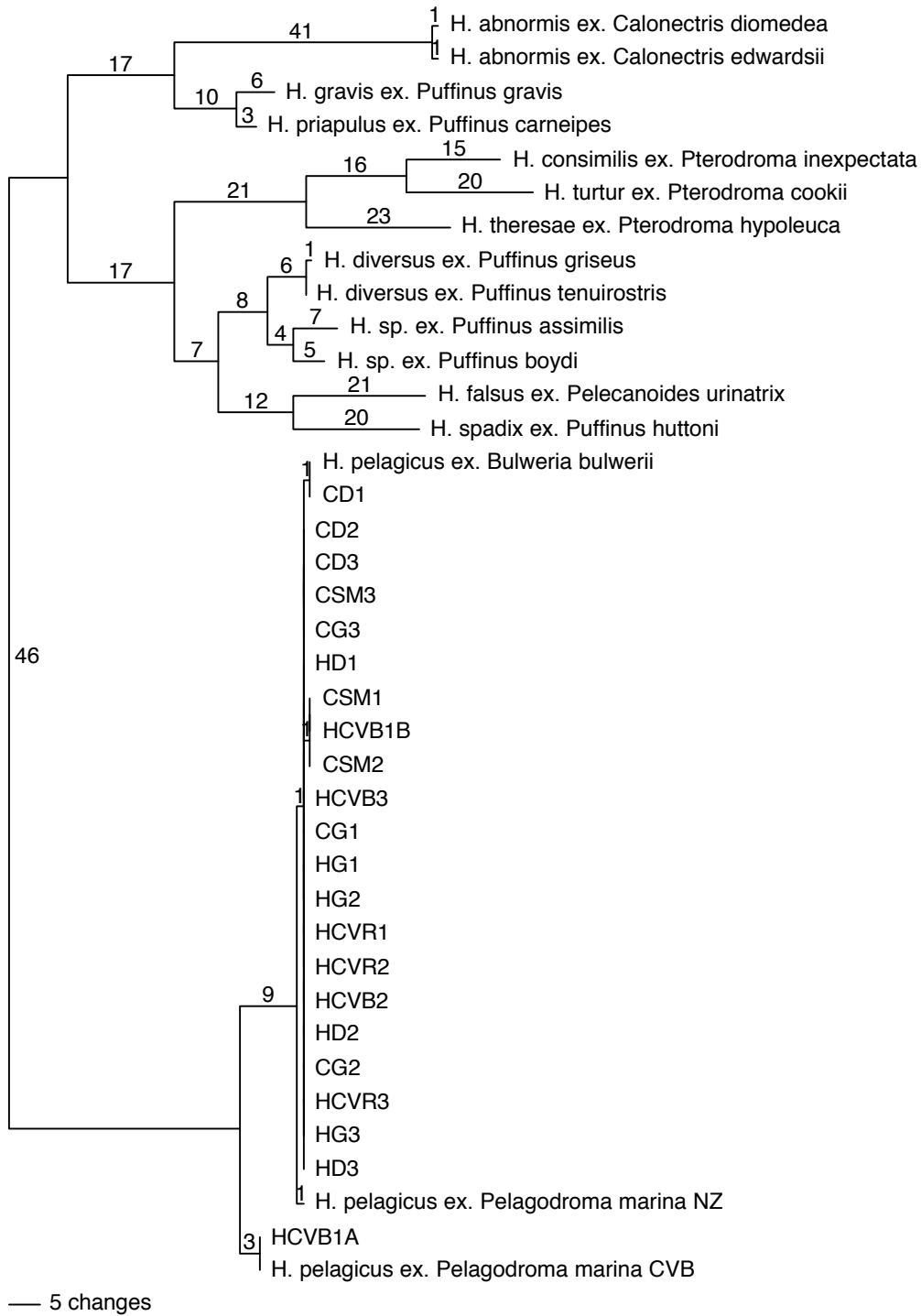


Fig 7. Neighbour joining phylogram for *H. pelagicus*, including outgroups, indicating no. of base changes. Abbreviation; CD1-3 are from Desertas (cool-season), HD1-3 are from Desertas (hot-season), CG1-3 are from Graciosa (cool-season), HG1-3 are from Graciosa (hot-season), CSM1-3 are from Santa Maria (cool-season), HCVB 1(A&B)-3 are from Cape Verde Branco (hot-season), and HCVR1-3 are from Cape Verde Raso (hot-season).

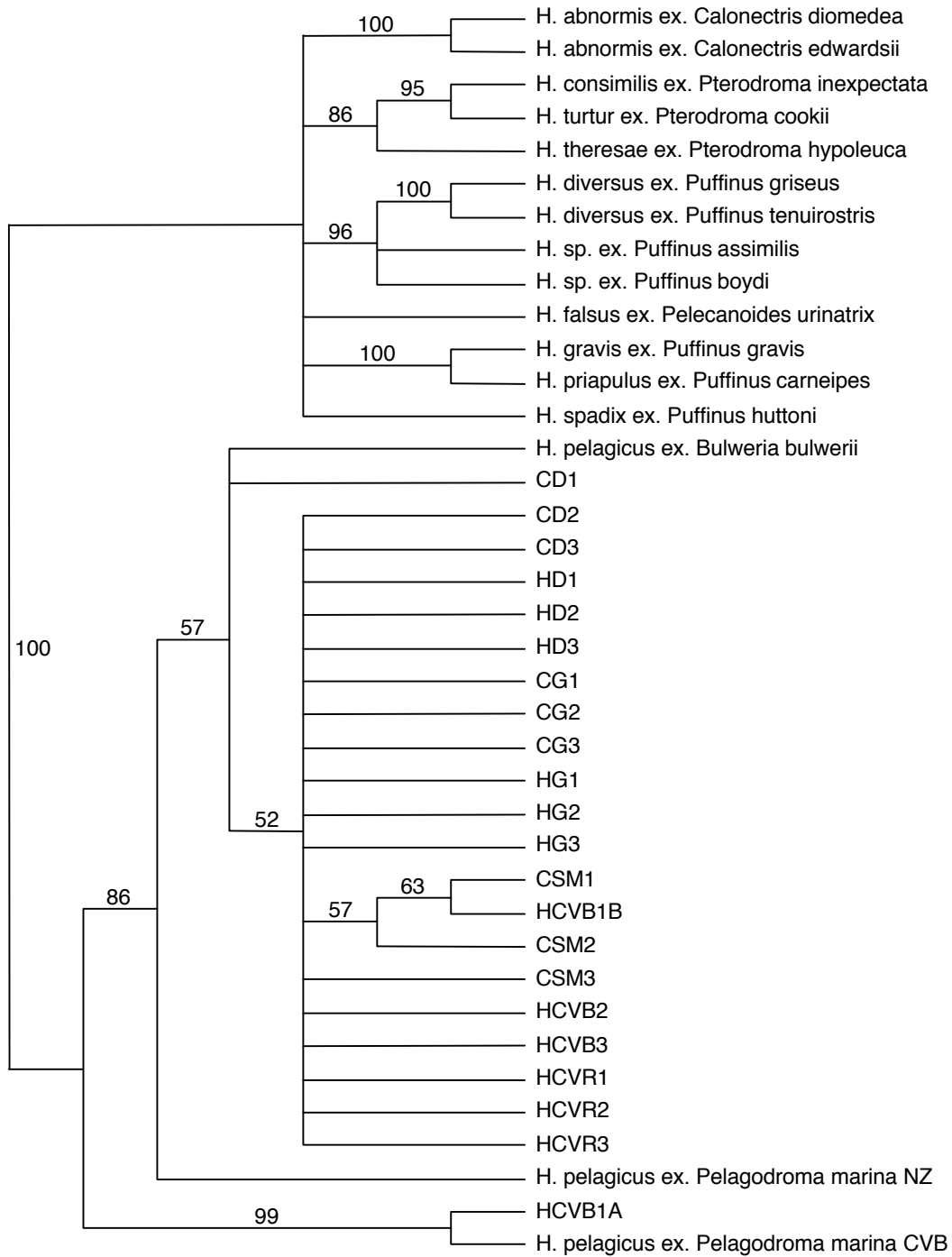


Fig 8. Neighbour joining bootstrap analysis for *H. pelagicus*, including outgroups, giving values for strength of support of each node. Abbreviations; CD1-3 are from Desertas (cool-season), HD1-3 are from Desertas (hot-season), CG1-3 are from Graciosa (cool-season), HG1-3 are from Graciosa (hot-season), CSM1-3 are from Santa Maria (cool-season), HCVB 1(A&B)-3 are from Cape Verde Branco (hot-season),and HCVR1-3 are from Cape Verde Raso (hot-season).

The neighbour joining bootstrap analysis for the data set (Fig 8.) shows that the data set completely supports (100%) the morphology of *H. pelagicus*. The data also strongly supports (99%) grouping HCVB 1A with *H. pelagicus* from *P. marina* (CVB). There is significant support (86%) for the morphology of all other *H. pelagicus*.

3.2. CO1

Gel electrophoresis for the CO1 region gave bands that were less clear than for 12S (Appendix 3) and sequencing was only successfully for a few individuals. These sequences were compared and found to be identical. These preliminary results supported the findings for the 12S region but it was decided not to continue with further analysis of this region due to lack of time.

3.3. Control Region

Sequencing of the control region was not successful. Gel electrophoresis produced apparently random banding patterns (Appendix 4) and found no bands consistent across all taxa. As with CO1, no further work was done on this region due to findings of a study by Shao (unpublished data) on the wallaby louse, *Heterodoxus macropus*, which will be discussed later.

3.4. RAPDs

Gel electrophoresis using the OPAB random primers produced many bands (Appendix 5a). However, when a PTP test was carried out on these data, it failed. Additionally, when the experiment was repeated using exactly the same procedure, different banding patterns were obtained (Appendix 5b). Therefore, the results from

the use of RAPDs cannot be used as they show no relationship and are not reproducible.

4. Discussion

4.1. 12S

4.1.1. Variation between breeding seasons

Results show that no genetic variation was detected between lice from hot- and cool-season breeders of *O. castro*. The implications of this finding are that, at present, the mechanism driving speciation in *O. castro* i.e. different breeding seasons, does not have the same affect on their lice. The lice are not reproductively isolated and there is no evidence that cospeciation is likely to occur. The form of separation driving speciation in *O. castro* may be an important factor. Speciation is thought to be sympatric (Monteiro & Furness, 1998). Therefore *H. pelagicus* is less likely to cospeciate with *O. castro* as lice from sympatric hosts are less likely to cospeciate than lice from allopatric hosts (Barker, 1994).

Lack of cospeciation between host and parasite is generally considered to be a direct result of widespread host switching (Barker, 1991). In this case, lice are switching between birds from the two different breeding seasons. This result opposes the idea that lice are very host specific (Furness and Palma, 1992). Their studies on Gough Island found a remarkable lack of shared louse species between seabirds despite the fact that they shared the same nesting habitats and compete for burrows. However, host switching is more likely to be successful if the new host is closely related to the host on which they evolved (Barker, 1994). Therefore, *H. pelagicus* should be able to colonise birds from either breeding season as they are very closely related, being classed as either a single species or sibling species (Monteiro & Furness, 1998), and have evolved on *O. castro*.

“Cospeciation will occur when opportunities for host switching do not exist” (Barker, 1994). Unlike gopher lice (Page & Hafner, 1996; Nadler *et al.*, 1990), ample opportunity exists for *H. pelagicus* to host switch. In colonies where two breeding seasons occur, birds from the two seasons are present on the island during the overlap in August and early September (Monteiro & Furness, 1998). In addition, they time-share nests, a factor that is thought to facilitate louse transfer among some species of birds (Eveleigh & Thelfall, 1976). Other studies also provide further evidence that supports the occurrence of host switching between the two sympatric species of *O. castro* through shared use of nest burrows. Clayton (1990) found that *Strigiphilus* owl lice were shared by hosts that were sympatric and nested in the same cavities. Price & Clayton (1983) found that the sympatric, syntopic (sharing a habitat) species of parrots, *Lorius hypoinochrous* and *Lorius lory*, share the same species of lice. This seems to be the situation for *O. castro*. The birds are at most sibling species and evidence suggests the lice that have evolved on them are able to host switch, with nest burrows the probable site of transfer.

A further explanation for the lack of between-season variation in the lice is that it may have not yet occurred. In general, speciation of parasitic lice occurs after the speciation of their host (Hafner & Nadler, 1988; Paterson *et al.*, 2000). If the two seasonal breeding populations of *O. castro* have only recently shown genetic variation, then the lice may follow in time. However, lice have been found to have a rate of evolution around 5.5 times the rate of seabirds (Paterson *et al.*, 2000) so any variation in the birds has already taken a long time to occur compared with the time it would take for it to occur in the lice.

4.1.2. Geographical variation

Results show that no genetic variation was detected between lice from different geographical populations of *O. castro*. This supports the results obtained for the birds (Lodha, unpublished data), that found birds from the same breeding season, but different island populations, were genetically similar. As no genetic variation was detected between island populations of *O. castro*, it was unlikely that any genetic variation would occur in their parasitic lice. This is because speciation in lice generally occurs after speciation in their hosts (Hafner & Nadler, 1988; Paterson *et al.*, 2000).

4.1.3. HCVB 1A

HCVB 1A was found to differ genetically from all other *H. pelagicus* lice sequenced in this study. There are two possible explanations for this finding. The first is that it may have been a straggler i.e. it is usually found on a different host and was only found on *O. castro* by chance. The neighbour joining phylogram (Fig 7) shows that *H. pelagicus* is found on bird species other than *O. castro*. This suggests that it may not be very host specific. The individual that was found on *B. bulwerii* was shown to be genetically similar to those sequenced in this study. It is probable that this louse is also a straggler that has been transferred to this bird from *O. castro*, as *B. bulwerii* is not usually a host for *H. pelagicus*. Therefore, it is conceivable that HCVB 1A is a straggler, most likely to have come from *P. marina* present on the island.

The second explanation is that this individual represents a second, genetically different population of *H. pelagicus* occurring on *O. castro*. This could be a population that is present in all colonies of *O. castro* but has not been detected in the

others due to the small number of lice involved in this study. Sampling more lice from this population (Cape Verde Branco), and the others, would give a clearer indication of whether this is a possible explanation.

4.2. Control Region

Results for the control region were not obtained due to failure to amplify this region. ND2 and 12Sair would normally be successful in priming the control region of insect mtDNA. However, during the course of this study, Shao (unpublished data) produced a map for the mtDNA of the wallaby louse, *Heterodoxus macropus*. This louse is only a distant relative of *Halipeurus pelagicus* but it is probably a more accurate indicator of the correct structure of mtDNA of *Halipeurus pelagicus* than maps produced for insects in general. Shaos' findings revealed that the control region of *Heterodoxus macropus* is not one large region but is split into two much smaller regions at different locations (Fig 9.).

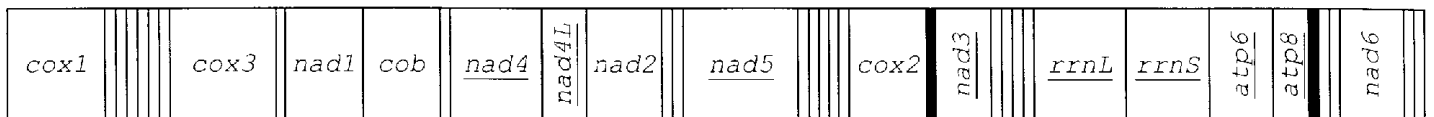


Fig. 9. Gene arrangement of mtDNA of the wallaby louse *Heterodoxus macropus*. Cytochrome oxidase subunits 1-3 (cox1-3), ATP synthase subunits 6 and 8 (atp6, atp8), NADH dehydrogenase subunits 1-6 and 4L (nadh1-6, nadh4L), 12S and 16S (rrnS, rrnL), and 22 tRNA genes, Control region represented by two black shaded areas. (Taken from Shao, unpublished data)

This probably explains why attempts to amplify the control region failed. This presents a problem in that it was the control region of *O. castro* that revealed genetic variation between hot- and cool-season breeders. As mentioned in the introduction, molecular studies allow disparate groups to be compared as homologous genes can be found in dissimilar species. However, if the mtDNA of *Halipeurus pelagicus* is like

that of *Heterodoxus macropus*, then no such comparison can be made. Even if the two smaller sections of control region were successfully sequenced, their small length means they are unlikely to show much variation.

4.3. RAPDs

Results for RAPDs analysis failed the PTP test. This indicates that no tree-like structure was found in the data obtained from the scoring of the gels. Therefore, it is probable that the lice involved in this study represent a panmictic population, i.e. each member of the population is equally likely to reproduce with any other, regardless of breeding season. This corresponds with the results for 12S which suggest that the lice populations are not isolated and that host switching probably occurs.

Results from RAPDs analysis also revealed that the experiment was not reproducible. This is a known problem with the RAPDs technique and several explanations have been put forward;

1. Differences in DNA purity between samples (Barker et al., 1990; Perez *et al.*, 1998)
2. Changes in reaction conditions such as, magnesium concentration, type of polymerase used and ambient temperature (Perez *et al.*, 1998)

It is unlikely that the DNA purity varied between the experiments in this study. Only a small volume of the total DNA extracted was used in the PCR. Therefore, the same extract was used in the original and repeat experiment. The most likely source of error is a change in reaction conditions. Although volumes are measured using pipettes, an error could have occurred that increased or reduced the volume of magnesium (or other reactants) used for PCR. Additionally, the repeat experiment

was carried out on a different day. It is possible that a variation in laboratory temperature could have affected the ramping between melting and annealing temperature. The fact that DNA and other PCR products are stored in a freezer could also be a source of variation. The products are defrosted and mixed thoroughly before use but the temperature they are at when used is unknown. If they were left to defrost for a longer period of time during one of the experiments, they could have reached a higher temperature before the PCR and this might affect the results.

4.4. Conclusion

In this study, no genetic variation was detected between lice from the two breeding populations of *O. castro*. This is most likely due to widespread host switching of the lice between birds from the hot- and cool-season, facilitated by the sharing of nest burrows. No genetic variation was found between lice from different localities. This result corresponds with the genetic data for *O. castro*, which found no variation between birds from different breeding colonies. Since between island variation is not found in their hosts, speciation of *H. pelagicus* is unlikely as lice generally speciate after their hosts.

Development of the project

This was a challenging project for someone with little genetic or evolutionary biology experience. A large amount of background reading was necessary before any papers on the subject could be properly understood. The work itself was also time consuming and very little would have been achieved without the help I received from Gareth Hood and Rob Cruickshank. Lack of time meant that only a limited number of lice could be sampled. Ideally, I would have liked to include a greater number of lice in the study. Further work on these populations of *H. pelagicus* could be done to explain the presence of HCVB 1A. Is it a straggler or does it provide evidence of a second genetically different population of *H. pelagicus* surviving on *O. castro*?

Since no genetic variation was detected by the methods used in this project, a population genetics approach, such as the use of allozymes or microsatellites, could be applied. This would sample a larger number of the population and would reveal any gene frequency variations which have not yet had time to become fixed in the population.

Another area of development investigating differences between hot- and cool-season breeders could involve a quantitative study of their lice. Are there a greater number of lice on birds from one of the seasons? It may also be possible to carry out a similar project to this one but with a different louse species. For example, *Austromenapon* is also present on *O. castro*. However, it was found to be present in much fewer numbers and not on all birds sampled, therefore it may not be suitable for such analysis.

C5M1
AAGTAAAAAA--ATTGG-TGCCGTAGAGTAGAGT-TTACCTAGTCCTTGAAAAATTAATAATTGGCCGCCCTATTAACTATACAGAGAACATGGCTTTTAAATGGAAAATTACACATAT
ATCTTACCCTATTGTAGAT-TG-ACATTTGGTTACACCCGCCCTCGGAAGAGTAGTTCCTAAGAACCTG-TTCTATTTT-CT--ATATTAAGAAATTAAGTCAAGTCAAAAGCACCGACTAT
CGATTAGGTTACCAATGTGTCTCAATATTTATTTATGTCTCAAAAATCTTAAATTAACAAAAGGTTTAAATACCTGAAAATT--AGAAATTAATGAAATCTGTAAAGTAAAGCTTTTAAAT
TTAAGCTGAAATGCTGCTCTAAATTAGTGT

C5M2
-----AGAGTAGAGT-TTACCTAGTCCTTGAAAAATTAATAATTGGCCGCCCTATTAACTATACAGAGAACATGGCTTTTAAATGGAAAATTACACATAT
ATCTTACCCTATTGTAGAT-TG-ACATTTGGTTACACCCGCCCTCGGAAGAGTAGTTCCTAAGAACCTG-TTCTATTTT-CT--ATATTAAGAAATTAAGTCAAGTCAAAAGCACCGACTAT
CGATTAGGTTACCAATGTGTCTCAATATTTATTTATGTCTCAAAAATCTTAAATTAACAAAAGGTTTAAATACCTGAAAATT--AGAAATTAATGAAATCTGTAAAGTAAAGCTTTTAAATTA

C5M3
AAGTAAAAAA--ATTGG-TGTGTAGAGTAGAGT-TTACCTAGTCCTTGAAAAATTAATAATTGGCCGCCCTATTAACTATACAGAGAACATGGCTTTTAAATGGAAAATTACACATAT
ATCTTACCCTATTGTAGAT-TG-ACATTTGGTTACACCCGCCCTCGGAAGAGTAGTTCCTAAGAACCTG-TTCTATTTT-CT--ATATTAAGAAATTAAGTCAAGTCAAAAGCACCGACTAT
CGATTAGGTTACCAATGTGTCTCAATATTTATTTATGTCTCAAAAATCTTAAATTAACAAAAGGTTTAAATACCTGAAAATT--AGAAATTAATGAAATCTGTAAAGTAAAGCTTTTAAATTA
TTAAGCTG-----

HCVB1A
AAGTAAAAAA--GTTTGG-TGTGTAGAGTAGAGT-TTATTTAGTCTTGAAAAATTAATAATTGGCCGCCCTATTAACTATACAGAGAACATGGCTTTTAAATGGAAAATTACACACA
TATCTTACCCTATGTTAGAT-TA-ACATTTGGTTACACCCGCCCTCGGAAGAGTAGTTCCTAAGAACCTG-TTCTAAGAACCTG--ATATTAAGAAATTAAGTCAAGTCAAAAGCACCGACTAT
ATCTAATAGGTTAAATAAGTGTCTCAATATTTACCTAATTAATGTTTCAAAAATCTTAAATTAATAAAGGTTTAAATACCTGAAAATT--AGAAATTAATGAAATCTGTAAAGTAAAGCTTTTAAATTA
AATTTAAGTTTGAATGCTGCTTAAATTAATTAATGTT

HCVB1B

ATCTTACCCTATTGTAGAT-TG-ACATTTGGTTACACCCGCCCTCGGAAGAGTAGTTCCTAAGAACCTG-TTCTAAGAACCTG--ATATTAAGAAATTAAGTCAAGTCAAAAGCACCGACTAT
CGATTAGGTTACCAATGTGTCTCAATATTTATTTATGTCTCAAAAATCTTAAATTAACAAAAGGTTTAAATACCTGAAAATT--AGAAATTAATGAAATCTGTAAAGTAAAGCTTTTAAATTA
TTAAGCTGAAATGCTGCTCTAAATTAGTGT-----

HCVB2
AAGTAAAAAA--ATTGG-TGTGTAGAGTAGAGT-TTACCTAGTCCTTGAAAAATTAATAATTGGCCGCCCTATTAACTATACAGAGAACATGGCTTTTAAATGGAAAATTACACATAT
ATCTTACCCTATTGTAGAT-TG-ACATTTGGTTACACCCGCCCTCGGAAGAGTAGTTCCTAAGAACCTG-TTCTAAGAACCTG--ATATTAAGAAATTAAGTCAAGTCAAAAGCACCGACTAT
CGATTAGGTTACCAATGTGTCTCAATATTTATTTATGTCTCAAAAATCTTAAATTAACAAAAGGTTTAAATACCTGAAAATT--AGAAATTAATGAAATCTGTAAAGTAAAGCTTTTAAATTA
TTAAGCTGAAATGCTGCTCTAAATTAGTGT

HCVB3
 AAGTAAAAAA--ATTGG-TGTGTAGAGTAGAGT-TTACCTAGTCCTTGAAAAATAAAAATTTGCTTTGGCCGCTATTAACTATACAGAGAAGCATGTCTTTTAAATGGAAAATTTACACATAT
 ATTCTTACCCTATTGTAGAT-TG-ACATTTGGTTACACCCGCGCTCGAAGAAGTAGT-TCTAAGAACTG-TTCTATTCT--ATTATTAAGAAATAAGTCAAGTCAAAACACCGACCTAT
 CGATTAGGTTACCAATGTGTCTCAATATTATTATGTCTCAAAAACTCTTAACTCAAAAAAGTTTAAATACCTGAAAT--AGAAATAATTGAATCTGTAAGTAAAGCTTTTATTATTAAAA
 TTAAGCTGAATGCTGCTCTA-----

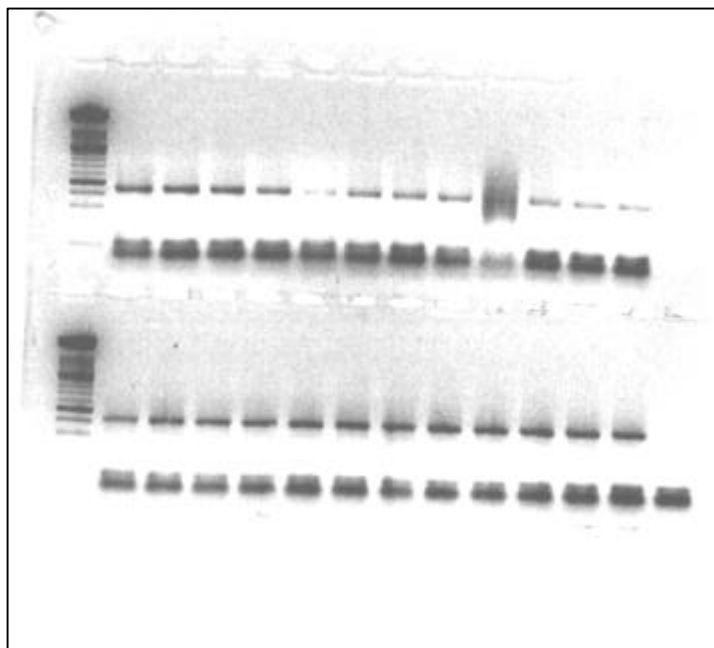
HCVR1
 AAGTAAAAAA--ATTGG-TGTGTAGAGTAGAGT-TTACCTAGTCCTTGAAAAATAAAAATTTGCTTTGGCCGCTATTAACTATACAGAGAAGCATGTCTTTTAAATGGAAAATTTACACATAT
 ATTCTTACCCTATTGTAGAT-TG-ACATTTGGTTACACCCGCGCTCGAAGAAGTAGT-TCTTAAAGAACTG-TTCTTAACTATTCT--ATTATTAAGAAATAAGTCAAGTCAAAACACCGACCTAT
 CGATTAGGTTACCAATGTGTCTCAATATTATTATGTCTCAAAAACTCTTAACTCAAAAAAGTTTAAATACCTGAAAT--AGAAATAATTGAATCTGTAAGTAAAGCTTTTATTATTAAAA
 TTAAGCTGAATGCTGCTCTAATTAGTGT

HCVR2
 AAGTAAAAAA--ATTGG-TGTGTAGAGTAGAGT-TTACCTAGTCCTTGAAAAATAAAAATTTGCTTTGGCCGCTATTAACTATACAGAGAAGCATGTCTTTTAAATGGAAAATTTACACATAT
 ATTCTTACCCTATTGTAGAT-TG-ACATTTGGTTACACCCGCGCTCGAAGAAGTAGT-TCTTAAAGAACTG-TTCTTAACTATTCT--ATTATTAAGAAATAAGTCAAGTCAAAACACCGACCTAT
 CGATTAGGTTACCAATGTGTCTCAATATTATTATGTCTCAAAAACTCTTAACTCAAAAAAGTTTAAATACCTGAAAT--AGAAATAATTGAATCTGTAAGTAAAGCTTTTATTATTAAAA
 TTAAGCTGAATGCTGCTCTAATTAGTGT

HCVR3
 AAGTAAAAAA--ATTGG-TGTGTAGAGTAGAGT-TTACCTAGTCCTTGAAAAATAAAAATTTGCTTTGGCCGCTATTAACTATACAGAGAAGCATGTCTTTTAAATGGAAAATTTACACATAT
 ATTCTTACCCTATTGTAGAT-TG-ACATTTGGTTACACCCGCGCTCGAAGAAGTAGT-TCTTAAAGAACTG-TTCTTAACTATTCT--ATTATTAAGAAATAAGTCAAGTCAAAACACCGACCTAT
 CGATTAGGTTACCAATGTGTCTCAATATTATTATGTCTCAAAAACTCTTAACTCAAAAAAGTTTAAATACCTGAAAT--AGAAATAATTGAATCTGTAAGTAAAGCTTTTATTATTAAAA
 TTA-----

Base sequences for all individuals sequenced. HCVB 1A is noticeably variable (differences have been enlarged and placed in bold).

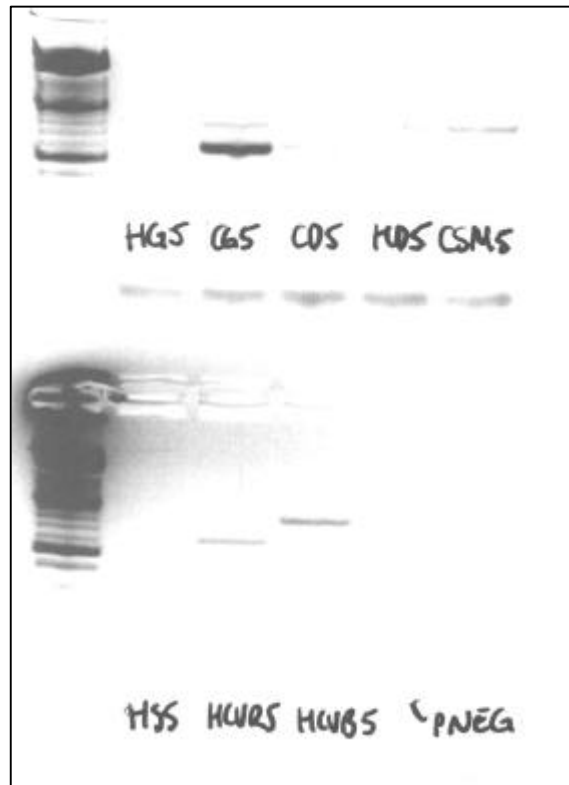
Appendix 3.



Gel plate showing the less intense, smudged band patterns produced for CO1. Each sample was run in 3 wells.

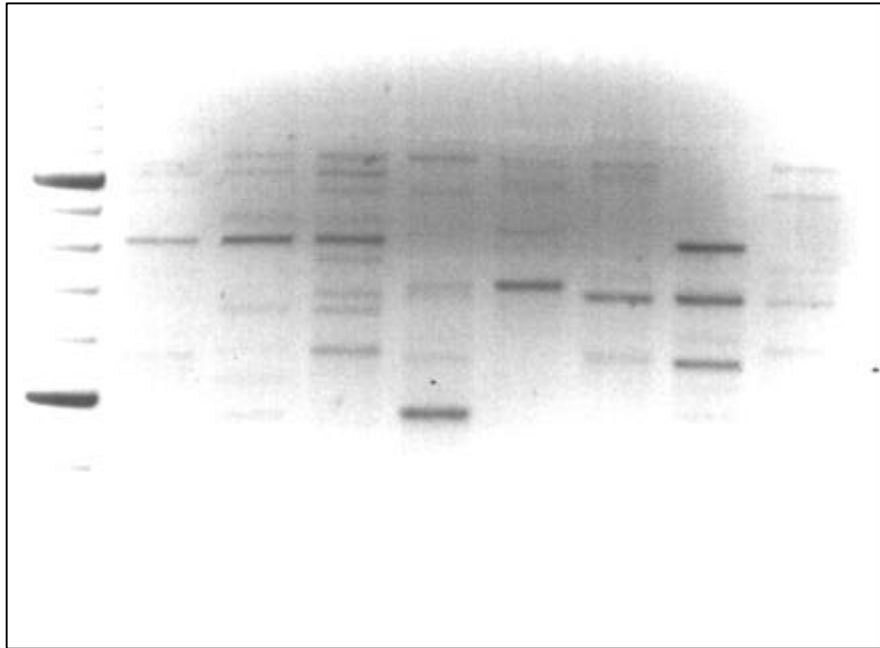
- Wells 1-3 = CSM
- Wells 4-6 = CG
- Wells 7-9 = HG
- Wells 10-12 = HCVB
- Wells 13-15 = HCVR
- Wells 16-18 = HS
- Wells 19-21 = CD
- Wells 22-24 = HD

Appendix 4.

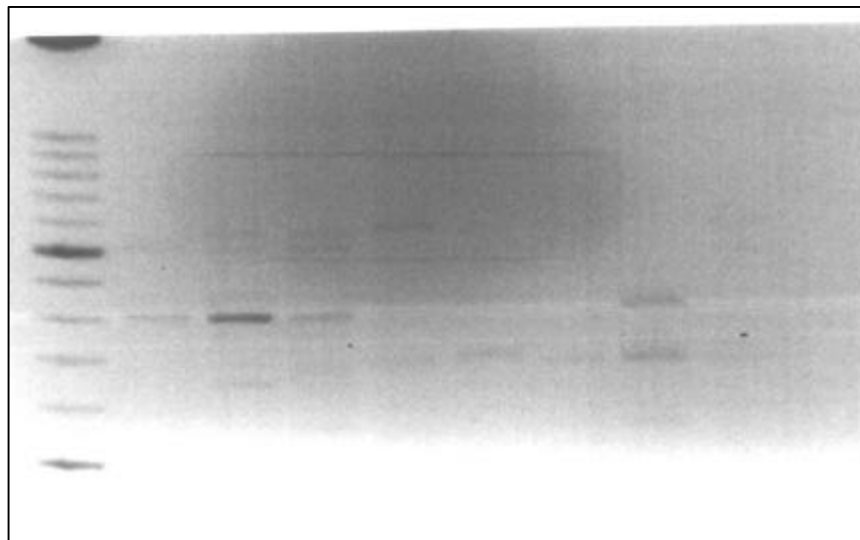


Gel plate for attempt to sequence the control region. Shows random banding patterns which were inconsistent across the groups.

Appendix 5.



5a.) Gel plate showing the random banding patterns produced with use of RAPDs. The photograph is of poor quality and so the gel was scored whilst under the UV light.



5b.) Gel plate showing banding patterns produced for repeat experiment carried out exactly the same as above. Note that the bands produced are much fainter, although photographing and scanning have further faded the appearance of the gel.