

Tese M/212

Ana Sofia Sona Tavares



**Análise da estruturação genética da sub-espécie *Patella candei gomesii* Drouet 1858 (Mollusca: Gastropoda: Patellidae) no arquipélago dos Açores**



Tese de Mestrado em Biodiversidade e Recursos Genéticos  
Faculdade de Ciências da Universidade do Porto  
Outubro, 2007

Ana Sofia Sona Tavares

**Análise da estruturação genética da sub-espécie *Patella  
candei gomesii* Drouet 1858 (Mollusca: Gastropoda:  
Patellidae) no arquipélago dos Açores**

DEPARTAMENTO DE ZOOLOGIA E ANATOMIA  
45780  
BIBLIOTECA

Faculdade de Ciências da Universidade do Porto

Porto, 2007



FC

Biblioteca  
Faculdade de Ciências  
Universidade do Porto



000055110

Dissertação de Mestrado em Biodiversidade e  
Recursos Genéticos apresentada à Faculdade  
de Ciências da Universidade do Porto.

Orientador: Prof. Doutor Paulo Alexandrino

## **Agradecimentos**

Gostaria de agradecer, pelo apoio, a todas as pessoas com quem trabalhei e que conheci durante o último ano e meio de estudo e trabalho no CIBIO e CTM.

Gostaria ainda de demonstrar a minha sincera gratidão:

Ao Professor Doutor Paulo Alexandrino por me ter proposto este projecto e aceitar orientar-me.

À Montse, pela verdadeira orientação que me deu e que fez com que o trabalho andasse para a frente. Gracias por tudo, sem ti ainda estaria longe do fim!

À Xana por algumas amostras, artigos, revisão do projecto de FCT, revisão da tese, cumplicidade, etc., etc.

Às "meninas" do CTM que, de uma maneira ou outra deram apoio, em especial à Teresa por me ensinar muito do que aprendi no laboratório e pela disponibilidade na confirmação da leitura dos géis manuais; à Paula pois, já no final e um pouco à pressa, também foi indispensável; à Tânia e à Joana pelas boleias e à última ainda pela amizade e pelos momentos de diversão no laboratório.

Aos colegas da Planeta Vivo (Carla, Rita e Duarte) pelo apoio, pelas opiniões (e paciência ao computador quando punha programas pesados de tratamento de dados a correr!), e ao Nuno pela flexibilidade horária permitida.

Por fim, mas não por último, à minha família (pai, mãe, Rosa, Pedro e Patrícia) e ao Daniel, numa palavra, por tudo. Desculpem a ausência e a impaciência nas alturas mais complicadas e obrigada por estarem sempre disponíveis.

<b>Índice</b>	<b>pág.</b>
Resumo	i
Abstract	ii
<b>1. Introdução e objectivos</b>	
1.1. Taxonomia	2
1.2. Aspecto ecológicos	4
1.3. Diferenciação genética	6
1.4. Importância dos estudos genéticos para a conservação de <i>P. candei</i> e principais ameaças à diversidade genética da espécie	9
1.5. Marcadores adequados a estudos populacionais: microssatélites	
1.5.1. Descrição dos microssatélites	11
1.5.2. Aplicabilidade dos microssatélites em estudos populacionais	12
1.5.3. Evolução dos microssatélites	13
1.6. Objectivos	16
<b>2. Metodologia</b>	
2.1. Área de estudo	18
2.2. Amostragem	19
2.3. Extracção de ADN	20
2.4. Selecção dos microssatélites e amplificação por PCR	21
2.5. Separação dos alelos e leitura dos genótipos individuais	23
2.5.1. Genotipagem em géis manuais	23
2.5.2. Genotipagem em sequenciador de gel	24
2.5.3. Erros de genotipagem e sua correcção	26
2.6. Análise da diversidade genética	
2.6.1. Frequências alélicas	29
2.6.2. Estimativas de diversidade genética	30
2.7. Análise da estruturação genética	33
2.8. Análise do gene mitocondrial Citocromo c Oxidade	38
<b>3. Resultados</b>	
3.1. Selecção de microssatélites e amplificação por PCR	40

<b>3.2. Detecção de erros de genotipagem</b>	41
<b>3.3. Análise da diversidade genética: frequências alélicas e estimativas da diversidade genética</b>	41
<b>3.4. Análise da estruturação genética</b>	45
<b>3.5. Distribuição geográfica dos polimorfismos de ADNmt</b>	49
<b>4. Discussão</b>	
<b>4.1. A utilização de marcadores microssatélites em <i>P. candei</i></b>	53
<b>4.2. Estimativas da diversidade genética e análise da estruturação</b>	54
<b>5. Conclusão e sugestões para futuros estudos</b>	58
<b>6. Referências bibliográficas</b>	60
<b>Anexos</b>	72

<b>Índice de figuras</b>	<b>pág.</b>
Fig.1.1 – Árvore filogenética obtida com base na análise de máxima parcimónia de caracteres morfológicos para a família Patellidae.	2
Fig.1.2 – Análise separada das espécies de <i>Patella</i> (não está desenhado o grupo externo), representada por uma árvore de máxima parsimónia obtida pela informação combinada das sub-unidades 12S e 16S do ADN mitocondrial.	3
Fig.1.3 - Vista dorsal e interna da concha das subespécies de <i>P. candei</i> .	4
Fig.1.4 - Dendrograma UPGMA construído com base na Identidade Genética de Nei (1972) a partir da análise de 15 <i>loci</i> aloenzimáticos entre populações de <i>Patella candei</i> dos Açores, Madeira e Canárias.	6
Fig.1.5 - Árvore de Wagner construída com base na Distância de Prevosti que mostra a diferenciação genética entre populações de <i>P. candei</i> dos Açores, Selvagens, Madeira e Canárias.	7
Fig.1.6 - <i>Loci</i> diagnósticos observados entre as subespécies de <i>Patella candei</i> na Macaronésia.	7
Fig.1.7 - Árvore de máxima verosimilhança ( <i>Maximum likelihood</i> ) baseada em sequências da subunidade I da Citocromo Oxidase.	8
Fig.2.1 - Enquadramento geográfico do arquipélago dos Açores.	18
Fig.2.2 - Representação das principais massas de água no Atlântico.	19
Fig.2.3 - Localização dos pontos de amostragem nos Açores e nas Canárias.	20
Fig.2.4 - Teste de extracção de ADN num gel de agarose 0,8% (p/v) corado em brometo de etídeo.	20
Fig.2.5 - Teste do sucesso da reacção de amplificação num gel de agarose 2% (p/v). Na primeira aplicação observa-se o marcador de peso molecular <i>Marker 5 (Eurogentec)</i> . Visualização feita por coloração com brometo de etídio.	23

Fig.2.6 - Visualização dos alelos obtidos para o marcador Pde3 em gel de poliacrilamida desnaturante 6% p/v por coloração com nitrato de prata.	24
Fig.2.7 – ALFexpressII e computador com o <i>software</i> de leitura.	24
Fig.2.8 – Curvas de absorvância do sequenciador ALFexpressII.	25
Fig.2.9 – Diagrama de ajuda na detecção de erros de genotipagem.	29
Fig.3.1 - Frequências alélicas do <i>locus</i> Pde1 para cada população.	42
Fig.3.2 - Frequências alélicas do <i>locus</i> Pde3 para cada população.	42
Fig.3.3 - Frequências alélicas do <i>locus</i> Pde6 para cada população.	42
Fig.3.4 - Frequências alélicas do <i>locus</i> Pde5 para cada população.	43
Fig.3.5 - Frequências alélicas do <i>locus</i> PrC5 para cada população.	43
Fig.3.6 - <b>A:</b> Representação gráfica do número de alelos por <i>locus</i> detectados no total de indivíduos (350); <b>B:</b> Representação gráfica da relação entre o número médio de alelos e o efectivo das populações consideradas.	44
Fig.3.7 – Resultados da análise com o programa BAPS.	49
Fig.3.8 - Rede de máxima parsimónia dos haplótipos obtidos a partir da análise de 172 sequências de uma porção da sub-unidade I do gene mitocondrial Citocromo c Oxidase, de populações dos Açores e das Canárias.	49
Fig.3.9 - Rede de máxima parsimónia dos haplótipos obtidos a partir da análise de sequências de uma porção da sub-unidade I do gene mitocondrial Citocromo c Oxidase para as populações dos Açores.	50
Fig.4.1 – Representação da diferenciação populacional entre as ilhas dos Açores.	55

<b>Índice de tabelas</b>	<b>pág.</b>
Tabela 2.1: Descrição dos microssatélites analisados: espécie onde foram isolados, motivo de repetição e tamanho do produto de PCR (pb) nessas espécies.	22
Tabela 3.1 - Condições de amplificação dos microssatélites.	40
Tabela 3.2 - Descrição da variação do tamanho de alelos detectados e consequente classificação dos <i>loci</i> quanto ao morfismo alélico.	41
Tabela 3.3 - Parâmetros de diversidade genética para 5 <i>loci</i> microssatélites analisados.	46
Tabela 3.4 - Valores da estimativa do coeficiente de diferenciação populacional entre pares de populações, $F_{ST}$ , na parte superior da matriz e da estimativa do fluxo génico, $Nm$ , na parte inferior.	45
Tabela 3.5: Valores do teste exacto de Fisher.	48
Tabela 3.6 - Resultados da AMOVA considerando 2 grupos: 1- Açores; 2- Canárias.	48
Tabela 3.7 - Resultados da AMOVA considerando 3 grupos: 1- Flores; 2-Faial, Pico, Graciosa e Terceira; 3- São Miguel e Santa Maria.	48
Tabela 3.8 - Resultados da AMOVA considerando 3 grupos: 1- Flores; 2-Faial, Pico, Graciosa e Terceira; 3- São Miguel e Santa Maria.	48

## Resumo

As lapas do género *Patella* (Gastropoda) constituem um recurso sobre-explorado em algumas ilhas Macaronésicas, como *P. candei* nas Ilhas Canárias. No entanto, não existe um conhecimento preciso da sua taxonomia e estrutura genética que permita levar a cabo uma gestão genética adequada para evitar as consequências de uma erosão genética por sobre-exploração. *P. candei* distribui-se pelos arquipélagos da Madeira, Açores e Canárias. Trabalhos recentes efectuados com ADN mitocondrial sugerem que se trata de uma espécie parafilética, e que *P. candei* das Canárias está filogeneticamente mais próxima de *P. lugubris* de Cabo Verde que *P. candei* dos Açores ou Madeira.

O objectivo deste estudo é a determinação da estrutura genética de *P. candei* nos Açores, utilizando uma bateria de marcadores microssatélites isolados em espécies próximas (*P. depressa* (4) e *P. rustica* (1)).

Para determinar o estatuto genético das populações de lapa *P. candei*, analisou-se uma população por ilha (Flores, Faial, Pico, Graciosa, Terceira, São Miguel e Santa Maria) e uma população das Canárias como grupo externo. Calcularam-se as frequências alélicas, a heterozigotia observada ( $H_o$ ) e esperada ( $H_e$ ), os testes exactos de ajuste a Hardy-Weinberg, assim como a partição do défice de heterozigóticos dentro ( $F_{IS}$ ) e entre populações ( $F_{ST}$ ), e a riqueza alélica ( $R_s$ ). A heterogeneidade global foi examinada mediante a Análise da Variância Molecular e a estruturação dos indivíduos foi obtida a partir do programa de análise bayesiana BAPS.

Observou-se um desajuste generalizado ao equilíbrio Hardy-Weinberg devido, provavelmente, à utilização de marcadores isolados em outras espécies. O índice  $F_{ST}$  apresentou valores significativos para quase todos os pares de locais de amostragem, o que, juntamente com a significância igualmente obtida com os testes exactos, parece indicar algum grau de restrição ao fluxo génico entre as ilhas dos Açores. Contudo, a AMOVA mostrou que 73-90% da mesma se encontra entre indivíduos. A análise de genótipos multilocus efectuada com o BAPS não permitiu definir nenhum agrupamento bem suportado dentro dos Açores, separando apenas os dois arquipélagos.

À vista destes resultados, parece haver alguma diferenciação entre ilhas mas é necessário prosseguir os estudos com maior número de microssatélites e específicos.

## Abstract

*Patella* limpets have become an over exploited resource in some Macaronesian islands, like *P. candei* in Canary Islands. However, there is not a precise knowledge about its taxonomy and genetic structure to allow the implementation of adequate genetic management in order to avoid the consequences of genetic erosion. *P. candei* can be found in the archipelagos of Madeira, Azores and Canary. Recent projects with mitochondrial DNA suggest that this is a paraphyletic species and that *P. candei* from Canary is closely related to *P. lugubris* from Cape Verde than *P. candei* from Azores or Madeira.

The main goal of this study is the determination of the genetic structure of *P. candei* in the Azores, using a battery of microsatellites isolated from related species (*P. depressa* (4) e *P. rustica* (1)).

To determine the genetic status of *P. candei* in the Azores, one population per island (Flores, Faial, Pico, Graciosa, Terceira, São Miguel and Santa Maria) and one population from Canary Islands as out-group were analysed. Allelic frequencies, observed ( $H_o$ ) and expected ( $H_e$ ) heterozygosity, Hardy-Weinberg exact tests, heterozygotes deficit within ( $F_{IS}$ ) and between ( $F_{ST}$ ) populations and allelic richness were calculated. The global heterogeneity was examined through the Analysis of Molecular Variance and the clustering of individuals was obtained using Bayesian inference software BAPS.

We observed a generalized deviation from Hardy-Weinberg equilibrium probably due to the use of microsatellites isolated from other species. The  $F_{ST}$  index presented significant values at almost every pair of sampling location, which, jointly with the significance equally obtained with the exact tests, seems to indicate some degree of restriction to the gene flow between Azores islands. However, the AMOVA showed that 73 to 90% of variation is found between individuals. The multilocus analysis with BAPS didn't allow the definition of any supported group within Azores, separating only the two archipelagos.

With these results, it seems to exist some differentiation between islands but it is necessary to prosecute this study with specific microsatellites and a larger number.

## **1. Introdução e objectivos**

---

## 1.1. Taxonomia

As espécies do género *Patella* são moluscos gastrópodes pertencentes à família Patellidae, ordem Patellogastropoda e comumente denominados lapas. De acordo com Ridgway e colaboradores (1998), a família Patellidae é dividida em quatro géneros: *Helcion* (Sul da África), *Cymbula* (Sudoeste de África, Este Atlântico e Sul do Oceano Índico), *Scutellastra* (Pacífico Indo-Oeste, Sul de África e Sul da Austrália) e *Patella* (Noroeste Atlântico e Mediterrâneo). Esta divisão dos Patellidae em quatro géneros resulta da filogenia obtida através de uma análise cladística de caracteres morfológicos, nomeadamente a forma e microestrutura da concha, pé e complexo paleal, rádula e esperma (fig.1.1; Ridgway *et al.*, 1998).

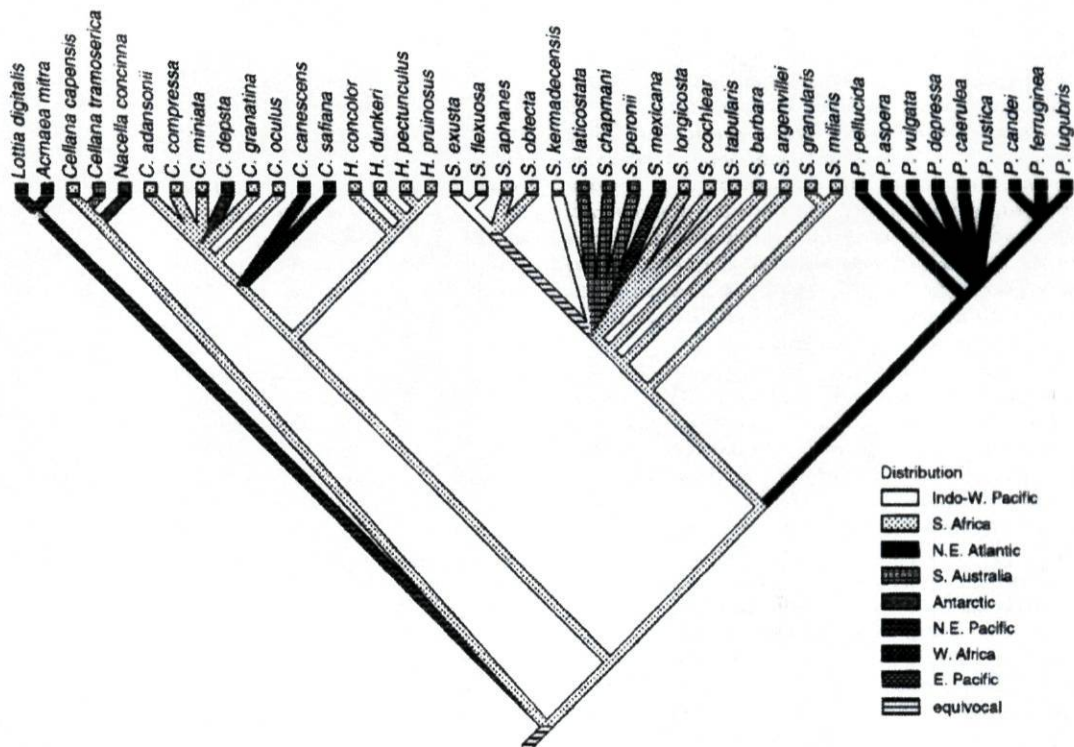


Fig.1.1 – Árvore filogenética obtida com base na análise de máxima parcimónia de caracteres morfológicos para a família Patellidae; C: *Cymbula*; H: *Helcion*; S: *Scutellastra*; P: *Patella* (retirada de Ridgway *et al.*, 1998).

Dentro do género *Patella*, Ridgway e colaboradores (1998) reconheceram nove espécies (fig.1.2): *P. depressa* (Pennant, 1777), *P. ulyssiponensis* (Gmelin, 1791), *P. rustica* (Linnaeus, 1758), *P. vulgata* (Linnaeus, 1758), *P. candei* (d'Orbigny, 1791), *P. ferruginea* (Gmelin, 1791), *P. caerulea* (Linnaeus, 1758), *P. pellucida* (Linnaeus, 1758) e *P. lugubris* (Gmelin, 1791). Baseados na análise de dois genes mitocondriais (12S e

16S), Koufopanou e colaboradores (1999) propõem a existência de duas espécies dentro de *P. ulyssiponensis*: a forma existente na Macaronésia, que estes autores denominam *P. aspera*, e a forma continental que mantém a denominação original. Estes resultados são posteriormente corroborados com a descoberta de diferenças entre populações da Macaronésia e populações continentais desta espécie com base na análise de loci aloenzimáticos (Weber *et al.*, 2005).

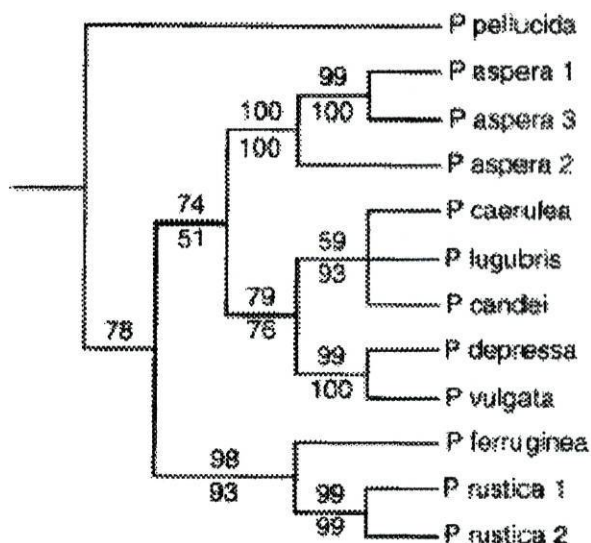


Fig.1.2 – Análise separada das espécies de *Patella* (não está desenhado o grupo externo), representada por uma árvore de máxima parsimónia obtida pela informação combinada das sub-unidades 12S e 16S do ADN mitocondrial (retirado de Koufopanou *et al.*, 1999).

Destas espécies, apenas quatro ocorrem actualmente na Macaronésia: *P. candei* e *P. ulyssiponensis*, presentes em todos os arquipélagos excepto em Cabo Verde; *P. rustica* que ocorre em todos os arquipélagos excepto nos Açores e *P. lugubris* apenas existente em Cabo Verde (revisto em Sá-Pinto *et al.*, *in press*).

*Patella candei* d'Orbigni 1840 é uma espécie endémica da Macaronésia, nomeadamente Açores, Madeira, Selvagens e Canárias (Weber & Hawkins, 2002). Na revisão do género *Patella*, Christiaens (1973) descreveu quatro subespécies dentro de *P. candei*: *P. candei gomesii* nos Açores, *P. candei ordinaria* na Madeira e Desertas, *P. candei crenata* nas Canárias e *P. candei candei* nas Selvagens e Canárias.

## 1.2. Aspectos ecológicos

As lapas habitam o litoral médio e superior, permanecendo fixas ao substrato rochoso durante o dia enquanto de noite se alimentam em zonas com algas, regressando sempre ao lugar de fixação (Lindner, 2000). Têm um pé mediano que lhes permite aderir fortemente à superfície da rocha através de dois mecanismos distintos: sucção em maré-alta e adesão em maré baixa (Little & Kitching, 1996). Possuem ainda uma concha que as protege das ondas e da predação (Little *et al.*, 1991). A concha tem forma cónica e tamanho variável, com uma base oval projectada até um ápice mais ou menos pronunciado de acordo com a espécie. Este ápice encontra-se muitas vezes desgastado nas zonas costeiras de rebentação. A superfície da concha é lisa ou sulcada radialmente, com bandas de crescimento concêntricas. O interior da concha é geralmente brilhante e vivamente colorido, observando-se uma impressão muscular em forma de ferradura aberta (fig.1.3) (revisto em Llavona, 2006).

A concha de *P. c. gomesii* (fig.1.3 A) é de tamanho médio, até 65 mm de comprimento e 20 mm de altura, tem uma forma comprimida e arredondada, com nervuras pouco marcadas e um ápice subcentral tendendo para a frente. Externamente, tem um tom rosado e cinza a castanho, com tonalidades esverdeadas e geralmente apresenta bandas radiais escuras; internamente, é acastanhada ou rosada com um tom azulado iridescente. O pé é normalmente cinzento ou branco-sujo e ocasionalmente amarelado. Os tentáculos paleais são cinzentos e púrpura (revisto em Christiaens, 1973 e Titselaar, 1998). Esta sub-espécie tem um comportamento intertidal e ocasionalmente subtidal superficial (Christiaens, 1973; Hawkins *et al.*, 1990).

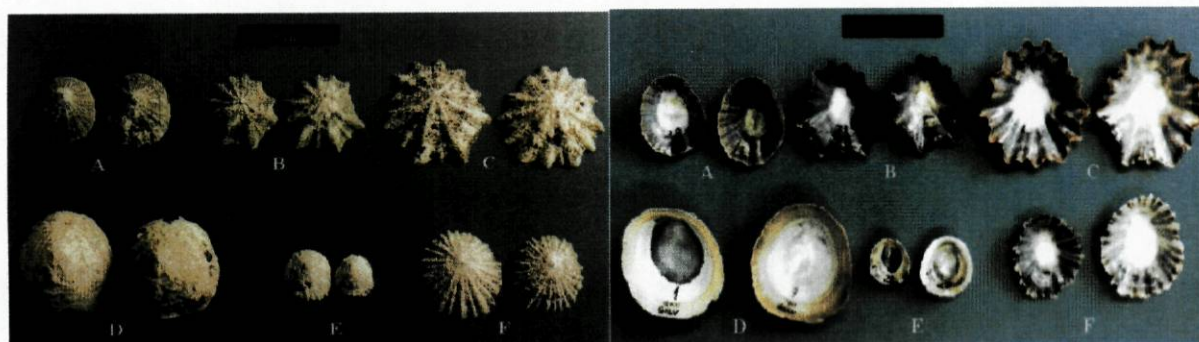


Fig.1.3 - Vista dorsal e interna da concha das subespécies de *P. candei*. A: *P. candei gomesii* (Açores); B: *P. c. ordinaria* (Madeira); C: *P. c. ordinaria* (Deserta grande); D: *P. c. candei* (Selvagens); E: *P. c. candei* (indivíduo raro de Fuerteventura, Canárias); F: *P. c. crenata* (Canárias) (Weber & Hawkins, 2002).

Na subespécie *P. c. gomesii* encontram-se descritas duas formas com morfologia distinta (Hawkins *et al.*, 1990). No entanto, a análise de dez loci enzimáticos feita por Pagarete e colaboradores (2005) não detectou diferenças genéticas entre estas duas ecoformas.

As lapas possuem ainda uma rádula denominada decoglosa, exibindo em cada fila dois dentes laterais e oito dentes marginais, dois com as extremidades endurecidas e seis sem extremidades endurecidas. A alimentação segue padrões regulares em que a rádula raspa lentamente o substrato deixando uma marca clara na rocha. A sua alimentação consiste essencialmente em micro e macro algas (Hill & Hawkins, 1991), cuja disponibilidade varia com as marés e fotoperíodo. Muitas espécies de lapas têm a capacidade de “voltar a casa”, isto é, alguns indivíduos têm um local próprio na rocha onde a sua concha encaixa perfeitamente e para onde regressam após cada viagem de alimentação.

A taxa de crescimento das lapas varia consideravelmente com a sua localização na costa, densidade e natureza da fauna associada (como mexilhões e balanídeos). A idade média de uma lapa num substrato com um nível alto de balanídeos varia de 12 a 15 anos, dependendo dos níveis da predação (Little *et al.*, 1991).

As espécies de *Patella* têm ciclos reprodutivos longos, com um período de dispersão de vários estados (Morais *et al.*, 2003). Têm fecundação externa e, dependendo da espécie, a reprodução pode ocorrer quase durante todo o ano, sendo que na maioria dos habitats alcança um máximo entre princípios de Setembro e finais de Outubro e atinge um mínimo no Verão seguinte (Little *et al.*, 1991). O início da reprodução parece estar relacionado com ventos fortes e mau tempo. Depois de um período planctónico, as larvas fixam-se às rochas da costa ou em fendas húmidas e vão subindo lentamente, variando a sua posição no extracto à medida que vão crescendo (Little *et al.*, 1991).

A existência de propágulos de dispersão ou fases larvares durante o ciclo de vida dos organismos intertidais contribui para estabelecer um fluxo génico entre populações de ilhas próximas, pois permitem a movimentação através de zonas de água mais profundas. A duração da fase pelágica do desenvolvimento larvar de *P. candei* não é conhecida, no entanto, está descrito para *P. vulgata* e *P. caerulea* que a larva torna-se completamente bentónica após metamorfose que tem lugar entre o nono e o décimo primeiro dia a contar da fertilização. Este período de dispersão de nove a onze dias

sugere capacidades de dispersão restritas (Smith, 1935). Quando a capacidade de dispersão de uma espécie é consideravelmente menor que o seu limite de distribuição geográfica, a diferenciação genética entre subpopulações deve aumentar com a distância entre elas. A ocorrência de picos de velocidade das correntes marítimas ou a extensão da fase larvar aquando da ausência de substrato adequado à fixação poderá no entanto conferir a estas espécies maior capacidade de dispersão (revisto em Weber *et al.*, 1998).

### 1.3. Diferenciação genética

Os estudos sobre a diferenciação genética das espécies de *Patella* têm-se baseado na análise de *loci* proteicos e de ADN mitocondrial, focando essencialmente análises a nível sub-específico e inferência de filogenias.

Usando informação obtida com base na análise de 15 *loci* aloenzimáticos, Côrte-Real e colaboradores (1996) observaram uma diferenciação significativa entre *P. candei gomesii*, subespécie dos Açores, e as subespécies que ocorrem na Madeira e nas Canárias (fig.1.4); entre as duas últimas foi obtido um valor de Identidade Genética de 0.969, típico de populações conspecíficas, enquanto o valor obtido com a subespécie dos Açores foi de 0,660, que se encontra dentro do limite esperado entre espécies congénicas. Baseando-se nessas diferenças, os autores sugerem que *P. candei gomesii* pode ser uma espécie diferente das formas de *P. candei* encontradas na Madeira e nas Canárias, contudo advertem para a necessidade de mais estudos.

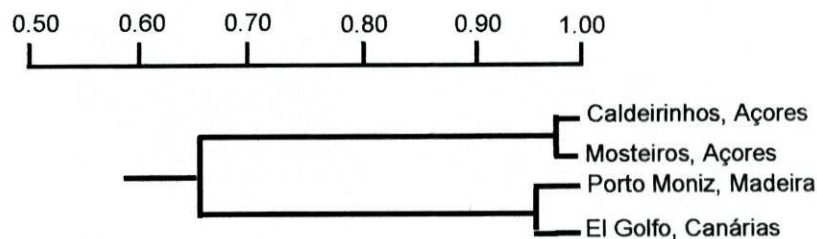


Fig.1.4 - Dendrograma UPGMA construído com base na Identidade Genética de Nei (1972) a partir da análise de 15 *loci* aloenzimáticos entre populações de *Patella candei* dos Açores, Madeira e Canárias (adaptado de Côrte-Real *et al.*, 1996).

Posteriormente, Weber e Hawkins (2002), analisando populações de *P. candei* dos Açores, Madeira, Selvagens e Canárias com base em 21 *loci* enzimáticos, detectaram a existência de dois grupos genéticos distintos: um incluindo *P. candei*

*gomesii* (Açores) e *P. candei candei* (Selvagens) e outro formado por *P. candei crenata* (Canárias) e *P. candei ordinaria* (Madeira) (fig.1.5).

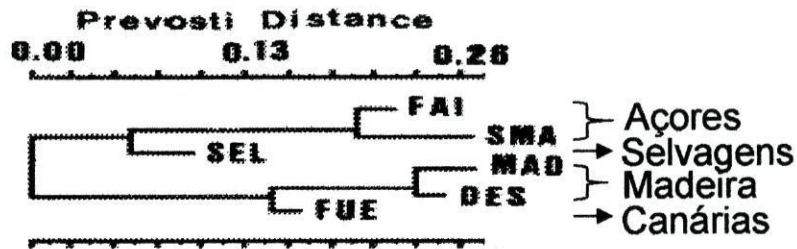


Fig.1.5 - Árvore de Wagner construída com base na Distância de Prevosti que mostra a diferenciação genética entre populações de *P. candei* dos Açores, Selvagens, Madeira e Canárias (Weber & Hawkins, 2002).

Estes autores observaram ainda que *P. c. gomesii* dos Açores não é geneticamente distinta de *P. c. candei* amostrada nas Selvagens, partilhando todos os alelos em dois *loci* que são diagnósticos em relação a *P. c. ordinaria* da Madeira (fig.1.6), devendo aquelas ser consideradas um complexo *P. c. candei*-*P. c. gomesii*. *P. c. crenata* não partilha nenhum *loci* diagnóstico com outras subespécies, sendo morfológicamente mais parecida com *P. c. ordinaria*. A ocorrência de hibridação detectada entre indivíduos das Canárias e das Selvagens sugere que os grupos *P. c. candei* - *P. c. gomesii* e *P. c. crenata* - *P. c. ordinaria* deveriam ser considerados como apenas duas subespécies de *P. candei*.

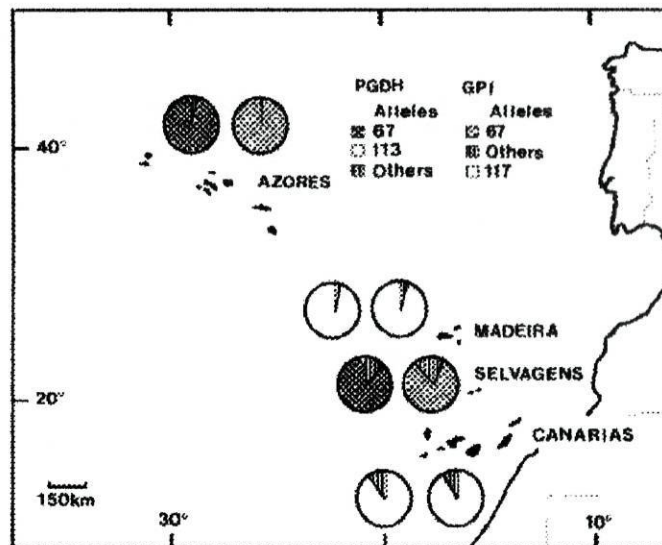


Fig.1.6 - *Loci* diagnósticos observados entre as subespécies de *Patella candei* na Macaronésia (Weber & Hawkins, 2002).

Com base na análise de três genes mitocondriais (subunidade I do gene da Citocromo c Oxidase, 12S ARNr e 16S ARNr), Sá-Pinto e colaboradores (2005) concluíram que *P. candei* é parafilética. Nas ilhas da Macaronésia foram observados dois grupos bem diferenciados, um que inclui *P. candei* dos Açores, Desertas e Madeira e outro que agrupa *P. candei* das Canárias e Selvagens com *P. lugubris* existente em Cabo Verde. O mesmo estudo demonstra que *P. caerulea* é a espécie mais próxima de *P. candei* e *P. lugubris* e que *P. depressa* se agrupa com estas três espécies com elevado suporte de *bootstrap* (fig.1.7). No mesmo estudo, verifica-se a ocorrência de cinco clados principais dentro do género *Patella*, dos quais três estão presentes nas ilhas da Macaronésia e apresentam diferenciação relativamente às populações do Continente.

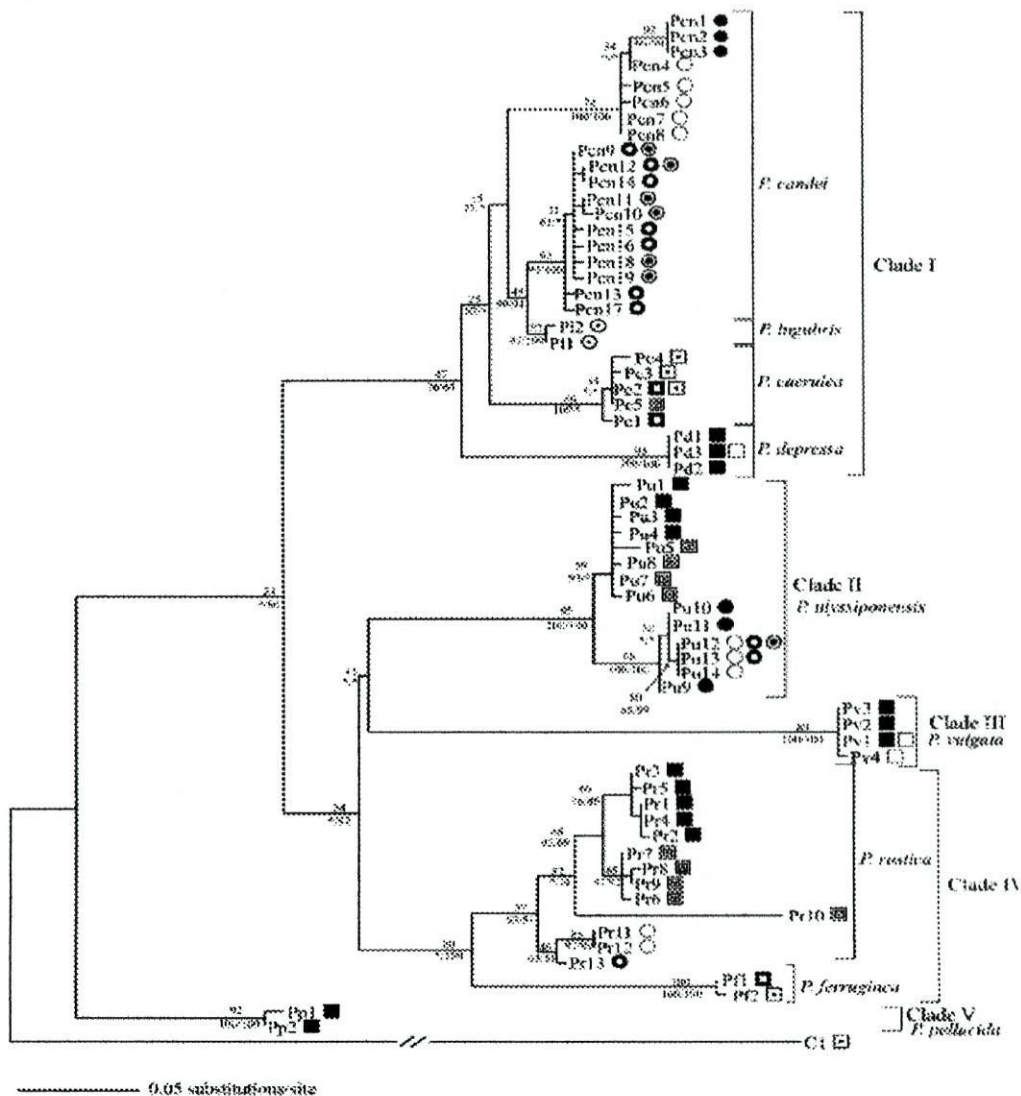


Fig.1.7 - Árvore de máxima verosimilhança (*Maximum likelihood*) baseada em sequências da subunidade I da Citocromo Oxidase. Os valores de *bootstrap* estão escritos por cima dos ramos; ● Açores; ○ Madeira e Desertas; ⊙ Selvagens; ⊗ Canárias; ○ Cabo Verde; □ Reino Unido; ■ Costa Atlântica da Península Ibérica; ■ Costa Mediterrânea da Península Ibérica; □ Tunísia; ▨ Grécia e Turquia (Sá-Pinto *et al.*, 2005).

Estes resultados não estão de acordo com a presente taxonomia, que reconhece duas espécies: *P. lugubris* em Cabo Verde e *P. candei* nas restantes ilhas da Macaronésia. Também não é suportada a divisão feita por Christiaens (1973), que descreveu quatro subespécies dentro de *P. candei*: *P. candei gomesii* nos Açores, *P. candei ordinaria* na Madeira e Desertas, *P. candei crenata* nas Canárias e *P. candei candei* nas Selvagens e Canárias, bem como a de Weber e Hawkins (2002), que observaram a existência de dois grupos genéticos distintos: um incluindo *P. candei gomesii* nos Açores e *P. candei candei* nas Selvagens e outro formado por *P. candei crenata* nas Canárias e *P. candei ordinaria* na Madeira. Possíveis explicações para as diferenças obtidas entre estes estudos com ADN mitocondrial e os estudos com enzimas, além do facto de estes marcadores contarem uma história evolutiva diferente, podem ser a ocorrência de uma introgressão histórica entre *P. candei* das Canárias e das Selvagens e um antepassado de *P. lugubris* ou a existência de homoplasia ou pressões selectivas semelhantes entre os dois grupos encontrados por Weber e Hawkins (2002) e por Corte-Real e colaboradores (1996) (Sá-Pinto *et al.*, 2005).

No complexo *P. candei*/*P. lugubris* verifica-se a ocorrência de quatro grupos diferenciados, reflectindo isolamento histórico entre quatro áreas geográficas: Açores, Madeira/Desertas, Selvagens/Canárias e Cabo Verde.

#### **1.4. Importância dos estudos genéticos para a conservação de *P. candei* e principais ameaças à diversidade genética da espécie**

Os recursos genéticos animais constituem por si só um património de valor inestimável. Este valor intrínseco é o fundamento da Genética da Conservação, que assenta sobre quatro aspectos (revisto em Diz, 2005): genético-reprodutivo, uma vez que a diversidade é necessária para manter a viabilidade das populações, a qual permite a adaptação a diferentes ambientes e à actuação da selecção artificial; científico, já que o conhecimento genético das populações permite caracterizar genes ou combinações génicas de relevância para investigação, no momento actual ou no futuro; ecológico-ambiental, pois os ecossistemas são resultado do equilíbrio entre clima, fauna e flora e qualquer factor que afecte algum destes domínios afectará também o resto dos seus processos de interacção; e histórico-cultural, dado que a conservação de determinadas espécies ou populações representa um património genético de uma região, com um desenvolvimento às vezes paralelo ao das populações humanas.

Diversos estudos têm realçado os efeitos da exploração humana, tanto para alimentação como para outros fins, sobre espécies intertidais, cuja redução de abundância e tamanho está bem documentada (Dulvy et al., 2003; Lindberg et al., 1998; Corte-Real et al., 1996; Navarro et al., 2005). A remoção de espécies grandes e/ou exploráveis pode ter efeitos profundos na estrutura da comunidade, particularmente quando se tratam de espécies chave (revisto em Casu et al., 2006).

Em 1983 foi descrita pela população local a ocorrência de uma “doença” nas lapas do grupo central. Hawkins e colaboradores (2000) defendem que a “doença” era apenas uma falsa justificação para a população sobre-explorar este recurso. Em consequência, as lapas tomaram-se muito mais raras e em 1988 as capturas em S. Miguel colapsaram.

Posteriormente, Navarro e colaboradores (2005) mostraram que os ecossistemas costeiros têm sido sujeitos a intensas perturbações humanas nas últimas décadas, sendo uma das mais importantes a colheita de macroinvertebrados para consumo e colecção. Nas Canárias, onde o estudo foi feito, os autores observaram que a pressão sobre as comunidades intertidais tem aumentado devido ao aumento da população nas ilhas e à gradual destruição dos habitats por alterações na zona costeira. Esta sobre-exploração pode ter sido a causa da extinção de *P. candei candei* em quase todo o arquipélago, o que mostra a importância de realizar estudos populacionais mais detalhados. É, assim, importante definir um estatuto de conservação adequado a cada espécie.

Existe pouca regulação das actividades marisqueiras sobre *P. candei*. A subespécie *P. c. candei*, endemismo da Macaronésia dos arquipélagos da Madeira, Canárias e Selvagens, é a única que tem um estatuto de conservação, estando incluída no Catálogo Nacional Espanhol de Espécies ameaçadas (9 Julho 1998, corrigida no BOE n.º 191 Agosto 1998) com a categoria de espécie em perigo de extinção (os factores negativos que incidem sobre esta subespécie fazem com que a sua sobrevivência seja pouco provável a curto prazo).

O desconhecimento sobre a estrutura populacional e a taxonomia das lapas no Atlântico Noroeste evidenciam a necessidade da realização de estudos genéticos mais aprofundados, que permitam definir medidas e gestão e conservação adequadas a esta espécie. Apesar de a aplicação de técnicas genéticas a populações exploradas de invertebrados ser semelhante à aplicada para vertebrados, têm sido realizados poucos estudos a este nível em gastrópodes comerciais. Os estudos sobre *Patella sp.* têm-se baseado na análise de aloenzimas (Lavie et al., 1987; Mauro et al., 2003; Weber &

Hawkins, 2005) e de ADN mitocondrial (Mauro *et al.*, 2003; Sá-Pinto *et al.*, 2005), estudos que mostram alguma discordância entre si. A baixa variabilidade das aloenzimas (algumas provavelmente sujeitas a selecção) e as particularidades do ADN mitocondrial, tomam aconselhável o desenvolvimento e aplicação de marcadores moleculares nucleares com maior variabilidade, como os microssatélites, que poderão potenciar a detecção de diferenciação populacional.

## **1.5. Marcadores adequados a estudos populacionais: microssatélites**

### **1.5.1. Descrição dos microssatélites**

Os microssatélites, também denominados *Short Tandem Repeats (STRs)* ou *Simple Sequence Repeats (SSRs)*, são pequenos fragmentos repetitivos abundantes no genoma eucariótico que exibem repetições em série de um motivo de dois a seis pares de bases (Tautz & Renz, 1984, Kashi & Soller, 1999). As variações no número de unidades de repetição, resultantes da perda ou ganho de uma ou mais unidades, traduzem-se num elevado grau de polimorfismo observado nestes *loci* em todos os organismos eucarióticos examinados até à data (Menotti-Raymond & O'Brien, 1995).

Com base no seu motivo de repetição, podem ser classificados como perfeitos – a sucessão das unidades de repetição não é interrompida - imperfeitos – a sequência de repetições está dividida por um ou mais nucleótidos - e compostos - quando são constituídos por motivos de repetição diferentes (Weber *et al.*, 1990).

Observa-se um padrão diferencial na distribuição dos microssatélites em regiões codificantes e não codificantes de moluscos que parece ser devido a uma selecção diferencial (Pérez *et al.*, 2005). Os microssatélites são menos abundantes em exões comparativamente às regiões não codificantes (Cruz *et al.*, 2005). Se bem que se tenha documentado a presença de microssatélites em regiões codificantes, intrões e ADN intergénico (Liu & Cordes, 2004), a maioria das sequências repetitivas de ADN não contribuem para o fenótipo dos indivíduos. Alguns microssatélites em moluscos apresentam as menores densidades genómicas observadas em eucariotas (ex. Cruz *et al.*, 2005). Ocorrendo maioritariamente em regiões não codificantes do genoma, são considerados selectivamente neutros (Ellegren, 2004).

Os marcadores neutrais têm provado ser uma ferramenta poderosa para investigar a existência de estruturação genética nas populações naturais, particularmente em organismos marinhos caracterizados por uma capacidade de dispersão limitada onde têm sido descritas diferenciações genéticas significativas a diferentes escalas espaciais (revisto em Casu *et al.*, 2006). Não se sabe ainda a função destes fragmentos, mas alguns investigadores sugerem uma intervenção em processos de recombinação, na evolução de novos genes, na condensação cromossómica, na formação de telómeros ou até na regulação dos genes (Stallings *et al.*, 1991) ou ainda na evolução dos genomas, ao criarem e manterem a variabilidade genética (Kashi & Soller, 1999).

### 1.5.2. Aplicabilidade dos microssatélites em estudos populacionais

Os microssatélites são excelentes marcadores para estudar relações entre subespécies e populações muito próximas devido à sua rápida taxa mutacional. Este tipo de estudos pode ser feito através da avaliação da diferenciação populacional, conhecimento do padrão de subdivisão geográfica e extensão do fluxo génico entre populações, inferência das flutuações demográficas, investigação de relações de proximidade genética entre populações e relacionamento de indivíduos com a população de origem mais provável (Jehle & Amtzen, 2002). O facto de a taxa mutacional variar entre *loci* permite ainda examinar diferentes escalas temporais.

Em geral, as regiões flanqueantes dos motivos de repetição apresentam-se altamente conservadas, permitindo a sua amplificação em espécies relacionadas (Brohede & Ellegren, 1999). São marcadores genéticos nucleares de hereditariedade mendeliana, multialélicos e codominantes, não revelando uma análise diferencial ao longo da ontogenia do indivíduo. A sua elevada variabilidade permite considerar cada indivíduo como uma única unidade taxonómica operativa (OTU- *Operational Taxonomic Unit*), possibilitando a aplicação de uma variedade de análises a nível individual.

A facilidade de amplificação por PCR (*Polimerase Chain Reaction*-Reacção em Cadeia da Polimerase) possibilita uma vasta utilização em estudos forenses e análises moleculares não-invasivas (Schotterer, 2004), uma vez que permite a amplificação em amostras de reduzida quantidade e/ou qualidade de ADN, mesmo quando se trata de ADN altamente degradado (Jame & Lagoda, 1996; Waser & Strobeck, 1998; Taberlet *et al.*, 1999).

A aplicação dos microssatélites apresenta no entanto algumas limitações. Mutações nas regiões flanqueantes são responsáveis pela ocorrência de alelos nulos ou silenciosos, isto é, alelos que não são amplificados na reacção de PCR (Jarne & Lagoda, 1996). Consequentemente, indivíduos heterozigóticos podem ser erradamente identificados como homozigóticos para *loci* onde a presença deste tipo de alelos se verifique; numa população não estruturada, esta diminuição na proporção de indivíduos heterozigóticos pode ser detectada pela observação de uma diminuição no valor da heterozigotia observada em relação à esperada considerando o Equilíbrio de Hardy-Weinberg (Bruford & Wayne, 1993).

Um outro problema na aplicação destes marcadores reside na ocorrência de homoplasia, ou seja, a existência de alelos idênticos como resultado de processos mutacionais convergentes e não ancestralidade comum. A homoplasia reduz o número de alelos observados por população, a proporção de indivíduos heterozigóticos e a diversidade genética (Estoup *et al.*, 1995; Jarne & Lagoda, 1996); é esperada em todos os modelos mutacionais excepto pelo IAM (*Infinite Allele Model*) de Kimura e Crow (1964), sendo normalmente mais elevada entre espécies do que entre populações da mesma espécie e rara dentro das populações. Não ter em consideração a homoplasia pode conduzir a uma subestimação da divergência real entre as populações estudadas. Contudo, a sua ocorrência não deverá constituir um problema significativo na maioria dos estudos de genética populacional, uma vez que a grande variabilidade encontrada nos microssatélites compensa, geralmente, a sua ocorrência (Jarne & Lagoda, 1996; Estoup *et al.*, 2002).

### 1.5.3. Evolução dos microssatélites

Apesar do seu uso corrente, o mecanismo responsável pela sua evolução é também pouco conhecido. As mutações serão resultado de um balanço entre a mutabilidade dos genomas e a eficiência dos mecanismos de reparação de emparelhamentos erróneos de cada espécie, selecção e limitação do tamanho (Pérez *et al.*, 2005; Cruz *et al.*, 2005). A variação dos microssatélites parece ser um fenómeno complexo influenciado principalmente pelo deslizamento (*slippage*) da polimerase durante o processo de replicação do ADN, originando a formação de ansas nucleotídicas responsáveis pelo aumento ou diminuição de uma ou mais unidades repetitivas (You-Chu *et al.*, 2002, Ellegren, 2004; Zane *et al.*, 2002).

A ocorrência de recombinação desigual entre duas cromátides pode também desempenhar um papel importante na evolução destes marcadores uma vez que os grandes saltos que levam à formação de alelos longos não são explicados pelo deslizamento da polimerase (Tóth *et al.*, 2000; Schlotterer, 2004). Esta teoria não é, no entanto, consensual (Ellegren, 2004, Huang *et al.*, 2002). A abundância diferencial de repetições que se observa entre exões, intrões e regiões inter-génicas, em numerosos eucariótas, sugere que um mecanismo de deslizamento e recombinação desigual é insuficiente para explicar a distribuição genómica dos microssatélites (Ellegren, 2000; Tóth *et al.*, 2000; Young *et al.*, 2000).

Outras causas poderão ser a ocorrência de fenómenos como a união entre hélices (*interhelical junctions*) durante o alinhamento cromossómico e a ocorrência de deleções ou inserções de bases, tanto no microssatélite como nas regiões flanqueantes (Wilder & Hollecher, 2001).

A taxa de mutação dos microssatélites pode ser determinada recorrendo a três métodos: a) contagem directa de mutações em pedigrees; b) estimativas indirectas de dados de ligação ou c) comparação dos valores observados e teóricos (Jame & Lagoda, 1996). Os valores observados variam entre  $10^{-6}$  e  $10^{-2}$ , duas a três ordens de magnitude superior aos valores conhecidos para as aloenzimas. A variação depende do organismo analisado (Jame & Lagoda, 1996; You Chu *et al.*, 2002), do tipo de unidade que se repete (di-, tri- ou tetra nucleótido), composição das bases e tipo de microssatélite (perfeito, composto ou interrompido) ((Bachtrog *et al.*, 2000; Balloux & Lugon-Moulin, 2002; Brohede *et al.*, 2002), e conseqüente nível de polimorfismo (numerosos alelos e elevada heterozigotia). De modo geral a taxa de mutação dos microssatélites está positivamente relacionada com o número de motivos de repetição (Pérez *et al.*, 2005; Jin *et al.*, 1996; Wierdl *et al.*, 1997; Schug *et al.*, 1998), ou seja, alelos com maior número de motivos de repetição estão sujeitos a mutações com mais frequência. Estudos recentes recorrendo à análise de mutações ocorridas em centenas de di e tetranucleotideos permitiram confirmar que existe um tamanho crítico para o qual as taxas de expansão e de contracção são idênticas, o que está relacionado com o facto de haver maior tendência para os alelos de menor tamanho aumentarem o número de unidades de repetição, enquanto os alelos com maior tamanho tenderem a diminuir, resultando num limite máximo para o tamanho dos alelos e uma aparente distribuição estável dos microssatélites (Xu *et al.*, 2000; Ellegren, 2000).

Estudos *in vitro* indicam que outros factores têm influência sobre a taxa de mutação dos microssatélites, para além do tamanho do motivo de repetição, como: a sequência da região flanqueadora (Kruglyak *et al.*, 1998; Santibanez-Koref *et al.*, 2001), a localização no cromossoma, o grau de heterozigotia e a substituição de bases na sequência do microssatélite (Schlotterer & Tautz, 1992; Goldstein & Clark, 1995; Harr *et al.*, 1998).

Três modelos evolutivos foram inicialmente adoptados de modo a traduzir e compreender melhor o comportamento mutacional dos microssatélites: o *Infinite Allele Model* (IAM), proposto por Kimura e Crow (1964); o *K-Allele Model* (KAM), de Crow e Kimura (1970) e o *Stepwise Mutation Model* (SMM), desenvolvido por Kimura e Ohta (1978).

Segundo o modelo IAM cada mutação envolve um número variável de unidades de repetição, sendo sempre responsável pela criação de um novo alelo (não presente na população) a uma taxa de mutação  $\mu$ ; estes diferentes estados podem não coexistir numa população pois alguns serão perdidos por deriva ou serão negativamente seleccionados, enquanto outros persistirão por serem benéficos em heterozigotia.

O modelo KAM assume que existe um determinado número ( $K$ ) de alelos possíveis na população e a mutação em cada um dos  $K-1$  alelos ocorre com a probabilidade  $\mu/(K-1)$ .

O terceiro modelo, SMM, presume que a mutação é responsável pelo aumento ou diminuição de uma única unidade de repetição ao alelo original, com igual e constante probabilidade  $\mu$ ; o alelo resultante poderá encontrar-se ou não previamente na população, havendo apenas dois estados adjacentes para os quais um alelo pode mutar.

Posteriormente, Di Rienzo e colaboradores (1994) desenvolveram um novo modelo, o *Two Phase Model* (TPM), que consiste num compromisso entre o SMM e o IAM. Este modelo sugere que a variação nestes marcadores resulta, principalmente, de alterações numa única unidade no número de repetições, contudo, outras mutações podem ocorrer implicando a alteração em mais do que uma unidade, sendo esta alteração explicada pelo modelo *Generalized Stepwise Model* (GSM). Deste modo, o alelo original é modificado em uma única unidade com a probabilidade aproximada  $P$  e em mais do que uma unidade com a probabilidade  $1-P$ .

Outros modelos têm vindo a ser desenvolvidos para explicar a evolução destes marcadores, no entanto, a controvérsia mantém-se e os métodos de análise estatística têm procurado incorporar três dos modelos referidos: o IAM, o SMM e o TPM, sendo na maioria dos estudos indicado o TPM ou GSM como os modelos mutacionais mais realistas (Ellegren, 2000; Schotterer, 2000)

## 1.6. Objectivos

Os principais objectivos deste trabalho consistiram na afinação para *P. candei gomesii* de marcadores microssatélites isolados em espécies próximas, como *P. depressa*, *P. caerulea* e *P. rustica*, de modo a analisar o nível de estruturação populacional desta espécie no arquipélago dos Açores.

Outro objectivo consistiu na comparação dos resultados obtidos neste trabalho com dados de ADN mitocondrial, existentes na bibliografia, de modo a obter maior suporte para o estabelecimento de um padrão de diferenciação.

Finalmente, pretendeu-se inferir os principais factores que afectam a diferenciação populacional e sua relação com aspectos de conservação e gestão desta sub-espécie.

## **2. Metodologia**

---

## 2.1. Área de estudo

O arquipélago dos Açores encontra-se situado a cerca de 1600 km do continente Europeu, distando cerca de 850 Km do arquipélago mais próximo, a Madeira. É um arquipélago Português com um grande número de ilhas, nove, que se distribuem em três grupos: Ocidental, constituído pelas ilhas Corvo e Flores; Central, composto pelas ilhas Faial, Pico, São Jorge, Graciosa e Terceira; e Oriental, do qual fazem parte as ilhas São Miguel e Santa Maria. Dentro do arquipélago dos Açores, a ilha do Corvo e a de Santa Maria distam entre si cerca de 600 km e o grupo ocidental dista aproximadamente 200 km do grupo central (fig.2.1). Todas as ilhas tiveram origem vulcânica, nunca tendo estado em contacto umas com as outras nem com qualquer outra massa de terra.

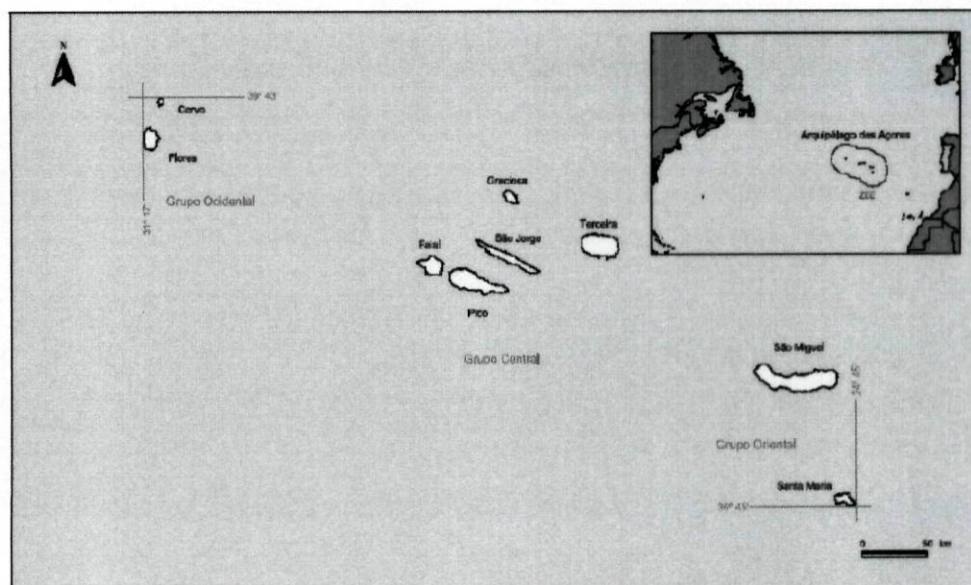


Fig.2.1 - Enquadramento geográfico do arquipélago dos Açores (retirado de Santos & Pinho, 2005).

A região dos Açores é fortemente influenciada por efeitos meteorológicos de diferentes proporções e com extremas variações ao longo do ano. Os ventos são a principal causa da existência de correntes superficiais nos oceanos. A Corrente dos Açores tem origem na ramificação da Corrente do Golfo que ocorre perto das coordenadas 40°N – 45° W e é uma das principais correntes de larga escala associadas à parte Este da Anticiclónica Subtropical Norte Atlântica. As correntes de superfície nos Açores movem-se geralmente em direcção à Madeira e depois ao Continente, com um padrão de deslocação de Oeste para Este (fig.2.2), apesar de em determinadas condições poderem sofrer inversão, principalmente no Outono e Inverno (revisto em

Hawkins *et al.*, 2000). Ao nível do arquipélago, a corrente dos Açores adquire uma velocidade média de  $10 \text{ m.s}^{-1}$  [1].

A Corrente das Canárias é formada por um braço da Corrente dos Açores que ruma a Este (Mason *et al.*, 2005).

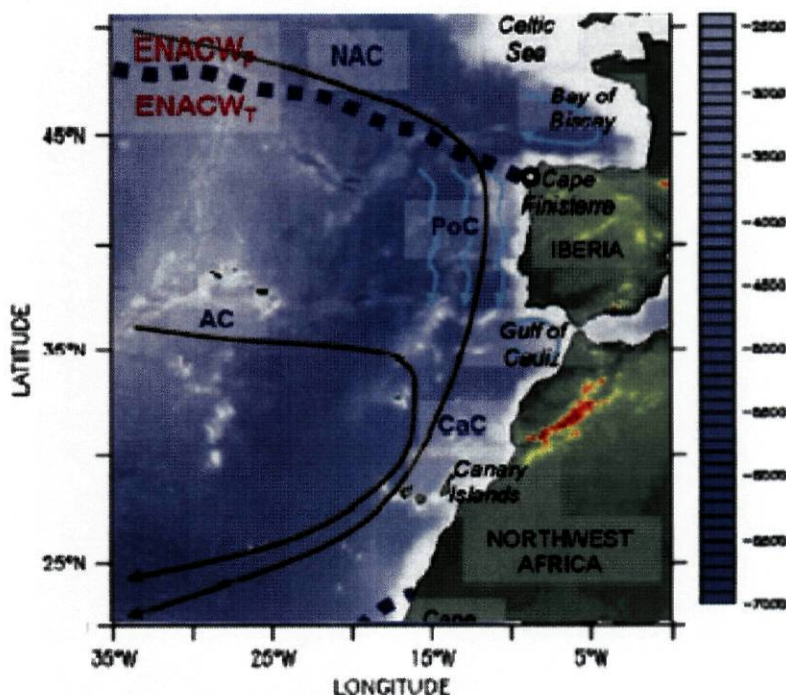


Fig. 2.2 - Representação das principais massas de água (ENACWP – Água Atlântico-Central de origem subpolar e ENACWT - Água Atlântico-Central de origem subtropical) e correntes da região Norte Atlântica Este (NAC – Corrente Norte Atlântica; PoC - Corrente de Portugal; AC – Corrente dos Açores e CaC – Corrente das Canárias) (Mason *et al.*, 2005).

## 2.2. Amostragem

Para a realização deste estudo foi utilizado um total de 310 indivíduos recolhidos num ponto por cada ilha (fig.2.3): Flores (52), Faial (41), Pico (31 indivíduos), Graciosa (47), Terceira (50), Santa Maria (50) e São Miguel (41), e 39 indivíduos recolhidos nas ilhas Canárias (analisadas como grupo externo), obtidas no âmbito do projecto POCTI/BSE/42003/2001 “As espécies do género *Patella* como modelo para o estudo dos padrões evolutivos e dinâmica de organismos marinhos”.

Todas as amostras de tecido foram preservadas em etanol 100% até ser efectuada a extracção de ADN.

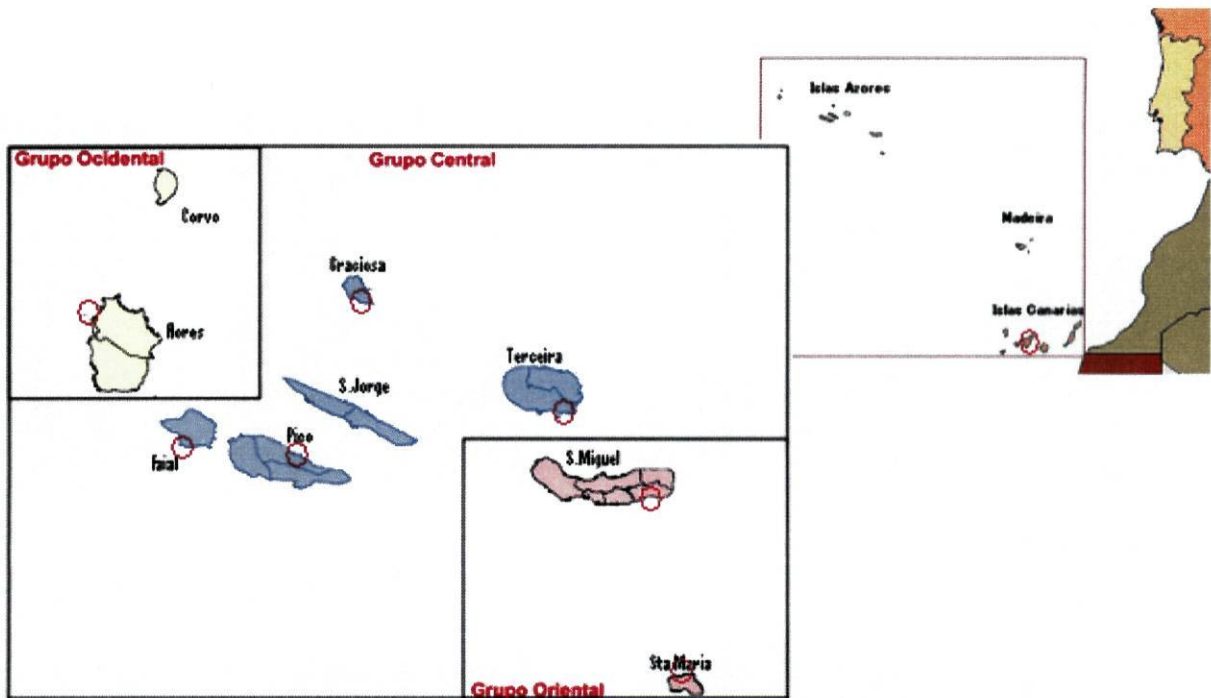


Fig.2.3 - Localização dos pontos de amostragem nos Açores e nas Canárias (círculos a vermelho).

### 2.3. Extração de ADN

A totalidade de ADN genómico foi extraída recorrendo a dois protocolos de extração: o Método Salino (Sambrook *et al.*, 1989) e um método derivado deste utilizando fenol-clorofórmio (Pérez, 2003; anexo 1).

O sucesso da extração de ADN foi avaliado por electroforese em gel 0.8% p/v a 300V em TBE 0.5x (Tampão Tris-borato-EDTA: Tris 89mM, ácido bórico 89mM, EDTA 2mM, pH 8.0), sendo aplicados 2µl de tecido misturados com 3µl de solução de azul de bromofenol (azul bromofenol 0,1%, EDTA 0,1M e sacarose 50g). A posterior coloração com brometo de etídio (0.05 % p/v) permitiu a visualização do ADN extraído por radiação ultravioleta e sua quantificação relativa (fig.2.4).

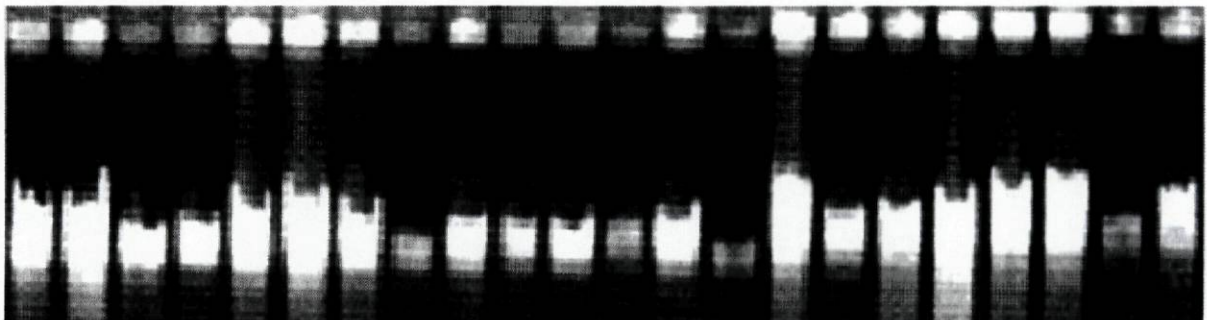


Fig.2.4 - Teste de extração de ADN num gel de agarose 0,8% (p/v) corado em brometo de etídeo.

De acordo com a concentração de ADN estimada, as amostras foram devidamente diluídas em água ultra-pura de modo a permitir uma boa amplificação por PCR e armazenadas a  $-20^{\circ}\text{C}$ , permanecendo nestas condições durante todo o trabalho laboratorial.

#### 2.4. Selecção de microssatélites e amplificação por PCR

A identificação genética individual através da utilização de microssatélites baseia-se no pressuposto de que, recorrendo a *loci* altamente polimórficos, se poderá traçar um perfil genético multilocus único para cada indivíduo. Existem casos em que as regiões flanqueantes de alguns microssatélites estão muito conservadas entre espécies, tendo sido amplificados com sucesso em organismos diferentes como Humanos e chimpanzés (4-6 milhões de anos) ou vaca e cabra (14-17 milhões de anos) (revisto em Llavona, 2006), excluindo a necessidade de clonar marcadores em cada espécie. Esta conservação interespecífica também se observa em invertebrados.

Deste modo, foi utilizado um conjunto de 20 microssatélites isolados em *P. depressa* (7) por Pérez (2007) e em *P. caerulea* (3) e *P. rustica* (10) por Pérez e colaboradores (não publicado). A descrição dos microssatélites apresenta-se na tabela 2.1.

A amplificação do ADN por PCR foi otimizada para cada um dos *locus* em análise, partindo de um gradiente de temperaturas ( $46-60^{\circ}\text{C}$ ) e de Magnésio (0.5 a 3.5 mM). Esta reacção dá-se em três fases: a) desnaturação – provoca a separação das duas cadeias de ADN, b) hibridação (*annealing*) – durante a qual ocorre a ligação dos *primers* à cadeia molde de ADN, e c) extensão – síntese de novos fragmentos de ADN mediante a actuação da polimerase do ADN (*Taq Polimerase*). A duração de cada etapa considerada, o número de ciclos, a temperatura de *annealing* e as quantidades dos diferentes reagentes da reacção foram determinadas de modo a otimizar as condições de PCR para cada um dos marcadores.

Em cada amplificação foi utilizado um controlo negativo sem ADN (branco), permitindo discriminar eventuais contaminações e um ou dois controlos positivos para despistar problemas relacionados com a reacção de PCR.

Tabela 2.1: Descrição dos microssatélites analisados: espécie onde foram isolados, motivo de repetição e tamanho do produto de PCR (pb) nessas espécies.

Espécie	Locus	Motivo de repetição	Tamanho
<i>P. depressa</i> <sup>1</sup>	Pde1	(GT) <sub>12</sub> (G) <sub>6</sub>	96-106
<i>P. depressa</i> <sup>1</sup>	Pde3	(GT) <sub>12</sub> G(GT) <sub>16</sub>	140-156
<i>P. depressa</i> <sup>1</sup>	Pde4	(CA) <sub>7</sub> ...(AC) <sub>6</sub> ...(CA) <sub>5</sub> ...(CA) <sub>1</sub>	193-205
<i>P. depressa</i> <sup>1</sup>	Pde5	(GT) <sub>4</sub> ...(GT) <sub>6</sub> ...(GT) <sub>9</sub>	176-200
<i>P. depressa</i> <sup>1</sup>	Pde6	(GTT) <sub>5</sub>	N.D.
<i>P. depressa</i> <sup>1</sup>	Pde7	(GT) <sub>9</sub>	173-185
<i>P. depressa</i> <sup>1</sup>	Pde9	(GT) <sub>22</sub>	162-208
<i>P. caerulea</i> <sup>2</sup>	Pca11	(AC) <sub>2</sub> GC(AC) <sub>7</sub>	100-134
<i>P. caerulea</i> <sup>2</sup>	Pca15	(TG) <sub>12</sub> (TATG) <sub>4</sub> (TG) <sub>12</sub>	350
<i>P. caerulea</i> <sup>2</sup>	Pca36	(CT) <sub>22</sub> ...(CA) <sub>12</sub>	130-170
<i>P. rustica</i> <sup>3</sup>	PrB7	(GT) <sub>19</sub> G <sub>11</sub>	144-150
<i>P. rustica</i> <sup>3</sup>	PrC2	C <sub>9</sub> (CA) <sub>16</sub>	157
<i>P. rustica</i> <sup>3</sup>	PrC5	(CGGA) <sub>n</sub>	160-200
<i>P. rustica</i> <sup>3</sup>	PrC12	(GATA) <sub>22</sub>	142-247
<i>P. rustica</i> <sup>3</sup>	PrD3	(GACA) <sub>7</sub>	173-183
<i>P. rustica</i> <sup>3</sup>	PrE2	(GACA) <sub>6</sub>	175-189
<i>P. rustica</i> <sup>3</sup>	PrE7	(TCTG) <sub>6</sub>	208-218
<i>P. rustica</i> <sup>3</sup>	PrF6	(GACA) <sub>9</sub>	247
<i>P. rustica</i> <sup>3</sup>	PrG6N	(CTGT) <sub>8</sub>	175-187
<i>P. rustica</i> <sup>3</sup>	PrG12	(GACG) <sub>6</sub>	265

<sup>1</sup> Pérez *et al.* 2007; <sup>2</sup> Pérez, dados não publicados; <sup>3</sup> Pérez *et al.*, dados não publicados.

A confirmação do sucesso de amplificação (fig.2.5) foi efectuada por corrida electroforética em gel de agarose 2% p/v. A electroforese e coloração foram realizadas de acordo com as condições referidas para o teste de extracção no caso de PCR's feitos com *Taq DNA Poly* e *Ecotaq*; para PCR's feitos com *GoTaq* tampão colorido – *Green buffer*) foram aplicados directamente no gel 2µl de amostra. A aplicação de um marcador de peso molecular conhecido (*Marker 5*, Eurogentec) permitiu inferir se o produto de amplificação correspondia ao esperado, de acordo com o tamanho médio documentado para os diferentes *loci*.

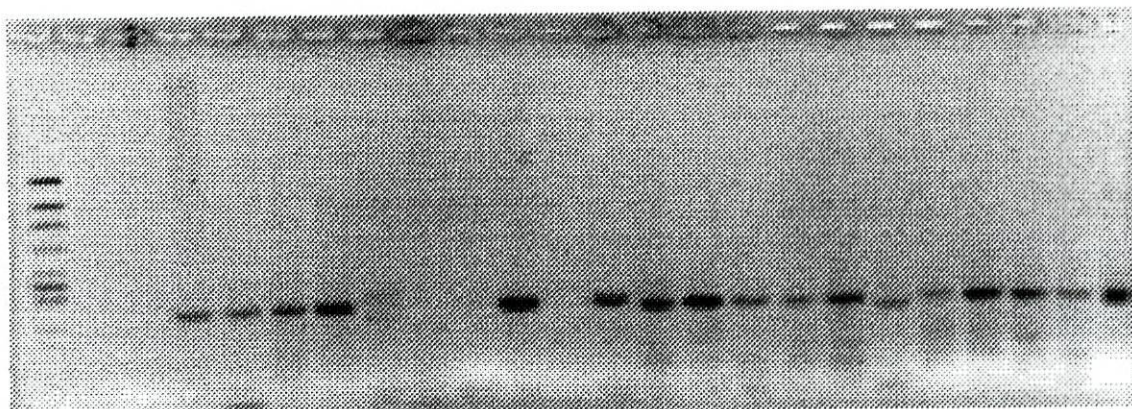


Fig.2.5 - Teste do sucesso da reacção de amplificação num gel de agarose 2% (p/v). Na primeira aplicação observa-se o marcador de peso molecular *Marker 5* (*Eurogentec*). Visualização feita por coloração com brometo de etídio.

## 2.5. Separação dos alelos e leitura dos genótipos individuais

Foram usados dois métodos de leitura dos alelos: sem marcação dos primers – géis manuais – e com marcação dos primers – sequenciador de gel.

### 2.5.1. Genotipagem em géis manuais

De acordo com a concentração de ADN estimada a partir do teste de amplificação, cada amostra foi diluída numa solução desnaturante (formamida 95%, hidróxido de sódio 10 mM, azul de bromofenol 0,05% e xilenocianol 0,05%), aquecida a 94°C durante três minutos para desnaturação e imediatamente conservadas em gelo; foram aplicados 4 µl de amostra no cátodo de um gel de poliacrilamida desnaturante 6% p/v (previamente aquecido a 48°C). Foram adicionados no gel marcadores de peso molecular conhecido (10bp *DNA ladder*, Invitrogen) de modo a permitir a leitura dos alelos. Após a genotipagem de alguns indivíduos, foram sempre colocadas amostras de referência nos géis seguintes para confirmar as leituras.

A separação dos fragmentos de ADN decorreu por electroforese vertical, a uma potência constante de 60 W em TBE 1x e à temperatura de 48°C. A duração da corrida dependeu do tamanho médio dos alelos em cada *locus*, variando entre 110 e 125 minutos.

A visualização dos alelos foi efectuada por coloração do gel em nitrato de prata, seguindo um protocolo de acordo com as indicações do fabricante (*Silver Sequence™*

*DNA Staining Reagents Kit, Promega*). A leitura dos tamanhos dos fragmentos foi feita com precisão de 1pb (fig.2.6).

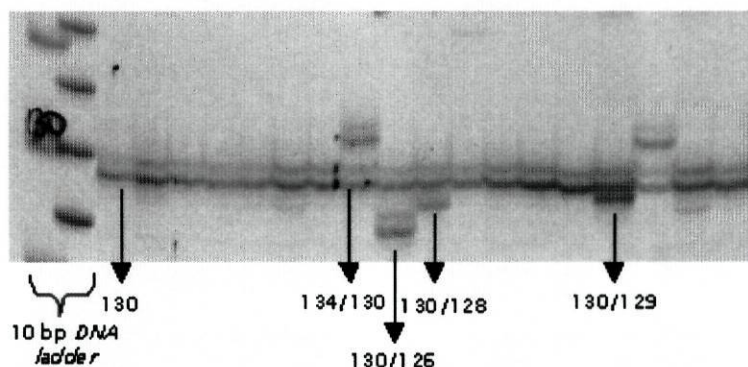


Fig.2.6 - Visualização dos alelos obtidos para o marcador Pde3 em gel de poliacrilamida desnaturante 6% p/v por coloração com nitrato de prata. Está indicada a leitura de alguns genótipos individuais.

### 2.5.2. Genotipagem em sequenciador de gel

Para a detecção dos fragmentos amplificados no sequenciador automático ALFexpressII (fig.2.7), foi marcado um dos primers específicos de cada *locus* com o fluoróforo Cy5. Este fluoróforo é sensível ao laser do sequenciador, sendo detectado pelos seus sensores.

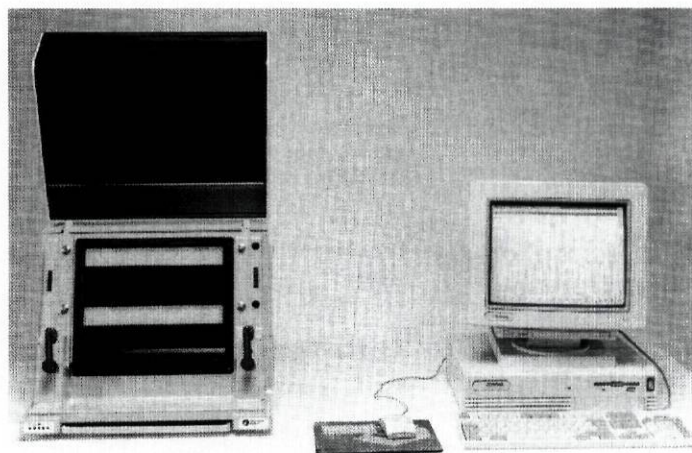


Fig.2.7 – ALFexpressII e computador com o *software* de leitura (Amersham Biosciences, 1998).

Este procedimento baseia-se na electroforese vertical em géis de acrilamida na qual os fragmentos de ADN têm um dos primers marcados com um fluoróforo para a sua detecção por laser. A electroforese realiza-se num recipiente vertical onde o controle da temperatura é feito por um circuito integrado de água destilada através de uma das

placas onde se coloca o gel. Durante a electroforese os fragmentos marcados migram através do gel. O laser incorporado no sistema atravessa o gel de um lado ao outro, na perpendicular à migração do gel. O gel usado no sequenciador tem 40 poços que são carregados com uma mistura constituída pelo produto de PCR da amostra, marcadores internos de tamanho conhecido adequado ao tamanho dos alelos e tampão de carga Azul de Dextrano (5mg/ml em formamida). Antes de serem carregadas, as amostras foram desnaturadas. Este sequenciador emprega um sistema monofluoróforo. Quando o fragmento de ADN passa pelo laser durante a sua migração, as bandas marcadas com o fluoróforo Cy5' excitam-se emitindo fluorescência que é detectada pelos diodos (componentes que conduzem a energia eléctrica) do sequenciador. Os sinais do diodo são enviados directamente ao computador para serem armazenados e posteriormente processados (Amersham Biosciences, 1998). A transmitância do laser para uma leitura correcta deve ser superior a 80%. Durante a migração formam-se deformações no gel que podem ser corrigidas pelo software interpolando as bandas das amostras problemáticas com a ajuda dos marcadores introduzidos.

As migrações foram feitas com as seguintes condições: 1500 V, 60 mA e 20W a 55°.

O computador gera imagens virtuais do resultado, uma delas correspondente ao gel, semelhante à que se observa nos géis manuais, e a outra é um conjunto de curvas de absorvância com os valores dos picos de emissão de fluorescência que nos indicam o tamanho dos diferentes alelos (fig.2.8).

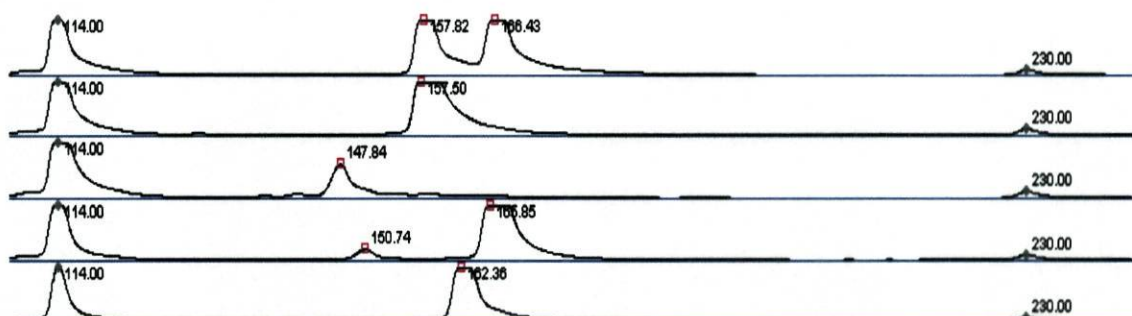


Fig.2.8 – Curvas de absorvância do sequenciador ALFexpressII.

A análise dos dados foi feita com o *software* ALFwin Fragment Analyser 1.00.36 (Amersham Pharmacia Biotech). Após introduzir no programa informação do tamanho dos picos que funcionam como marcadores internos, com tamanho conhecido, e dos picos de cada amostra, o programa identifica os últimos em relação ao tempo de

migração. Posteriormente alinha-se o gel de forma a corrigir deformações. Para a leitura devemos ter em conta diversos factores tais como possíveis contaminações dos poços adjacentes ou erros típicos do próprio sistema.

### 2.5.3. Erros de genotipagem e sua correcção

Os erros de genotipagem ocorrem quando o genótipo determinado após a análise molecular não corresponde ao genótipo real do indivíduo considerado. Na prática, são definidos como as diferenças entre dois ou mais genótipos obtidos independentemente para a mesma amostra. Estes erros podem ocorrer em qualquer dos passos da genotipagem (Bonin *et al.*, 2004; Pompanon *et al.*, 2005) e consistirem no aparecimento de alelos nulos, alelos intermédios, alelos semelhantes aos marcadores internos, bandas *stutter*, picos +A e alelos *drop-out* (Castro *et al.*, 2004; Matsumoto *et al.*, 2004).

Os alelos nulos são alelos não amplificados que, quando em heterozigotia, resultam num aparente homozigótico. Para os microssatélites, os alelos nulos podem aparecer quando ocorrem mutações nas regiões flanqueantes, impedindo um ou os dois *primers* de se ligarem. A presença de alelos nulos em marcadores microssatélites é frequente, especialmente em grandes populações naturais e é um fenómeno comum em populações de moluscos marinhos (Bruford & Wayne, 1993; Li *et al.*, 2003; Karlsson & Mork, 2005). O seu aparecimento produz um erro importante nas proporções genotípicas, de modo que se observa um claro desvio do equilíbrio de Hardy-Weinberg devido a um excesso de homozigóticos. Existem várias estatísticas para estudar a frequência de alelos nulos. Baseiam-se na frequência de indivíduos que não amplificaram, no número de alelos nulos que se inferem de análises de parentesco e na relação entre presença de alelos nulos e défice de heterozigóticos (Dakin & Avise, 2004).

Os alelos intermédios são aqueles que não correspondem a uma unidade de repetição do microssatélite, tendo um tamanho intermédio para o que seriam séries alélicas perfeitas. Estes alelos podem corresponder a mutações pontuais (delecções ou inserções) dos marcadores ou a desvios do gel durante a migração (Pino Querido-Ferreira, 2004).

Os alelos de tamanho próximo dos marcadores internos podem confundir-se com estes, subestimando-se a sua frequência. Normalmente são reconhecíveis pela sua maior intensidade.

As bandas *stutter* (sombra) são produzidas por deslizamentos da ADN polimerase durante a reacção de PCR de modo que se originam fragmentos que são uma ou várias repetições mais curtos ou mais longos que o verdadeiro alelo. Após migração, estas bandas apresentam-se como uma cadeia de picos seguidos sobressaindo um que será o verdadeiro alelo (Matsumoto *et al.*, 2004).

Os picos +A são artefactos um nucleotídeo maiores que o verdadeiro alelo. Estes são criados pela ADN polimerase ao adicionar um nucleotídeo, geralmente Adenina, à extremidade 3' do fragmento amplificado (Matsumoto *et al.*, 2004).

Os alelos *drop-out* são alelos que aparecem bastante afastados da série principal (Castro *et al.*, 2004). O pico de absorvância detectado é em geral menor que o pico prévio em indivíduos heterozigóticos.

No caso do sequenciador ALF as diferenças de migração entre géis podem ser elevadas, de modo que é muito importante realizar a leitura seguindo os mesmos critérios em todos os géis para poder comparar os dados, por exemplo repetir em todos os géis a mesma amostra controle.

Em estudos de genética populacional, os erros de genotipagem afectam as estimativas das frequências alélicas e a discriminação dos diferentes genótipos. Falsas estimativas da frequência alélica podem levar à consideração de um excesso de homozigóticos, errar na atribuição do Equilíbrio de Hardy-Weinberg, sobrestimar hibridação ou inferências erróneas da estruturação populacional.

Devido à potencialidade infinita das fontes de erro, é impossível erradica-las completamente. Uma das aproximações mais usadas para reduzir estes erros é a baseada na replicação da amplificação do ADN para a mesma amostra mas acarreta custos adicionais, pelo que não pode ser usada para a totalidade das amostras quando se trabalha com uma quantidade considerável de indivíduos.

Após finalizada a tipagem dos 5 *loci*, os dados foram analisados de modo a detectar possíveis erros. Esta análise foi feita em diferentes passos, tendo sido iniciada por uma primeira análise das homozigotias e heterozigotias e seguindo o diagrama de ajuda (fig. 2.9) do programa MICRO-CHECKER 2.2.3 (Van Oosterhout *et al.*, 2004). Este consiste numa aplicação desenvolvida para procurar alelos nulos, bandas *stutter* e alelos *drop-out*, que ocorrem durante o processo de PCR e que vão afectar a leitura dos microssatélites.

Esta aplicação usa um método baseado em simulações de Monte-Carlo (*bootstrap*) para gerar frequências esperadas das diferenças de tamanho dos alelos em homozigotia e heterozigotia. A teoria de equilíbrio de Hardy-Weinberg é usada para calcular frequências alélicas esperadas e a frequência de quaisquer alelos nulos detectados. A frequência esperada de homozigóticos para cada classe é calculada com base na frequência de heterozigóticos para essa classe. O número total esperado de homozigóticos é então calculado e comparado com o número observado.

Após a detecção dos erros, foi feita a análise para um intervalo de confiança de 95% incluindo os dados considerados suspeitos pelo programa. Se a frequência estimada de alelos nulos for significativa ( $> 0,05$ ), o programa cria genótipos ajustados, onde o número de genótipos homozigóticos em cada classe de tamanho é ajustado de modo a reflectir o número "real" estimado de homozigóticos. O valor zero substitui um dos alelos homozigóticos que, no ficheiro de entrada para os próximos programas é substituído por um alelo que não se encontra nos dados e de tamanho superior ao maior detectado. Estes genótipos ajustados são ordenados de acordo com o tamanho dos alelos, não correspondendo à ordem original das amostras. Consequentemente, não pode ser feita qualquer análise multilocus com estes dados.

Existem quatro estimadores de alelos nulos, desenvolvidos por Chakraborty e colaboradores (1992), por Brookfield (1996) e por Van Oosterhout e colaboradores (2004), cada um mais apropriado para diferentes situações (*MICRO-CHECKER guide book*):

- Estimador Brookfield 1: quando todas as amostras analisadas amplificaram, mostrando uma ou duas bandas; ou após eliminar amostras que não amplificaram quando se pensa que estas podem representar alelos nulos, ADN degradado ou problemas com a PCR;
- Estimador Brookfield 2: quando se tem a certeza que as amostras não amplificadas representam alelos nulos;
- Estimador Chakraborty e colaboradores: quando as amostras não amplificaram devido a degradação do ADN (não amplificaram para os outros *loci*).
- Estimador Oosterhout: assume que não existem homozigóticos nulos e considera todas as amostras não amplificadas como ADN degradado, erro humano, etc.; a frequência dos alelos nulos é estimada a partir do cálculo prévio das frequências alélicas de todos os *loci* visíveis.

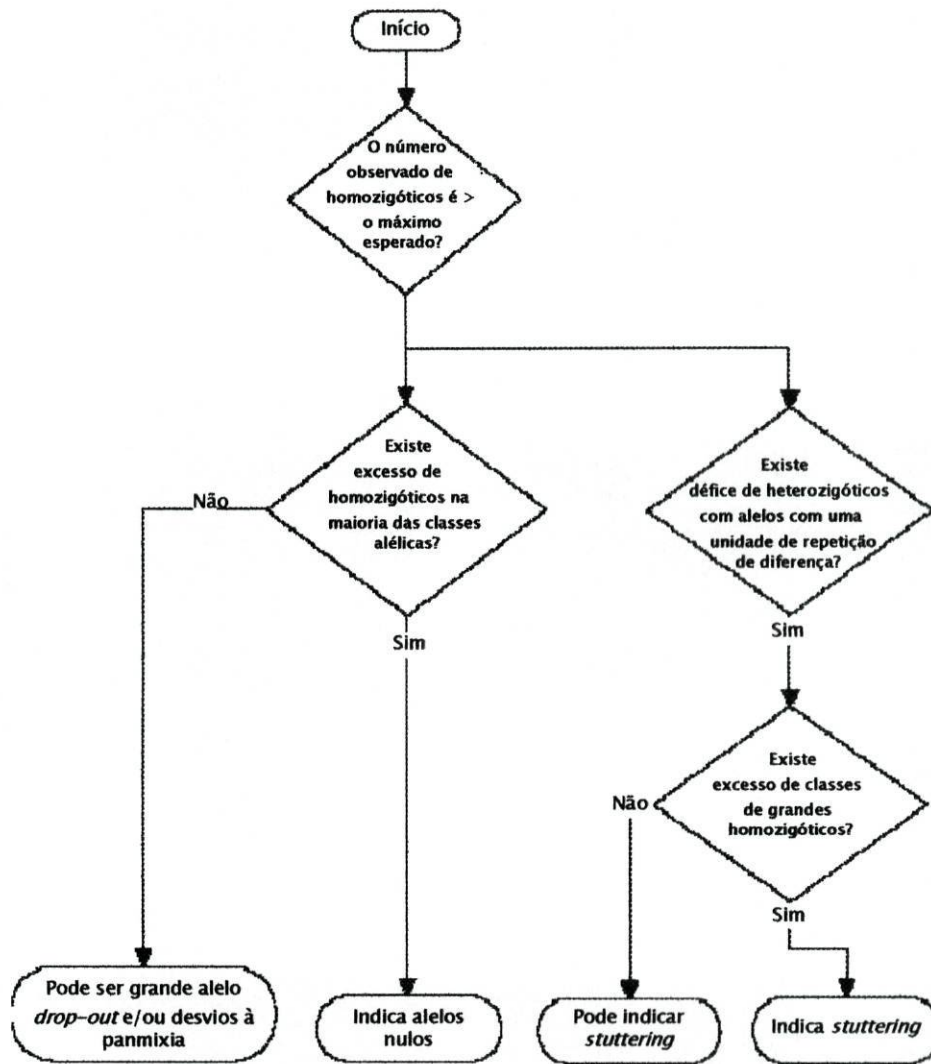


Fig. 2.9 – Diagrama de ajuda na detecção de erros de genotipagem (adaptado de MICRO-CHECKER *guide book*).

Estas equações estimam a frequência de alelos nulos a partir da aparente deficiência em heterozigóticos. Assumem panmixia, isto é, assumem que a deficiência heterozigótica observada é devida a alelos nulos e não causada por efeitos de Wahlund, *inbreeding*, cruzamentos direccionados, etc..

## 2.6. Análise da diversidade genética

### 2.6.1. Frequências alélicas

A constituição alélica é definida pelos alelos que existem para cada *locus* numa população. As frequências alélicas informam-nos da abundância de cada

um dos alelos e são por isso o primeiro parâmetro representável que indica diferenças de diversidade genética entre populações.

As frequências dos alelos de cada *locus* foram calculadas directamente a partir dos dados genotípicos usando o programa GENEPOP 3.4 (Raymond & Rousset, 1995).

## 2.6.2. Estimativas de diversidade genética

As estimativas de diversidade genética foram calculadas com a ajuda dos programas GENEPOP 3.4 (Raymond & Rousset, 1995) e FSTAT 2.9.3.2 (Goudet, 1995).

### - Número de alelos por *locus*

Trata-se do número total de alelos distintos que cada *locus* apresenta em todas as populações analisadas. No caso de microssatélites com um motivo repetido dinucleotídeo, os tamanhos dos microssatélites serão geralmente pares ou ímpares. Isto pode variar quando ocorre algum tipo de mutação em que não se repete o motivo completo ou em microssatélites que apresentem, por exemplo, uma cauda poli-A, em que a variação pode ser mononucleotídica.

### - Número médio de alelos por *locus*

É a média aritmética do número de alelos distintos que cada *locus* apresenta, calculada para cada *loci* em todas as populações.

### - Riqueza alélica ( $R_s$ )

A riqueza alélica é uma medida do número de alelos independente do tamanho da amostra, o que permite comparar amostras com diferentes tamanhos. Foi calculada com o programa FSTAT 2.9.3.2 (Goudet, 1995) utilizando a seguinte expressão (Petit *et al.*, 1998):

$$R_s = \sum_{i=1}^{N_i} \left[ 1 - \frac{\binom{2N - N_i}{2n}}{\binom{2N}{2n}} \right]$$

O cálculo baseia-se na estimativa do número de alelos numa subamostra de  $2n$  genes dos  $2N$  genes que foram amostrados ( $N \geq n$ ). No FSTAT,  $n$  é definido como o número mais pequeno de indivíduos para um *locus* numa amostra.  $N_i$  é o número de alelos de tipo  $i$  nos  $2N$  genes. Cada termo do somatório corresponde à probabilidade de

amostrar o alelo  $i$  pelo menos uma vez numa amostra de tamanho  $2n$ . A riqueza alélica permite-nos calcular o tamanho efectivo da população (Petit *et al.*, 1998).

$R_s$  depende do tamanho da amostra retirada da população, especialmente nas populações que não apresentam efeito gargalo de garrafa (*bottleneck*) (Leberg, 2002), uma vez que nestas se conseguem detectar todos os alelos amostrando um menor número de indivíduos. A expressão aplicada neste estudo apresenta uma correcção a este problema, de modo que para tamanhos amostrais suficientemente grandes (30 indivíduos como neste caso) a riqueza alélica descrita por Petit e colaboradores (1998) seria independente do tamanho da amostra.

#### **- Número de alelos por locus e por população ( $A$ )**

Trata-se do número de alelos distintos que cada população apresenta para cada um dos *loci*. Este cálculo dá-nos uma ideia do estado de uma população uma vez que quanto maior for o número de alelos por *locus* e por população maior diversidade apresenta essa população.

#### **- Número médio de alelos por locus e por população**

É a media aritmética do número de alelos distintos que cada população apresenta para cada um dos *loci*.

#### **- Alelos privados ( $A$ privados)**

São os alelos presentes exclusivamente numa população, normalmente com frequências baixas. Estes alelos permitem melhorar a designação de um indivíduo a uma das populações amostradas, isto se existir uma estrutura clara na população.

#### **- Tamanho alélico modal ( $A$ modal)**

Corresponde ao tamanho alélico mais frequente na população. Na maioria dos casos as populações apresentam uma distribuição unimodal, podendo ocasionalmente ser bimodal.

#### **- Amplitude alélica ( $A$ amplitude)**

Indica a amplitude dos tamanhos dos alelos de cada *locus* em cada população. Expressa-se como o valor máximo e mínimo obtidos para cada *locus* em cada população.

### - Diversidade genética ou heterozigotia

A heterozigotia refere-se à proporção de indivíduos heterozigóticos numa população. Esta proporção apresenta um valor observado ( $H_o$ ) que pode diferir do seu valor esperado ( $H_e$ ). As diferenças entre ambas reflectem-se na estatística  $F_{IS}$ .

### - Índice $F_{IS}$

A estatística  $F$ , ou Índice de Fixação, foi descrita por Wright (1921) para caracterizar a distribuição genotípica num *locus* de dois alelos. Posteriormente, o mesmo autor desenvolveu este conceito descrevendo três novos índices:  $F_{IS}$ ,  $F_{ST}$  e  $F_{IT}$ .

As frequências genotípicas observadas podem comparar-se com as frequências esperadas de acordo com a formulação de Hardy-Weinberg. Estes autores demonstraram que em gerações consecutivas a distribuição dos genótipos na população não se altera. Os desvios ao equilíbrio de HW indicam que pelo menos um dos requerimentos não se cumpre. O grau de desvio é geralmente formulado como:

$$F_{IS} = (F_{IT} - F_{ST}) / (1 - F_{ST})$$

donde  $F_{IT}$  é o cruzamento (*inbreeding*) ou a correlação de genes dentro do indivíduo;  $F_{ST}$  é a correlação de genes de diferentes indivíduos da mesma população (coancestry) e  $F_{IS}$  é a correlação de genes dentro de indivíduos dentro de populações (Wright, 1951, 1965).

Desenvolvendo esta fórmula obtém-se a expressão:

$$F_{IS} = (H_e - H_o) / H_e = 1 - (H_o / H_e)$$

donde  $H_o$  é a heterozigotia observada (número observado de heterozigóticos) e  $H_e$  a heterozigotia esperada (número de heterozigóticos esperados a partir das frequências alélicas estimadas).

Esta estatística descreve a intensidade e o sinal do desvio ao equilíbrio. É positivo quando existe uma deficiência de heterozigóticos (ou excesso de homozigóticos) na população. Quando  $F_{IS}$  é negativo, existe um excesso de heterozigóticos. Em alguns casos,  $F_{IS}$  pode ter um valor positivo muito elevado devido a factores como presença de alelos nulos, cruzamentos preferenciais, efeito de Wahlund, etc..

O valor desta estatística obteve-se pelo método descrito por Weir & Cockerham (1984) com ajuda do programa GENEPOP 3.4 (Raymond & Rousset, 1995). Este programa baseia-se no uso de testes exactos, ou seja, não afectados pelo conhecimento imperfeito de alguns parâmetros. Baseia-se em técnicas de permutação. O valor de  $P$  obtém-se da soma de probabilidades de todas as amostras e pode ser estimado com o GENEPOP 3.4 (Raymond & Rousset, 1995) considerando uma subamostra ao acaso das amostras iniciais mediante duas técnicas distintas: a permutação ou os algoritmos das Cadeias de Markov baseadas em réplicas de Monte Carlo. Em ambos, a probabilidade de amostrar um elemento particular é igual à sua probabilidade de ocorrência segundo a hipótese nula. Os algoritmos da Cadeia de Markov usam-se com a mesma finalidade que as permutações mas são mais exactos e rápidos quando se trabalha com grandes tamanhos amostrais e com número elevado de *loci* (Raymond & Rousset, 1997).

## 2.7. Análise da estruturação genética

### - Índice $F_{ST}$

$F_{ST}$  define-se como a correlação entre gâmetas tirados ao acaso da mesma subpopulação em relação ao total (Wright, 1951, 1965). Também se pode definir como a variância estandardizada das frequências génicas ou coeficiente efectivo de hibridação (Weir & Cockerham, 1984). Os subíndices desta estatística significam subpopulação ( $S$ ) e população total ( $T$ ).  $F_{ST}$  é uma medida do grau de diferenciação entre subpopulações. Para cada *locus*, pode ser definida combinando as frequências de todos os alelos desse *locus* (Chakraborty & Leimar, 1987).

Esta estatística estima a conectividade e os padrões de fluxo génico entre populações. Inicialmente foi descrito para estudar o efeito de cruzamentos dentro das populações, e posteriormente ampliou-se o conceito a populações interrelacionadas.  $F_{ST}$  proporciona uma medida da distância genética e descreve-se como:

$$F_{ST} = 1 - (H_S / H_T)$$

donde

$$H_S = (\sum_i H_i) / n \text{ é a média de } H_i$$

$H_i = 1 - \sum_k p_{ik}^2$ , é a heterozigotia esperada segundo o equilíbrio de HW para a subpopulação, ou a medida da probabilidade de não identificação dos alelos amostrados na mesma subpopulação  $i$ .

$$H_T = 1 - \sum_k p_k^2$$

$$p_k = (\sum_i k p_{ik}) / n$$

sendo  $p$  e  $k$  as respectivas frequências dos alelos de um *locus* (Chakraborty & Leimar, 1987).

$F_{ST}$  foi descrito como a correlação entre dois alelos escolhidos ao acaso dentro de subpopulações e os alelos amostrados no total das populações (Wright, 1951, 1965). Weir & Cockerham (1984) descrevem esta estatística como:

$$F_{ST} = \sigma_p^2 / p(p-1)$$

donde  $p$  é a frequência média da amostra de alelo  $A$  e  $\sigma_p^2$  é a variância.

Um valor nulo de  $F_{ST}$  indica-nos que amostramos uma unidade panmítica, isto é, uma população em que os indivíduos se reproduzem ao acaso, e as frequências genotípicas podem ser calculadas a partir das gaméticas. Neste caso há ausência de diferenciação. Quando o valor de  $F_{ST}$  se aproxima da unidade significa que existe divergência entre as populações e que pelo menos duas das populações amostradas estão fixadas para diferentes alelos.  $F_{ST}$  é sensível ao alto polimorfismo quando a migração é escassa (Raymond & Rousset, 1995; Balloux & Lugon-Moulin, 2002).

Os valores da estatística  $F_{ST}$  variam então de 0 (toda a diversidade génica é compartilhada igualmente dentro de uma espécie e não há diferenças genéticas entre as populações) a 1 (toda a diversidade génica dentro de uma espécie é encontrada como diferenças fixas entre populações e sem diversidade génica dentro das populações).

Este valor é estimado no programa ARLEQUIN 3.01 (Excoffier *et al.*, 2005).

#### - Estimativa do fluxo génico ( $Nm$ )

Esta estimativa é inversamente proporcional à proporção de variância dentro de uma espécie devida à subdivisão populacional, isto é, ao valor de  $F_{ST}$ . Segundo Balloux e

Lugon-Moulin (2002), a estrutura genética de uma espécie reflecte o número de alelos intercambiados entre populações e estes fluxos génicos tornam homogêneas as frequências de alelos entre essas populações, determinando os efeitos relativos da selecção e da deriva genética e, conseqüentemente, a composição genética dos indivíduos. Um fluxo génico alto evita a adaptação local, reduzindo a fixação de alelos que são favorecidos sob condições locais e impedindo o processo de especiação. Por outro lado, o fluxo génico gera novos polimorfismos nas populações e aumenta o tamanho efectivo da população local e sua habilidade de resistir a mudanças aleatórias nas frequências génicas, opondo-se à deriva genética e gerando novas combinações de genes nos quais a selecção natural pode actuar. Esta estimativa é determinada com o programa GENETIX (Belkhir *et al.*, 1996).

#### - Testes exactos de Fisher

O teste exacto de Fisher é um teste estatístico usado para determinar se existem associações não ao acaso entre duas variáveis dicotómicas, através de uma matriz 2X2. O programa calcula a diferença entre os dados observados e os dados esperados, assumindo um modelo de independência. Para calcular o valor de probabilidade do teste, as tabelas devem estar ordenadas segundo um critério que meça a dependência, e as tabelas que apresentem desvios de independência iguais ou maiores que os da tabela observada ficam com as suas probabilidades unidas. Existe uma série de critérios que podem ser utilizados para medir a dependência. No caso do teste exacto de Fisher são tipicamente usados o qui-quadrado de Pearson ou diferenças nas proporções, que são equivalentes (Weisstein, *on-line*).

Para a análise estatística foi usado o método AMOVA, Análise de Variância Molecular. É um método para estimar diferenças populacionais partindo de dados genéticos, dando uma ideia da estrutura genética populacional medida através da partição da variância genética. Os dados dos marcadores microssatélites podem ser analisados desta forma para comprovar se a diferenciação estimada por permutações ao acaso é significativa. No entanto, não existem algoritmos que introduzam os padrões de desequilíbrio na análise genética da estrutura populacional, de modo que se assume uma população em equilíbrio (Excoffier *et al.*, 1992, Michalakis & Excoffier, 1996).

## - AMOVA

Para realizar a AMOVA empregou-se o programa ARLEQUIN 3.01 (Excoffier *et al.*, 2005). Este programa permite definir grupos de populações nos quais se quer comprovar a existência de diferenciação genética. Para o cálculo desta estatística parte-se de uma matriz inicial de genótipos. A AMOVA trata os dados moleculares como um vector booleano,  $p_i$ , ou seja, uma matriz  $1 \times n$  de zeros e uns que indicam a ausência ou presença de um marcador. Calculam-se as distâncias euclídeas entre pares de vectores, calculam-se as distâncias quadradas euclídeas e constrói-se uma matriz dividida em submatrizes correspondentes às divisões dentro da população (Excoffier *et al.*, 1992). As somas dos quadrados podem ser analisadas por uma análise de variância aninhada. Os dados moleculares são distâncias euclídeas derivadas de vectores de uns e zeros de modo que é improvável que sigam uma distribuição normal. Computa-se uma distribuição nula reamostrando os dados. Em cada permutação atribui-se ao acaso cada indivíduo a uma população mantendo o tamanho da amostragem constante. Estas permutações repetem-se várias vezes até construir uma distribuição nula. É importante que os dados derivem de um número adequado de marcadores. Neste estudo foram usados 5 marcadores microssatélites. Foram aplicados os valores 1000000 para o número de passos da Cadeia de Markov e 50000 para o número de passos de memorização e 1500 permutações para testar a significância dos Índices de Fixação. Foram feitas duas análises com agrupamentos de populações distintos com base na distribuição geográfica. Na primeira análise foram definidos quatro grupos, de acordo com a localização geográfica, tendo a seguinte composição:

- Grupo 1: Açores (Flores, Faial, Pico, Graciosa, Terceira, São Miguel e Santa Maria);
- Grupo 2: Canárias.

Na segunda análise optou-se por usar apenas as populações dos Açores e definir três grupos:

- Grupo 1: Flores;
- Grupo 2: Faial, Pico, Graciosa e Terceira;
- Grupo 3: São Miguel e Santa Maria.

## - Inferência Bayesiana

Os métodos de análise Bayesiana apresentam inúmeras vantagens relativamente aos anteriores, nomeadamente o facto de serem baseados num modelo genético, incorporarem a incerteza inerente ao facto das frequências alélicas serem estimadas e, conseqüentemente, conhecidas com um determinado nível de erro e permitirem uma modelação e aplicação computacional facilmente adaptadas a diferentes aproximações (Anderson & Thompson, 2002). Assumem, ainda, particular importância quando as populações parentais não possuem diferenças alélicas fixas nos marcadores genéticos analisados.

Por definição, o paradigma Bayesiano baseia-se no facto da probabilidade ser uma medida directa de incerteza relativamente a um determinado parâmetro fixo, mas desconhecido. Deste modo, a análise Bayesiana procura fazer inferências relativamente ao valor desse parâmetro (Schoemaker *et al.*, 1999), através da obtenção de uma distribuição de probabilidade posterior (*posterior distribution*) de acordo com os dados obtidos e o conhecimento *a priori*, ainda que incompleto, da distribuição de um ou mais parâmetros do modelo (*prior distribution*). Esta inferência consiste na especificação dessa distribuição *a priori* (que representa uma certeza relativamente aos valores dos parâmetros) e utilização do teorema de Bayes para a transformar na distribuição *a posteriori*, que representa a certeza após observação dos dados (Dawson & Belkhir, 2001).

De modo a testar o poder dos microssatélites estudados na alocação de indivíduos nas populações foi feita uma análise Bayesiana com o programa BAPS - *Bayesian Analysis of Population Structure* - versão 4.14 (Corander *et al.*, 2003), que estima estruturação populacional escondida agrupando populações (isto é, localizações geográficas de amostragem) em grupos panmíticos. Esta versão do BAPS usa optimização estocástica para inferir a estrutura genética, ao contrário do algoritmo MCMC usado pelo programa STRUCTURE (Pritchard *et al.*, 2000), também de análise bayesiana, dando resultados 100 a 1000 vezes mais depressa e com um nível de exactidão comparável. O agrupamento das populações foi feito usando um modelo não-espacial e outro espacial (que incorpora coordenadas dos locais de amostragem), definindo o número máximo de populações ( $K$ ) igual a 8. As coordenadas utilizadas na análise espacial encontram-se no Anexo 2; estas não correspondem aos verdadeiros pontos de amostragem, uma vez que não existe essa informação, mas a pontos próximos na ilha respectiva.

## 2.8. Análise do gene mitocondrial Citocromo c Oxidase

Foram analisadas 172 sequências da subunidade I do gene mitocondrial Citocromo c Oxidase, provenientes de 27 indivíduos das Flores, 18 da Graciosa, 24 do Pico, 19 de S. Miguel, 23 do Faial, 24 de Santa Maria, 20 da Terceira e 16 das Canárias, recorrendo ao programa NETWORK 4.0.2.1. (Fluxus Technology Ltd., *on-line*), de modo a obter uma distribuição dos diferentes haplótipos. As sequências foram obtidas a partir do trabalho de Sá-Pinto e colaboradores (*in press*).

O programa NETWORK gera árvores evolucionárias a partir de informação genotípica. Pode, assim, fornecer estimativas de idade para qualquer ancestral na árvore. É um programa usado para reconstruir redes e árvores filogenéticas, inferir tipos ancestrais ou potenciais, ramificações e variantes evolucionárias e estimar datações. Os algoritmos foram desenhados para bio-moléculas não-recombinantes, como por exemplo o ADN mitocondrial.

Este *software* foi desenvolvido para reconstruir todas as árvores de máxima parsimónia para os dados inseridos. Apenas considera informação variável, isto é, posições que permitam separar os indivíduos em pelo menos dois grupos. Inclui duas opções para a construção das redes; foi escolhida a opção que utiliza um algoritmo *Median-Joining* (Bandelt *et al.*, 1999) e que permite que a informação tenha posições ambíguas ou várias possibilidades (por exemplo, em determinada posição pode haver um nucleotídeo A, C, G, ou N).

### **3. Resultados**

---

### 3.1. Selecção de microssatélites e amplificação por PCR

Dos 20 *loci* utilizados inicialmente, 5 não amplificaram; para os restantes, a temperatura de *annealing* variou entre os 46°C para o Pca11 e os 58 para o Pca15 e a concentração de MgCl<sub>2</sub> variou entre 1 mM para o Pde5 e 3 para o Pde1, sendo as condições mais comuns uma temperatura de 50°C e uma concentração de MgCl<sub>2</sub> de 1,5 mM. A enzima *Gotaq* apenas não amplificou 3 *loci*: o Pde1 e o Pde7 que foram amplificados com a polimerase *Taq DNA Poly* e o Pde5 para o qual foi utilizada a enzima *EcoTaq*. As condições para todos os *loci* estão descritas na tabela 3.1.

Tabela 3.1 - Condições de amplificação dos microssatélites.

Locus	Sequências dos primers F e R (5´-3´)	T. (°C)	MgCl <sub>2</sub> (mM)	Enzima
Pde1	GGA TTA CCC GTA GAA ACC ACC TGG TTA GCA TTG AGC	50	3	Taq DNA Poly
Pde3	GTA TTA TGT CGT CGC CCC TC ATG CCC CAA AAA CAA TAG G	55	2,7	GoTaq
Pde4	CTG TAG GCG AAC CAC AGA C GCA TCC CAG GTT AGT ATG TAG		Não amplifica	
Pde5	CGG AGC AGA ACA GTA GGA C CAA TCA AAT GGG AAT AAA CAT	48	1	EcoTaq
Pde6	AAC CGA ACA CAT TAC GAC G AGC TTA TAC CCA CAA CAA A	57	1,5	Gotaq
Pde7	CAA CAA CCA GGG GCG TAT T AGA AAT GCG GCC TGT AGG T	55	2,5	Taq DNA Poly
Pde9	CGC GAG AGA ATT TGT GCC TA ATC AAG TGG CCT GTG GTT TC	50	2	Gotaq
Pca11	AAG CCC CAA CTT CAC TAC ATG GAC ATT CAA AGC CAG	46	2	Gotaq
Pca15	CTT CTT CAT GGG GAC TTC GGA TCT TCC TTG GGG TAT	58	1,5	Gotaq
Pca36	GAA CTA GCC GTG CCA ATA TCA ACT AGG TCG CTT CTG AG	50	2	Gotaq
PruB7	CAA ACG AAAA ATC TGT GA AAT AAT GCC TTT CTA CGC	55	1,5	Gotaq
PruC2	AAT AAT GCC TTT CTA CGC CGA CGA CTT AGA CAA ACG	50	2	Gotaq
PruC5	ACC CTA CAA GTC CAA GAA CC TTT TAT ACG CCC ACA TTT TT	50	1,5	Gotaq
PruC12	TGT CTT CAT GCT TAT TA TTT TCA TCC ACT TTT AT		Não amplifica	
PruD3	AAT AAA CAC CTA CAC CAA GT GTC ACC CTC CTC TCC TAC		Não amplifica	
PruE2	GTA CGC TCT AAC AAA AT TTC TAT ACA ATC CGA TT		Não amplifica	
PruE7	ATC TCA GCA TTG GCT C AAG TTT GAA CAT CCT AGC		Não amplifica	
PruF6	CGT TAT AGT GTC CCA AAA AGT ATT CTC TCA AAC CCC ATC CAG	54	1,5	Gotaq
PruG6N	TAG CTT CGA CAG GCA ACC CGT GTC CCA AAA AGT GTG	50	1,5	Gotaq
PruG12	TGA ATC ATC CAG TTA CTG TA CGT AGT ATT AGA TGT TTT GG	53	1,5	Gotaq

Após obter sucesso na amplificação, foram tipados 5 indivíduos por população ou uma população inteira para detectar o polimorfismo dos microssatélites. Dos 15 *loci* que amplificaram 8 revelaram-se monomórficos (tabela 3.2). Dos microssatélites polimórficos foram seleccionados 5 de acordo com o número de alelos que apresentavam, dando-se preferência aos *locus* mais variáveis: Pde1, Pde3, Pde5, Pde6 e PrC5. Na tabela 3.2 está descrito o intervalo de tamanho dos alelos detectados para cada *loci*, que variam entre um mínimo de 94 pb (Pde1) e um máximo de 350 pb (Pca15).

Tabela 3.2 - Descrição da variação do tamanho de alelos detectados e consequente classificação dos *loci* quanto ao morfismo alélico; P: polimórfico; M: monomórfico.

<i>Locus</i>	Pde1	Pde3	Pde5	Pde6	Pde7	Pde9	Pca11	Pca15	Pca36	PrB7	PrC2	PrC5	PrF6	PrG6N	PrG12
Tamanho	94-121	120-138	211-259	117-129	172	130	118	340-350	168	155	146	116-196	200	250-261	134
Morfismo	P	P	P	P	M	M	M	P	M	M	M	P	M	P	M

### 3.2. Detecção de erros de genotipagem

Face à observação da existência de um elevado défice de heterozigóticos para a maioria dos *loci* e populações, foi utilizado o programa MICRO-CHECKER para detectar a presença de alelos nulos. Foi aplicada o estimador Oosterhout por ser o que melhor se aplica. Foram detectados erros que possivelmente correspondem a alelos nulos em todas as populações para os *loci* Pde5 e PrC5; para o *locus* Pde3 apenas não foram detectados alelos nulos nas populações Pico e Graciosa e para o *locus* Pde1 não foram detectados estes alelos nas populações das Flores, Pico e São Miguel; para o *locus* Pde6 apenas foram detectados alelos nulos para a população da Graciosa.

### 3.3. Análise da diversidade genética: frequências alélicas e estimativas de diversidade genética

As frequências alélicas encontram-se representadas na forma de gráficos nas figuras 3.1 a 3.5 e descritas em tabela no anexo 3. À excepção de Pd6, que possui dois alelos quase equifrequentes em todos os locais de amostragem, todos os *loci* apresentam um alelo muito frequente, comum a todos os locais de amostragem. A frequência estimada de alelos nulos chega por vezes a ser mais elevada que a frequência do alelo não nulo mais comum. Para as Canárias, os *loci* Pde1, Pde5 e PrC5 apresentam uma distribuição mais homogénea das frequências alélicas por diferentes classes de tamanho.

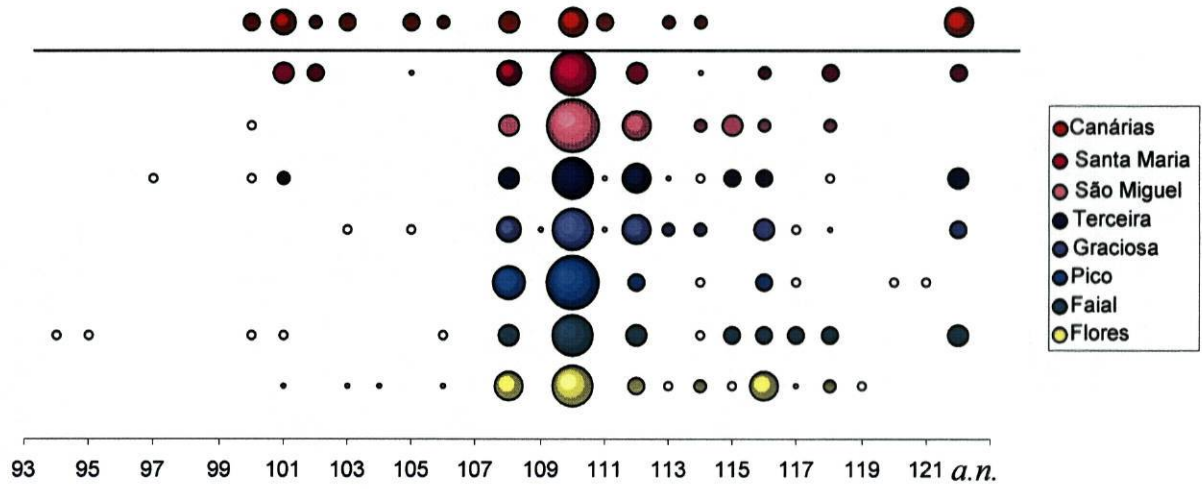


Fig. 3.1 - Frequências alélicas do locus Pde1 para cada população (a.n.- alelo nulo).

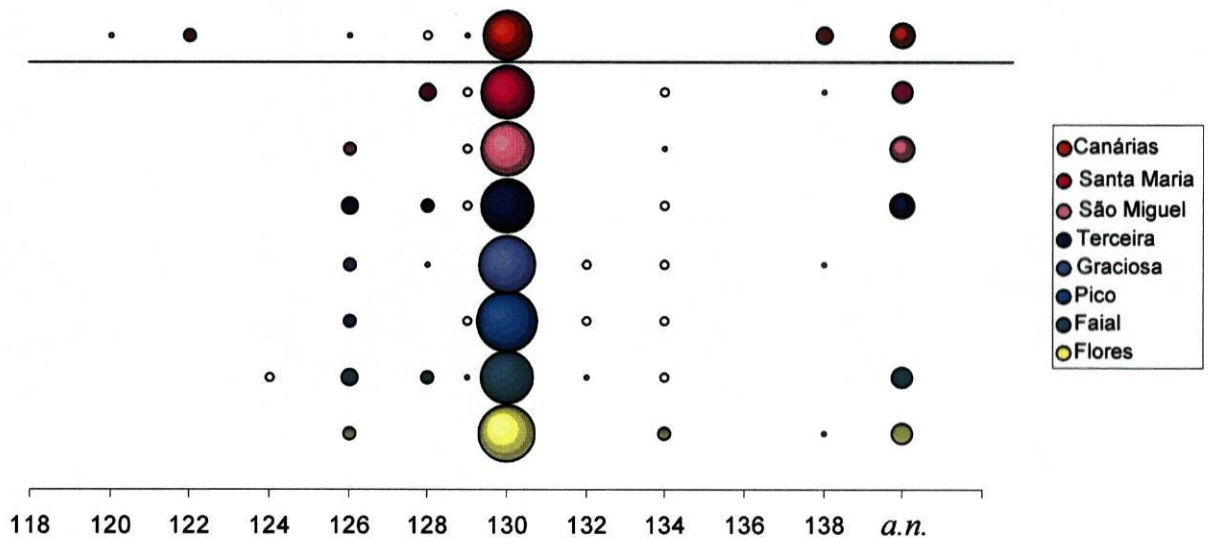


Fig. 3.2 - Frequências alélicas do locus Pde3 para cada população (a.n.- alelo nulo).

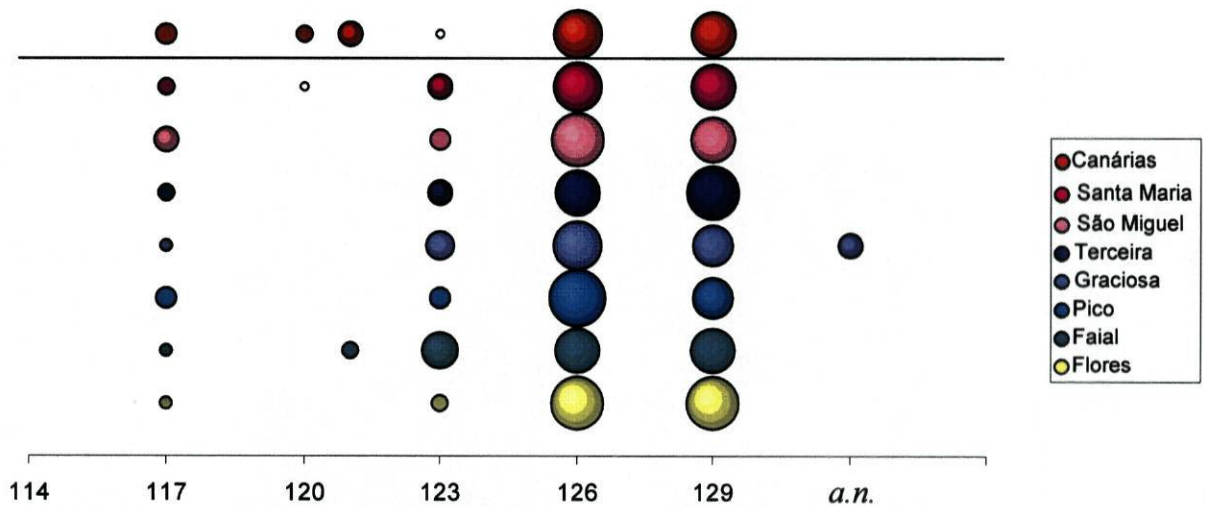


Fig. 3.3 - Frequências alélicas do locus Pde6 para cada população (a.n.- alelo nulo).

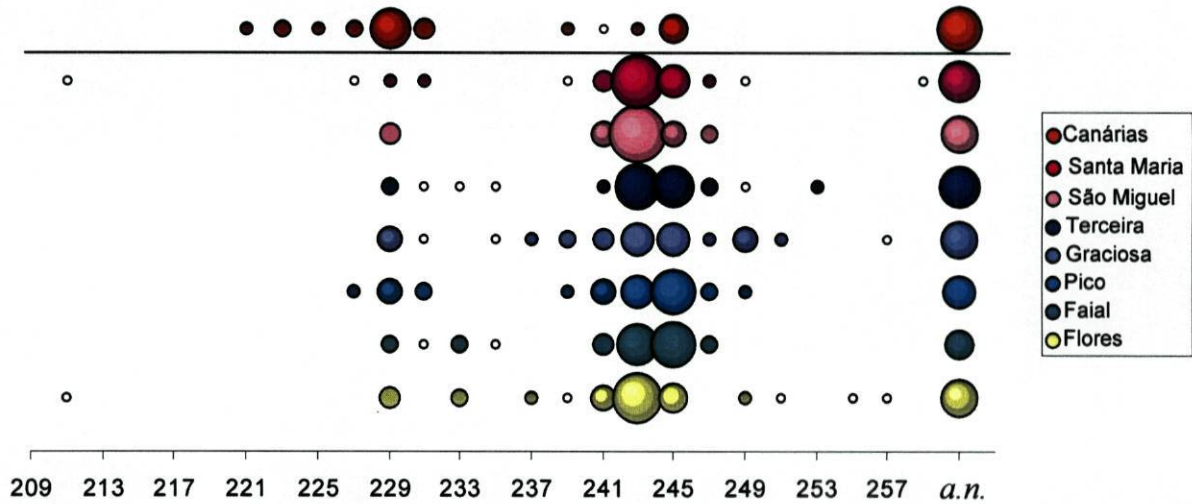


Fig. 3.4 - Frequências alélicas do locus Pde5 para cada população (a.n. - alelo nulo).

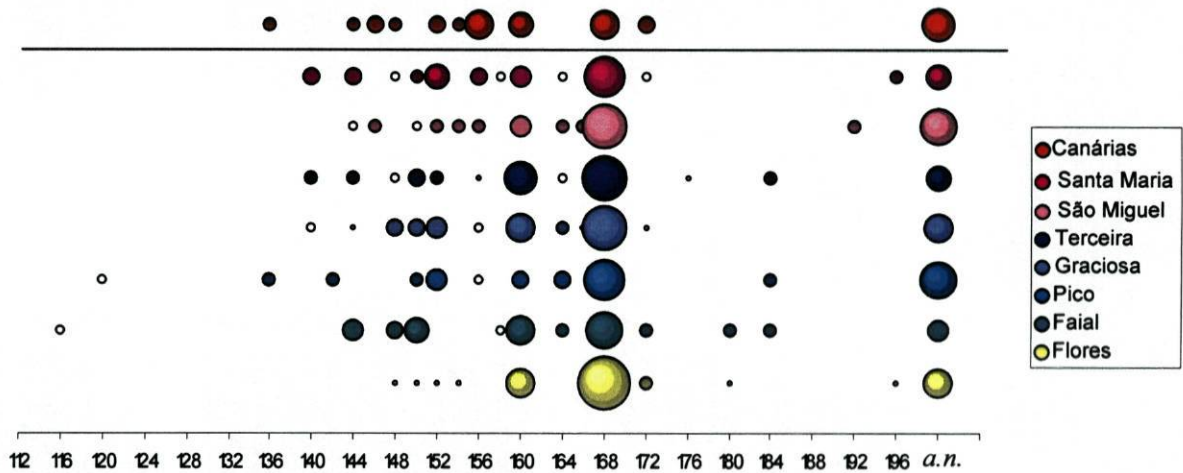


Fig. 3.5 - Frequências alélicas do locus PrC5 para cada população (a.n. - alelo nulo).

Nos 5 *loci* tipados o número de alelos por *locus* variou entre 7 para o Pde6 e 25 para o Pde1 (fig. 3.6 A). O maior valor de número médio de alelos foi o obtido para a população da Graciosa com um valor na ordem dos 9,6 alelos para 47 indivíduos analisados; o menor valor aparece para a população de São Miguel com um valor na ordem dos 7 alelos para 41 indivíduos tipados (fig. 3.6 B).

Os parâmetros da diversidade genética determinados para os 5 *loci* analisados estão representados na tabela 3.3, estando apenas contabilizado o alelo considerado nulo nos parâmetros “número de alelos” e “riqueza alélica” apesar de estar incluído na determinação dos restantes parâmetros. Dos *loci* analisados, o Pde1 revelou-se o mais polimórfico, com um número de alelos por população entre 8 e 14. Os microssatélites Pde5 e PrC5 apresentam entre 6 a 13 e 10 a 13 alelos diferentes por população,

respectivamente, já incluindo um alelo nulo. Para os restantes *loci* foram detectados entre 5 a 8 alelos para o Pde3 (alelo nulo incluído), e 4 a 6 alelos para o Pde6. Em termos de riqueza alélica, a relação mantém-se, apresentando o Pde1 o maior valor (12,383) e o Pde6 o menor valor (3,860).

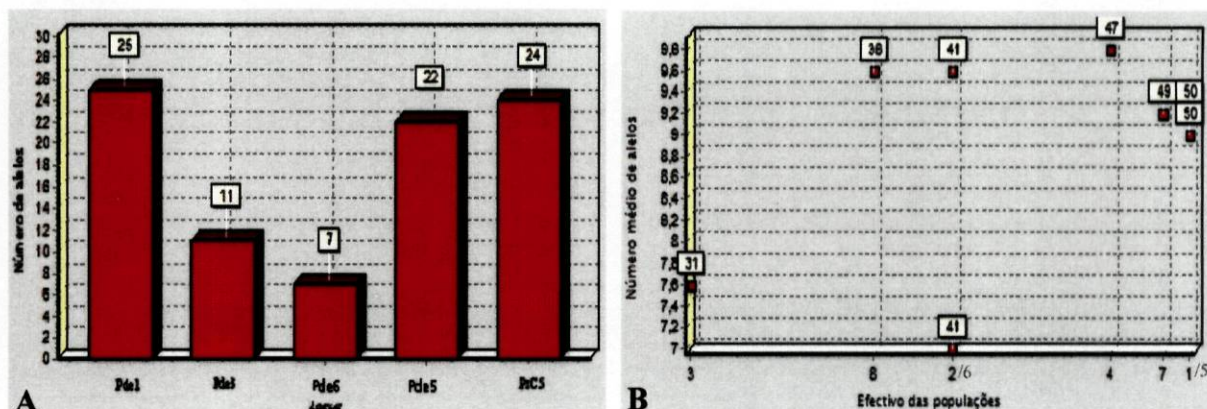


Fig. 3.6 - A: Representação gráfica do número de alelos por *locus* detectados no total de indivíduos (350); B: Representação gráfica da relação entre o número médio de alelos e o efectivo das populações consideradas (1 – Flores; 2 – Faial; 3 – Pico; 4 – Graciosa; 5 – Terceira; 6 – São Miguel; 7 – Santa Maria; 8 – Canárias).

O *locus* Pde6 apresenta apenas um alelo privado na população das Canárias, enquanto o Pde1 apresenta 8 alelos privados distribuídos por 5 populações.

Os *locus* Pde1, Pde3 e PrC5 apresentam um valor modal igual para todas as populações (110, 130 e 168, respectivamente); o Pde6 apresenta uma distribuição unimodal para o alelo 129 nas ilhas Flores, Faial e Terceira e 126 nas restantes ilhas. O *locus* Pde5 apresenta um valor bimodal para a população da Graciosa (243-245) e valores unimodais para os alelos 245 nas populações de Faial e Pico, 229 para a população das Canárias e 243 para as restantes populações.

O parâmetro amplitude alélica é bastante variável de população para população excepto para o *locus* Pde6, que vai de 117 a 129 pb em todas as populações. Para os restantes *loci*, o Pde3 apresenta a menor diferença de pares de bases (8) enquanto o PrC5 apresenta o maior valor (80).

Os valores mais baixos de heterozigotia, tanto esperada como observada, foram obtidos para o *locus* Pde3 (0,168 a 0,549). Os *loci* Pde1, Pde5 e PrC5 apresentaram valores altos de heterozigotias, entre 0,562 e 0,888. O *locus* Pde6 apresentou valores intermédios (0,575 a 0,705).

Para o *locus* Pde1 todas as populações apresentaram valores significativos para o índice  $F_{IS}$ . Igualmente, à excepção das populações das Flores, Faial e Pico para o Pde5 e Flores e Santa Maria para o PrC5, todas as populações apresentam desvios significativos ao equilíbrio de Hardy-Weinberg para estes dois *loci*. Um padrão diferente foi detectado nos *loci* Pde6 e Pde3, para os quais apenas um e dois locais de amostragem respectivamente apresentam desvios significativos ao equilíbrio de Hardy-Weinberg.

### 3.4. Análise da estruturação genética

Na tabela 3.4 estão representados os valores do índice  $F_{ST}$  e da estimativa do número de migrantes por geração. As comparações envolvendo a amostra das Canárias apresentam os valores de  $F_{ST}$  mais elevados ( $0,05 < F_{ST} < 0,08$ ). Entre as populações dos Açores, os valores deste índice são inferiores a 0,05 para todos os pares de populações. Os valores de  $F_{ST}$  são significativos para a maioria dos pares (79%), sendo significativos para todos os pares Canárias-Açores.

O número de migrantes estimados entre as Canárias e as ilhas Açoreanas varia entre 2,8 e 4,5. Dentro dos Açores, as estimativas de migração mais elevadas foram obtidas entre Terceira-Santa Maria (50,1), São Miguel-Santa Maria (25,2) e Terceira-São Miguel (20,5).

Tabela 3.4 - Valores da estimativa do coeficiente de diferenciação populacional entre pares de populações,  $F_{ST}$ , na parte superior da matriz e da estimativa do fluxo génico,  $Nm$ , na parte inferior; ■ significativo ( $P < 0,01$ ), □ não significativo.

População	Flores	Faial	Pico	Graciosa	Terceira	São Miguel	Santa Maria	Canárias
Flores	---	0,024	0,023	0,015	0,024	0,023	0,030	0,078
Faial	10,141	---	0,025	0,020	0,016	0,033	0,028	0,057
Pico	10,797	9,714	---	0,014	0,032	0,023	0,023	0,081
Graciosa	16,642	12,200	17,932	---	0,014	0,025	0,018	0,063
Terceira	10,263	15,603	7,659	17,607	---	0,012	0,005	0,052
São Miguel	10,667	7,224	10,787	9,798	20,497	---	0,010	0,064
Santa Maria	7,974	8,621	10,573	13,293	50,052	25,234	---	0,076
Canárias	2,939	4,102	2,846	3,722	4,532	3,643	3,038	---

Os valores de probabilidade do Teste Exacto (tabela 3.5) são significativos ( $p < 0,01$ ) para todos os pares de locais de amostragem estudados excepto dois: Faial-Terceira e Terceira-Santa Maria.

Tabela 3.3 - Parâmetros de diversidade genética para 5 loci microsatélites analisados.

População	Flores		Falal		Pico		Graciosa		Terceira		S. Miguel		Sta. Maria		Canárias		
	50	40	40	30	50	30	50	50	50	40	40	50	50	40	40	40	
Pde1																	
A (ni)	14 (45)	14* (35)	8 (26)	13* (45)	13* (44)	8 (32)	10* (48)	12* (24)									
Rs	10,708	12,383	7,403	10,646	10,987	7,490	8,772	11,996									
A privados	104,119	94,95	120,121	109	97	-	-	-									
A modal	110	110	110	110	110	110	110	110									
A amplitude	101-119	94-118	108-121	103-118	97-118	100-118	101-118	100-114									
H <sub>e</sub>	0,781	0,831	0,592	0,811	0,821	0,641	0,758	0,888									
H <sub>o</sub>	0,795	0,730	0,667	0,778	0,711	0,562	0,653	0,750									
F <sub>is</sub> (W&C)	-0,007**	0,135**	-0,108**	0,053**	0,145**	0,138**	0,148**	0,176**									
Pde3																	
A (ni)	5* (49)	8* (37)	5 (27)	6 (44)	6* (45)	5* (36)	6* (43)	8* (34)									
Rs	4,355	6,820	4,435	4,668	5,434	4,468	5,157	6,817									
A privados	-	124	-	-	-	-	-	120,122									
A modal	130	130	130	130	130	130	130	130									
A amplitude	126-138	124-134	126-134	126-138	126-134	126-134	128-138	120-138									
H <sub>e</sub>	0,352	0,515	0,168	0,230	0,486	0,407	0,426	0,549									
H <sub>o</sub>	0,347	0,461	0,179	0,204	0,456	0,389	0,400	0,514									
F <sub>is</sub> (W&C)	0,024	0,116	-0,042	0,123	0,071	0,059**	0,072	0,078**									
Pde6																	
A (ni)	4 (43)	5 (29)	4 (28)	5* (44)	4 (43)	4 (37)	5 (46)	6 (25)									
Rs	3,860	4,992	3,999	4,895	3,951	3,994	4,724	5,920									
A privados	-	-	-	-	-	-	-	121									
A modal	129	129	126	126	129	126	126	126									
A amplitude	117-129	117-129	117-129	117-129	117-129	117-129	117-129	117-129									
H <sub>e</sub>	0,575	0,705	0,606	0,701	0,619	0,629	0,644	0,704									
H <sub>o</sub>	0,651	0,742	0,655	0,704	0,682	0,649	0,609	0,680									
F <sub>is</sub> (W&C)	-0,122	0,036**	-0,064	0,007	-0,090	-0,018	0,065	0,054									

(continua)

Tabela 3.3 - Parâmetros de diversidade genética para 5 loci microsatélites analisados (continuação)

População	Flores		Falal	Pico	Graciosa	Terceira	S. Miguel	Sta. Maria	Canárias
	50	50	40	30	50	50	40	50	40
<b>Pde5</b>									
A (ni)	13* (37)		9* (27)	10* (23)	13* (42)	11* (50)	6* (39)	12* (48)	11* (37)
Rs	10,771		8,697	10,000	10,992	8,474	5,934	8,916	9,908
A privados	255		-	-	-	253	-	259	221,223,225
A modal	243		245	245	243, 245	243	243	243	229
A amplitude	211-257		229-247	227-249	229-257	229-253	229-247	211-249	221-245
H <sub>e</sub>	0,799		0,791	0,822	0,861	0,775	0,710	0,753	0,795
H <sub>o</sub>	0,784		0,741	0,739	0,809	0,735	0,710	0,787	0,789
F <sub>IS</sub> (W&C)	0,032		0,083	0,123	0,072**	0,062**	0,012**	-0,035**	0,020**
<b>PrC5</b>									
A (ni)	10* (49)		12* (34)	11* (25)	12* (45)	12* (43)	12* (36)	13* (39)	11* (36)
Rs	6,738		10,999	10,722	9,966	10,381	10,720	11,657	10,476
A privados	-		116	120,142	-	176	192	-	-
A modal	168		168	168	168	168	168	168	168
A amplitude	148-180		116-196	120-184	140-172	140-184	144-192	140-196	136-172
H <sub>e</sub>	0,638		0,846	0,788	0,783	0,787	0,770	0,828	0,856
H <sub>o</sub>	0,646		0,742	0,731	0,727	0,714	0,750	0,744	0,784
F <sub>IS</sub> (W&C)	-0,001		0,136**	0,093**	0,082**	0,104**	0,040**	0,115	0,098**

Nota: Para cada locus e população está representado o tamanho da amostra (N), o número de indivíduos genotipados (ni), o número de alelos (A, estando indicado \* quando apresenta alelo nulo), a riqueza alélica (Rs) baseada num número mínimo de 23 indivíduos diplóides, o número de alelos privados (A privados), o tamanho modal dos alelos (A modal), a amplitude de tamanhos em cada população (A amplitude), a heterozigotia esperada (H<sub>e</sub>) e a observada (H<sub>o</sub>) e o índice de fixação F<sub>IS</sub> e os desvios ao esperado por Hardy-Weinberg (\*\*=P < 0.01).

Tabela 3.5: Valores do teste exacto de Fisher; ■ significativo ( $P < 0.01$ ), ■ não significativo.

Local de amostragem	Faial	Pico	Graciosa	Terceira	São Miguel	Santa Maria	Canárias
Flores							
Faial	-			n.s.			
Pico	-	-					
Graciosa	-	-	-				
Terceira	-	-	-	-		n.s.	
São Miguel	-	-	-	-	-		
Santa Maria	-	-	-	-	-	-	

Os resultados da AMOVA indicam que, nas duas análises, a maior percentagem de variação é devida à variação entre indivíduos: 75,95% para a análise com 2 grupos, 77,47% para a análise com 3 grupos e 89,59% para a análise sem grupos (tabelas 3.6, 3.7 e 3.8); a percentagem de variação entre grupos é muito maior na análise que inclui a população das Canárias, observando-se aproximadamente 5% de variação entre arquipélagos; na análise feita apenas com as populações dos Açores agrupadas geograficamente, a variação entre grupos não é significativa.

Tabela 3.6 - Resultados da AMOVA considerando 2 grupos: 1- Açores; 2- Canárias; \* -  $P < 0.01$ .

Fonte de Variação	d.f.	Soma dos Quadrados	Componentes de Variação	Percentagem de Variação
Entre grupos	1	17,883	0,097 *	4,990
Entre populações dentro de grupos	6	30,366	0,032 *	1,670
Entre indivíduos dentro de populações	338	475,654	0,389 *	19,990
Entre indivíduos	346	494,000	1,428 *	73,34
Total	691	1287,903	1,836	

Tabela 3.7 - Resultados da AMOVA considerando 3 grupos: 1- Flores; 2-Faial, Pico, Graciosa e Terceira; 3- São Miguel e Santa Maria; \* -  $P < 0.01$ ; <sup>ns</sup> - não significativo.

Fonte de Variação	d.f.	Soma dos Quadrados	Componentes de Variação	Percentagem de Variação
Entre grupos	2	12,489	0,008 <sup>ns</sup> .	0,460
Entre populações dentro de grupos	4	17,877	0,027 *	1,470
Entre indivíduos dentro de populações	301	655,614	0,378 *	20,590
Entre indivíduos	308	438,000	1,422 *	77,470
Total	615	1123,981	1,836	

Tabela 3.8 - Resultados da AMOVA não considerando nenhum grupo e utilizando apenas indivíduos dos Açores; \* -  $P < 0.01$ .

Fonte de Variação	d.f.	Soma dos Quadrados	Componentes de Variação	Percentagem de Variação
Entre populações	6	37,806	0,04910*	2,61
Entre indivíduos dentro de populações	302	598,438	0,14694*	7,80
Entre indivíduos	309	521,500	1,68770*	89,59
Total	617	1157,744	1,88374	

O método de agrupamento Bayesiano deu resultados semelhantes às análises anteriores, apontando para a existência de dois grupos diferenciados, um correspondendo à população das Canárias e o outro ao conjunto das populações dos Açores (fig.3.7 A). A análise espacial do programa BAPS deu um resultado concordante (fig.3.7 B).

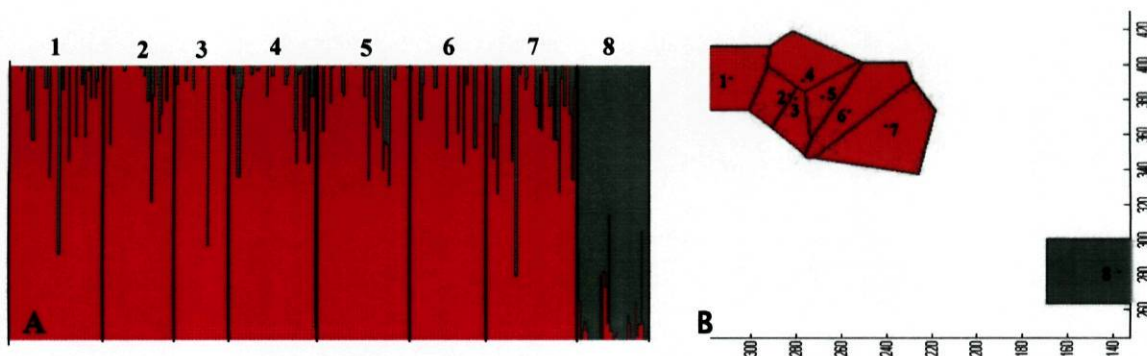


Fig. 3.7 – Resultados da análise com o programa BAPS; **A**: análise não espacial; **B**: análise espacial.; 1: Flores; 2: Faial; 3: Pico; 4: Graciosa; 5: Terceira; 6: São Miguel; 7: Santa Maria; 8: Canárias.

### 3.5. Distribuição geográfica dos polimorfismos de ADNmt

Nas 172 sequências correspondentes a uma porção da subunidade I do gene mitocondrial Citocromo c Oxidase analisadas detectou-se um total de 23 haplótipos e 56 posições variáveis. A rede de haplótipos obtida com o programa NETWORK encontra-se apresentada na figura 3.9 e a representação dos haplótipos por ilha para os Açores na fig. 3.10.

Os haplótipos 1 a 12 são exclusivos de sequências provenientes dos Açores e os restantes são exclusivos de amostras das Canárias. A sub-rede dos Açores apresenta uma configuração em estrela, com uma haplótipo central muito frequente e vários haplótipos mais raros que diferem do alelo central num máximo de quatro mutações.

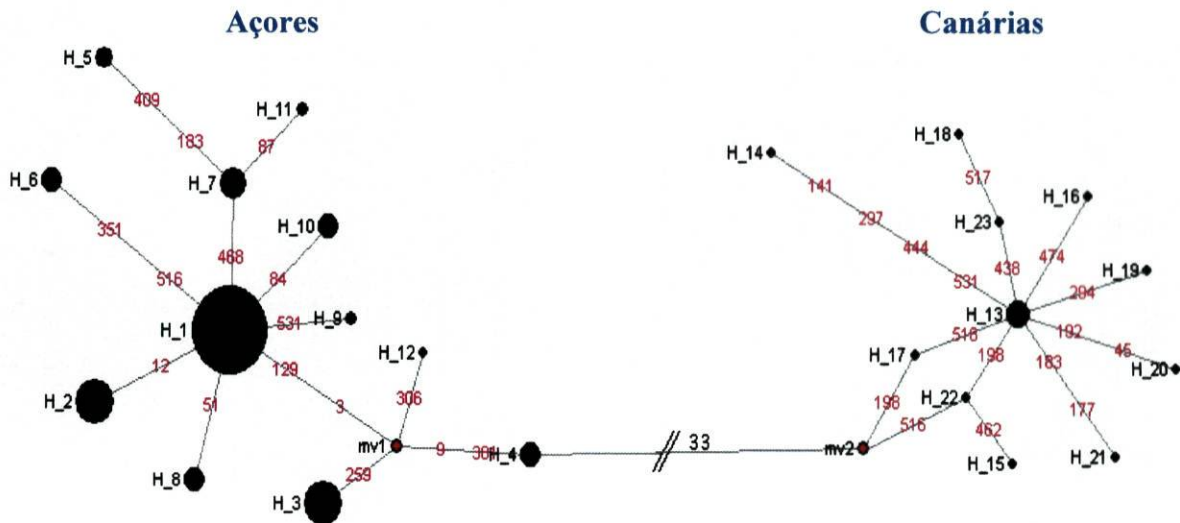


Fig. 3.8: Rede de máxima parsimónia dos haplótipos obtidos a partir da análise de 172 sequências de uma porção da sub-unidade I do gene mitocondrial Citocromo c Oxidase, de populações dos Açores e das Canárias.

Dentro do arquipélago dos Açores, o haplótipo 1 é o mais frequente, sendo comum a 76 sequências e presente em todas as ilhas; existem haplótipos exclusivos de algumas ilhas, como o H5 exclusivo das Flores (comum a 5 indivíduos), H6 da Terceira (6 indivíduos), H11 do Faial (2 indivíduos) e H12 exclusivo da Graciosa (apenas 1 indivíduo). Pode-se ainda observar que as ilhas do grupo oriental, São Miguel e Santa Maria, partilham os mesmos haplótipos, e as ilhas Terceira e Graciosa, geograficamente próximas, partilham apenas o haplótipo mais comum e são as ilhas que apresentam menor variedade haplotípica. Nas Flores, o haplótipo mais comum é o H3, que só se volta a encontrar na ilha do Faial, a mais próxima geograficamente.

Dentro do arquipélago dos Açores, o haplótipo 1 é o mais frequente, sendo comum a 76 sequências e presente em todas as ilhas; existem haplótipos exclusivos de algumas ilhas, como o H5 exclusivo das Flores (comum a 5 indivíduos), H6 da Terceira (6 indivíduos), H11 do Faial (2 indivíduos) e H12 exclusivo da Graciosa (apenas 1 indivíduo). Pode-se ainda observar que as ilhas do grupo oriental, São Miguel e Santa Maria, partilham os mesmos haplótipos, e as ilhas Terceira e Graciosa, geograficamente próximas, partilham apenas o haplótipo mais comum e são as ilhas que apresentam menor variedade haplotípica. Nas Flores, o haplótipo mais comum é o H3, que só se volta a encontrar na ilha do Faial, a mais próxima geograficamente.

#### **4. Discussão**

---

#### 4.1. A utilização de marcadores microssatélites em *P. candei*

Pode-se verificar que os microssatélites isolados em *P. depressa* funcionaram melhor, isto é, apenas não amplificou 1 em 7, e, dos que amplificaram, 4 revelaram-se polimórficos e puderam ser utilizados neste trabalho; pelo contrário, os microssatélites isolados em *P. rustica* tiveram uma taxa de amplificação baixa (60%) e apenas 1 amplificou para a quase totalidade dos indivíduos, revelando-se polimórfico e podendo ser utilizado. Este resultado é concordante com o trabalho de Pérez e colaboradores (2007), que analisaram 6 loci para 10 a 23 indivíduos, obtendo valores de amplificação dos microssatélites de *P. depressa* em *P. candei* de 78% sendo em *P. rustica* apenas 44%, o que era esperado do ponto de vista evolutivo, uma vez que se encontram filogeneticamente mais próximas, (Sá-Pinto *et al.*, 2005), esperando-se por isso maior similaridade nas regiões flangeantes. Apesar de a espécie *P. caerulea* estar mais próxima filogeneticamente de *P. candei* do que *P. depressa* (Sá-Pinto *et al.*, 2005), não houve melhor amplificação dos microssatélites isolados naquela espécie mas pode dever-se a terem sido experimentados apenas 3 loci.

Existe uma variação do tamanho dos alelos entre a espécie onde o microssatélite foi isolado e a espécie em estudo, o que se poderá dever tanto ao maior ou menor número de repetições em tandem como à existência de variação nas regiões flangeantes do microssatélite (Colson & Goldstein, 1999).

Excepto o Pde3, os loci apresentaram valores de heterozigotia esperada entre 0,575 e 0,888, valores semelhantes aos obtidos para outros moluscos (Kenchington *et al.*, 2006; Lado-Insua, 2006).

Para quase todos os microssatélites analisados observou-se a existência de um défice de heterozigóticos na maioria das populações analisadas. Embora a mistura de mais do que uma local de amostragem ou espécie num dado local de amostragem (efeito de Wahlund), a ocorrência de cruzamentos consanguíneos (*inbreeding*) ou a acção da selecção natural possam resultar em défice de heterozigóticos, as características dos microssatélites e da sua análise tornam mais provável a hipótese de este défice se dever à ocorrência de alelos nulos e/ou erros de leitura provocados pelo aparecimento de bandas *stutter* e efeito *drop-out*. É improvável que o défice de heterozigóticos observado se deva a um erro de amostragem dado que se analisou um número apropriado de indivíduos por local de amostragem não se verificando uma relação directa entre o efectivo amostral e o número médio de alelos, isto é, os locais com maior número de

indivíduos amostrados não são os que apresentam o valor mais elevado do número médio de alelos. Na leitura dos géis de acrilamida prestou-se especial atenção à presença de alelos *drop-out* e a bandas sobrepostas, estabelecendo-se regras de genotipagem comuns para todos os loci e populações. A baixa heterozigotia parece dever-se então à presença de alelos nulos.

A presença de alelos nulos em marcadores microssatélites é frequente, especialmente em grandes populações naturais, sendo um fenómeno comum em populações de moluscos marinhos (Bruford & Wayne, 1993; Li *et al.*, 2003; Karlsson & Mork, 2005). Para determinar a presença de alelos nulos nos loci analisados deveria fazer-se um estudo sobre a linha germinal. Sendo detectada a existência de alelos nulos seria necessário desenhar novos *primers* e, se o problema persistisse, seria aconselhável procurar outros marcadores, dado que a presença de alelos nulos subestima a heterozigotia e a riqueza alélica, produzindo erros nos cálculos de desequilíbrio de ligamento (Diz, 2005). Com a aplicação do programa MICRO-CHECKER foi detectado um elevado número de alelos nulos nos microssatélites analisados; foi escolhido o estimador Oosterhout para corrigir os genótipos por ser o que melhor se adequou, permitindo uma melhor aproximação ao equilíbrio de Hardy-Weinberg. As correcções genotípicas propostas por este estimador foram usadas para calcular os parâmetros de diversidade genética, para a análise estatística e para estimar distâncias genéticas. Contudo, estas correcções também têm implicações, limitando a análise de sub-estruturação genética, uma vez que, para cada *loci*, é impossível saber se existe um ou mais alelos nulos e se estes diferem ou não entre locais de amostragem.

Assim, os *loci* utilizados devem ser vistos como um ponto de partida para o estudo das relações populacionais neste arquipélago mas será indicado utilizar *loci* específicos para tentar evitar estes problemas; o *locus* Pde6 poderá ser utilizado uma vez que apresenta alelos nulos apenas numa população e com uma frequência baixa. A correcção dos genótipos com o MICRO-CHECKER não eliminou totalmente os desvios ao equilíbrio, contudo permitiu reduzir os valores significativos do índice de fixação para 55%.

#### **4.2. Estimativas da diversidade genética e análise da estruturação**

Para organismos intertidais, as ilhas são modelos particularmente adequados que permitem testar predições relacionadas com a eficácia de dispersão das larvas pelágicas,

uma vez que o fluxo génico tem de ser assegurado pela dispersão larvar (Sá-Pinto *et al.*, *in press*).

Apesar desta localização isolada, foram encontradas poucas espécies marinhas endémicas, o que pode ser explicado pelo facto de se supor que a fauna existente é muito recente, tendo resultado de uma repovoação das ilhas durante a pós-glaciação (10000 a 12000 anos) (Titselaar, 1998). Além deste facto, tem vindo a ser reportado um aumento das espécies em extinção pelo que é indispensável aumentar o conhecimento sobre as espécies aí existentes.

O conjunto de resultados obtidos permitiu individualizar facilmente o arquipélago dos Açores do arquipélago das Canárias, não permitindo individualizar populações bem definidas nos Açores, a partir dos locais de amostragem. Contudo, os valores significativos de  $F_{ST}$  ( $P < 0,01$ ) e os resultados do teste exacto observados indicam que os diferentes locais de amostragem provavelmente não fazem parte de uma população panmíctica, parecendo existir alguma restrição ao fluxo génico em algumas ilhas, nomeadamente entre a ilha das Flores, no grupo Ocidental, e as restantes ilhas (fig. 4.1).

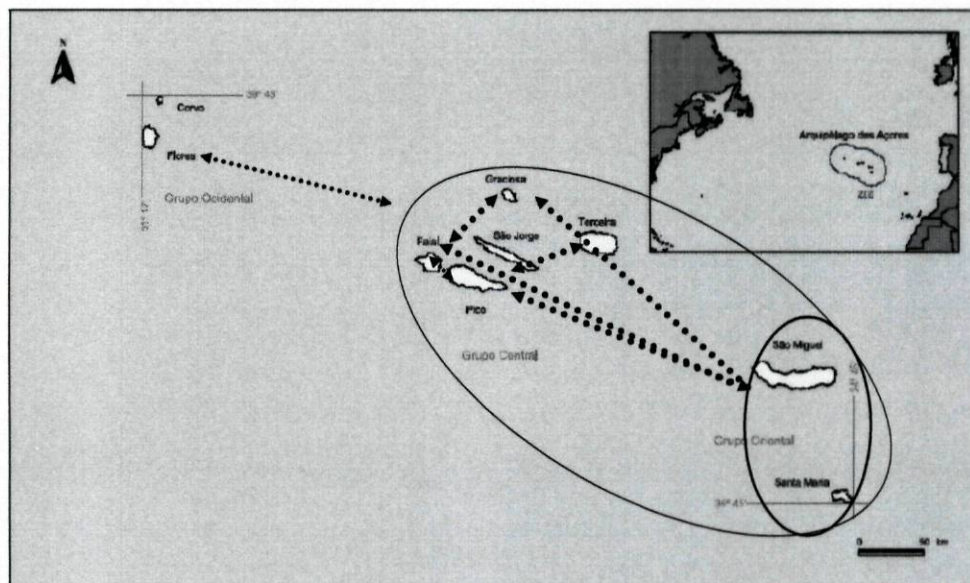


Fig. 4.1 – Representação da diferenciação populacional entre as ilhas dos Açores;  
...► diferenciação significativa; ausência de símbolo – diferenciação não significativa.

Foram obtidos baixos valores absolutos de  $F_{ST}$ , inferiores a 0,05 unidades, entre todos os pares de populações açorianas; estes baixos níveis de estruturação podem dever-se à existência de algum fluxo génico entre locais, contudo menor do que o esperado numa população panmíctica pois os valores de  $F_{ST}$  são significativos, tal como o teste exacto, ou a uma recente origem das diversas populações que ainda não tiveram

tempo de se diferenciar e que consiste também numa das possíveis hipóteses para explicar a rede em estrela do ADNmt. Os maiores valores de  $F_{ST}$  foram obtidos entre Açores e Canárias, entre 0,05 e 0,08 unidades, tal como esperado de acordo com os dados de ADNmt e de outros estudos já referidos que evidenciam a existência de subespécies diferentes em cada arquipélago. No entanto, estes valores deveriam ser de uma ordem de grandeza maior, traduzindo a diversidade genética de forma mais próxima da realidade; esta redução pode ser devida a fenómenos como homoplasia e/ou ocorrência de alelos nulos.

Estes resultados não são totalmente concordantes com os valores de  $F_{ST}$  obtidos no trabalho de Sá-Pinto e colaboradores (*in press*) em que utilizaram ADN mitocondrial. Os autores obtiveram valores de  $F_{ST}$  altos e muito significativos em todas as comparações envolvendo a ilha das Flores ( $0,11 < F_{ST} < 0,38$ ) ou a ilha da Graciosa ( $0,13 < F_{ST} < 0,23$ ); para as restantes ilhas obtiveram valores mais baixos (inferiores a 0,09). A diferenciação envolvendo a ilha das Flores pode ser explicada, segundo os autores, pela grande distância geográfica que separa esta ilha das restantes. No entanto, não foi detectado um padrão significativo de isolamento pela distância, o que parece indicar que outros factores além da distância, como por exemplo o padrão das correntes, poderão estar envolvidos na diferenciação populacional. Estes factores poderão explicar a elevada diferenciação observada entre a Graciosa e as restantes ilhas do grupo central. Os autores defendem que os microssatélites poderiam revelar melhor os padrões de estruturação uma vez que a falta de diferenciação observada pode ser devida a baixa resolução do fragmento de ADN mitocondrial analisado naquele estudo.

Os resultados da AMOVA mostram que a maior percentagem de variação é obtida entre indivíduos (73 a 90%). Pode-se, no entanto, observar que a percentagem de variação entre grupos é cerca de 11 vezes superior na análise que incluiu a população das Canárias como um grupo (4,99), por comparação à análise feita sem esta população (0,46), o que mais uma vez mostra a separação da população deste arquipélago com os locais de amostragem do arquipélago dos Açores mas não permite inferir qualquer diferenciação dentro deste.

Os resultados obtidos com o programa de análise Bayesiana BAPS fizeram igualmente distinção entre a população das Canárias e os locais de amostragem açorianos mas não mostram qualquer estruturação dentro do arquipélago dos Açores. A análise espacial permitida pelo BAPS, que incorpora coordenadas dos pontos de amostragem, não permitiu recuperar mais nenhum grupo, concordando com a análise

não-espacial. No entanto, o método bayesiano implementado no BAPS foi testado por Waples & Gaggiotti (2006) e provou ser muito conservativo a identificar estruturação populacional; no estudo feito por estes autores, diferentes grupos genéticos só foram detectados sob uma migração muito restrita ( $Nm=1$ ;  $F_{ST}>0,13$ ). Segundo os autores, as razões para isto acontecer não são muito claras mas sugerem possíveis explicações: (i) penalização muito severa para postular populações adicionais (e daqui estimar parâmetros adicionais); (ii) migração recente que pode atenuar diferenças entre populações.

A rede de haplótipos de ADNmt obtida (figs. 3.8 e 3.9) está, de modo geral, de acordo com os resultados já discutidos. Obteve-se um total de 23 haplótipos, 12 dos quais exclusivos dos indivíduos provenientes do arquipélago dos Açores e os restantes exclusivos de indivíduos das Canárias. Dentro dos Açores, foi possível observar diferenças entre a ilha das Flores e as restantes, obtendo-se uma distribuição dos haplótipos em estrela, estando o haplótipo mais frequente presente em todas as ilhas e irradiando deste os restantes haplótipos excepto para o caso das Flores, com um haplótipo mais frequente diferente. Os haplótipos exclusivos de cada ilha são apenas partilhados por alguns indivíduos (entre 1 a 6).

A diferenciação encontrada para as Flores, tanto com a análise mitocondrial como nuclear, pode ser devida à grande distância que separa esta ilha das restantes estudadas. Tendo em conta o valor de velocidade média da Corrente dos Açores de  $10 \text{ m.s}^{-1}$  [1], numa situação normal seriam precisos 23 dias para atravessar uma distância de aproximadamente 200 km e a duração da fase pelágica não é suficiente.

A menor variabilidade no ADN mitocondrial nos Açores face ao observado noutras regiões como as Canárias pode ter várias origens como efeitos de fundador (colonização a partir de um pequeno número de indivíduos), *bottlenecks* históricos ou *selective sweep* (fenómeno que leva a uma perda de biodiversidade por arrastar a sequência de ADNmt que está a ser estudada; Sá-Pinto *et al.*, *in press*). Neste estudo, pela análise do parâmetro de riqueza alélica observa-se o mesmo padrão: nos Açores a média dos valores de riqueza alélica é sempre menor que o valor obtido para a população das Canárias, sendo esta diferença acentuada para três dos *loci* (Pde1-1,79; Pde3-1,72; Pde6-1,73; Pde5-0,94; PrC5-0,39). Pode-se assim dizer que a baixa variabilidade encontrada nos microssatélites dá maior suporte às hipóteses que invocam a ocorrência de um *bottleneck* ou colonização recente dos Açores em relação à hipótese de ocorrência de um *selective sweep* para explicar a baixa variabilidade no mitocondrial.

## **5. Conclusão e sugestões para futuros estudos**

---

Após os estudos com aloenzimas e ADN mitocondrial, a incorporação de microssatélites pretendia ajudar a resolver o problema dos *loci* submetidos a selecção e que apresentam baixa variabilidade, como as aloenzimas, ou a tendência da herança materna do ADN mitocondrial. Apesar de todos os problemas detectados, como alelos nulos e desvios ao equilíbrio de Hardy-Weinberg, foi possível diferenciar perfeitamente as populações dos Açores das populações das Canárias. Dentro do arquipélago dos Açores foi possível observar alguma diferenciação entre a ilha das Flores e as restantes ilhas, o que deve ser tido em conta em eventuais planos de gestão.

Sugere-se, então, para futuros estudos:

- aumentar o número de marcadores na análise, acrescentando microssatélites entretanto isolados noutras espécies do género *Patella* e também isolar microssatélites em *P. candei*, de modo a tentar diminuir a quantidade de alelos nulos;
- alargar a análise feita com ADNmt, pois apesar de pouco variável, mostrou ser informativo;
- analisar sequências de taxa de evolução intermédia, como por exemplo sequências intrónicas, sequências não codificantes do ADN nuclear;
- aumentar o número de locais amostrados, principalmente incluir uma amostra da ilha Corvo, que estando geograficamente muito próximo das Flores permitirá comprovar ou não que a diferenciação desta ilha se deve à distância geográfica relativamente às restantes ilhas.

## **6. Referências bibliográficas**

---

[1] [www.oceancurrents.rsmas.miami.edu/atlantic/azores\\_1.html](http://www.oceancurrents.rsmas.miami.edu/atlantic/azores_1.html), acesso a 16-10-2007.

**Amersham Biosciences, 1998.** ALFexpressII User Manual.  
<http://www.amershambiosciences.com>

**Anderson, E.C. & Thompson, E.A., 2002.** A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* 160:1217-1219.

**Bachtrog, D., Weiss, S., Zangerl, B., Brem, G. & Schlotterer, C., 2000.** Distribution of dinucleotide microsatellites in the *Drosophila melanogaster* ome. *Molecular Biology and Evolution* 16:602-610.

**Balloux, F. & Lugon-Moulin, N., 2002.** The estimation of population differentiation with microsatellite markers. *Molecular Ecology* 11:155-165.

**Bandelt, H.J., Forster, P. & Röhl, A., 1999.** Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16:37-48.

**Belkhir, K., Borsa, P., Goudet, J., Chikhi, L. & Bonhomme, 1996.** Genetix, logiciel sous Windows™ pour la génétique des populations. Laboratoire Génome et Population, CNRS UPR 9060, Université de Montpellier II, Montpellier, France.

**Bonin, A., Bellemain, E., Bronken Eidesen, P., Pompanon, F., Brochmann, C. & Taberlet, P., 2004.** How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* 13:3261-3273.

**Brohede, J. & Ellegren, H., 1999.** Microsatellite evolution: polarity of substitutions within repeats and neutrality of flanking sequences. *Proceedings of the Royal Society: Biological Sciences* 266: 825-833.

**Brohede, J., Primer, C.R., Moller, A. & Ellegren, H., 2002.** Heterogeneity in the rate and pattern of germline mutation at individual microsatellite loci. *Nucleic Acids Research* 30 (9): 1997-2003.

**Brookfield, J.F.Y., 1996.** A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol. Ecol.* 5: 453-455.

- Bruford, M.W. & Wayne, R.K., 1993.** Microsatellites and their application to populations genetics studies. *Curr. Opin. Genet. Develop.* 3:939-943.
- Castro, J., Bouza, C., Presa, P., Pino-Querido, A., Rianza, A., Ferreiro, I., Sánchez, L. & Martínez, P., 2004.** Potential sources of error in parentage assessment of turbot (*Scophthalmus maximus*) using microsatellite loci. *Aquaculture* 242: 119-135.
- Casu, M., Casu, D., Lai, T., Cossu, P. & Curini-Galletti, M., 2006.** Inter-simple sequence repeat markers reveal strong genetic differentiation among populations of the endangered mollusc *Patella ferruginea* (Gastropoda: Patellidae) from two Sardinian marine protected areas. *Marine Biology*. In press.
- Cavalli-Sforza, L.L. & Edwards, A.W.F., 1967.** Phylogenetic analysis: models and estimation procedures. *American Journal of Human Genetics* 19: 233-257.
- Chakraborty, R., De Andrade, M., Daiger, S.P. & Budowle, B., 1992.** Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. *Annals of Human Genetics* 56: 45-47.
- Chakraborty, R. & Leimar, O., 1987.** Genetic variation within a subdivided population. Em: *Population Genetics and Fishery Management*. Ryman, N. & Utter, F., ed., pp. 89-120. University of Washington Press. Seattle.
- Christiaens, J., 1973.** Revision du genre *Patella* (Mollusca, Gastropoda). *Bull. Mus. Natl. Hist. Nat.*, (Sér 3) 182, 1305-1392.
- Colson, I. & Goldstein, D.B., 1999.** Evidence for Complex Mutations at Microsatellite Loci in *Drosophila*. *Genetics* 152: 617-627.
- Corander, J., Waldman, P. & Sillanpää, M.J., 2003.** Bayesian analysis of genetic differentiation between populations. *Genetics* 163: 367-374.
- Côrte-Real, H.B., Hawkins, S.J. & Thorpe, J.P., 1996.** Population differentiation and taxonomic status of the exploited limpet *Patella candei* in the Macaronesian islands (Azores, Madeira, Canaries). *Marine Biology* 125:141-152.
- Crow, J.F. & Kimura, M., 1970.** An introduction to population genetics. Harper and Row, New York, USA.

**Cruz,F., Pérez,M. & Presa,P., 2005.** Distribution and abundance of microsatellites in the genome of bivalves. *Gene* 346:241-247.

**Dawson,K.J. & Belkhir,K., 2001.** A bayesian approach to the identification of panmictic populations and the assignment of individuals. *Genetic Research* 78: 59-77.

**Dakin,E. & Avise,J., 2004.** Microsatellite null alleles in parentage analysis. *Heredity* 93 :504-509.

**Di Rienzo,A., Peterson,A.C., Garza,J.C., Valdes,A.M., Slatkin,M. & Freimer,N.B., 1994.** Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences* 91: 3166-3170.

**Diz, A.E., 2005.** Aplicación de microsatélites en estimativas de flujo génico y estructura poblacional de *Mytilus galloprovincialis*. Tesis Doctoral. Departamento de Bioquímica, Genética e Inmunología, Universidade de Vigo.

**Dulvy,N.K., Sadovy,Y. & Reynolds,J.D., 2003.** Extinction vulnerability in marine populations. *Fish and Fisheries* 4:25-64.

**Ellegren, H., 2000.** Microsatellite mutations in the germline: implications for the evolutionary inference. *Trends in Genetics*, 16, 551-558.

**Ellegren, H., 2004.** Microsatellites: simple sequences with complex evolution. *Genetics* 5: 435-445.

**Estoup,A., Jarne,P. & Cornuet,J.M., 2002.** Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* 11:1591-1604.

**Estoup,A., Taillez,C., Cornuet,J.M. & Solignac,M., 1995.** Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Molecular Biology and Evolution*, 12, 1074-1084.

**Excoffier,L., Laval, L.G. & Schneider,S., 2005.** Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1: 47-50.

**Excoffier, L., Simouse, P.E. & Quattro, J.M., 1992.** Analysis of molecular variance inferred from metric distances among ADN haplotypes: application to human mitochondrial ADN restriction data. *Genetics* 131: 479-491.

**Felsenstein, J., 2005.** PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

**Goldstein, D.B. & Clark, A.G., 1995.** Microsatellite variation in North American populations of *Drosophila melanogaster*. *Nucleic Acids Research* 23:3882-3886.

**Goudet J., 1995.** Fstat (version 2.9.3.): a computer program to calculate F-statistics. *Journal of heredity*, 86:485-486.

**Harr, B., Zangerl, B., Brem, G., & Schlotterer, C., 1998.** Conservation of locus specific microsatellite variability across species: a comparison of two *Drosophila* sibling species *D. melanogaster* and *D. simulans*. *Mol. Biol. Evol.* 15:176-184.

**Hawkins, S.J., Côrte-Real, H.B., Martins, H.R., Santos, R.S. & Frias-Martins, A.M., 1990.** A note on the identity of *Patella* in the Azores. *Açoreana (Suplemento)* 167-173.

**Hawkins, S.J., Côrte-Real, H.B., Pannacciulli, F.G., Weber, L.C. & Bishop, J.D., 2000.** Thoughts on the ecology and evolution of the intertidal biota of the Azores and other Atlantic islands. *Hydrobiologia* 440:3-17.

**Hill, A.S. & Hawkins, S.J., 1991.** Seasonal and spatial variation of aphylithic microalgal distribution and abundance and its ingestion by *Patella vulgata* on a moderately exposed rocky shore. *J. Mar. Biol. Ass. UK* 71:403-423.

**Huang, Q.Y., Xu, F.H., Shen, H., Deng, H.Y., Liu, Y.J., Liu, Y.Z., Li, J.L., Recker, R.R. & Deng, H.W., 2002.** Mutation patterns at dinucleotide microsatellites loci in humans. *American Journal of Human Genetics* 70: 625-634.

**Jarne, P., & Lagoda, P., 1996.** Microsatellites: from molecules to populations and back. *Trends in Ecology and Evolution* 11: 424-429.

**Jehel, R. & Arntzem, J.W., 2002.** Review: microsatellite markers in amphibian conservation genetics. *Herp. Journal* 12:1-9.

- Jin,L., Macaubas,C., Hallmayer,J., Kimura,A. & Mignot,E., 1996. Mutation rate varies among alleles at a microsatellite locus: phylogenetic evidence. *Proceedings of the National Academy of Sciences U.S.A.*, 93:15285-15288.
- Karlsson,S. & Mork,J., 2005. Deviation from Hrdy-Weinberg equilibrium, and temporal instability in allele frequencies at microsatellite loci in a local population of Atlantic cod. *ICES Journal of Marine Science* 62(8): 1588-1896.
- Kashi,Y. & Soller,M., 1999. Function roles of microsatellites and minisatellites. In *Microsatellite: evolution and applications* (Goldstein,D.B. & Schlottere,C. eds.) Oxford Univ. Press, Oxford, U.K..
- Kenchington,E.L., Patwary,M.U., Zouros & Bird,J.C., 2006. Genetic differentiation in relation to marine landscape in broadcast-spawning bivalve mollusc (*Placopecten magellanicus*). *Molecular Ecology* 15:1781-1796.
- Kimura,M. & Crow,J.F., 1964. The number of alleles that can be maintained in a finite population. *Genetics* 49: 725-738.
- Kimura,M. & Ohta,T., 1978. Stepwise mutation model and distribution of allele frequencies in a finite population. *Proceedings of the National Academy of Sciences* 75: 2868-2872.
- Koufopanou,V., Reid,D.G., Ridgway,E.A. & Thomas,R.H., 1999. A molecular phylogeny of the Patellid limpets (Gastropoda: Patellidae) and its implications for the origins of their antitropical distribution. *Molecular Phylogenetics and Evolution* 11 (1): 138-156.
- Kruglyak,S., Durrett,R.T., Schug,M.D. & Aquadro,C.F., 1998. Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. *Proceedings of the National Academy of Sciences U.S.A.* 95:10774-10778.
- Lado-Insua, T., 2006. Influencia del vertido del burque *Prestige* sobre la diversidad genética de *Mytilus galloprovincialis*. Tesis de Licenciatura. Universidad de Vigo – Dpto. Bioquímica, Genética e Inmunología.
- Lavie,B., Noy,R. & Nevo,E., 1987. Genetic variability in the marine gastropods *Patella coerulea* and *Patella aspera*: patterns and problems. *Marine Biology* 96:367-370.

- Li,Q., Park,C., Kobayashi,T. & Kijima,A., 2003. Inheritance of microsatellite DNA markers in the Pacific abalone *Haliotis discos hannai*. *Marine Biotechnology* 3:331-338.
- Lindberg,D.R., Estes,J.A. & Warheit,K.I., 1998. Human influences on trophic cascades along rocky shores. *Ecological Applications* 8 (3): 880-890.
- Lindner, G., 2000. Moluscos y caracolas de los mares del mundo. Manual de identificación. Ed. Omega.
- Little,C. & Kitching,J.A., 1996. The biology of rocky shores. Oxford University Press.
- Little,C., Partridge,J.C. & Teagle,L., 1991. Foraging activity of limpets in normal and abnormal tidal regimes. *J. Mar. Ass. UK* 71:537-554.
- Liu,Z.J. & Cordes,J.F., 2004. DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* 238: 1-37.
- Llavona, A.V., 2006. Caracterización interespecífica de marcadores microsatélites en el género *Patella* spp. Tese de Licenciatura. Departamento de Bioquímica, Genética e Imunologia, Universidade de Vigo, 36810 Espanha.
- Luikart,G. & England,P.R., 1999. Statistical analysis of microsatellite DNA data. *Tree* 14 (7): 253-256.
- Mason,E., Coombs,S. & Oliveira,P.B., 2005. An overview of the literature concerning the oceanography of the eastern North Atlantic region. *Relat. Cient. Téc. IPIMAR, Série digital* (<http://ipimar-iniap.ipimar.pt>) nº 33, 58 pp.
- Matsumoto,T., Yukawa,W., Nozari,Y., Nakashige,R., Shinya,M., Makino,S., Yagura,M., Ikuta,T., Imanishi,T., Inoko,H., Tamiya,G. & Gojobori,T., 2004. Novel algorithm for automated genotyping of microsatellites. *Nucleic Acids Research* 32 (20): 6069-6077.
- Mauro,A., Arculeo,M. & Parrinello,N., 2003. Morphological and molecular tools in identifying the Mediterranean limpets *Patella coerulea*, *Patella aspera* and *Patella rustica*. *J. Exp. Mar. Biol. Ecol.* 295:131-143.
- Menotti-Raymond, M.A. & O'Brien,S., 1995. Evolutionary conservation of ten microsatellites loci in four species of FELIDAE. *Journal of Heredity* 86: 319-322.

- Michalakis, Y. & Excoffier, L., 1996.** A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* 142:1061-1064.
- Morais, S., Boaventura, D., Narciso, L., Ré, P. & Hawkins, S.J., 2003.** Gonad development and fatty acid composition of *Patella depressa* Pennant (Gastropoda: Prosobranchia) populations with different patterns of spatial distribution, in exposed and sheltered sites. *J. Exp. Mar. Biol. Ecol.* 294:61-80.
- Navarro, P.G., Ramírez, R., Tuya, F., Fernandez-Gil, G., Snchez-Jerez, P. & Haroun, R.J., 2005.** Hierarchical analysis of spatial distribution patterns of patellid limpets in the Canary islands. *J. Moll. Stud.* 71:67-73.
- Pagarete, A. J., Costa, M.J. & Brito, C., 2005.** Genetic conspecificity in two habitat morphs of the limpet *Patella candei gomesii* Drouet (Mollusca: Gastropoda: Patellidae) reinforces limpet morphological plasticity. *The journal of the Young Investigators* Vol 12 (1) on line.
- Page, R.D.M., 1996.** TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* 12: 357-358.
- Pérez, M., 2003.** Calibrado y aplicación de marcadores moleculares para el diagnóstico genético de stocks, productos y subproductos de merluza, *Merluccius spp.* Tesis doctoral. Universidad de Vigo.
- Pérez, M., Branco, M., Llavona, A., Ribeiro, P.A., Santos, A.M., Hawkins, S.J., Dávila, J.A., Presa, P. & Alexandrino, P., 2007.** Development of microsatellite loci for the black-footed limpet, *Patella depressa*, and cross-amplification in two other *Patella* species. *Conserv. Genet.* 8:739-742.
- Pérez, M.R., Cruz, F. & Presa, P., 2005.** Distribution properties of polymononucleotide repeats in molluscan genomes. *Journal of Heredity* 96:40-51.
- Petit, R., El Mousadik, A. & Pons, O., 1998.** Identifying Populations for Conservation on the Basis of Genetic Markers. *Conservation Biology*, 12 (4): 844-855.
- Pino Querido-Ferreira, A., 2004.** Estudio dos parámetros xenéticos esenciais para a aplicación de marcadores microsatélite na asignación de parentescos en rodaballo

(*Scophtalmus maximus* L.). Tese de Licenciatura. Universidade de Santiago de Compostela, Espanha.

**Pompanon,F., Bonin,A., Bellemain,E. & Taberlet,P., 2005.** Genotyping errors: causes, consequences and solutions. *Nature Reviews – Genetics*, 6, 847-859.

**Pritchard,J., Stephens,M. & Donnelly,P., 2000.** Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

**Raymond,M. & Rousset,F., 1995.** GENEPOP (Version 1.2). Population genetics software for exact tests and ecumenicism. *Journal of Heredity* 43:223-225.

**Raymond,M. & Rousset,F., 1997.** Statistical analysis of population genetic data: new tools, old concepts. *Trends in Ecology and Evolution* 12 (8):313-317.

**Ridgway,S.A., Reid,D.G., Taylor,J.D., Branch,G.M. & Hodgson,A.N., 1998.** A cladistic phylogeny of the family Patellidae (Mollusca: Gastropoda). *Phil. Trans. R. Soc. Lond. B* 353: 1645-1671.

**Sá-Pinto,A., Branco,M., Harris,J.H. & Alexandrino,P., 2005.** Phylogeny and phylogeography of the genus *Patella* based on mitochondrial ADN sequence data. *Journal of Experimental Marine Biology and Ecology* 325: 95-110.

**Sá-Pinto,A., Branco,M., Sayanda,D. & Aldrino,P., in press.** Patterns of colonisation, evolution and gene flow in species of the genus *Patella* in the Macaronesian islands.

**Saitou,N. & Nei,M., 1987.** The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406-425.

**Sambrook,J., Fritsch,E.F. & Maniatis,T., 1989.** Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York.

**Santibanez-Koref,M.F., Gang,R. & Hancock,J.M., 2001.** A relationship between lengths of microsatellites and nearby substitution rates in mammalian genomes. *Molecular Biology and Evolution* 18:2119-2123.

**Santos,J.P & Pinho,J.L., 2005.** Estudo das correntes oceânicas na região envolvente da Ilha Terceira no Arquipélago dos Açores. *Engenharia Civil – Universidade do Minho* N.º 23: 31-41.

- Schlotterer, S., 2000.** Evolutionary dynamics of microsatellite ADN. *Chromossoma*, 109, 365-371.
- Schlotterer, S., 2004.** The evolution of molecular markers-just a matter of fashion? *Genetics* 5: 63-69.
- Schlotterer, C. & Tautz,D., 1992.** Slippage synthesis of simple sequence ADN. *Nucleic Acids. Res.* 20:211-215.
- Schoemaker,J.S., Painter,I.S. & Weir,B.S., 1999.** Bayesian statistics in genetics-a guide for the uninitiated. *Trends Genet.* 15: 354-358.
- Schug,M.D., Wetterstrand,K.A., Gaudette,M.S., Lim,R.H., Hutter,C.M. & Aquadro,C.F., 1998.** The distribution and frequency of microsatellite loci in *Drosophila melanogaster*. *Molecular Ecology* 7:57-70.
- Smith,F.G., 1935.** The development of *Patella vulgata*. Huxley Research Laboratory, Imperial College of Science and Technology. Cap. III, vol. CCXXV.
- Stallings,R.L., Ford,A.F., Nelson,D., Torney,D.C., Hildebrand,C.E. & Moysis,R.K., 1991.** Evolution and distribution of (GT)<sub>n</sub> repetitive sequences in mammalian genomes. *Genomics* 10:807-815.
- Taberlet,L. Waits,L.P. & Luikart,G., 1999.** Non-invasive genetic sampling: look before you leap. *Trends Ecol. Evol.* 14:323-327.
- Tautz,D. & Renz,M., 1984.** Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research* 12:4127-4138.
- Titselaar,F.F., 1998.** A revision of the recent European Patellidae (Mollusca: Gastropoda) Part. 1 The Patellidae of the Azores, Madeira, the Selvagens and the Canary Islands. *Vita Marina* 45 (3-4) 21-62.
- Tóth,G., Gáspari, Z. & Jerzy, J., 2000.** Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research* 10: 967-981.
- Van Oosterhout,C.V., Hutchinson,W.F., Wills,D.P.M. & Shipley,P., 2004.** Micro-checker: software for identifying and correcting genotyping errors in microsatellite data, *Molecular Ecology Notes* 4:535-538.

- Waser, P.M. & Strobeck, C., 1998.** Genetic signatures of inter population dispersal. *Trends Ecol. Evol.* 13:43-44.
- Weber, L.I. & Hawkins, S.J., 2002.** Evolution of the limpet *Patella candei* d'Orbigny (Mollusca: Patellidae) in Atlantic archipelagos: human intervention and natural processes. *Biol. J. Linn. Soc.* 77, 341-353.
- Weber, L.I. & Hawkins, S.J., 2005.** *Patella aspera* and *P. ulyssiponensis*: genetic evidence of speciation in the North-east Atlantic. *Marine Biology* 147:153-162.
- Weber, J.L., Kitek, A.E., May, P.E., Polymeropoulos, M.H. & Ledbetter, S., 1990.** Dinucleotide repeat polymorphism at the DXS453, DXS454 and DXS458 loci. *Nucleic Acids Research* 18:4037.
- Weber, L.I., Thorpe, J.P., Santos, R.S. & Hawkins, S.J., 1998.** Identification of stocks of the exploited limpets *Patella aspera* and *Patella candei* at Madeira Archipelago by allozyme electrophoresis. *Journal of Shellfish Research* Vol. 1 n. 4, 945-953.
- Weir, B.S. & Cockerham, C.C., 1984.** Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.
- Wierdl, M., Dominska, M. & Petes, T.D., 1997.** Microsatellite instability in yeast: dependence on the length of the microsatellite. *Genetics* 146:769-779.
- Weisstein, E.W., O-line.** Fisher's Exact Test. From *MathWorld*--A Wolfram Web Resource. <http://mathworld.wolfram.com/FishersExactTest.html>
- Wilder, J. & Hollocher, H., 2001.** Mobile elements and the genesis of microsatellites in dipterans. *Molecular Biology and Evolution* 18:384-392.
- Wilson, A.C., Cann, R., Carr, S., George, M., Gyllenstein, U.B., Helmbychowski, K.M., Higuchi, R.G., 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.
- Wright, S., 1921.** Systems of mating. *Genetics* 6:111-178.
- Wright, S., 1951.** The genetical structure of populations. *Annals of Eugenetics* 15:323-354.

**Wright, S., 1965.** The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* 19:395-420.

**Xu, X., Peng, M., Fang, Z. & Xu, X., 2000.** The direction of microsatellites mutations is dependent upon allele length. *Nature Genet.* 24:396-399.

**You-Chu, L., Korol, A.B., Tzion, F., Beiles, A. & Nevo, E., 2002.** Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology* 11: 2453-2465.

**Young, E.T., Sloan, J.S. & Riper, K.V., 2000.** Trinucleotide repeats are clustered in regulatory genes in *Saccharomyces cerevisiae*. *Genetics* 154:1053-1068.

**Zane, L., Barcelloni, L. & Patarnello, Y., 2002.** Strategies for microsatellite isolation: a review. *Mol. Ecol.* 11:1-6.

## **Anexos**

---

## Anexo 1

Protocolo de extracção de ADN: método salino/ fenol-clorofórmio (Pérez, 2003)

1-Cortar um bocado de tecido (~100mg), lavá-lo em água e cortá-lo em pedacinhos.

2-Colocar o tecido num eppendorf de 1,5ml com 300 µl de tampão de lise (0.1M Tris; 0.01M EDTA; 0.1M NaCl; 1% SDS; pH=0.8). Adicionar 30µl de Pronase- PK (20mg/ml). Vortexar uns segundos e incubar a 40°C durante 15h ou até o tecido se encontrar completamente dissolvido.

3-Adicionar 150µl de NaCl 5M e misturar fortemente durante 2min.

4-Centrifugar a 12000rpm durante 5min( os mucossacarídeos precipitam juntamente com as proteínas).

5-Transferir o sobrenadante para um eppendorf limpo e adicionar 150µl de NaCl 5M; mexer fortemente durante 2min; centrifugar a 12000rpm durante 5min.

6-Cuidadosamente, transferir o sobrenadante para um eppendorf limpo e precipitar o ADN com uma concentração final de 0.2M de NaCl (60µl) e etanol puro até encher o eppendorf.

7-Centrifugar a 14000rpm durante 15min; eliminar o sobrenadante vertendo o eppendorf se a pellet estiver bem presa e resuspender com 300µl de água do esguicho usando a pipeta.

8-Adicionar 0.5µl de RNase (10mg/ml) e deixar à temperatura ambiente no rotamix durante 1h.

A partir daqui é feito na hote!

9-Misturar bem com 1 volume de Phenol:Chlorophorme:Isoamyl alcohol (25:24:1) (agitar antes de usar) (300µl). Colocar no rotamix durante 30min.

10-Centrifugar a 8000 rpm durante 5min.

11-Transferir o sobrenadante para um eppendorf limpo e misturar com 0.5 volumes de Chlorophorme:Isoamyl alcohol (49:1) (150 $\mu$ l). Misturar bem durante 10min.

12-Centrifugar a 8000 rpm durante 5min.

13-Transferir o sobrenadante para um eppendorf limpo (definitivo).

14-Adicionar 2 volumes de etanol puro (600 $\mu$ l) Misturar. Incubar 10min a  $-20^{\circ}\text{C}$  (congelador). Centrifugar a 12000 rpm a  $4^{\circ}\text{C}$  e durante 12min.

15-Eliminar o sobrenadante vertendo o eppendorf e lavar a pellet com 750 $\mu$ l de álcool a 70%. Centrifugar a 14000 rpm a  $4^{\circ}\text{C}$  e durante 5 min.

16-Eliminar o álcool vertendo e com a ajuda da pipeta na parte final. Deixar a pellet a secar de um dia para o outro.

17-Ressuspender em 50 $\mu$ l de AE (buffer) e colocar no balancé o resto do dia.

## Anexo 2

Coordenadas utilizadas na análise espacial do programa BAPS.

População	Coordenadas	
	N	O
Flores	39,2	31,0
Faial	38,3	28,3
Pico	38,2	28,1
Terceira	39,0	27,6
Graciosa	38,3	27,1
São Miguel	37,4	25,4
Santa Maria	36,5	24,5
Canárias	28,2	14,1

## Anexo 3

Frequências alélicas para os locus Pde1, Pde3, Pde6, Pde5 e PrC5.

Locus	População							
	Flores	Faial	Pico	Graciosa	Terceira	S.Miguel	Sta.Maria	Canárias
<b>Pde1</b>								
94	0	0,0135	0	0	0	0	0	0
95	0	0,0135	0	0	0	0	0	0
97	0	0	0	0	0,0222	0	0	0
100	0	0,027	0	0	0,0222	0,0156	0	0,0625
101	0,0114	0,027	0	0	0,0444	0	0,102	0,125
102	0	0	0	0	0	0	0,051	0,0417
103	0,0114	0	0	0,0222	0	0	0	0,0625
104	0,0114	0	0	0	0	0	0	0
105	0	0	0	0,0222	0	0	0,0102	0,0625
106	0,0114	0,0135	0	0	0	0	0	0,0417
108	0,1818	0,0811	0,2222	0,1222	0,1111	0,0781	0,1429	0,1042
109	0	0	0	0,0111	0	0	0	0
110	0,375	0,3514	0,5926	0,3444	0,3444	0,5625	0,4388	0,1875
111	0	0	0	0,0111	0,0111	0	0	0,0625
112	0,0568	0,1081	0,0556	0,1889	0,1556	0,1563	0,0816	0
113	0,0227	0	0	0,0333	0,0111	0	0	0,0417
114	0,0341	0,027	0,0185	0,0444	0,0222	0,0313	0,0102	0,0417
115	0,0227	0,0541	0	0	0,0667	0,0938	0	0
116	0,1932	0,0676	0,0556	0,1	0,0556	0,0313	0,0408	0
117	0,0114	0,0541	0,0185	0,0222	0	0	0	0
118	0,0341	0,0541	0	0,0111	0,0222	0,0313	0,051	0
119	0,0227	0	0	0	0	0	0	0
120	0	0	0,0185	0	0	0	0	0
121	0	0	0,0185	0	0	0	0	0
122	0	0,1081	0	0,0667	0,1111	0	0,0714	0,1667
<b>Pde3</b>								
120	0	0	0	0	0	0	0	0,0143
122	0	0	0	0	0	0	0	0,0429
124	0	0,0256	0	0	0	0	0	0
126	0,051	0,0769	0,0357	0,0455	0,0652	0,0417	0	0,0143
128	0	0,0513	0	0,0114	0,0435	0	0,0778	0,0286
129	0	0,0128	0,0179	0	0,0217	0,0278	0,0222	0,0143
130	0,7959	0,6795	0,9107	0,875	0,6957	0,75	0,7444	0,6429
132	0	0,0128	0,0179	0,0227	0	0	0	0
134	0,0408	0,0256	0,0179	0,0341	0,0217	0,0139	0,0333	0
138	0,0102	0	0	0,0114	0	0	0,0111	0,0714
140	0,102	0,1154	0	0	0,1522	0,1667	0,1111	0,1714
<b>Pde6</b>								
117	0,0349	0,0323	0,0862	0,0341	0,0455	0,1081	0,0543	0,08
120	0	0	0	0	0	0	0,0217	0,06
121	0	0,0484	0	0	0	0	0	0,1
123	0,0465	0,2258	0,069	0,1364	0,1023	0,0676	0,0978	0,02
126	0,4535	0,3387	0,5345	0,4091	0,375	0,5	0,4457	0,4
129	0,4651	0,3548	0,3103	0,3182	0,4773	0,3243	0,3804	0,34
132	0	0	0	0,1023	0	0	0	0

Anexo 3 (continuação)

Frequências alélicas para os locus Pde1, Pde3, Pde6, Pde5 e PrC5.

Locus	População							
	Flores	Faial	Pico	Graciosa	Terceira	S.Miguel	Sta.Maria	Canárias
<b>Pde5</b>								
211	0,0135	0	0	0	0	0	0,0106	0
221	0	0	0	0	0	0	0	0,0263
223	0	0	0	0	0	0	0	0,0526
225	0	0	0	0	0	0	0	0,0263
227	0	0	0,0217	0	0	0	0,0106	0,0395
229	0,0676	0,0556	0,087	0,1071	0,051	0,0658	0,0319	0,2763
231	0	0,0185	0,0435	0,0119	0,0102	0	0,0319	0,0658
233	0,0405	0,0556	0	0	0,0102	0	0	0
235	0	0,0185	0	0,0119	0,0102	0	0	0
237	0,027	0	0	0,0238	0	0	0	0
239	0,0135	0	0,0217	0,0476	0	0	0,0106	0,0263
241	0,0946	0,0741	0,1087	0,0714	0,0306	0,1053	0,0638	0,0132
243	0,3649	0,2778	0,1739	0,1667	0,3061	0,4605	0,3936	0,0263
245	0,1216	0,3148	0,3043	0,1667	0,2551	0,1053	0,1596	0,1316
247	0	0,0556	0,0435	0,0238	0,0408	0,0395	0,0213	0
249	0,027	0	0,0217	0,1071	0,0102	0	0,0106	0
251	0,0135	0	0	0,0238	0	0	0	0
253	0	0	0	0	0,0306	0	0	0
255	0,0135	0	0	0	0	0	0	0
257	0,0135	0	0	0,0119	0	0	0	0
259	0	0	0	0	0	0	0,0106	0
261	0,1892	0,1296	0,1739	0,2262	0,2449	0,2237	0,2447	0,3158
<b>PrC5</b>								
116	0	0,0143	0	0	0	0	0	0
120	0	0	0,0192	0	0	0	0	0
136	0	0	0,0385	0	0	0	0	0,027
140	0	0	0	0,0227	0,0357	0	0,0513	0
142	0	0	0,0385	0	0	0	0	0
144	0	0,0857	0	0,0114	0,0476	0,0139	0,0641	0,0405
146	0	0	0	0	0	0,0278	0	0,0541
148	0,0104	0,0571	0	0,0568	0,0238	0	0,0128	0,027
150	0,0104	0,1286	0,0385	0,0568	0,0714	0,0139	0,0385	0
152	0,0104	0	0,0769	0,0795	0,0357	0,0417	0,1154	0,0541
154	0,0104	0	0	0	0	0,0417	0	0,027
156	0	0	0,0192	0,0227	0,0119	0,0417	0,0513	0,1757
158	0	0,0143	0	0	0	0	0,0256	0
160	0,1771	0,1714	0,0577	0,1477	0,2024	0,0833	0,0897	0,1351
164	0	0,0429	0,0577	0,0341	0,0238	0,0417	0,0128	0
166	0	0	0	0,0114	0	0,0278	0	0
168	0,5417	0,2857	0,3462	0,3977	0,381	0,3889	0,3462	0,1892
172	0,0313	0,0286	0	0,0114	0	0	0,0256	0,0541
176	0	0	0	0	0,0119	0	0	0
180	0,0104	0,0286	0	0	0	0	0	0
184	0	0,0429	0,0385	0	0,0357	0	0	0
192	0	0	0	0	0	0,0278	0	0
196	0,0104	0	0	0	0	0	0,0385	0
200	0,1875	0,1	0,2692	0,1477	0,119	0,25	0,1282	0,2162