

**Universidade de Lisboa**

Faculdade de Ciências  
Departamento de Biologia Vegetal



**Caracterização e Estrutura Genéticas da Cultivar  
de Oliveira 'Cobrançosa' e sua Relação com o  
Zambujeiro**

Mara Lisa Vieira Alves

Mestrado em Biologia Celular e Biotecnologia

2007

**Universidade de Lisboa**

Faculdade de Ciências  
Departamento de Biologia Vegetal



**Caracterização e Estrutura Genéticas da Cultivar  
de Oliveira ‘Cobrançosa’ e sua Relação com o  
Zambujeiro**

**Dissertação de Mestrado**

**Realizado sob a orientação do Professor Doutor Pedro Fevereiro**

Mara Lisa Vieira Alves

Mestrado em Biologia Celular e Biotecnologia

2007

## Agradecimentos

Este último ano tem sido repleto de muito trabalho mas fundamentalmente de muita amizade. A todas as pessoas que me ajudaram agradeço do fundo do meu coração:

Ao meu orientador, Professor Doutor Pedro Fevereiro, pela orientação, pela amizade, por me aturar e até por, de vez em quando, se rir das minhas piadas.

Um agradecimento também à Doutora Manuela Veloso, ao Doutor Rogério Tenreiro, à Doutora Vitória Gemas, à Doutora Isabel Raposo e à Ana Catarina, pela preciosa ajuda e pela disponibilidade demonstrada.

A toda a equipa do laboratório de Biotecnologia de Células Vegetais (ITQB) por me ter acolhido e suportado, mesmo quando eu andava de “mau humor”.

Às minhas manas adoptivas, Inyfffeas e Mary John (vocês estão aqui... para sempre!), ao meu Fairy, ao Dudu, e à aquisição mais recente (Hugooo, és o maior!). Sem vocês não teria sobrevivido!

Ao grupo das ‘jolas pós-laborais (Sassoeiros *rules!*).

Ao meu “cantinho” especial no Campo Grande (Luisinha, Alex, Susie, Micas), por estarem sempre de portas abertas.

À minha colega, amiga e *flat mate*, Moni, por me suportar diariamente (para ti um chioração do tamanho da Cruz Quebrada! “E sim, eu até tenho alguma piada”).

E finalmente, às pessoas mais importantes da minha vida, os meus pais e irmã. Amo-vos!

Obrigada a todos!

## Resumo

A oliveira é uma das espécies cultivadas com maior importância cultural e económica na bacia do Mediterrâneo (Ribeiro 1991). A 'Cobrançosa' é uma cultivar Portuguesa com muita importância na região de Trás-os-Montes encontrando-se actualmente disseminada por 6 regiões de Denominação de Origem Protegida (MADRP 2007). De modo a garantir a qualidade do azeite e aumentar a sua produção, é necessário seleccionar e certificar as árvores desta cultivar.

Foram utilizados 10 *loci* associados a microssatélites para caracterizar e avaliar a estrutura genética de seis populações de 'Cobrançosa' e sua relação com outras cultivares e com o zambujeiro. Os parâmetros de diversidade como a heterozigosidade esperada e observada foram calculados. Determinou-se a distância genética entre acessos de acordo com  $[1 - \text{proporção de alelos partilhados}] (Dps)$  (Bowcock *et al.* 1994). A matriz resultante foi utilizada para estabelecer as relações fenéticas entre os acessos e para avaliar a eventual estrutura genética entre as diferentes populações.

Não se detectou diferenciação entre as seis populações de 'Cobrançosa' avaliadas. Nesta cultivar foram detectados em média  $5.4 \pm 1.7$  alelos por *locus*, sendo que, em média,  $2.6 \pm 1.3$  tinham uma frequência inferior a 5% que poderão ser úteis, se à sua presença estiverem associadas determinadas características de valor agronómico. O valor médio de diversidade genética foi de  $0.511 \pm 0.203$  o que permite perspectivar um processo de melhoramento desta cultivar de modo a fazer face às exigências do mercado do azeite.

Embora a maior parte das árvores da cultivar 'Cobrançosa' seja geneticamente muito semelhante, verificou-se que esta cultivar parece ter uma origem policlonal.

Com este trabalho foi possível desenvolver padrões de microssatélites que permitem a identificação dos acessos da cultivar 'Cobrançosa' e inclusive a detecção de sete árvores que possivelmente estarão incorrectamente identificadas como pertencendo a essa cultivar.

**Palavras-chave:** *Olea europaea* L., 'Cobrançosa', diversidade genética, relações fenéticas, microssatélites.

## Abstract

Olive is one of the most economically important species in the Mediterranean Basin (Ribeiro 1991). 'Cobrançosa' is a Portuguese cultivar with great importance in Trás-os-Montes region and is currently being cultivated in all six regions of Protected Origin Denomination (MADRP 2007). To guarantee olive oil quality and to enhance its productivity it is necessary to select and certify the trees from this cultivar.

In this study 10 microsatellite loci were used to characterize and evaluate the genetic structure of six populations of 'Cobrançosa' and its relationships with other cultivars and with its wild relatives, oleasters. Diversity parameters like expected and observed heterozygosity were calculated. The genetic distance between the accessions was calculated according to  $[1 - \text{proportion of shared alleles}]$  (Dps) (Bowcock *et al.* 1994). The resulting matrix was used to establish the phenetic relationships among accessions and to evaluate the partitioning of the genetic variance observed.

No differentiation among the 'Cobrançosa' populations was detected. A mean of  $5.4 \pm 1.7$  alleles per locus were detected of which  $2.6 \pm 1.3$  had a frequency below 5%. These rare alleles can be useful if some attractive agronomic characteristics are associated to their presence. The mean value of genetic diversity was  $0.511 \pm 0.203$  which allows for the establishment of an improvement program to face the demands of the olive oil industry.

Although most of the 'Cobrançosa' accessions evaluated were genetically similar, it was demonstrated that this cultivar has indeed a polyclonal origin.

With this work it was possible to develop microsatellite profiles to identify the accessions of the 'Cobrançosa' cultivar including the detection of trees that were incorrectly identified as belonging to this cultivar.

**Keywords:** *Olea europaea* L., 'Cobrançosa', genetic diversity, phenetic relationships, microsatellites.

# Índice

Agradecimentos .....	ii
Resumo .....	iii
Abstract .....	iv
1. Introdução .....	1
1.1. <i>Olea europaea</i> L. – Taxonomia .....	1
1.2. <i>Olea europaea</i> L. subsp. <i>europaea</i> – Dados históricos .....	1
1.3. <i>Olea europaea</i> L. subsp. <i>europaea</i> – Biologia .....	2
1.4. Domesticação – Consequências .....	2
1.5. A oliveira em Portugal .....	3
1.5.1. Denominação de origem protegida (DOP) .....	3
1.5.2. ‘Cobrançosa’ .....	4
1.6. Identificação das cultivares de oliveira .....	5
1.7. Marcadores moleculares .....	5
1.8. Utilização de marcadores moleculares em oliveira .....	6
1.9. Microssatélites ou SSRs ( <i>Simple Sequence Repeats</i> ) .....	7
1.9.1. Definição .....	7
1.9.2. Modelos mutacionais em microssatélites .....	7
1.9.3. Porquê os microssatélites? .....	8
1.9.4. Descrição Técnica .....	9
1.10. Objectivos .....	10
2. Materiais e Métodos .....	11
2.1. Material Vegetal .....	11
2.2. Extração de DNA .....	13
2.3. Delineamento Experimental .....	14
2.4. Condições de Amplificação .....	14
2.5. Análise dos Fragmentos .....	15
2.6. Cálculo da probabilidade de erro e reprodutibilidade .....	16
2.7. Cálculo dos parâmetros de diversidade e das relações genéticas .....	17
2.7.1. Distância genética .....	17
2.7.2. Estrutura genética .....	17
2.7.3. Relações fenéticas .....	18
2.7.3.1. Métodos hierárquicos .....	18
2.7.3.2. Análise das coordenadas principais .....	19
2.7.4. Diversidade genética .....	19

2.7.4.1. Frequência alélica .....	19
2.7.4.2. Heterozigosidade .....	19
2.7.4.3. Poder de discriminação .....	19
3. Resultados e Discussão .....	20
3.1. Amplificação dos microssatélites identificados noutras espécies .....	20
3.2. Reprodutibilidade .....	20
3.3. Estrutura genética inter- e intrapopulacional .....	21
3.4. Nível de polimorfismo detectado e diversidade genética em <i>Olea europaea</i> L. ....	22
3.5. Diversidade genética na cultivar ‘Cobrançosa’ .....	24
3.6. Relações fenéticas .....	25
4. Conclusões .....	28
Referências bibliográficas .....	29
Anexos .....	I
Tabelas .....	I
Figuras .....	III

## 1. Introdução

### 1.1. *Olea europaea* L. – Taxonomia

A espécie, *Olea europaea* L., pertence à família Oleaceae, sub-família Oleideae. A família inclui cerca de 30 géneros (Johnson 1957), contendo desde espécies arbustivas ornamentais como o jasmim (*Jasminum fruticans* L.) a espécies de porte arbóreo, como o freixo (*Fraxinus excelsior* L. e *F. angustifolia* Vahl.) e o lentisco (*Phyllirea angustifolia* L.). O género *Olea*, inclui dois sub-géneros: *Olea* e *Paniculatae*. O sub-género *Olea* é dividido em duas secções: *Olea*, que contém apenas a espécie *Olea europaea*, e *Ligustroides*. De acordo com a revisão mais recente da taxonomia de *Olea europaea* L. (Green e Wickens 1989, Green 2002) esta espécie engloba seis subespécies, de acordo com características morfológicas e a sua distribuição geográfica. A subespécie *europaea* inclui duas variedades botânicas: *europaea* (oliveiras cultivadas) e *sylvestris* (oliveiras selvagens), vastamente distribuídas pela Bacia do Mediterrâneo.

### 1.2. *Olea europaea* L. subsp. *europaea* – Dados históricos

A oliveira é possivelmente a fruteira com maior importância da bacia do Mediterrâneo. Conjuntamente com a vinha, pertence ao grupo das plantas que fundou a horticultura no Velho Mundo. Desde a Idade do Bronze que a riqueza de muitos povos do Mediterrâneo se centrou em torno do cultivo das oliveiras, as quais providenciavam o azeite e a azeitona. O azeite era usado na cozinha como unguento e para a iluminação. Dada a sua grande capacidade de conservação, este produto era facilmente comercializado. Os frutos também eram armazenados para consumo. O pão e o azeite eram – e ainda são – a base da alimentação de muitas comunidades de camponeses da Bacia do Mediterrâneo (Zohary e Hopf 2000).

Assume-se que as oliveiras cultivadas tiveram origem nas oliveiras selvagens (também chamadas de zambujeiros), tendo-se disseminado por todo o Mediterrâneo por acção do homem. É também assumido que possíveis cruzamentos entre as oliveiras cultivadas e as oliveiras selvagens possam ter levado à criação de novas cultivares em diferentes partes do Mediterrâneo (Besnard *et al.* 2001a).

Existem evidências que parecem demonstrar que a domesticação da oliveira se deu em ambas as extremidades do Mediterrâneo. No médio oriente a domesticação parece ter começado no início da Idade do Bronze, enquanto que no Mediterrâneo Ocidental essa domesticação parece ter-se iniciado apenas no final da Idade do Bronze. No século VI antes de Cristo o cultivo da oliveira espalhou-se por todo o Mediterrâneo, chegando à Tunísia e

Sicília, e mais tarde ao norte de Itália. No século XV, com a chegada dos europeus à América, as oliveiras chegaram ao “novo mundo”. No entanto, só recentemente é que o seu cultivo ganhou alguma importância económica, para além da área do mediterrâneo, nomeadamente, na Austrália, América do Sul (Argentina e Chile) e África do Sul (Doveri e Baldoni 2007).

### **1.3. *Olea europaea* L. subsp. *europaea* – Biologia**

As oliveiras são maioritariamente diplóides ( $2n = 2x = 46$ ) (Falistocco e Tosti 1996, Minelli *et al.* 2000), mas já foi relatada a existência de indivíduos tetraplóides (Rugini *et al.* 1996). O conteúdo de ácido desoxirribonucleico (DNA – *deoxyribonucleic acid*) por núcleo diplóide é de 2.90 picogramas, o que equivale a um genoma com o tamanho de 2.8 Gbp (Loureiro *et al.* 2007).

A oliveira é uma árvore com um crescimento relativamente lento. A produção começa 5 a 6 anos após a plantação mas geralmente só se atinge o máximo de produtividade após 10 anos de cultivo (Doveri e Baldoni 2007). Se bem tratadas, as árvores podem continuar a providenciar frutos durante centenas de anos. As oliveiras sob cultivo apresentam uma variação considerável no tamanho, forma, e conteúdo em óleo dos seus frutos. Existem centenas de variedades distintas, e diferentes partes da bacia do mediterrâneo são frequentemente caracterizadas por variedades específicas. As oliveiras cultivadas podem ser grosseiramente divididas em dois tipos principais: as variedades utilizadas para a produção de azeite; e as variedades utilizadas para a produção de azeitonas de mesa. No entanto, existem variedades que são usadas para os dois fins, denominadas de “dupla-aptidão” Zohary e Hopf (2000).

### **1.4. Domesticação – Consequências**

Sob um processo de domesticação a biologia de propagação das árvores alterou-se. As variedades cultivadas são mantidas e multiplicadas através da propagação vegetativa dando origem a populações clonais. Este método está dependente primariamente da utilização de botões que se desenvolvem na base do tronco, e que enraízam facilmente quando cortados. As árvores também podem ser propagadas por estacaria e enxertia (Zohary e Hopf 2000).

Na oliveira, assim como noutras árvores de fruto, a passagem para a propagação vegetativa é a forma que o agricultor encontrou para resolver os problemas da segregação genética causados pela reprodução sexual das espécies com polinização cruzada. As oliveiras cultivadas são, de um modo geral, extremamente heterozigóticas dando origem a progénies com características segregantes. Na verdade, a maioria das plantas resultantes das sementes assemelham-se morfológicamente às formas selvagens e não são próprias

para a produção de frutos de qualidade. Consequentemente, a propagação a partir da semente é impraticável em olivicultura. De forma a fixar genótipos interessantes o produtor tem que fazer a propagação do olival através da propagação clonal (Zohary e Hopf 2000).

## **1.5. A oliveira em Portugal**

A oliveira (*Olea europaea* L. subsp. *europaea*) é uma espécie cultivada de grande importância económica na Bacia do Mediterrâneo (Ribeiro 1991). Em 2005, esta região foi responsável por cerca de 96% da produção mundial (MADRP 2007). A oliveira é cultivada em Portugal pelo menos desde as invasões romanas (Ribeiro 1991) e é uma das culturas mais importantes no país, produzindo o segundo produto agrícola no que se refere ao valor económico, sendo actualmente Portugal o oitavo produtor mundial de azeite.

As variedades portuguesas têm sido maioritariamente mantidas e propagadas vegetativamente através de estacas usando o material vegetal disponível localmente (Gemas *et al.* 2004). Os agricultores seleccionaram o material com base nos traços agronómicos resultantes da interacção entre os genótipos e as condições ambientais existentes (Gemas *et al.* 2004).

Em território nacional, são maioritariamente utilizadas 22 cultivares para a produção de azeite e azeitonas de mesa (Leitão *et al.* 1986), destacando-se o Alentejo e Trás-os-Montes como as principais regiões produtoras (MADRP 2007). A cultivar mais utilizada no país é a 'Galega' (actualmente representando cerca de 60% de todas as oliveiras no país). Existem ainda outras cultivares que ocupam áreas significativas como a 'Cobrançosa' na região de Trás-os-Montes, correspondendo a 20% da área de olival (MADRP 2007).

Em 1999, 74% do olival Português era constituído por árvores com mais de 50 anos (MADRP 2007). No entanto, de modo a satisfazer as exigências do mercado do azeite, os terrenos de oliveiras têm de ser renovados usando árvores certificadas, seleccionadas a partir das cultivares Portuguesas (Gemas *et al.* 2004).

### **1.5.1. Denominação de origem protegida (DOP)**

A União Europeia desenvolveu o conceito de Denominação de Origem Protegida (DOP) para promover os azeites regionais (produtos provenientes de uma área geográfica delimitada) com características únicas, satisfazendo um determinado número de especificações, como as cultivares que podem fazer parte de um determinado azeite. Estas regras encontram-se enumeradas num caderno de especificações, Regulamento (CE) n.º 510/2006 (MADRP 2007).

Os produtos de denominação de origem protegida (DOP) possuem um valor económico superior comparativamente com os que não têm essa denominação (Doveri e Baldoni 2007).

Consequentemente, de modo a garantir aos consumidores a autenticidade dos produtos, é importante evitar possíveis misturas ou substituições dos materiais brutos e de elaborar métodos eficazes para permitir o controlo das cultivares utilizadas durante o processamento do material (Pasqualone et al. 2007). O International Olive Oil Council (1993) e a Codex Alimentarius Commission (1993) determinaram os padrões de qualidade que os diferentes tipos de azeite devem apresentar.

Em Portugal existem seis Denominações de Origem Protegida para o mercado do azeite - Azeite de Trás-os-Montes, Azeites da Beira Interior, Azeites do Ribatejo, Azeites do Norte Alentejo, Azeite de Moura e Azeites do Alentejo Interior (MADRP 2007). Esta certificação baseia-se em características agro-ecológicas dessas áreas e no uso das cultivares locais. Com a reabilitação dos olivais Portugueses é necessário garantir que a renovação seja feita usando árvores certificadas, escolhidas a partir de cultivares Portuguesas seleccionadas de modo a garantir a autenticidade do produto obtido. Considerando o actual contexto no melhoramento das oliveiras em Portugal, a análise da diversidade genética é essencial para uma correcta aquisição, manutenção e uso dos recursos genéticos disponíveis (Gemas et al. 2004).

### 1.5.2. 'Cobrançosa'

A 'Cobrançosa' é uma cultivar Portuguesa utilizada na produção de azeite. Actualmente é uma das cultivares cujo cultivo está autorizado em todas as regiões agrárias (Tabela 1), com maior incidência na região de Trás-os-Montes, para a produção de azeites DOP (MADRP 2007). É auto-compatível, tem uma capacidade de produção média mas constante e o seu fruto tem uma capacidade média para a produção de azeite. Apresenta grande tolerância ao frio, à clorose do ferro e ao patógeno *Pseudomonas syringae* subsp. *savastanoi* responsável pela tuberculose em oliveira (Bartolini et al. 2005).

Tabela 1 – Variedade de oliveira autorizadas nas várias regiões agrárias, em Portugal. ( Adaptado de Ministério da Agricultura do Desenvolvimento Rural e Pescas, 2007)

Região	Variedades Autorizadas
Algarve	Maçanilha, Galega, <u>Cobrançosa</u>
Alentejo	Galega, Cordovil, Verdeal, <u>Cobrançosa</u> , Maçanilha, Carrasquenha, Bical de Évora, Redondil, Conserva de Elvas, Azeiteira
Ribatejo	Galega, <u>Cobrançosa</u> , Lentisca
Beiras	Galega, Cordovil, Bical, <u>Cobrançosa</u> , Carrasquinha, Cornicabra
Trás-os-Montes	<u>Cobrançosa</u> , Verdeal, Madural, Negrinha, Santulhana

## 1.6. Identificação das cultivares de oliveira

Actualmente, os métodos oficiais usados para caracterizar e identificar variedades em fruteiras ainda se baseiam na caracterização morfológica e observações do fenótipo de acordo com os descritores da UPOV (*Unité pour la Protection des Obtentions Végétales*) e do IPGRI (*International Plant Genetic Resources Institute*). Esta metodologia requer tempo e as características morfológicas podem estar dependentes de influências ambientais (Bouhadida *et al.* 2007).

Existem, pelo menos, 1275 cultivares de oliveiras descritas (Bartolini *et al.* 1998) mas muitas outras variedades e ecótipos contribuem para a riqueza do germoplasma de oliveira. Poucas são as cultivares dispersas globalmente, sendo a maioria cultivado ao nível regional/nacional (Doveri e Baldoni 2007). Esta grande variabilidade traz consigo a necessidade de se desenvolverem ferramentas que permitam distinguir as diferentes cultivares.

Os métodos tradicionais utilizados para a identificação de cultivares e de porta-enxertos baseiam-se em observações fenotípicas; no entanto este é um processo moroso devido à longa fase juvenil das árvores e está sujeito às influências ambientais (Wünsch e Hormaza 2002).

As características morfológicas e agronómicas são muitas vezes usadas para uma caracterização inicial, dado que esta informação é muito útil para os utilizadores dos recursos genéticos vegetais. Este tipo de dados não é só interessante para os utilizadores directos desses recursos mas também permite a interpretação da relação entre os génotipos e as condições ambientais.

As técnicas baseadas em marcadores moleculares têm-se tornado ferramentas muito poderosas na análise da diversidade genética. Se à presença de determinados marcadores moleculares estiverem associadas características interessantes, estes marcadores tornam-se valiosos na aceleração dos programas de melhoramento (Hammer *et al.* 2003).

## 1.7. Marcadores moleculares

Em anos recentes tem-se vindo a assistir a avanços críticos na aplicação de marcadores moleculares para a genotipagem de plantas cultivadas. As vantagens a eles associadas têm sido altamente reconhecidas, sendo estes sistemas ainda são mais importantes em lenhosas perenes devido às particularidades destas espécies, como seja o seu longo tempo de juvenilidade e o facto da sua propagação ser feita vegetativamente (Wünsch e Hormaza 2002).

A produção de plantas certificadas requer a aplicação de técnicas rápidas e fiáveis para a identificação dos génotipos. A incorporação de novas metodologias permite a genotipagem

dos indivíduos em qualquer fase do seu desenvolvimento e independentemente de factores ambientais que possam influenciar o fenótipo (Wünsch e Hormaza 2002).

Entre outras aplicações, os marcadores moleculares, quando associados a genes de interesse, são úteis na selecção e introgressão de genes (Durman e Korban 1994), na construção de mapas genéticos (por exemplo, de la Rosa *et al.* 2003 e Wu *et al.* 2004), na identificação de cultivares e no estabelecimento de relações genéticas entre elas (por exemplo, Lopes *et al.* 2004).

Existem vários métodos disponíveis para analisar a diversidade genética. Dado que cada um deles nos dá diferentes tipos de informação, a sua escolha depende da informação necessária à sua aplicação, tal como os recursos e as infra-estruturas necessários (Hammer *et al.* 2003).

### **1.8. Utilização de marcadores moleculares em oliveira**

Os estudos genómicos em oliveira são ainda muito preliminares, estando longe de produzir resultados efectivos na utilização de ferramentas moleculares para selecção de novas cultivares (revisto por Doveri e Baldoni 2007).

Por essa razão, e considerando a falta de conhecimento em relação à variabilidade já presente nas oliveiras selvagens e cultivadas, nos últimos 10 anos tem-se avaliado esse germoplasma composto por um grande número de cultivares e populações selvagens (Doveri e Baldoni 2007).

A análise molecular, usando quer marcadores nucleares ou citoplasmáticos, mostrou que as populações de zambujeiros do Mediterrâneo Oriental e Ocidental se encontram fortemente diferenciadas (Besnard *et al.* 2002, Lumaret *et al.* 2004). Pelo contrário, as oliveiras cultivadas não apresentam esta estrutura geográfica, embora a sua variabilidade seja elevada. Têm sido repetidamente apresentadas evidências moleculares da selecção multilocal para a maioria das cultivares (Besnard *et al.* 2001a, Rotondi *et al.* 2003).

O estabelecimento dos perfis individuais, como se se tratasse de uma impressão digital, é uma preciosa ajuda na identificação de cultivares, já que pode ser utilizado na geração de um perfil específico para um dado genótipo. Na última década os marcadores moleculares têm sido muito utilizados na identificação e caracterização de diferentes cultivares. Estes estudos têm utilizado RAPD (*Random Amplified Polimorphic DNA*) (Belaj *et al.* 2001, Besnard *et al.* 2001b, Ganino *et al.* 2007), AFLPs (*Amplified Fragment Length Polymorphisms*) (Angiolillo *et al.* 1999, Rotondi *et al.* 2003), ISSRs (*Inter Simple Sequence Repeats*) (Gemas *et al.* 2004, Martins-Lopes *et al.* 2007) e SSRs (*Simple Sequence Repeats*) (Rallo *et al.* 2000, Sefc *et al.* 2000, Carriero *et al.* 2002, Cipriani *et al.* 2002, Díaz *et al.* 2006).

## 1.9. Microssatélites ou SSRs (*Simple Sequence Repeats*)

### 1.9.1. Definição

Os microssatélites são uma classe de marcadores moleculares, também chamados de SSRs (*Simple Sequence Repeats*) isto porque o seu polimorfismo se baseia na variação do número de repetições de uma sequência curta de DNA (com 1 a 6 pares de bases) repetida sequencialmente, encontrando-se espalhados pelo genoma de todos os organismos (Cipriani *et al.* 2002). A maioria destes *loci* são neutros o que os torna compatíveis com as premissas da maior parte das teorias de genética populacional (Murray *et al.* 1996).

Os SSRs são marcadores moleculares multi-alélicos, hipervariáveis e co-dominantes de grande robustez e reprodutibilidade, e como a sua produção se faz por reacção da polimerase de DNA em cadeia (DNA *Polimerase Chain Reaction* – PCR), é necessária pouca quantidade de DNA para se proceder à sua análise (Díaz *et al.* 2006). Além disso podem ser transferidos para espécies próximas já que as sequências que flanqueiam os microssatélites são altamente conservadas (de la Rosa *et al.* 2002; Rallo *et al.* 2003). Esta vantagem não é satisfeita por mais nenhum dos marcadores usados no estabelecimento de perfis individuais (RAPDs, RFLPs, ISSRs ou AFLPs) (Wünsch e Hormaza 2002). Devido a estas características os SSRs têm vindo a ser os marcadores de eleição para a caracterização molecular de diferentes espécies de plantas (Gupta e Varshney 2000).

A maioria dos microssatélites tem uma taxa de mutação muito elevada (entre  $10^{-2}$  e  $10^{-6}$  mutações por *locus* e por geração) o que gera grandes níveis de diversidade alélica (Schlötterer 2000). Essa taxa de mutação é dependente do número e do tipo de repetições bem como da região flanqueante e da sua localização no cromossoma (Goldstein e Clark 1995).

### 1.9.2. Modelos mutacionais em microssatélites

O processo mutacional associado aos microssatélites pode ser muito complexo (Schlötterer 2000).

O mecanismo mais simples para explicar o processo mutacional dos microssatélites baseia-se no deslizamento da polimerase durante a replicação do DNA sendo que também as mutações se podem dever a uma troca de segmentos de cromossomas durante o emparelhamento dos cromossomas homólogos na meiose ou ainda à ocorrência de mutações pontuais (Richard e Paques 2000). Existem dois modelos fundamentais associados aos processos mutacionais dos microssatélites: o modelo dos alelos infinitos (*infinite allele model* – IAM) e o modelo das mutações passo-a-passo (*stepwise mutation model* – SMM). O primeiro estipula que todos os eventos mutacionais dão origem a um

novo alelo, e que este alelo pode ter uma diferença de uma ou mais repetições em relação ao alelo que lhe deu origem. O modelo SMM prevê que um evento mutacional resultará num alelo que difere apenas numa repetição. Neste caso, a diferença de tamanho dos diferentes alelos dá-nos informação adicional acerca das relações de ancestralidade entre eles. Quanto maior for a diferença, maior será o número eventos mutacionais que ocorreram desde o ancestral comum. Existe assim um registro de processos mutacionais do passado (Hardy *et al.* 2003).

O modelo dos alelos infinitos, em que cada mutação leva à criação de um novo alelo, tem sido tradicionalmente o modelo adoptado para a análise em genética populacional, e por ser o mais simples e mais geral é muitas vezes utilizado por defeito (Selkoe e Toonen 2006). Comparativamente com o modelo de mutações passo-a-passo, o modelo dos alelos infinitos é geralmente mais robusto e mais fiável (Landry *et al.* 2002). Em oliveira, o tamanho dos alelos entre espécies e mesmo dentro da mesma espécie não está de acordo com o modelo mutacional passo-a-passo, indicando assim que são fontes de variação alterações, quer no número de repetições, quer nas regiões flanqueantes (Rallo *et al.* 2003).

### **1.9.3. Porquê os microssatélites?**

Os microssatélites tornaram-se uma das classes de marcadores moleculares mais utilizados no melhoramento de plantas, na identificação de variedades, em testes de paternidade e para o mapeamento do genoma (Wünsch e Hormaza 2002).

Os microssatélites são especialmente atractivos quando estudamos espécies com um nível baixo de variação genética, populações com grande taxa de cruzamento entre si e populações recentemente formadas ou muito próximas geograficamente cuja diferenciação pode ser mais difícil utilizando outra metodologia (Rakoczy-Trojanowska e Bolibok 2004).

Cada tipo de marcador tem as suas vantagens e desvantagens sendo necessário avaliar a sua utilidade antes da sua aplicação. Entre os marcadores de DNA desenvolvidos nas últimas décadas, os microssatélites são os marcadores de escolha para o estabelecimento de perfis individuais em árvores de fruto. Uma das suas vantagens é o grande grau de polimorfismo por eles revelado devido à grande taxa de mutação associada a estas sequências repetitivas. Outra vantagem reside no facto de eles serem marcadores co-dominantes permitindo a distinção entre indivíduos homozigóticos e heterozigóticos, ao mesmo tempo que seguem a segregação mendeliana que permite a sua utilização em análises de paternidade. Para além disso, a grande reprodutibilidade demonstrada pelos microssatélites, quando comparados com outros marcadores moleculares, permite a normalização de protocolos e comparação de resultados entre laboratórios (McGregor *et al.* 2000).

#### 1.9.4. Descrição Técnica

Como os microssatélites correspondem a sequências curtas de DNA (100-300 bp) eles podem ser amplificados por PCR, mesmo que o DNA esteja algo degradado (Taberlet *et al.* 1999). Esta característica permite que se utilizem métodos de extração de DNA rápidos e de baixo custo (Taberlet *et al.* 1999).

Como as regiões que flanqueiam as sequências repetitivas são altamente conservadas entre indivíduos da mesma espécie e por vezes entre diferentes espécies, o *locus* de um microssatélite pode ser identificado através das suas regiões flanqueantes. Pequenas sequências de DNA denominadas oligonucleótidos podem ser desenhadas para depois amplificar o *locus* do microssatélite através de PCR (Selkoe e Toonen 2006). Alternativamente, podemos pesquisar na bibliografia por sequências de oligonucleótidos iniciadores de amplificação já descritos para a espécie em questão ou em espécies próximas (Selkoe e Toonen 2006). A base de dados GenBank, dirigido pelo National Center of Biotechnology Information, pode ser utilizada para esse efeito pesquisando em <http://www.ncbi.nlm.nih.gov/> pelo nome da espécie, família ou género e seleccionando a base de dados correspondente aos nucleótidos (Selkoe e Toonen 2006).

Ao contrário das regiões flanqueantes, as sequências repetitivas dos microssatélites sofrem frequentemente mutações através de processos como o deslizamento da polimerase ou da revisão da sequência durante a replicação do DNA, alterando assim o número de repetições e consequentemente o tamanho do *locus* (Eisen 1999). Devido a diferenças no tamanho dos alelos, estes podem ser facilmente detectados numa electroforese de alta resolução utilizando, por exemplo, géis de poliacrilamida. Esta técnica permite a rápida genotipagem de muitos indivíduos e de muitos *loci* (Schlötterer 2000).

Antes de se tirarem conclusões sobre os resultados obtidos pela implementação desta técnica, três pontos principais devem ser testados: a neutralidade de cada *locus*; a presença de alelos nulos; e a independência entre os diferentes *loci* (revisto por Murray *et al.* 1996).

A presença de selecção num determinado *locus* pode estabilizar as frequências alélicas em todas as subpopulações levando assim a uma estimativa deficitária da estrutura da população ou da distância genética (Murray *et al.* 1996). A comparação entre as frequências genotípicas observadas com as esperadas segundo o equilíbrio de Hardy-Weinberg pode detectar se algum dos *loci* analisados está sob selecção (Murray *et al.* 1996).

Convém também ter em atenção aos potenciais erros que podem afectar os resultados obtidos. Os erros associados à utilização de microssatélites devem-se fundamentalmente a erros humanos ou do equipamento, erros introduzidos durante a amplificação dos *loci* e

erros associados ao DNA (por exemplo, a existência de alelos nulos) (revisto por Hoffman e Amos 2005). A presença de alelos nulos acontece quando ocorrem mutações na região de ligação dos oligonucleótidos iniciadores de amplificação e consequentemente os alelos de determinados indivíduos poderão falhar a amplificação (Paetkau e Strobeck 1995).

A escolha de *loci* com baixa ou alta diversidade genética depende da questão a que queremos responder. Questões de paternidade e de estrutura clonal são mais bem estudadas usando *loci* com altos níveis de diversidade genética. Neste caso a utilização de um número reduzido de loci pode ser suficiente para estabelecer perfis genéticos únicos dos diferentes indivíduos (revisto por Selkoe e Toonen 2006).

### **1.10. Objectivos**

Os objectivos deste trabalho são:

- 1) Caracterizar a diversidade molecular da cultivar Portuguesa 'Cobrançosa' recorrendo a microssatélites;
- 2) Avaliar a estrutura genética da cultivar 'Cobrançosa' e dos zambujeiros;
- 3) Desenvolver padrões de microssatélites que permitam a identificação e certificação da cultivar 'Cobrançosa';
- 4) Avaliar as relações fenéticas entre a cultivar 'Cobrançosa', os zambujeiros e outras cultivares.

## 2. Materiais e Métodos

### 2.1. Material Vegetal

A população estudada incluiu 77 indivíduos da cultivar Portuguesa ‘Cobrançosa’ recolhidas pelo Professor Antero Martins, de 9 regiões de Trás-os-Montes. Estas árvores fazem parte de um ensaio de proveniência situado em Mirandela, Herdade do Escaramunheiro (Martins *et al.* 1998) e estão catalogadas como sendo árvores antigas de forma a garantir que estarmos a cobrir o máximo de diversidade genética desta cultivar (Figura 1).

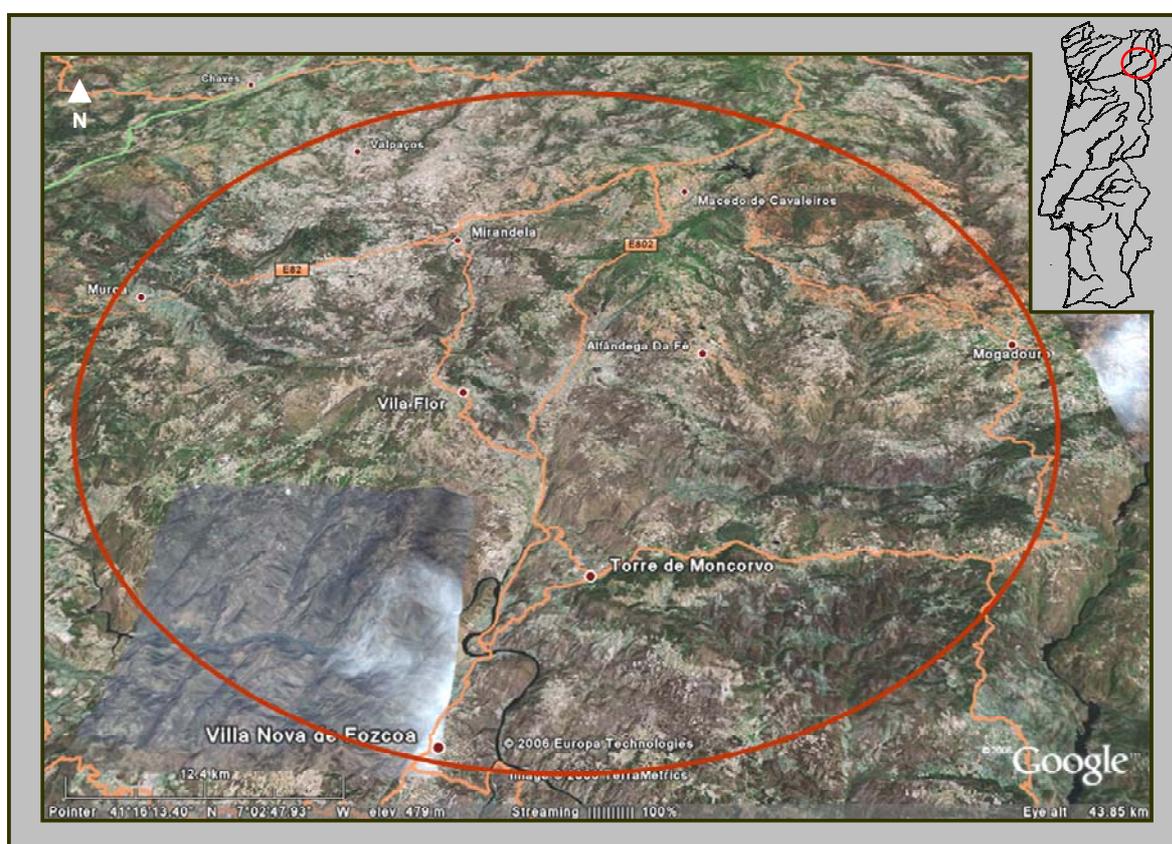


Figura 1 – Localização geográfica do local de origem dos acessos da cultivar ‘Cobrançosa’ utilizados neste trabalho.

Também fizeram parte deste estudo 3 indivíduos da cultivar Portuguesa ‘Galega’, 3 indivíduos da cultivar ‘Conserva de Elvas’, 1 indivíduo da cultivar ‘Carrasquenha’, 1 indivíduo da cultivar ‘Maçanilha de Tavira’, 1 indivíduo da cultivar ‘Madural’, 1 indivíduo da cultivar ‘Negrinha’, 1 indivíduo da cultivar ‘Redondil’, 1 indivíduo da cultivar ‘Verdeal de Trás-os-Montes’, 1 indivíduo da cultivar ‘Cordovil de Castelo Branco’, 1 indivíduo da cultivar ‘Azeiteira’, 12 indivíduos da cultivar Espanhola ‘Arbequina’, 3 indivíduos da cultivar Espanhola ‘Arbosana’, 2 indivíduos da cultivar Grega ‘Koroneiki’ e ainda 30 zambuzejros de

5 populações de 4 regiões Portuguesas (Tabela 2). O germoplasma, representado por folhas jovens, foi armazenado a -80°C até ser extraído o DNA.

Tabela 2 – Lista do germoplasma de *Olea europaea* subsp. *europaea* utilizado. As letras, entre parênteses, correspondem às abreviaturas utilizadas neste trabalho para identificar os diferentes acessos. DO-ENME corresponde ao Departamento de Olivicultura da Estação Nacional de Melhoramento de Elvas.

Variedade botânica	Cultivar	Nº de acessos	Local de Colheita	País de origem
<i>sylvestris</i> (zambujeiro – Z)	–	4	Aguda	Portugal
		7	Alcaria	Portugal
		3	S. Olaia	Portugal
		3	Ferestelo	Portugal
		3	Magoito	Portugal
		10	S. Arrábida	Portugal
<i>europaea</i>	‘Cobrançosa’ (C)	8	Alfândega da Fé	Portugal
		8	Foz Côa	Portugal
		24	Macedo de Cavaleiros	Portugal
		22	Mirandela	Portugal
		5	Mogadouro	Portugal
		2	Moncorvo	Portugal
		1	Murça	Portugal
		6	Valpaços	Portugal
	1	Vila Flor	Portugal	
	‘Galega’ (G)	3	Elvas (DO-ENME)	Portugal
	‘Conserva de Elvas’ (CE)	3	Elvas (DO-ENME)	Portugal
	‘Arbequina’ (A)	12	Santarém	Espanha
	‘Arbosana’ (B)	3	Santarém	Espanha
	‘Carrasquenha’ (CR)	1	Elvas (DO-ENME)	Portugal
	‘Maçanilha de Tavira’ (MÇ)	1	Elvas (DO-ENME)	Portugal
	‘Madural’ (MD)	1	Elvas (DO-ENME)	Portugal
	‘Negrinha’ (N)	1	Elvas (DO-ENME)	Portugal
	‘Redondil’ (R)	1	Elvas (DO-ENME)	Portugal
	‘Verdeal de Trás-os-Montes’ (V)	1	Elvas (DO-ENME)	Portugal
	‘Cordovil de Castelo Branco’ (CB)	1	Elvas (DO-ENME)	Portugal
‘Azeiteira’ (AZ)	1	Elvas (DO-ENME)	Portugal	
‘Koroneiki’ (K)	2	Santarém	Grécia	

## 2.2. Extracção de DNA

O DNA foi extraído segundo um protocolo modificado no laboratório de Biotecnologia de Células Vegetais no Instituto de Tecnologia Química e Biológica (ITQB) baseado no método de Doyle e Doyle (1987).

Para cada amostra, maceraram-se cerca de 100mg de folhas jovens de oliveira em azoto líquido, com o auxílio de uma almofariz, pilão e areia quartzosa, estéreis e previamente arrefecidos em azoto líquido. Adicionou-se o material vegetal moído, num tubo de 2 mL, a 750  $\mu$ L de tampão de extracção (2% CTAB, 1.4 M NaCl, 100 mM Tris-HCl pH 8.0, 20 mM EDTA, 1% PVP 40000 MW, 2%  $\beta$ -mercaptoetanol, 0.5 mg/mL proteinase K), previamente aquecido a 60°C. O tubo foi incubado a 60°C durante uma hora, invertendo-se o tubo em intervalos de 15 minutos. Adicionou-se 700  $\mu$ L de FCI [fenol:clorofórmio: álcool isoamílico a 25:24:1 (v/v)] e agitou-se o tubo cuidadosamente por inversão cerca de 100 vezes. Centrifugou-se a 10000g durante 5 minutos para separar a fase orgânica da fase aquosa. A fase aquosa (sobrenadante) foi transferida para um novo tubo de 2 mL descartando-se a fase orgânica. Ao novo tubo adicionou-se 700  $\mu$ L de CI [clorofórmio: álcool isoamílico a 24:1 (v/v)]. Inverteu-se 100 vezes e centrifugou-se a 10000g durante 5 minutos. O sobrenadante foi transferido para um novo tubo de 2 mL e foi-lhe novamente adicionado 700 $\mu$ L de CI. Após inversão do tubo 50 vezes e centrifugação a 10000g durante 5 minutos, retirou-se o sobrenadante para um tubo de 1.5 mL. Adicionou-se 2/3 do volume de isopropanol a -20°C e agitou-se até a “aranha” de DNA estar completamente formada. Deixou-se repousar na bancada entre 15 e 30 minutos, seguindo-se uma centrifugação durante 5 minutos a 6000g. Verteu-se o sobrenadante e lavou-se o precipitado com uma solução de etanol 70% e NaCl 0.15% durante uma hora a 37°C ou durante a noite a 4°C. Secou-se o precipitado à temperatura ambiente e juntou-se 25 a 50 $\mu$ L de TE pH8.0 (10 mM de Tris pH8.0 + 1 mM EDTA pH8.0) para dissolver o precipitado de DNA. Seguidamente procedeu-se à digestão do RNA da amostra, para tal adicionou-se à solução de DNA 1 a 2  $\mu$ L de RNase (10 mg/mL) (Roche Applied Science) e incubou-se o tubo durante uma hora a 37°C. O DNA foi posteriormente guardado e mantido a -20°C até à sua utilização.

Em algumas amostras foi necessário efectuar, após a extracção, um processo adicional de precipitação diferencial de polissacáridos utilizando o protocolo desenvolvido por Michaels *et al.* (1994).

A avaliação da qualidade e quantidade de DNA extraído foi feita através de uma electroforese em gel de agarose a 0.8% corado com brometo de etídio (0.5 $\mu$ g/mL), a uma voltagem de 4 V/cm. A quantificação foi feita por comparação da espessura das bandas do DNA das amostras com aquelas obtidas com o marcador de peso molecular derivado do fago Lambda (Fermentas) usando diferentes concentrações conhecidas. A aquisição das imagens foi feita com o Gel-Doc 1000, utilizando o software Quantity One da Biorad. Diluiu-

se as amostras de DNA em TE pH8.0 (10mM de Tris pH8.0 + 1mM EDTA pH8.0) de forma a que todas ficassem com uma concentração aproximada de 50 ng/μL de DNA.

### 2.3. Delineamento Experimental

Para este trabalho foi escolhido um conjunto inicial de 31 pares de oligonucleótidos iniciadores de amplificação desenvolvidos previamente em *Olea europaea* L., *Fraxinus excelsior* e *Phillyrea angustifolia* (Tabela I, em anexo) devido, fundamentalmente, ao grande grau de polimorfismo apresentado, à não existência de alelos nulos e por apresentarem um poder de discriminação superior a 50% (Brachet *et al.* 1999; Carriero *et al.* 2002; Cipriani *et al.* 2002; Lefort *et al.* 1999; de la Rosa *et al.* 2002; Saumitou-Laprade *et al.* 2000 e Sefc *et al.* 2000).

Foi escolhido aleatoriamente um sub-grupo de 12 indivíduos para testar o grau de polimorfismo detectado pelo conjunto inicial de oligonucleótidos iniciadores de amplificação na população em estudo. Esta primeira aproximação permitiu-nos reduzir para 12 o número de *loci* avaliados na totalidade dos indivíduos com base no nível de polimorfismo detectado, facilidade de avaliar o tamanho dos fragmentos obtidos.

Para detectar o tamanho dos fragmentos dos indivíduos analisados utilizamos fazer uma electroforese por capilar utilizando o sequenciador automático de fluorescência CEQ8000 da Beckman & Coulter. Para tal foi necessário marcar a extremidade 5' do oligonucleótido iniciador de amplificação *forward* com um fluoróforo para que os fragmentos sejam detectados pelo aparelho. Existem vários fluoróforos possíveis sendo que se utilizou o D4 dado possuir um sinal de emissão mais forte em relação aos outros fluoróforos disponíveis para o sequenciador da Beckman & Coulter.

A reprodutibilidade dos resultados obtidos foi avaliada através da repetição da amplificação por reacção da polimerase em cadeia (PCR) e subsequente análise dos fragmentos de 10% da população estudada (Hoffman e Amos 2005), escolhida aleatoriamente para cada *locus*.

### 2.4. Condições de Amplificação

As reacções de amplificação foram feitas para um volume final de 25 μL contendo aproximadamente 50 ng de DNA, 1x do tampão de PCR (Invitrogen), 2.5 mM de MgCl<sub>2</sub>, 200 μM de dNTP's (Invitrogen), 100 pmol de cada oligonucleótidos iniciadores de amplificação (*reverse* e *forward*) (Proligo), 1 Unidade (U) de Taq polimerase (Invitrogen) e completou-se o volume com água bidestilada estéril. A reacção de amplificação por PCR realizou-se num termociclador T-Gradient da Biometra, programado para as seguintes condições de amplificação: 1º – pré-desnaturação das cadeias de DNA durante 5 minutos a 94°C; 2º – 30

ciclos de desnaturação do DNA a 94°C durante 45 segundos, seguido da ligação dos oligonucleótidos iniciadores de amplificação às sequências complementares de DNA à temperatura de ligação ao DNA de cada par de oligonucleótidos iniciadores de amplificação (Tabela I, em anexo) durante 45 segundos, e extensão dos oligonucleótidos iniciadores de amplificação a 72°C durante 1 minuto; 3º – extensão final a 72°C durante 10 minutos, pausa a 4°C. Confirmou-se a amplificação através da electroforese em gel de agarose 1.5% corado com brometo de etídio (0.5 µg/mL), a uma voltagem de 4 V/cm. A aquisição das imagens foi feita com o Gel-Doc 1000, utilizando o software Quantity One da Biorad.

## 2.5. Análise dos Fragmentos

Para se proceder à genotipagem dos indivíduos da população em estudo fez-se uma electroforese por capilar usando o sequenciador automático de fluorescência CEQ8000 da Beckman & Coulter seguindo o protocolo standardizado e usando a aplicação informática específica para a análise de fragmentos (CEQ 8000 DNA *Fragments Analysis* – Beckman & Coulter 2002a). Para cada amostra utilizou-se entre 0.5 e 2 µL do produto de amplificação, 24 µL de formamida deionizada (Sigma) e 0.5 µL do padrão de tamanho molecular de DNA (DNA Size Standard Kit - 400) da Beckman & Coulter que permite fazer a calibração dos tempos de detecção dos fragmentos em pares de bases (bp). As condições de corrida da electroforese por capilar foram as definidas pelo fabricante (Frag3) para o padrão de tamanho molecular utilizado (A-1876B: Highly Precise DNA Sizing on the CEQ™ Genetic Analysis System – Beckman & Coulter 2002b). O DNA migra nos capilares de acordo com o seu tamanho sendo depois detectado o seu tempo de migração dado que o fluoróforo ligado ao fragmento de DNA emite um sinal que é detectado pelo aparelho.

A aplicação informática associada ao sequenciador transforma os sinais detectados num electroferograma. A posição do pico ao longo do eixo dos xx corresponde ao tamanho do produto de amplificação medido em pares de bases (bp) e a sua intensidade (altura – eixo dos yy) corresponde à concentração do produto de amplificação da amostra injectada. O tamanho do fragmento detectado pelo sequenciador e aplicação informática associada reflecte também o tamanho do iniciador utilizado. Como a espécie em estudo é diplóide, no electroferograma podem estar representados dois picos (indivíduos heterozigóticos, isto é, com duas formas alélicas) ou um só pico (indivíduo homozigóticos, isto é, com apenas uma forma alélica).

Foi construída uma base de dados com os tamanhos dos alelos observados para cada *locus* nos indivíduos em estudo. A interpretação dos electroferogramas para a construção deste ficheiro foi sujeita a algumas directivas: 1) foram descartados os picos cuja altura era 80% inferior à altura do pico principal (Upadhyaya 2006); 2) devido ao erro associado ao aparelho ( $\pm 1$  bp) não foram considerados como sendo alelos diferentes fragmentos que

apenas diferiam em 1 par de base (A-1929A: Strategies for Automating the Review of Data – Beckman & Coulter 2002c). Podemos ver na Figura 2 um exemplo de um electroferograma gerado pelo sequenciador automático de fluorescência da Beckman & Coulter.

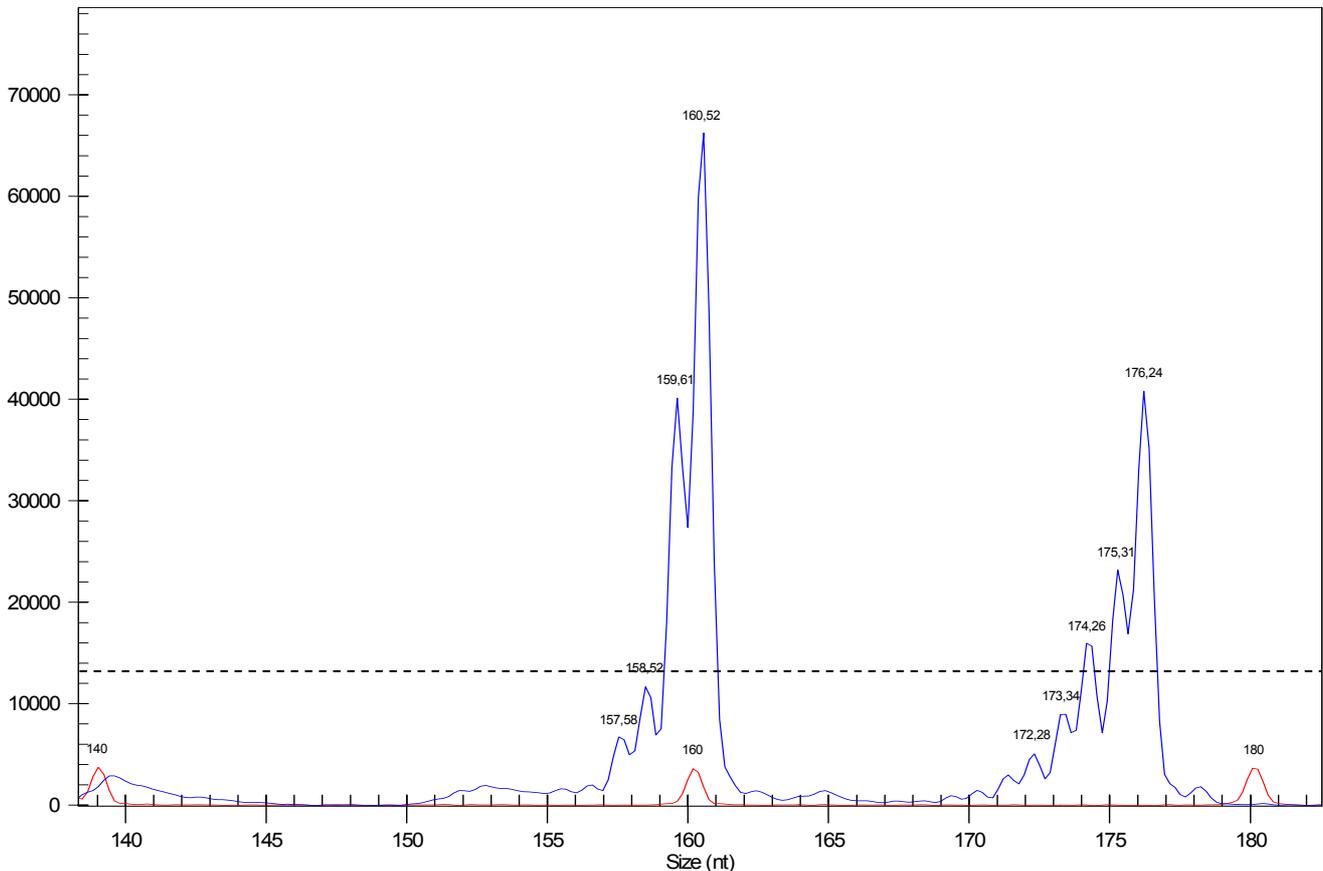


Figura 2 – Exemplo de um electroferograma gerado pelo sequenciador automático de fluorescência da Beckman & Coulter. O padrão de tamanho molecular de DNA está a vermelho; os fragmentos de DNA amplificado estão a azul. A linha a tracejado corresponde a 20% da intensidade do pico maior. Neste caso está representado um indivíduo heterozigótico (1º alelo – 160 bp; 2º alelo – 176 bp). Os picos que precedem os picos principais resultam do deslizamento da polimerase de DNA durante a amplificação.

## 2.6. Cálculo da probabilidade de erro e reprodutibilidade

O cálculo da probabilidade de erro e da reprodutibilidade dos resultados foram determinadas segundo Sneath e Johnson (1972) utilizando as seguintes fórmulas:

$$S_i^2 = n/2t \text{ (variância por locus);}$$

$$S^2 = 1/N \sum_i^N S_i^2 \text{ (variância total);}$$

$$p = \frac{1}{2} (1 - \sqrt{1 - 4S^2}) \text{ (probabilidade de erro);}$$

$$R = 1 - p \text{ (reprodutibilidade),}$$

em que  $n$  corresponde ao número de discrepâncias alélicas detectados para cada *locus*,  $t$  corresponde ao número de comparações feitas e  $N$  corresponde ao número total de *loci*.

Este cálculo serviu para verificar se os *loci* utilizados asseguravam um valor mínimo de reprodutibilidade de 85-90% de modo a minimizar as distorções (Sneath e Johnson 1972). Nas situações em que isso não aconteça será aconselhável ponderar a remoção dos *loci*

que apresentem os maiores níveis de erro associados de modo a melhorar o poder da estatística (Selkoe e Toonen 2006).

## 2.7. Cálculo dos parâmetros de diversidade e das relações genéticas

### 2.7.1. Distância genética

Nenhuma das medidas usadas mais frequentemente, Dice, Jaccard, coeficiente simples de divergência (ou a distância Euclidiana ao quadrado), é apropriada para indivíduos diplóides com marcadores codominantes com sejam os microsatélites (Kosman e Leonard 2005).

A distância entre os indivíduos foi calculada segundo a proporção de alelos partilhados (Bowcock *et al.* 1994) utilizando a aplicação informática Microsat (Minch *et al.* 1997). Esta distância pode ser utilizada para calcular a distância genética entre indivíduos ou entre populações (Chakraborty e Jin 1993).

A proporção de alelos partilhados foi calculada segundo a fórmula:

$$P_s = \frac{\sum_u S}{2u}$$

em que  $S$  é o número de alelos partilhados em todos os *loci* e  $u$  corresponde aos *loci*. E a distância genética entre indivíduos ( $D_{PS}$ ) foi estimada por:

$$D_{PS} = 1 - P_s$$

A determinação da verdadeira dissemelhança (distância) genética entre indivíduos é um ponto importante e decisivo para o agrupamento e análise da diversidade dentro de e entre as populações em estudo dado que a utilização de diferentes medidas de dissemelhança podem dar resultados diferentes (Kosman e Leonard 2005). A distribuição do número de alelos partilhados entre indivíduos é um bom indicador do grau de proximidade existente entre indivíduos (Chakraborty e Jin 1993).

### 2.7.2. Estrutura genética

A matriz de distâncias foi sujeita a uma análise de variância molecular (AMOVA) usando a aplicação informática GenAlEx (Peakall e Smouse 2006). Este procedimento permite a partição hierárquica da variação genética entre regiões e entre populações.

As árvores da cultivar 'Cobrançosa' e os zambujeiros constituíram as regiões a analisar e o local de origem de cada árvore constituíram as diferentes populações. Apenas foram utilizadas, por razões estatísticas, as populações que continham mais do que 3 indivíduos. Foi feita uma avaliação não paramétrica, através de 1000 permutações do nível de significância da AMOVA.

A AMOVA também permitiu analisar a estrutura hierárquica nas duas regiões consideradas, considerando a diversidade existente entre e dentro das diferentes

populações. O nível de significância da estatística foi testado, não-parametricamente, através de 1000 permutações. A estatística  $\Phi$ , é a análoga aos parâmetros medidos pela estatística  $F$  e da mesma forma, permite medir a diferenciação genética entre pares de populações, em que 1 corresponde a duas populações completamente distintas e 0 corresponde a duas populações indistintas. Este modelo estatístico pode ser interpretado como a correlação entre as frequências alélicas de diferentes indivíduos na mesma população, sob o modelo IAM (Weir e Cockerham 1984), e pensa-se que seja uma medida mais apropriada para populações que divergiram recentemente (Olsen e Schaal 2001).

### 2.7.3. Relações fenéticas

As relações fenéticas entre todos os indivíduos presentes no estudo foram avaliadas por meio de métodos hierárquicos e através da análise das coordenadas principais tendo por base uma matriz de distâncias genéticas, utilizando a aplicação informática NTSYSpc versão 2.11a (Rohlf 1998).

Dado que os métodos hierárquicos irão sempre constituir grupos, mesmo que eles não existam na realidade, é importante verificar por outros métodos se as relações encontradas pelos métodos hierárquicos são uma boa representação das verdadeiras distâncias genéticas entre os acessos. O procedimento mais simples será primeiro utilizar o módulo *SAHN* para constituir os grupos e depois usar o módulo *COPH* para, a partir da árvore produzida pelo módulo *SAHN*, produzir uma matriz de valores cofenéticos. O módulo *MXCOMP* é então utilizado para comparar a matriz dos valores cofenéticos com a matriz de distâncias. Também é útil proceder a uma análise das coordenadas principais de modo a verificar se estão representados os mesmos agrupamentos obtidos pelo módulo *SAHN*.

#### 2.7.3.1. Métodos hierárquicos

Partindo da matriz de distâncias, utilizamos o módulo *SAHN* para efectuar uma análise de grupos usando o algoritmo que procede aos agrupamentos através de médias aritméticas não ponderadas (UPGMA – Unweighted Pair Group Method using Arithmetic Average) (Sneath e Sokal 1973). Este método define a distância entre dois grupos como a média de todas as semelhanças entre os membros dos dois grupos, dois a dois (Sneath e Sokal 1973). De modo a testar a consistência dos agrupamentos formados, o fenograma obtido foi utilizado para obter uma matriz simétrica com os valores cofenéticos (módulo *COPH*) e depois comparou-se esta matriz com a matriz de distâncias genéticas usando para isso o módulo *MXCOMP*. Este módulo calcula o coeficiente de correlação cofenética entre as duas matrizes. A representação das distâncias genéticas entre os indivíduos dada pelo fenograma é muito boa se o coeficiente for superior ou igual a 0.90.

### 2.7.3.2. Análise das coordenadas principais

O módulo *Dcenter* foi utilizado para transformar a matriz de distâncias em produtos escalares para que se possa calcular os valores próprios e os vectores próprios utilizados na análise das coordenadas principais (Gower 1966). Durante essa transformação todos os elementos da matriz triangular ( $d_{ij}$ ), excepto os da diagonal principal, são substituídos por  $-1/2d_{ij}^2$  e depois a média de cada linha e a média de cada coluna é subtraída ao valor de cada elemento e por fim soma-se a média da matriz. Obteve-se assim a matriz centrada. Esta matriz, através do módulo *EIGEN*, foi utilizada no cálculo dos valores próprios e dos vectores próprios, que permitem a representação das coordenadas principais, em duas e em três dimensões. A consistência dos resultados foi avaliada pela percentagem da variância total explicada pelos primeiros vectores das coordenadas principais.

### 2.7.4. Diversidade genética

Os parâmetros de diversidade genética foram analisados recorrendo à aplicação informática Identity 1.0 (Wagner e Sefc 1999). Todas estas medidas estão dependentes do número de alelos identificados em cada *locus* e da sua frequência nos indivíduos analisados.

#### 2.7.4.1. Frequência alélica

A frequência alélica corresponde à quantidade de vezes que cada alelo está representado no conjunto dos indivíduos analisados. Consideram-se alelos raros todos os alelos cuja frequência seja igual ou menor a 5% (Belaj *et al.* 2007).

#### 2.7.4.2. Heterozigosidade

A heterozigosidade observada ( $H_o$ ) foi calculada para cada *locus* como a razão entre o número de indivíduos heterozigóticos e o número total de indivíduos analisados.

A heterozigosidade esperada ( $H_e$ ), também chamada de diversidade genética, foi calculada para cada *locus* segundo a fórmula (Nei 1987):  $H_e = 1 - \sum p_i^2$ , em que  $p_i$  é a frequência de cada um dos alelos.

#### 2.7.4.3. Poder de discriminação

O poder de discriminação (PD) para cada *locus* foi avaliado pela seguinte fórmula:  $PD = 1 - \sum g_i^2$ , em que  $g_i$  é a frequência de cada genótipo (Kloosterman *et al.* 1993). O valor obtido permite avaliar o poder de discriminação que cada *locus* tem na distinção dos indivíduos.

### 3. Resultados e Discussão

#### 3.1. Amplificação dos microssatélites identificados noutras espécies

Todos os 31 *loci* amplificados (Tabela I, em anexo) se mostraram polimórficos no sub-grupo de 12 árvores de teste utilizadas. Dos 12 *loci* escolhidos para avaliar todos os acessos, três foram descritos inicialmente noutras espécies que não em oliveira: os *loci* FEMSATL4 e M230aa, descritos inicialmente em *Fraxinus excelsior*, e o *locus* PA(GA)5, descrito inicialmente em *Phillyrea angustifolia*. Estes três *loci* já tinham sido utilizados com sucesso por Khadari e colaboradores (2003) em *Olea europaea*.

A escolha dos 12 *loci* teve como base o alto nível de polimorfismo detectado por este conjunto de *loci* e verificou-se que é suficiente utilizar um número reduzido de *loci* altamente polimórficos para discriminar as diferentes cultivares em estudo (Figura III e IV, em anexo).

#### 3.2. Reprodutibilidade

Dos 12 pares de iniciadores de amplificação usados nos 138 acessos, os que mostraram resultados mais fiáveis foram o GAPU71B, o UDO99-028, o EMO3 e o PA(GA)5, todos sem nenhuma discrepância entre a primeira leitura e a leitura obtida na réplica (Tabela 4). A probabilidade de erro associado à determinação do tamanho dos alelos foi de 16.01%, o que corresponde a uma reprodutibilidade de 83.99% (Tabela 4). Os *loci* que mais contribuíram para este valor foram o UDO99-019 (variância 0.286), e o FEMSATL4 (variância 0.308) (Tabela 4). A remoção desses dois *loci* faz com que a probabilidade de erro passa a ser de 11.54% e consequentemente a reprodutibilidade passa a ser de 88.46% (Tabela 5) o que, segundo Sneath e Johnson (1972), assegura que a representação das relações entre indivíduos não seja alterada mesmo na ausência de dados totalmente reprodutíveis (reprodutibilidade 100%). Por isso, e de forma a melhorar a precisão dos resultados, optou-se por retirar os dois *loci* com uma maior probabilidade de erro associada (*loci* UDO99-019 e FEMSATL4).

O erro experimental não terá tanto efeito no estabelecimento das relações fenéticas já que fazendo a comparação entre o fenograma com os 12 *loci* escolhidos inicialmente e entre o fenograma sem os 2 *loci* com uma maior probabilidade de erro associado se obteve um coeficiente de correlação cofenética de 0.99, o que demonstra a equivalência entre ambas as matrizes. No entanto, o mesmo já não se passa quando estamos a avaliar os parâmetros descritivos da diversidade genética que são mais sensíveis à incorrecta determinação dos tamanhos dos alelos.

Tabela 4 - Probabilidade de erro e reprodutibilidade associados à determinação do tamanho dos alelos.  $N_{DO}$  - número de discrepâncias observadas entre alelos;  $N_{TC}$  - número total de comparações alélicas feitas;  $S^2$  - variância por *locus*.

<i>Locus</i>	$N_{DO}$	$N_{TC}$	$S^2$
GAPU101	5	24	0.208
GAPU89	6	28	0.214
GAPU71B	0	26	0.000
UDO99-015	3	22	0.136
<b>UDO99-019</b>	8	28	<b>0.286</b>
UDO99-028	0	26	0.000
EMO3	0	24	0.000
ssrOeUA-DCA15	1	26	0.038
ssrOeUA-DCA16	7	26	0.269
<b>FEMSATL4</b>	8	26	<b>0.308</b>
PA(GA)5	0	22	0.000
M230a	4	26	0.154
<b>TOTAL</b>	<b>42</b>	<b>304</b>	<b>Média 0.134</b>
Probabilidade de erro (%)			16.01
Reprodutibilidade (%)			83.99

Tabela 5 - Erro e reprodutibilidade associados à determinação do tamanho dos alelos, sem os *loci* UDO99-019 e FEMSATL4.  $N_{DO}$  - número de discrepâncias observadas entre alelos;  $N_{TC}$  - número total de comparações alélicas feitas;  $S^2$  - variância por *locus*.

<i>Locus</i>	$N_{DO}$	$N_{TC}$	$S^2$
GAPU101	5	24	0.208
GAPU89	6	28	0.214
GAPU71B	0	26	0.000
UDO99-015	3	22	0.136
UDO99-028	0	26	0.000
EMO3	0	24	0.000
ssrOeUA-DCA15	1	26	0.038
ssrOeUA-DCA16	7	26	0.269
PA(GA)5	0	22	0
M230a	4	26	0.154
<b>TOTAL</b>	<b>42</b>	<b>304</b>	<b>Média 0.102</b>
Probabilidade de erro (%)			11.54
Reprodutibilidade (%)			89.56

### 3.3. Estrutura genética inter- e intrapopulacional

A AMOVA, baseada na matriz de distâncias calculada a partir de *1-proporção de alelos partilhados* (Dps), demonstrou que a maior percentagem de variação genética se deve a diferenças entre os indivíduos dentro das populações (57%), e que as diferenças entre as duas variedades consideradas ('Cobrançosas' vs. zambujeiros) explica 41% da variação genética observada, sendo a restante variação genética (2%) devida a diferenças entre as várias populações das duas variedades ( $p$  0.031). Não se detectou diferenciação entre as populações de 'Cobrançosa' (Tabela 6). Esta observação poderá ser explicada pelo facto de os 77 acessos desta cultivar provirem de uma área com um raio inferior a 40 Km, o que poderá ter levado à troca de material entre os agricultores, diminuindo assim a possibilidade de se estabelecerem diferenças significativas entre o local de origem das diferentes árvores. Ao contrário, entre as populações de zambujeiro obtiveram-se diferenças significativas entre as populações de Alcaria e Aguda ( $\Phi_{st}$  0.127,  $p$  0.003), entre as populações de Alcaria e S.Olaia ( $\Phi_{st}$  0.162,  $p$  0.012), Alcaria e Ferestelo ( $\Phi_{st}$  0.086,  $p$  0.045) e as populações de S.Olaia e S.Arrábida ( $\Phi_{st}$  0.100,  $p$  0.019) (Tabela 6). Um nível de diferenciação semelhante entre populações de zambujeiro foi descrito recentemente por Belaj *et al.* (2007) entre populações de zambujeiro espanholas e italianas ( $\Phi_{st}$  0.082,  $p$  0.001). Lumaret e colaboradores (2004) demonstraram, utilizando isozimas, que as populações de zambujeiro

do Mediterrâneo Ocidental são geneticamente mais diversas do que as populações do Mediterrâneo Oriental. Assim, o facto de os zambujeiros utilizados neste estudo provirem de um *pool* genético tão diverso poderá explicar o facto de se ter detectado uma ligeira diferenciação entre as populações, mesmo a nível nacional.

Tabela 6 – Diferenciação das diferentes populações baseada na estatística  $\phi$  (1 corresponde a duas populações completamente distintas e 0 corresponde a duas populações indiferenciadas). Nível de significância estatística: \*\*\*  $p = 0.001$ ; \*\*  $0.001 < p < 0.01$ ; \*  $0.01 < p < 0.05$ . (A → Alfândega da Fé; B → Foz Côa; C → Macedo de Cavaleiros; D → Mirandela; E → Mogadouro; F → Valpaços; G → Aguda; H → Alcaria; I → S. Olaia; J → Magoito; K → Ferestelo; L → S. Arrábida).

		'Cobrançosa'					Zambujeiro						
		A	B	C	D	E	F	G	H	I	J	K	L
'Cobrançosa'	A	0.000											
	B	0.000	0.000										
	C	0.000	0.000	0.000									
	D	0.000	0.000	0.000	0.000								
	E	0.000	0.030	0.034	0.013	0.000							
	F	0.000	0.000	0.000	0.000	0.064	0.000						
Zambujeiro	G	<b>0.287*</b>	<b>0.198*</b>	<b>0.190*</b>	<b>0.292*</b>	<b>0.307*</b>	0.101	0.000					
	H	<b>0.535***</b>	<b>0.462***</b>	<b>0.451***</b>	<b>0.545***</b>	<b>0.549***</b>	<b>0.413***</b>	<b>0.127**</b>	0.000				
	I	<b>0.617***</b>	<b>0.519**</b>	<b>0.472**</b>	<b>0.586***</b>	<b>0.659*</b>	<b>0.420*</b>	0.063	<b>0.162*</b>	0.000			
	J	<b>0.579**</b>	<b>0.478**</b>	<b>0.436**</b>	<b>0.556**</b>	<b>0.618*</b>	<b>0.386*</b>	0.000	0.020	0.088	0.000		
	K	<b>0.576**</b>	<b>0.479**</b>	<b>0.423**</b>	<b>0.542**</b>	<b>0.627*</b>	<b>0.380*</b>	0.000	<b>0.087*</b>	0.094	0.000	0.000	
	L	<b>0.427***</b>	<b>0.354***</b>	<b>0.375***</b>	<b>0.455***</b>	<b>0.431***</b>	<b>0.311***</b>	0.047	0.000	<b>0.100*</b>	0.000	0.000	0.000

### 3.4. Nível de polimorfismo detectado e diversidade genética em *Olea europaea*

#### L.

Os 12 *loci* utilizados para avaliar os 138 acessos utilizados neste estudo eram todos polimórficos, sendo que o *locus* com um menor número de alelos detectados foi o PA(GA)5, com quatro alelos, e no extremo oposto, o *locus* GAPI101, com 26 alelos (Tabela 7).

Um grande nível de polimorfismo foi detectado, no trabalho de Bandelj *et al.* (2004), entre as 19 cultivares com 14 marcadores SSR. O número médio de alelos detectados por *locus* foi de 6.8, valor este que é inferior ao encontrado por Sefc *et al.* (2000) para o mesmo conjunto de *loci*. A análise deste grupo incluía um grande número de cultivares (47 no total, 9 italianas e 38 da Península Ibérica) o que explica o maior número médio de alelos detectados (8.3). Uma grande variabilidade dos *loci* associados a microssatélites de oliveira também foram demonstrados em estudos de Carriero *et al.* (2002), com uma média de 5.7

alelos detectados em 10 *locus* em 20 cultivares e por Rallo *et al.* (2000), que desenvolveram 5 marcadores SSR que testaram em 48 cultivares, obtendo um número médio de 5.2 alelos por *locus*. Neste trabalho, a média de alelos detectados foi de  $11.2 \pm 6.2$  e no total foram detectados 112 alelos nos 138 acessos (Tabela 7), valor este superior ao obtido pelos autores anteriormente referidos. Este facto está relacionado com a inclusão dos zambujeiros, que são geneticamente mais diversos que as oliveiras cultivadas. Belaj *et al.* (2007) utilizando 8 marcadores SSR para caracterizar molecularmente diversas populações de zambujeiro do Mediterrâneo Ocidental detectaram, em média, 18.5 alelos por *locus*.

Os valores de  $H_e$  obtidos nos 12 *loci* analisados compreendiam valores entre 0.449 (UDO99-015) e 0.866 (GAPU101) (Tabela 7), com um valor médio de  $0.676 \pm 0.110$ . Na generalidade, os microssatélites desenvolvidos em oliveira parecem ser caracterizados por níveis médios de heterozigidade, isto é, nunca superiores a 0.66 (Rallo *et al.* 2000, Cipriani *et al.* 2002, de la Rosa *et al.* 2002, Díaz *et al.* 2006). Neste trabalho, o valor obtido também é muito próximo desse valor médio de heterozigidade ( $H_e$  0.676).

Os *loci* UDO99-015, UDO99-028 e *ssrOeUA-DCA16* apresentaram um valor de heterozigidade inferior ao esperado (Tabela 7), segundo o equilíbrio de Hardy-Weinberg. Tal pode ser em parte explicado pelo facto de a maioria dos acessos considerados pertencerem a cultivares que foram sujeitas a um processo de domesticação, sendo a diferença entre  $H_o$  e  $H_e$  o reflexo da selecção feita pelo homem de determinados genótipos.

Tabela 7 – *Loc*i analisados nos 138 acessos, tamanho dos fragmentos amplificados e parâmetros descritivos. ( $H_o$  heterozigidade observada;  $H_e$  heterozigidade esperada; PD poder de discriminação)

Locus	Tamanho fragmentos (bp)	Nº de alelos	$H_o$	$H_e$	PD (%)
GAPU101	84–220	26	0.944	0.866	83.8
GAPU89	158–207	9	0.826	0.633	69.8
GAPU71B	114–144	10	0.948	0.718	67.9
UDO99-015	92–141	12	0.206	0.449	54.1
UDO99-028	107–174	14	0.536	0.677	78.0
EMO3	204–216	6	0.625	0.693	68.3
<i>ssrOeUA-DCA15</i>	233–268	6	0.919	0.701	70.9
<i>ssrOeUA-DCA16</i>	123–188	14	0.382	0.610	67.6
PA(GA)5	113–127	4	0.863	0.638	51.4
M230a	196–236	11	0.816	0.779	79.6

Todos os 12 *loci* apresentaram na generalidade um bom poder de discriminação, com valores compreendidos entre 51.4%, no *locus* PA(GA)5, e 83.8%, no *locus* GAPU101 (Tabela 7). O valor médio do poder de discriminação do grupo de *loci* analisados foi de  $69.14 \pm 10.28$ . Vários autores (Cipriani *et al.* 2002, Díaz *et al.* 2006) verificaram que o poder de discriminação de cada locus estava correlacionado com a diversidade genética ( $H_e$ ) e que a simples análise do valor da heterozigidade esperada seria um bom indicador da

capacidade de discriminação de cada *locus*. No entanto, os dados obtidos neste estudo não permitem afirmar tão veementemente esta afirmação já que se observou uma fraca correlação entre o poder de discriminação de cada *locus* e  $H_e$  ( $R^2$  0.621) (Figura VII, em anexo). Assim, será conveniente avaliar sempre ambos os parâmetros.

### 3.5. Diversidade genética na cultivar ‘Cobrançosa’

Em relação aos 77 acessos da cultivar ‘Cobrançosa’ analisados, foram detectados 54 alelos, dos quais 26 correspondiam a alelos raros (frequência < 5%) (Tabela II, em anexo). Comparativamente com os valores obtidos na totalidade dos acessos analisados, houve uma redução substancial no número de alelos detectados por *locus* nos acessos de ‘Cobrançosa’ (Tabelas 7 e 8). Em ‘Cobrançosa’ foram detectados em média  $5.4 \pm 1.7$  alelos por *locus*, sendo que, em média,  $2.6 \pm 1.3$  tinham uma frequência inferior a 5% o que demonstra a capacidade deste tipo de marcadores moleculares na identificação de alelos raros e que poderão ser de extrema utilidade se à presença desses alelos estiverem associados determinadas características de valor agronómico. O nível de diversidade genética decresceu em praticamente todos os *loci*, especialmente nos *loci* UDO99-028 e *ssrOeUA-DCA16* (Tabelas 7 e 8) e o valor médio de diversidade genética foi de  $0.511 \pm 0.203$  o que permite perspectivar um processo de melhoramento de modo a fazer face às exigências do mercado do azeite.

Tabela 8 – Parâmetros descritivos da diversidade genética em ‘Cobrançosa’. ( $H_e$  heterozigidade esperada ou diversidade genética).

Locus	Nº de alelos	Nº de alelos raros	$H_e$
GAPU101	8	3	0.741
GAPU89	5	2	0.597
GAPU71B	7	4	0.620
UDO99-015	8	5	0.312
UDO99-028	4	2	0.240
EMO3	4	1	0.658
<i>ssrOeUA-DCA15</i>	4	1	0.617
<i>ssrOeUA-DCA16</i>	4	3	0.139
PA(GA)5	4	2	0.568
M230	6	3	0.617
Média $\pm$ Desvio padrão	$5.4 \pm 1.7$	$2.6 \pm 1.3$	$0.511 \pm 0.203$

Existem dois grupos de ‘Cobrançosas’ com genótipos muito idênticos ( $Dps$  0.036), englobando a maioria dos acessos analisados nesta cultivar (58.44% dos acessos de ‘Cobrançosa’) (Figura I, em anexo) e que na análise das coordenadas principais aparece no mesmo agrupamento (grupo A na Figura II, em anexo). Este grau de diferenciação entre árvores da mesma cultivar poderá ser explicado pela ocorrência de mutações somáticas que ocorrem em árvores com um tempo de vida tão extenso e que são propagadas vegetativamente, como sugerido por Cipriani e colaboradores (2002). No grupo A da análise

das coordenadas principais (Figura II, em anexo) estão associadas outras árvores desta cultivar cuja distância genética se encontra entre 0.036 e 0.250 (Figura I, em anexo). Na mesma figura podemos verificar que existe um segundo grupo de acessos da cultivar 'Cobrançosa' que é geneticamente mais distante do primeiro (Dps 0.504) e geneticamente mais diverso, aparecendo na análise das coordenadas principais como um grupo isolado do primeiro (grupo B na Figura II, em anexo). A ocorrência deste segundo grupo de acessos da cultivar 'Cobrançosa', geneticamente mais distante e mais diverso, pode ser explicado pela ocorrência de fenómenos de reprodução sexuada e, como as oliveiras são propagadas vegetativamente, é razoável assumir que parte da variabilidade genética detectada corresponde a uma herança antiga (Gemas et al. 2004), o que implica que os indivíduos desta cultivar foram seleccionados a partir de diferentes clones. A incorporação de indivíduos originados por propagação sexuada no germoplasma usado para propagação tem sido considerada como o principal responsável da diversidade genética detectada nas cultivares utilizadas nos sistemas tradicionais de cultivo (Cipriani et al. 2002).

Nos grupos A e B (Figura II, em anexo) estão representados 92.21% dos acessos analisados da cultivar 'Cobrançosa', encontrando-se os restantes 7.79% dispersos entre as diferentes cultivares (Figura IV, em anexo), a uma distância genética média de  $0.653 \pm 0.013$  da maioria dos acessos de 'Cobrançosa'. Nomeadamente, 4 acessos que formam um pequeno grupo à parte e dois acessos que se encontram misturados com os zambujeiros (Figura VI, em anexo). Estas sete árvores constituirão possivelmente acessos mal identificados como pertencendo à cultivar 'Cobrançosa'. Aliás, situações de incorrecta identificação já foram descritas anteriormente por diversos grupos (Cipriani et al. 2002, Lopes et al. 2004) e pode ser explicado pelo facto de muitas das cultivares serem muito semelhantes morfologicamente.

### 3.6. Relações fenéticas

A utilização dos 10 *loci* presentes na Tabela 5, através de uma análise hierárquica utilizando o método UPGMA, permite distinguir, na sua quase totalidade, as árvores da cultivar 'Cobrançosa' das restantes oliveiras estudadas (Figura I, em anexo). Somente o acesso da cultivar 'Verdeal de Trás-os-Montes' parece agrupar-se conjuntamente com as 'Cobrançosas' (grupo B na Figura II, em anexo).

Os acessos pertencentes às cultivares 'Galega', 'Conserva de Elvas', 'Koroneiki', 'Arbequina' e 'Arbosana' encontram-se agrupados de acordo com a cultivar a que pertencem (Figuras III e IV, em anexo). De todos os acessos pertencentes a *Olea europaea* subsp. *europaea* var. *europaea*, os que são geneticamente mais diferentes dos acessos de 'Cobrançosa' são os indivíduos pertencentes às cultivares 'Galega', 'Conserva de Elvas' e 'Koroneiki' (Figuras III e IV, em anexo). Vários autores (Díaz et al. 2006, Gemas et al. 2004)

já tinham verificado, usando diferentes sistemas de marcadores moleculares, que a 'Galega' é uma cultivar feneticamente distinta das restantes cultivares analisadas. Gemas *et al.* (2004) sugerem que tal é devido ao facto desta cultivar ter uma origem diferente das outras cultivares portuguesas analisadas. A posição da cultivar 'Koroneiki' pode ser explicada simplesmente pela maior distância geográfica em relação ao seu local de origem.

Tal como descrito por Díaz e colaboradores (2006), os acessos da cultivar 'Cobrançosa' parecem ficar preferencialmente agrupados com as cultivares Espanholas (Figuras III e IV, em anexo).

Exceptuando a cultivar 'Cobrançosa', em todas as outras cultivares foram apenas avaliados um número reduzido de indivíduos mas mesmo assim foi detectada alguma variabilidade genética (cultivares com mais do que um acesso) (Figura IV, em anexo) o que demonstra a capacidade destes marcadores moleculares na detecção de variabilidade intra-cultivar. A grande diversidade encontrada entre diferentes cultivares de oliveira deve-se possivelmente às diferentes origens do germoplasma, que resulta provavelmente de cruzamentos entre oliveiras cultivadas e oliveiras selvagens (zambujeiros), originando novas cultivares em diferentes partes do Mediterrâneo, e de uma fraca pressão selectiva (Besnard *et al.* 2001a, Contento *et al.* 2002). E, como foi adiantado por Martins-Lopes e colaboradores (2007), tal pode explicar o facto de não se observar uma clara distinção entre cultivares Portuguesas e estrangeiras. O mesmo foi observado por Khadari *et al.* (2003), usando cultivares francesas e estrangeiras, tal como por Besnard *et al.* (2001a) que usando 102 génotipos de diversos países do Mediterrâneo obtiveram resultados semelhantes.

Tal como já tinha sido verificado no trabalho de Gemas *et al.* (2004), utilizando marcadores RAPD, a cultivar 'Negrinha' apareceu como sendo geneticamente muito próxima da 'Azeiteira' (Figuras III e IV, em anexo), não se confirmando assim o resultado obtido por Lopes *et al.* (2004), usando marcadores SSR, em que a cultivar 'Cobrançosa' também se agrupava com estas duas cultivares. O facto de que a 'Azeiteira' e a 'Negrinha' serem as únicas cultivares, presentes neste estudo, utilizadas exclusivamente para a produção de azeitonas de mesa (Bartolini *et al.* 2005) pode explicar as semelhanças genéticas entre ambas.

Pela análise do fenograma e da representação em 3 dimensões da análise das coordenadas principais dos acessos de 'Cobrançosa' e de zambujeiros (Figuras V e VI, em anexo) podemos verificar que a utilização do grupo de 10 *loci* permite estabelecer uma clara distinção entre 'Cobrançosas' e zambujeiros, embora hajam alguns acessos da cultivar que se encontram dissociados do grupo 'Cobrançosas'. Este resultado seria expectável pois a análise da variância molecular mostrou que, embora a maior percentagem de variação genética fosse explicada por diferenças entre os indivíduos dentro das populações (57%), 41% da variação genética observada se devia às diferenças entre 'Cobrançosas' e

zambujeiros, que correspondiam à maior parte dos acessos avaliados. Apenas alguns zambujeiros aparecem misturados com as cultivares (Figura V, em anexo). Poderão ser estas árvores formas ferais resultantes do cruzamento entre árvores cultivadas e selvagens? Geralmente a ocorrência de formas ferais resulta de cruzamentos entre oliveiras selvagens e cultivadas ou do desenvolvimento das sementes de oliveiras cultivadas (Zohary D e Hopf M 2000). Como os zambujeiros analisados provêm de populações localizadas em Parques Naturais é admissível pensar que estas possíveis formas ferais sejam o produto do cruzamento entre indivíduos cultivados e selvagens. De outro modo, é razoável pensar que a presença de alelos comuns em zambujeiros e árvores de diferentes cultivares seja o resultado da existência de antepassados comuns. Tal estaria de acordo com algumas perspectivas que consideram que as cultivares autóctones resultam do cruzamento de cultivares importadas com zambujeiros locais e posterior selecção (Besnard et al. 2001a, Contento et al. 2002).

## 4. Conclusões e Perspectivas

Concluindo, a análise da partição hierárquica da variação genética (AMOVA) não detectou diferenciação entre as populações de 'Cobrançosa'. Ao contrário, entre as populações de zambujeiro obtiveram-se diferenças significativas entre as seis populações avaliadas.

Foi estabelecida uma matriz de distância genética (*1-proporção de alelos partilhados*) e procedeu-se à análise de grupos usando o algoritmo que estabelece os agrupamentos com base nas médias aritméticas não ponderadas (UPGMA) e à análise das coordenadas principais. Embora a maior parte das árvores da cultivar 'Cobrançosa' seja geneticamente muito semelhante, verificou-se que esta cultivar parece ter uma origem policlonal. Além disso, o valor médio de diversidade genética detectado nesta cultivar ( $H_e$  0.511) permite perspectivar um processo de melhoramento para fazer face às exigências do mercado do azeite. A detecção de alelos raros demonstra a capacidade dos marcadores moleculares SSR na identificação de alelos cuja frequência é inferior a 5% e que poderão ser de extrema utilidade se à sua presença estiverem associadas determinadas características de valor agronómico. Foi possível desenvolver padrões de microssatélites que permitem a identificação de árvores da cultivar 'Cobrançosa' e inclusive detectar sete árvores que constituirão possivelmente acessos mal identificados como pertencendo a essa cultivar.

Apesar da sua importância económica em todos os países do Mediterrâneo pouco tem sido feito no que diz respeito ao melhoramento da oliveira apesar da necessidade de se melhorar a produtividade e as qualidades agronómicas. A maioria dos programas de selecção tem-se baseado na selecção clonal. A exploração da variabilidade fenotípica tem permitido a identificação de clones valiosos dentro de inúmeras cultivares.

Experiências relacionadas com a transformação genética estão em progresso com o objectivo de seleccionar cultivares resistentes a doenças ou de introduzir genes envolvidos em vias metabólicas importantes (revisto por Doveri e Baldoni 2007).

Dado que a maioria das características desejáveis são quantitativas, a localização de QTL (*quantitative trait loci*) parece ser a estratégia adequada a seguir. Isto implica o desenvolvimento de progénies adequadas usando linhas parentais com fenótipos contrastantes e suficientemente heterozigóticas.

Os objectivos elementares dos programas de melhoramento das oliveiras tem sido dirigidos no sentido de ultrapassar os factores que limitam a produtividade. Estes incluem: encurtamento da fase juvenil, aumento do tamanho e número de frutos, aumento do conteúdo e qualidade do azeite (composição em ácidos gordos, conteúdo em polifenóis), produtividade constante, diminuição da altura das árvores e outras manipulações na

arquitectura das árvores de modo a facilitar a poda e colheita e diminuir os custos de cultivo, aumentar a resistência a pestes (particularmente a mosca da oliveira, *Bactrocera oleae*) e a doenças (*Spilocaea oleagina*, *Verticillium dahliae*, *Pseudomonas savastanoi*). Outros objectivos importantes estão relacionados com o melhoramento da tolerância ao frio (de modo a permitir o cultivo de oliveira em zonas mais a norte). Embora as oliveiras sejam normalmente consideradas como uma espécie tolerante à seca, a sua produtividade é fortemente afectada sob condições de secura, por isso há interesse na pesquisa de cultivares tolerantes. Um outro aspecto que também pode lucrar com o estudo genético das cultivares é a pesquisa de porta-enxertos com uma maior tolerância ao stress abiótico e biótico.

## Referências bibliográficas

- Angiolillo A, Mencuccini M e Baldoni L. (1999). Olive genetic diversity assessed using amplified fragment length polymorphisms. *Theor Appl Genet* **98**:411–421.
- Bandelj D, Jakše J e Javornik B (2004). Assessment of genetic variability of olive varieties by microsatellite and AFLP markers. *Euphytica* **136**:93–102.
- Bartolini G, Prevost G, Messeri C e Carignani G (2005). Olive Germplasm: Cultivars and World-Wide Collections. FAO, Rome, Italy.
- Bartolini G, Prevost G, Messeri C, Carignani G e Menini UG (1998). Olive germplasm. Cultivars and World-Wide collections. FAO, Rome, Italy.
- Beckman & Coulter (2002a). CEQ™8000 Genetic Analysis System User's Guide. Disponível em <www.beckmancoulter.com>, a 10 de Novembro, 2007.
- Beckman & Coulter (2002b). A-1876B: Highly Precise DNA Sizing on the CEQ™ Genetic Analysis System. Disponível em <www.beckmancoulter.com>, a 10 de Novembro, 2007.
- Beckman & Coulter (2002c). A-1929A: Strategies for Automating the Review of Data. Disponível em <www.beckmancoulter.com>, a 10 de Novembro, 2007.
- Belaj A, Muñoz-Diez C, Baldoni L, Porceddu A, Barranco D e Satovic Z (2007). Genetic diversity and population structure of wild olives from the North-western Mediterranean by SSR markers. *Annals of Botany* 1-10.
- Belaj A, Trujillo I, De la Rosa R e Rallo L (2001). Polymorphism and discrimination capacity of randomly amplified polymorphic markers in an olive germplasm bank. *J Am Soc Hortic Sci* **126**:64–71.
- Besnard G, Baradat P e Bervillé A. (2001a). Genetic relationships in the olive (*Olea europaea* L.) reflect multilocal selection of cultivars. *Theor Appl Genet* **102**:251–258.
- Besnard G, Baradat P, Chevalier D, Tagmout A e Bervillé A (2001b). Genetic differentiation in the olive complex (*Olea europaea* L.) revealed by RAPDs and RFLPs in the rRNA genes. *Genet Resour Crop Evol* **48**:165-182.
- Besnard G, Khadari B, Baradat P e Bervillé A (2002). *Olea europaea* L. (Oleaceae) phylogeography based on chloroplast DNA polymorphism. *Theor Appl Genet* **104**:1353-1361.
- Bouhadida M, Casas AM, Moreno MA e Gogorcena Y (2007). Molecular characterization of Miraflores peach variety and relatives using SSRs. *Scientia Horticulturae* **111**:140-145.
- Bowcock AM, Ruíz-Linares A, Tomfohrde J, Minch E, Kidd JR e Cavalli-Sforza LL (1994). High resolution human evolutionary trees with polymorphic microsatellites. *Nature* **368**:455-457.

- Brachet S, Jubier MF, Richard M, Jung-Muller B e Frascaria-Lacoste N (1999). Rapid identification of microsatellite loci using 50 anchored PCR in the common ash *Fraxinus excelsior*. *Mol Ecol* **8**:160–163.
- Carriero F, Fontanazza G, Cellini F e Giorio G (2002). Identification of simple sequence repeats (SSRs) in olive (*Olea europaea* L.). *Theor. Appl. Genet.* **104**:301–307.
- Chakraborty R e Jin L (1993). A unified approach to study hypervariable polymorphisms: Statistical considerations of determining relatedness and population distances. In: DNA fingerprinting: State of the science. Pena SDJ, Chakraborty R, Epplen JT e Jeffreys AJ eds. Birkhauser Verlag Basel/Switzerland.
- Cipriani G, Marrazzo MT, Marconi R, Cimato A e Testolin R (2002). Microsatellite markers isolated in olive (*Olea europaea* L.) are suitable for individual fingerprinting and reveal polymorphism within ancient cultivars. *Theor Appl Genet* **104**:223–228.
- Codex Alimentarius Commission (1993). Report of the Fourteenth Session of the Codex Committee on Fats and Oils, London, Sept 27–Oct 1, 1993, Alinorm 95/17, Rome FAO/WHO, Italy.
- Contento A, Ceccarelli M, Gelati MT, Maggini F, Baldoni L e Cionini PG (2002). Diversity of *Olea* genotypes and the origin of cultivated olives. *Theor Appl Genet* **104**:1229–1238.
- de la Rosa R, James CM e Tobutt KR (2002). Isolation and characterization of polymorphic SSRs in olive (*Olea europaea* L.) and their transferability to other genera in the Oleaceae. *Mol Ecol Notes* **2**:265–267.
- de la Rosa R, Angiolillo A, Rallo L, Guerrero C, Pellegrini M, Besnard G, Bervillé A, Martin A e Baldoni L (2003). A first genetic linkage map of olive (*Olea europaea* L.) cultivars using RAPD and AFLP markers. *Theor Appl Genet* **106**:1273–1282.
- Díaz A, De la Rosa R, Martín A e Rallo P (2006). Development, characterization and inheritance of new microsatellites in olive (*Olea europaea* L.) and evaluation of their usefulness in cultivar identification and genetic relationship studies. *Tree Genetics & Genomes* **2**:165–175.
- Doveri S e Baldoni L (2007). Olive. In: Fruit and Nuts. Kole C (ed). Springer-Verlag, Berlin Heidelberg, pp 254-264.
- Doyle JJ e Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Focus* **12**:13–15.
- Durman, RE e Korban SS (1994). Evidence of gene introgression in apple using RAPD markers. *Euphytica* **79**:109–114.
- Eisen JA (1999). Mechanistic basis for microsatellite instability. In: Microsatellites: Evolution and Applications (eds Goldstein, D.B. e Schlötterer, C.). Oxford University Press, Oxford, UK, pp. 34–48.

- Falisticco E e Tosti N (1996). Cytogenetic investigation in *Olea europaea* L. *J Genet Breed* **50**:235-238.
- Ganino T, Beghè D, Valenti S, Nisi R e Fabbri A (2007). RAPD and SSR markers for characterization and identification of ancient cultivars of *Olea europaea* L. in the Emilia region, Northern Italy. *Genet Resour Crop Evol* **54**:1531-1540.
- Gemas VJV , Almadanim MC, Tenreiro R, Martins A e Fevereiro P (2004). Genetic diversity in the Olive tree (*Olea europaea* L. subsp. *europaea*) cultivated in Portugal revealed by RAPD and ISSR markers. *Genetic Resources and Crop Evolution* **51**:501–511.
- Goldstein D B e Clark AG (1995). Microsatellite variation in North American populations of *Drosophila melanogaster*. *Nucleic Acids Res.* **23**:3882-3886.
- Gower J (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* **53**:325-338.
- Green PS (2002). A revision of *Olea* L. *Kew Bull* **57**:91-140
- Green PS e Wickens GE (1989). The *Olea europaea* complex. In: Tan K (ed) *The Davis and Hedge Festschrift*. Edinburgh Univ press, UK, pp 287-299.
- Gupta PK e Varshney RK (2000). The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. *Euphytica* **113**:163–185.
- Hammer K, Arrowsmith N e Gladis T (2003). Agrobiodiversity with emphasis on plant genetic resources. *Naturwissenschaften* **90**:241–250.
- Hardy OJ, Charbonnel N, Fréville H e Heuertz M (2003). Microsatellite Allele Sizes: A Simple Test to Assess Their Significance on Genetic Differentiation. *Genetics* **163**:1467-1482.
- Hoffman JI e Amos W (2005). Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology* **14**:599–612.
- International Olive Oil Council (1993). International trade standard applying to olive oils and olive-pomace oils. COI/T. 15/NC, No. 1/Rev. 6, Madrid, Spain.
- Johnson LAS (1957). A review of the family *Oleaceae*. *Contr New South Wales Nat Herbar* **2**:397-418.
- Khadari B, Breton C, Moutier N, Roger JP, Besnard G, Bervillé A e Dosba F (2003). The use of molecular markers for germoplasm management in a French olive collection. *Theor Appl Genet* **106**:521-529.
- Kloosterman AD, Budowle B e Daselaar P (1993). PCR amplification and detection of the human D1S80 VNTR locus. Amplification conditions, population genetics and application forensic analysis. *Legal Medicine* **105**:257-264.

- Kosman E e Leonard KJ (2005). Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species. *Molecular Ecology* **14**:415-424.
- Landry PA, Koskinen MT e Primmer CR (2002). Deriving evolutionary relationships among populations using microsatellites and  $\Delta\mu^2$ : all loci are equal, but some are more equal than others. *Genetics* **161**:1339–1347.
- Lefort F, Brachet S, Frascaria-Lacoste N, Edwards KJ e Douglas GC (1999). Identification and characterization of microsatellite loci in ash (*Fraxinus excelsior* L.) and their conservation in the olive family (Oleaceae). *Mol Ecol* **8**:1075–1092.
- Leitão F, Pontes MF, Calado ML e Almeida FJ (1986). Descrição de 22 variedades de oliveira cultivadas em Portugal. Ministério da Agricultura, Pescas e Alimentação, Direcção Geral do Planeamento e Agricultura, Lisboa.
- Lopes MS, Mendonça D, Sefc KM, Gil FS e da Câmara Machado A (2004). Genetic evidence of Intra-cultivar Variability within Iberian Olive Cultivars. *HortScience* **39** (7):1562-1565.
- Loureiro J, Rodriguez E, Costa A e Santos C (2007). Nuclear DNA content estimations in wild olive (*Olea europaea* L. ssp. *europaea* var. *sylvestris* Brot.) and Portuguese cultivars of *O. europaea* using flow cytometry. *Genet Resour Crop Evol* **54**:21–25.
- Lumaret R, Ouazzani N, Michuad H, Vivier G, Deguilloux M-F e Di Giusto F (2004). Allozyme variation of oleaster populations (wild olive tree) (*Olea europaea* L.) in the Mediterranean basin. *Heridity* **92**:343-351.
- Martins A, Santos L, Lopes J e Gouveia J (1998). Primeiros resultados da selecção da variedade de oliveira Cobrançosa. *Revista de Ciências Agrárias* **21**:35-46.
- Martins-Lopes P, Lima-Brito J, Gomes S, Meirinhos J, Santos L e Guedes-Pinto H (2007). RAPD and ISSR molecular markers in *Olea europaea* L.: Genetic variability and molecular cultivar identification. *Genetic Resources and Crop Evolution* **54**:117–128.
- McGregor CE, Lambert CA, Greyling MM, Louw JH e Warnich L (2000). A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*Solanum tuberosum* L.) germplasm. *Euphytica* **113**:135–144.
- Minch E, Ruiz-Linares A, Goldstein D, Feldman M e Cavalli-Sforza LL (1997). MICROSAT: a computer program for calculating various statistics on microsatellite allele data, ver. 1.5d. Stanford, CA: Stanford University. Disponível em <<http://hpgl.stanford.edu/projects/microsat>>, a 10 de Novembro, 2007.
- Minelli S, Maggini F, Gelati MT, Angiolillo A e Cionini PG (2000). The chromosome complement of *Olea europaea* L.: characterization by differential staining of the chromatin and *in situ* hybridisation of highly repeated DNA sequences. *Chrom Res* **8**:615-619.

- Ministério da Agricultura, do Desenvolvimento Rural e das Pescas (MADRP) (2007). Olivicultura – Diagnóstico Sectorial. Gabinete de Planeamento e Políticas. Lisboa, Portugal.
- Murray BW (1996). The Estimation of Genetic Distance and Population Substructure from Microsatellite allele frequency data. McMaster University. Disponível em <<http://helix.mcmaster.ca/brent/brent.html>> a 10 de Novembro, 2007.
- Nei M (1987) DNA polymorphism within and between populations. In: Nei M (ed) Molecular evolutionary genetics. Columbia University Press, New York, pp 254-286.
- Olsen KM e Schaal BA (2001). Microsatellite variation in cassava (*Manihot esculenta*, Euphorbiaceae) and its wild relatives: further evidence for a southern Amazonian origin of domestication. *Am J Bot* **88**:131–142.
- Pasqualone A, Montemurro C, Summo C, Sabetta W, Caponio F e Blanco A (2007). Effectiveness of Microsatellite DNA Markers in Checking the Identity of Protected Designation of Origin Extra Virgin Olive Oil. *J. Agric. Food Chem.* **55**:3857-3862.
- Peakall R e Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*. **6**: 288-295.
- Paetkau D e Strobeck C (1995). The molecular-basis and evolutionary history of a microsatellite null allele in bears. *Mol. Ecol.* **4**:519–520.
- Rakoczy-Trojanowska M e Bolibok H (2004). Characteristics and a comparison of three classes of microsatellite-based markers and their application in plants. *Cellular & Molecular Biology Letters* **9**:221-238.
- Rallo P, Dorado G e Martin A (2000). Development of simple sequence repeats (SSRs) in olive tree (*Olea europaea* L.). *Theor Appl Genet* **101**:984-989.
- Rallo P, Tenzer I, Gessler C, Baldoni L, Dorado G e Martín A (2003). Transferability of olive microsatellite loci across the genus *Olea*. *Theor Appl Genet* **107**:940-946.
- Ribeiro O (1991). Portugal, o Mediterrâneo e o Atlântico. Livraria Sá Costa Editora, Lisboa.
- Richard GF e Paques F (2000). Mini- and microsatellite expansions: the recombination connection. *EMBO Rep.* **1**:122-126.
- Rohlf FJ (1998). NTSYSpc version 2.0: Numerical Taxonomy and Multivariate Analysis System. Department of Ecology and Evolution, State University of New York, Stony Brook, New York.
- Rotondi A, Magli M, Ricciolini C e Baldoni L (2003). Morphological and molecular analysis for the characterization of a group of Italian olive cultivars. *Euphytica* **132**:129-137.
- Rugini E, Pannelli G, Ceccarelli M e Muganu M (1996). Isolation of triploid and tetraploid olive (*Olea europaea* L.) plants from mixoploid cv Frantoio and Leccino mutants by *in vivo* and *in vitro* selection. *Plant Breed* **115**:23-27.

- Saumitou-Laprade P, Vassiliadis C, Epplen JT e Hardt C. (2000). Isolation of microsatellite loci for paternity testing in *Phillyrea angustifolia* L. (Oleaceae). *Molecular Ecology* **9** (1):112–114.
- Sefc KM, Lopes MS, Mendonça D, Rodrigues Dos Santos M, Laimer Da Câmara Machado M e Da Câmara Machado A (2000) Identification of microsatellite loci in olive (*Olea europaea*) and their characterization in Italian and Iberian olive trees. *Mol Ecol* **9**:1171–1173.
- Selkoe K e Toonen RJ (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**:615-629.
- Schlötterer C (2000) Evolutionary dynamics of microsatellite DNA. *Chromosoma* **109**:365–371.
- Sneath PHA e Johnson R (1972). The influence on numerical taxonomic similarities of errors in microbiological tests. *J Gen Microbiol* **72**:377-392.
- Sneath PHA e Sokal RR (1973). Numerical taxonomy — the principles and practice of numerical classification. W. H. Freeman: San Francisco.
- Taberlet P, Waits LP e Luikart G (1999) Noninvasive genetic sampling: look before you leap. *Trends in Ecology and Evolution* **14**:323–327.
- Upadhyaya HD (2006). Genotyping Groundnut (*Arachis hypogaea* L.) Composite Collection (ICRISAT). In: Generation Challenge Program 2006 Annual Research Meeting 12-16 September 2006, Sao Paulo, Brazil.
- Wagner HW e Sefc KM (1999). IDENTITY 1.0. Centre for Applied Genetics, University of Agricultural Sciences Vienna.
- Weir BS e Cockerham CC (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**:1358–1370.
- Wu SB, Collins G e Sedgley M (2004). A molecular linkage map of olive (*Olea europaea* L.) based on RAPD, microsatellite, and SCAR markers. *Genome* **47**:26–35.
- Wünsch A e Hormaza JI (2002). Cultivar identification and genetic fingerprinting of temperate fruit tree species using DNA markers. *Euphytica* **125**:59–67.
- Zohary D e Hopf M (2000). Domestication of plants in the old world: The origin and spread of cultivated plants in West Asia, Europe and the Nile Valley. 3rd ed. Oxford Univ. Press.

## Anexos

## Tabelas

Tabela I – *Loci* utilizados. (1) Carriero *et al.* 2002; (2) Cipriani *et al.* 2002 ; (3) de la Rosa *et al.* 2002 ; (4) Sefc *et al.* 2000 ; (5) Lefort *et al.* 1999 ; (6) Saumitou-Laprade *et al.* 2000 ; (7) Brachet *et al.* 1999. T<sub>L</sub> corresponde à temperatura de ligação, em graus Celsius, dos oligonucleótidos iniciadores ao DNA durante a reacção de amplificação.

Locus	Sequência repetida de DNA	Oligonucleótidos iniciadores (5'-3')	T <sub>L</sub> (°C)
GAPU101 (1)	(GA) <sub>8</sub> (G) <sub>3</sub> (AG) <sub>3</sub>	F:CATGAAAGGAGGGGGACATA R:GGCACTTGTGTGCAGATTG	57
GAPU103 A (1)	(TC) <sub>26</sub>	F:TGAATTTAACTTTAAACCCACACA R:GCATCGCTCGATTTTTATCC	57
GAPU47 (1)	(CT) <sub>16</sub>	F:GATCAGCTTAGTCTCATATTCTCTCTC R:CCTCGACTGATTTACACACCA	57
GAPU89 (1)	(AG) <sub>16</sub> (G) <sub>3</sub> (GA) <sub>9</sub>	F:GATCATTCCACACACGAGAG R:AACACATGCCACAAACTGA	57
GAPU71B (1)	GA(AG) <sub>6</sub> (AAG) <sub>8</sub>	F:GATCAAAGGAAGAAGGGGATAAA R:ACAACAAATCCGTACGCTTG	57
GAPU71 A (1)	(AG) <sub>10</sub>	F:GATCATTTAAATATTAGAGAGAGAGA R:TCCATCCATGCTGAACTT	57
UDO99–015 (2)	(TG) <sub>12</sub>	F:ATTCATCTATGGGCCGCTTC R:TCAACACAACCTACTAGCCTACCA	57
UDO99–017 (2)	(TG) <sub>11</sub>	F:GCCACAAACTCTTTGAACC R:GCGATTTTTCCCTGTATTTAGGT	57
UDO99–006 (2)	(GT) <sub>5</sub> (AT) <sub>6</sub> G(GT) <sub>9</sub>	F:TCAGTTTGTGCTTTAGTGGA R:TTGTAATATGCCATGTAACCTCGAT	57
UDO99–024 (2)	(CA) <sub>11</sub> (TA) <sub>2</sub> (CA) <sub>4</sub>	F:GGATTTATTAAGCAAACATACAAA R:CAATACAAATGAGCATGATAAGACA	57
UDO99–019 (2)	(GT) <sub>20</sub> (AT) <sub>5</sub>	F:TCCCTTGTAGCCTCGTCTTG R:GGCCTGATCATCGATACCTC	57
UDO99–014 (2)	(GT) <sub>10</sub>	F:TTCCCTTATTCAATGTGAACC R:ACTGCAGTTTGGGAATCAAA	57
UDO99–028 (2)	(CA) <sub>23</sub> (TA) <sub>3</sub>	F:CTGCAGCTTCTGCCCATAC R:GCAGATCATCATTTGGCACT	57
UDO99–044 (2)	(TG) <sub>7</sub> G(TG) <sub>2</sub> (G) <sub>3</sub> (TG) <sub>2</sub> (TGCG) <sub>5</sub>	F:AATTCGACAAGTTGTGTGTG R:CACAGCACCCAACCAGATTT	57
UDO99–011 (2)	(CT) <sub>7</sub> (CA) <sub>10</sub> (CT) <sub>2</sub> (CA) <sub>2</sub> CT(CA) <sub>2</sub> CT(CA) <sub>9</sub>	F:TGACTCCCTTTAAACTCATCAGG R:TGCGCATGTAGATGTGAATATG	57
UDO99–012 (2)	(GT) <sub>10</sub>	F:TCACCATTCTTAACTTCACACCA R:TCAAGCAATTCACGCTATG	57
UDO99–031 (2)	(TG) <sub>21</sub> (TATG) <sub>6</sub>	F:TATCCTCTATGTGGCGATG R:TTGGTTAAAAGGATTGATACA	57
EMO90 (3)	(CA) <sub>10</sub>	F:CATCCGGATTTCTTGCTTTT R:AGCGAATGTAGCTTTGCATGT	55
EMO2 (3)	(AG) <sub>5</sub> -G-(GA) <sub>10</sub>	F:CTCGCACTTTAAATTCATATGGGTAGGT R:GCGTGCTTGGGTGCTTGTGTTG	60
EMO3 (3)	(CA) <sub>7</sub>	F:GGTGTAGCCCAAGCCCTTAT R:TGCATGACCGTGGTGTAAAGT	60
ssrOeUA-DCA15 (4)	(CA) <sub>3</sub> G(AC) <sub>14</sub>	F:GATCTTGTCTGTATATCCACAC R:TATACCTTTCCATCTTGACGC	50

Tabela I – (continuação)

Locus	Sequência repetida de DNA	Oligonucleótidos iniciadores (5'-3')	T <sub>L</sub> (°C)
ssrOeUA-DCA8 (4)	(GA) <sub>18</sub>	F:ACAATTCAACCTCACCCCATACCC R:TCACGTCAACTGTGCCACTGAACTG	50
ssrOeUA-DCA18 (4)	(CA) <sub>4</sub> CT(CA) <sub>3</sub> (GA) <sub>19</sub>	F:AAGAAAGAAAAAGGCAGAATTAAGC R:GTTTTCGTCTCTCTACATAAGTGAC	50
ssrOeUA-DCA3 (4)	(GA) <sub>19</sub>	F:CCCAAGCGGAGGTGTATATTGTTAC R:TGCTTTTGTGCGTGTGGAGATGTTG	50
ssrOeUA-DCA9 (4)	(GA) <sub>23</sub>	F:AATCAAAGTCTTCCTTCTCATTTCG R:GATCCTTCCAAAAGTATAACCTCTC	50
ssrOeUA-DCA16 (4)	(GT) <sub>13</sub> (GA) <sub>29</sub>	F:TTAGGTGGGATTCTGTAGATGGTTG R:TTTTAGGTGAGTTCATAGAATTAGC	50
FEMSATL4 (5)	(CA) <sub>2</sub> (AG) <sub>24</sub>	F:TTCATGCTTCTCCGTGTCTC R:GCTGTTTCAGGCGTAATGTG	52
FEMSATL19 (5)	(CA) <sub>6</sub> CGGC(CA) <sub>13</sub>	F:CTGTTCAATCAAAGATCTCA R:TGCTCGCATATGTGCAGATA	52
PA(GA)1 (6)	(AG) <sub>32</sub>	F:TTTTGAGGATCATTAGCAGCC R:AAATTGGAATGGATTCTCT	55
PA(GA)5 (6)	(GA) <sub>12</sub>	F:TCTTTGCTTCGTTGCTTTTG R:TCTTGCTCCCTCGACATTTT	55
M230a (7)	(TG) <sub>15</sub> (AG) <sub>23</sub>	F:CGCACGTTCTTTCTATTTG R:GACCGGCTGACTATTTTCTC	50

Tabela II – Alelos e sua frequência nos acessos da cultivar 'Cobrançosa'. (A – alelos; f – frequência). Os alelos raros (frequência &lt; 5%) encontram-se a sublinhado.

Locus	A	172	180	184	192	198	200	218	220
GAPU101	A	<u>172</u>	<u>180</u>	<u>184</u>	192	198	200	218	220
	f	0.00685	0.00685	0.00685	0.40411	0.08904	0.08904	0.17808	0.21918
GAPU89	A	160	<u>171</u>	176	<u>179</u>	205			
	f	0.49351	0.00649	0.11039	0.00649	0.38312			
GAPU71B	A	<u>119</u>	122	125	<u>127</u>	<u>138</u>	141	<u>144</u>	
	f	0.00667	0.40667	0.08667	0.00667	0.03333	0.45333	0.00667	
UDO99-015	A	<u>92</u>	101	<u>103</u>	105	<u>126</u>	<u>133</u>	<u>139</u>	141
	f	0.01299	0.05844	0.01299	0.82468	0.01299	0.01299	0.00649	0.05844
UDO99-028	A	<u>125</u>	133	<u>140</u>	149				
	f	0.01299	0.11688	0.00649	0.86364				
EMO3	A	208	210	212	<u>216</u>				
	f	0.40132	0.18421	0.38158	0.03289				
ssrOeUA-DCA15	A	<u>233</u>	244	255	264				
	f	0.01316	0.46711	0.38158	0.13816				
ssrOeUA-DCA16	A	123	<u>125</u>	<u>155</u>	<u>188</u>				
	f	0.92667	0.02667	0.04000	0.00667				
PA(GA)5	A	<u>113</u>	117	<u>121</u>	127				
	f	0.04000	0.46000	0.03333	0.46667				
M230a	A	<u>211</u>	<u>214</u>	228	<u>230</u>	232	236		
	f	0.00658	0.03947	0.47368	0.01316	0.38816	0.07895		

## Figuras

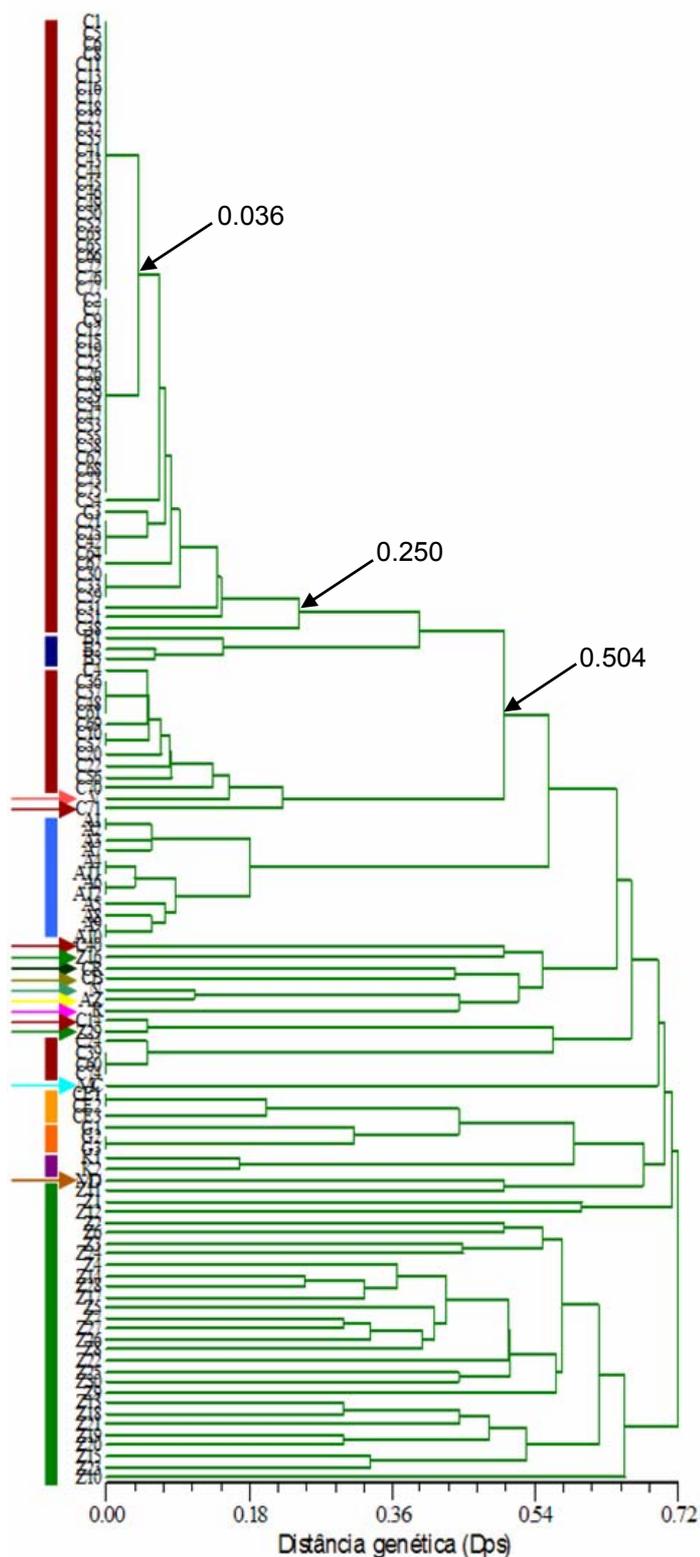


Figura I – Fenograma da distância baseada na proporção dos alelos partilhados, obtido a partir do método UPGMA, de todos os acessos analisados. (Coeficiente de correlação cofenética = 0.948) ■ Zambujeiro; ■ Madural; ■ Koroneiki; ■ Galega; ■ Conserva de Elvas; ■ Maçanilha de Tavira; ■ Cobrançosa; ■ Redondil; ■ Azeiteira; ■ Negrinha; ■ Cordovil de Castelo Branco; ■ Arbequina; ■ Verdeal de Trás-os-Montes; ■ Carrasquenha; ■ Arbosana'.

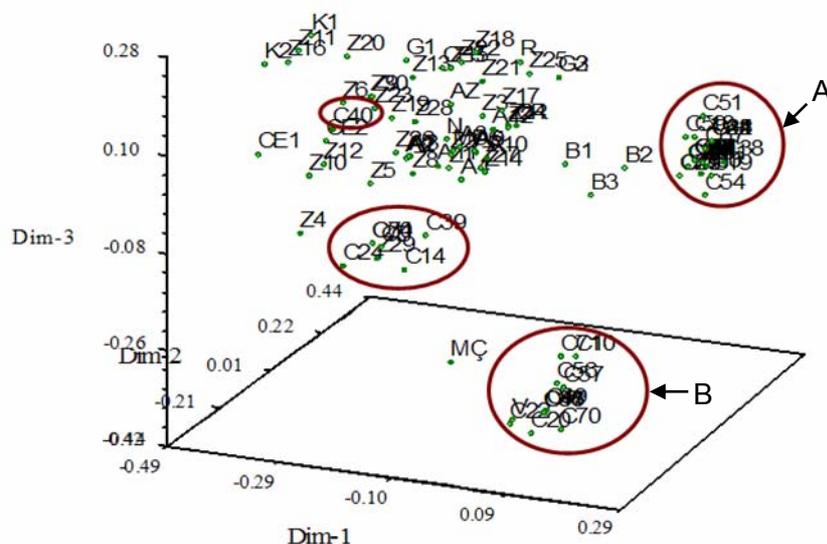
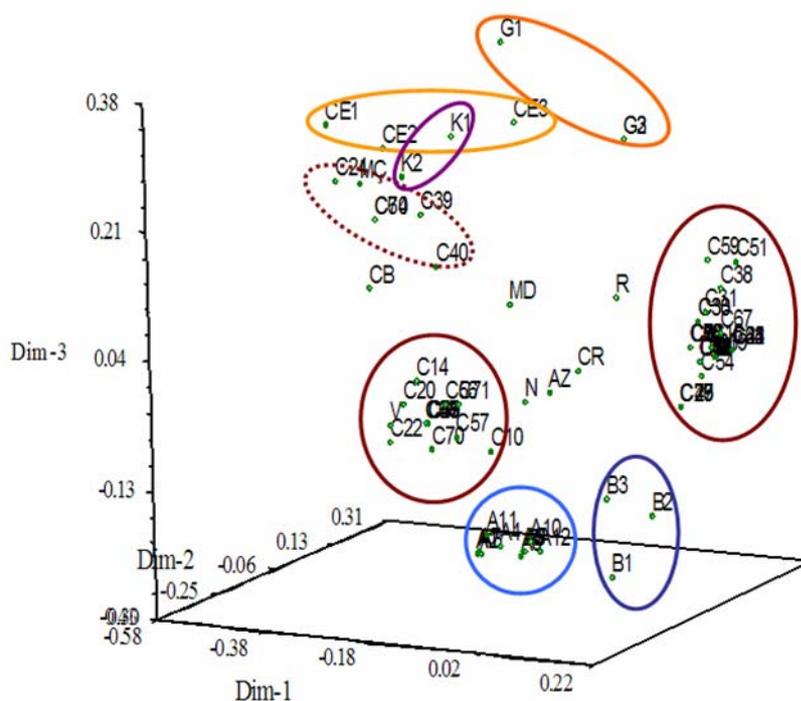


Figura II – Análise das coordenadas principais dos 138 acessos analisados. Representação gráfica a 3 dimensões (1ª dimensão 29.50%; 2ª dimensão 18.16%; 3ª dimensão 13.50%). Grupos de ‘Cobrançosa’ circunscritos a vermelho.



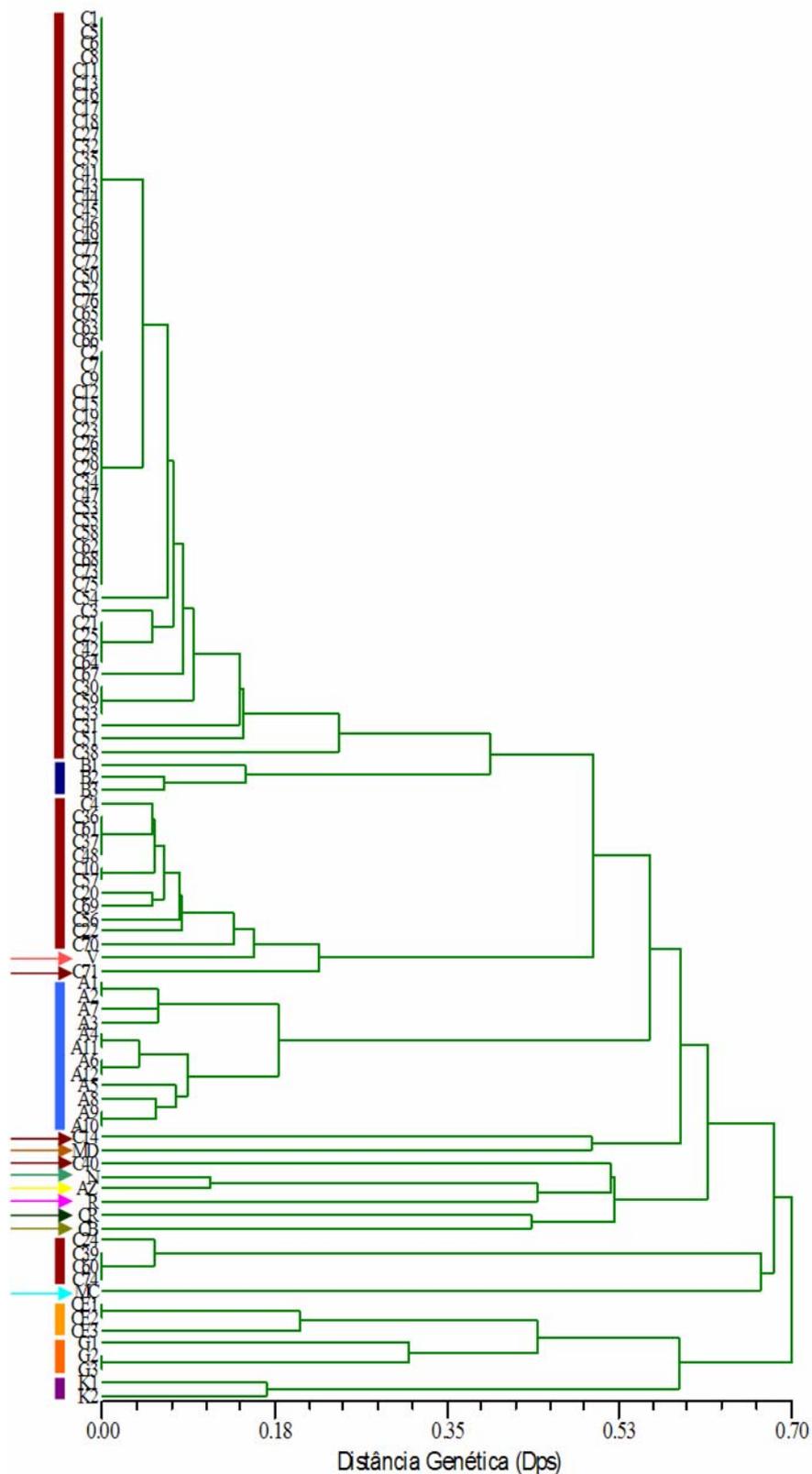


Figura IV - Fenograma da distância baseada na proporção dos alelos partilhados, obtido a partir do método UPGMA, dos acessos de 'Cobraçosa' e das outras cultivares. (Coeficiente de correlação cofenética = 0.961)

■ 'Madural'; ■ 'Koroneiki'; ■ 'Galega'; ■ 'Conserva de Elvas'; ■ 'Maçanilha de Tavira';  
 ■ 'Cobraçosa'; ■ 'Redondil'; ■ 'Azeiteira'; ■ 'Negrinha'; ■ 'Cordovil de Castelo Branco';  
 ■ 'Carrasquenha'; ■ 'Arbequina'; ■ 'Arbosana'; ■ 'Verdeal de Trás-os-Montes'.

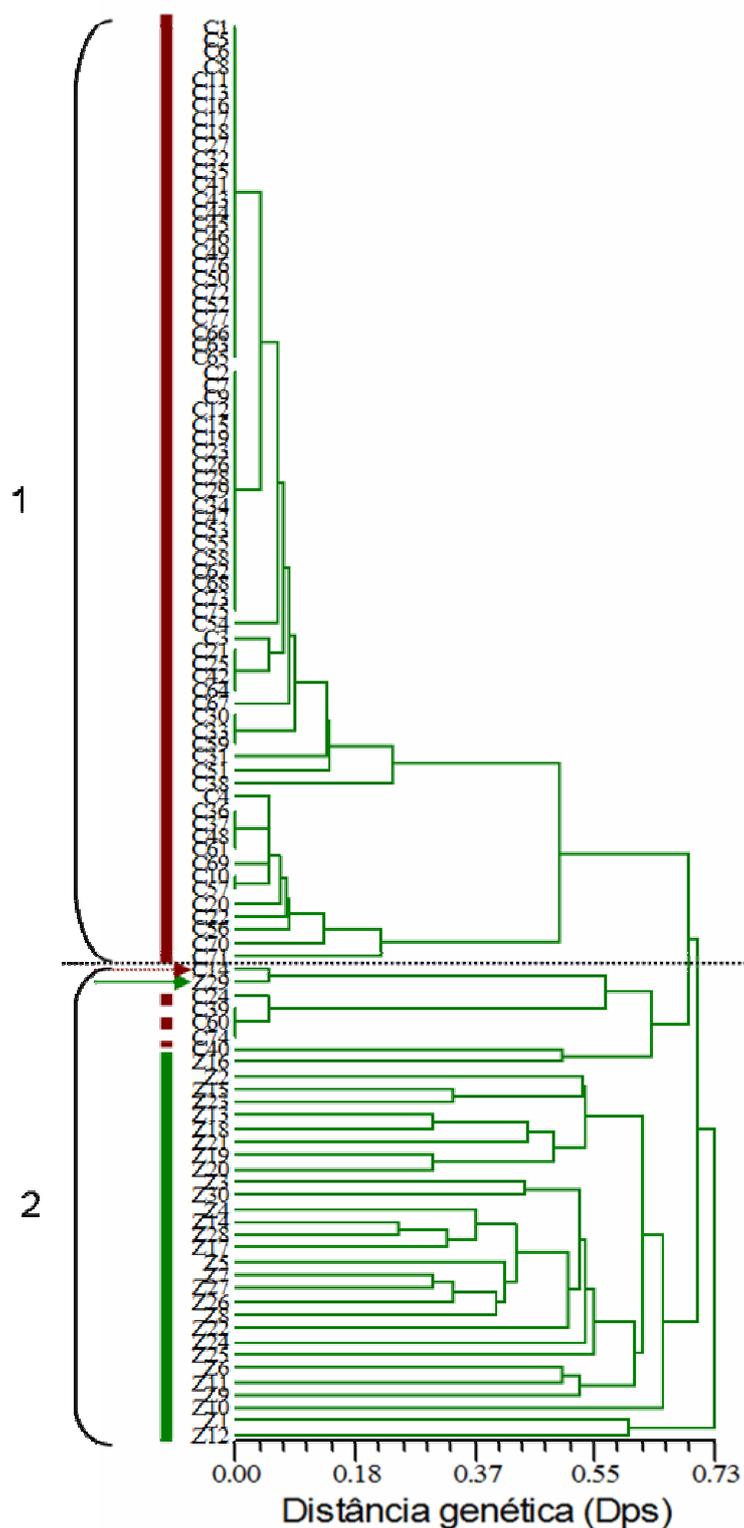


Figura V - Fenograma da distância baseada na proporção dos alelos partilhados, obtido a partir do método UPGMA, dos acessos de 'Cobraçosa' e dos acessos de zambujeiro. (Coeficiente de correlação cofenética = 0.975). Grupo 1- 'Cobraçosa'; Grupo 2- maioritariamente zambujeiros. Dps entre os dois grupos, 0.694. ■ 'Cobraçosa'; ■ Zambujeiro.

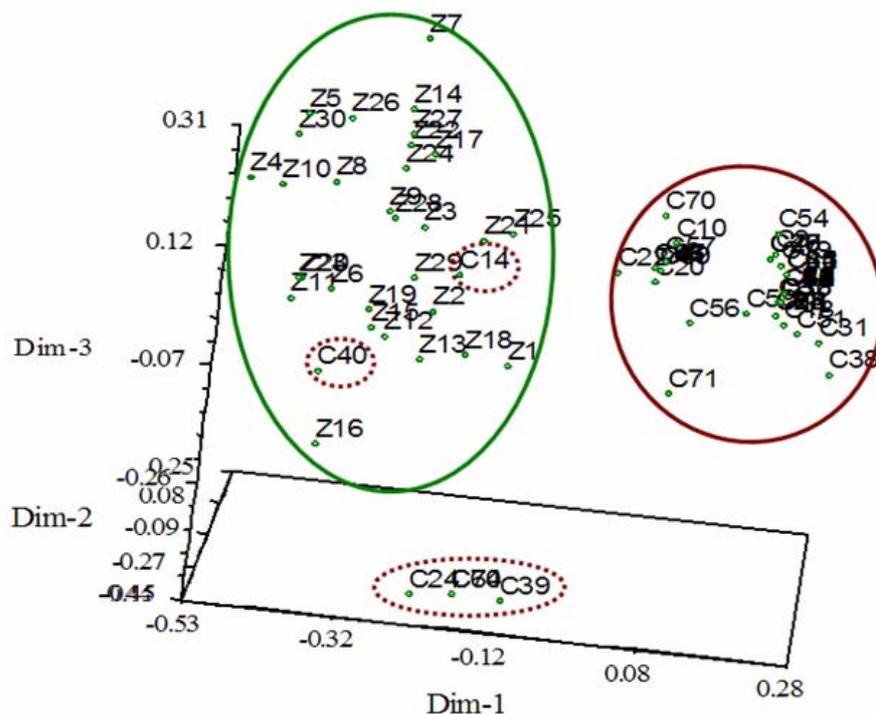


Figura VI - Análise das coordenadas principais dos acessos de 'Cobrançosa' e zambujeiro. Representação gráfica a 3 dimensões (1ª dimensão 40.80%; 2ª dimensão 18.97%; 3ª dimensão 11.18%). Grupos de 'Cobrançosa' circunscritos a vermelho; zambujeiros circunscritos a verde.

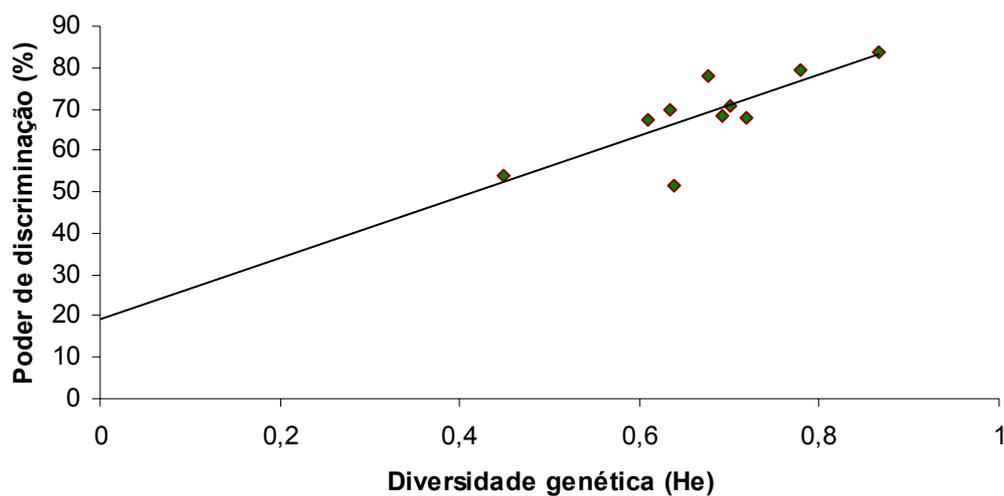


Figura VII - Correlação entre diversidade genética e o poder de discriminação de cada *locus*, nos 138 acessos analisados.  $R^2 = 0.621$ .