

UNIVERSIDADE DO ALGARVE
FACULDADE DE ENGENHARIA DE RECURSOS NATURAIS

Mapeamento genético do ananaseiro
(*Ananas comosus* (L.) Merrill)

(Tese para a obtenção do grau de Doutor no ramo de biologia, especialidade de genética)

Jorge Daniel Dias Carlier

Orientador: Doutor José Manuel Peixoto Teixeira Leitão, Professor Catedrático da Faculdade de Engenharia de Recursos Naturais da Universidade do Algarve

Constituição do Júri:

Presidente: Reitor da Universidade do Algarve

Vogais: Doutor Henrique de Pinho Guedes Pinto, Professor Catedrático da Universidade de Trás-os-Montes e Alto Douro;

Doutora Deborah Mary Power, Professora Catedrática da Faculdade de Engenharia de Recursos Naturais da Universidade do Algarve;

Doutor Gustavo Nuno Barbosa Nolasco, Professor Associado com agregação da Faculdade de Engenharia de Recursos Naturais da Universidade do Algarve;

Doutor Manuel Pedro Salema Fevereiro, Professor auxiliar com agregação da Faculdade de Ciências da Universidade de Lisboa;

Doutora Cristina Maria Pereira Marques, Investigadora do Instituto de Investigação da Floresta e Papel – RAIZ.

FARO
2006

Nome: Jorge Daniel Dias Carlier

Faculdade: Faculdade de Engenharia de Recursos Naturais

Orientador: Doutor José Manuel Peixoto Teixeira Leitão

Data: 27 de Fevereiro de 2006

Título da tese: Mapeamento Genético do Ananaseiro (*Ananas comosus* (L.) Merrill)

RESUMO

O objectivo desta tese foi a construção de um mapa genético de referência para o ananaseiro (*Ananas comosus* (L.) Merrill). Com base numa população F1 (*A. comosus* var. *comosus* X *Ananas comosus* var. *bracteatus*) foi construído um mapa para cada progenitor. Subsequentemente, com base na população F2 foi construído um mapa que integra marcadores dos dois progenitores. A análise de marcadores que segregam nas duas descendências, F1 e F2, permitiu a construção de um mapa integrado contendo 783 marcadores moleculares (588 AFLP, 137 RAPD, 57 ISSR, *locus* isoenzimático PGM) e 1 *locus* determinante da característica morfológica “piping”, distribuídos em 77 grupos de ligamento. A existência no mapa integrado de 31 grupos de ligamento (um número relativamente próximo do número de cromossomas do ananaseiro, $n=25$) com dimensões superiores a 50cM, aliada a uma cobertura do genoma nos diferentes mapas entre 60.7 e 91.4% e uma densidade moderadamente elevada de marcadores que se encontram a uma distância média de 4.16 cM, são factos que apontam para que estes primeiros mapas genéticos do ananaseiro possam ser um auxiliar indispensável na construção de mapas genómicos físicos, na sequenciação genómica e no isolamento de genes desta espécie via clonagem baseada em mapas genéticos.

PALAVRAS-CHAVE:

mapa genético, mapa de ligamento, marcadores moleculares, *Ananas, bracteatus, comosus*

Título da tese em inglês: Genetic Mapping of Pineapple (*Ananas comosus* (L.)

Merrill)

ABSTRACT

The main aim of this thesis was the construction of a Pineapple (*Ananas comosus* (L.) Merrill) genetic reference map. Based on an (*A. comosus* var. *comosus* X *Ananas comosus* var. *bracteatus*) F1 population, two genetic maps, one for each progenitor, were assembled. Subsequently, a new map integrating markers from both progenitors was constructed on the basis of an F2 population. Using markers segregating simultaneously among the F1 and F2 populations, an integrated map was constructed gathering 783 molecular markers (588 AFLP, 137 RAPD, 57 ISSR, and the PGM isozyme *locus*) and the morphological trait *locus* “piping” into 77 linkage groups. In view of the facts that (1) in the integrated map, 31 linkage groups (a number close to the 25 pineapple chromosomes) are longer than 50cM; (2) that the different maps span over 60.7% to 91.4% of the genome, and; (3) that marker density is moderately high, with 4.16 cM average distance between markers, it is legitimate to suppose that these first genetic maps of pineapple will be very useful for construction of physical maps, for genome sequencing and for gene isolation via map-based cloning in this crop species.

KEY-WORDS:

genetic map, linkage map, molecular markers, *Ananas, bracteatus, comosus*

Este trabalho foi financiado pela bolsa de doutoramento SFRH/BD/12652/2003 da Fundação para a Ciência e a Tecnologia.

Agradecimentos

À Alexandra e aos nossos filhos Isaac e Maria.

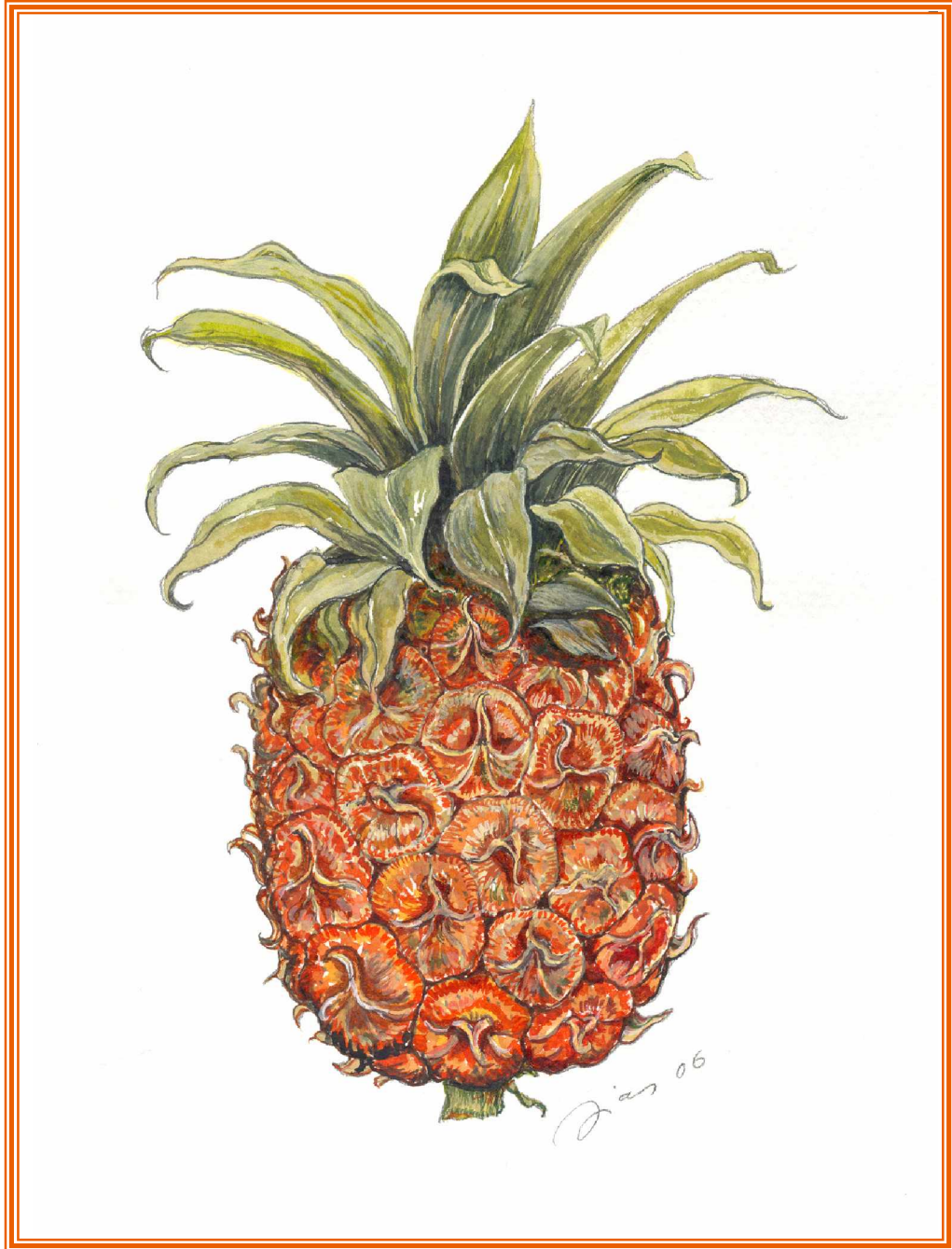
À minha mãe.

Ao meu pai.

Aos meus irmãos.

Ao meu orientador, pela oportunidade e apoio dados para a realização deste doutoramento e pela total disponibilidade com que sempre me atendeu para o esclarecimento de dúvidas.

Aos colegas de laboratório com quem partilhei o dia-a-dia de trabalho durante a concretização do doutoramento.



Fruto ananás (*Ananas comosus* var. *comosus* ((L.) Merrill) Coppens & Leal).
Aquarela de Fernanda Dias (minha mãe).

Índice geral

ÍNDICE DE TABELAS	8
ÍNDICE DE FIGURAS	9
LISTA DE ABREVIATURAS	10
1- INTRODUÇÃO	13
2- REVISÃO BIBLIOGRÁFICA	15
2.1- GENÓMICA ESTRUTURAL	15
2.1.1- MAPEAMENTO FÍSICO	16
2.1.2- MAPEAMENTO GENÉTICO	20
2.1.2.1- POPULAÇÕES E ESTRATÉGIAS DE MAPEAMENTO GENÉTICO	22
2.1.2.1.1- POPULAÇÕES DE CRUZAMENTOS CONTROLADOS	22
2.1.2.1.2- POPULAÇÕES NATURAIS	26
2.1.2.2- MARCADORES GENÉTICOS	29
2.1.2.3- ANÁLISE DE LIGAMENTO	36
2.1.2.4- QUALIDADE DOS MAPAS GENÉTICOS	37
2.1.2.5- IMPORTÂNCIA E APLICAÇÕES DOS MAPAS GENÉTICOS	41
2.1.3- GENÓMICA ESTRUTURAL NOS PRINCIPAIS FRUTOS TROPICAIS	45
2.2- ANANASEIRO	48
2.2.1- IMPORTÂNCIA ECONÓMICA	48
2.2.2- TAXONOMIA	49
2.2.3- HABITAT NATURAL E ORIGEM	52
2.2.4- NÚMERO DE CROMOSSOMAS E TAMANHO DO GENOMA	53
2.2.5- FLUXO GENÉTICO E PROPAGAÇÃO	54
2.2.6- DIVERSIDADE MOLECULAR	55
2.2.7- MELHORAMENTO	57
2.2.8- RECURSOS GENÉTICOS E GENÓMICOS	58
2.3- IMPORTÂNCIA DO MAPEAMENTO GENÉTICO NO ANANASEIRO	61
3- MATERIAIS E MÉTODOS	64
3.1- MATERIAL VEGETAL	64
3.2- EXTRACÇÃO DE DNA	64
3.3- MARCADORES ANALISADOS	65

3.3.1- MARCADOR MORFOLÓGICO	65
3.3.2- ANÁLISES ISOENZIMÁTICAS	66
3.3.3- ANÁLISES RAPD	67
3.3.4- ANÁLISES AFLP	70
3.3.5- ANÁLISES ISSR	74
3.4- ANÁLISE DA SEGREGAÇÃO DOS MARCADORES	76
3.5- CONSTRUÇÃO DOS MAPAS	77
3.6- IDENTIFICAÇÃO DOS GRUPOS E DOS MARCADORES	79
3.7- COMPRIMENTO E COBERTURA DO MAPA	80
3.8- DISTRIBUIÇÃO DOS MARCADORES NO MAPA	81
3.9- CLONAGEM E SEQUENCIAÇÃO DE MARCADORES	81
3.10- ANÁLISE DE SEQUÊNCIAS	83
4- RESULTADOS	85
4.1- MAPEAMENTO COM A POPULAÇÃO F1	85
4.1.1- MARCADORES ANALISADOS	85
4.1.2- MAPA DA VAR. <i>BRACTEATUS</i>	89
4.1.3- MAPA DA VAR. <i>COMOSUS</i>	92
4.1.4- INTEGRAÇÃO DOS MAPAS DA VAR. <i>BRACTEATUS</i> E DA VAR. <i>COMOSUS</i> ...	95
4.1.5- COMPRIMENTO E COBERTURA DOS MAPAS	96
4.1.6- DISTRIBUIÇÃO DOS MARCADORES NOS MAPAS	97
4.2- MAPEAMENTO COM A POPULAÇÃO F2	99
4.2.1- MARCADORES ANALISADOS	99
4.2.2- MAPA	101
4.2.3- COMPRIMENTO E COBERTURA DO MAPA	106
4.2.4- DISTRIBUIÇÃO DOS MARCADORES NO MAPA	107
4.3- INTEGRAÇÃO DOS MAPAS F1 E F2	107
4.4- MARCADORES SEQUENCIADOS	113
5- DISCUSSÃO	119
6- CONSIDERAÇÕES FINAIS	127
7- REFERÊNCIAS BIBLIOGRÁFICAS	128

Índice de tabelas

Tabela 1	Funções de mapeamento: a mais simples (linear) e as duas mais utilizadas (Haldane e Kosambi).	21
Tabela 2	Características de mapas construídos para algumas espécies de plantas com um elevado número de cromossomas.	40
Tabela 3	Referências de mapas genéticos para os 20 frutos mais produzidos no mundo segundo estatísticas da FAO para o ano 2004.	42
Tabela 4	Taxonomia do ananaseiro segundo Smith e Downs (1979), e Leal (1990).	50
Tabela 5	Taxonomia do ananaseiro por Coppens d'Eeckenbrugge e Leal (2003).	52
Tabela 6	Soluções de revelação dos sistemas isoenzimáticos analisados.	67
Tabela 7	“Primers” RAPD analisados.	68
Tabela 8	Condições de amplificação dos marcadores RAPD.	69
Tabela 9	Combinações de “primers” AFLP analisados e respectivo número de código.	71
Tabela 10	Adaptadores e “primers” utilizados nas análises AFLP.	71
Tabela 11	Condições da pré-amplificação dos marcadores AFLP.	72
Tabela 12	Condições da amplificação selectiva dos marcadores AFLP.	73
Tabela 13	“Primers” ISSR analisados e respectivo número de código.	75
Tabela 14	Mistura reactiva para amplificação de ISSRs marcados com [$\gamma^{33}\text{P}$].	75
Tabela 15	Ciclos de amplificação para os marcadores ISSR.	76
Tabela 16	Condições da electroporação bacteriana.	82
Tabela 17	Características genómicas de 15 espécies de plantas (adaptado de Chagné et al. 2002).	97

Tabela 18	Número de “primers” e de marcadores analisados na população F2.	99
Tabela 19	Sequências de marcadores.	115
Tabela 20	Entradas das sequências com similaridades mais elevadas com o marcador OPD15-700 identificadas por Blastn no NCBI.	118
Tabela 21	Entrada da sequência com maior similaridade ao marcador OPD15-700 identificada por Blastx na base de dados “nr” do NCBI.	118
Tabela 22	Exemplos de cobertura de mapas genéticos construídos para diferentes espécies.	120

Índice de figuras

Figura 1	Zimogramas dos sistemas isoenzimáticos PGM e MDH de 10 plantas da população F1.	85
Figura 2	Padrões de amplificação RAPD de 18 plantas da população F1 obtidos com o “primer” OPB15 e padrões AFLP de 32 plantas F1 gerados com a combinação de “primers” M-CTA/E-AAG, com pares de marcadores cuja segregação em alternância sugere serem alelos do mesmo <i>locus</i> .	87
Figura 3	Mapa genético obtido com a população F1 para a var. <i>bracteatus</i> .	90
Figura 4	Mapa genético obtido com a população F1 para a var. <i>comosus</i> .	93
Figura 5	Padrões de amplificação ISSR de 19 plantas da população F2 obtidos com o “primer” (GA)8YT, com marcadores a segregarem.	100
Figura 6	Mapa genético obtido com a população F2.	103
Figura 7	Mapa integrado dos mapas F1 e F2.	109

Lista de abreviaturas

ACC -	1-Aminociclopropano-1-carboxilato
AFLP -	“Amplified Fragment Length Polymorphism”
ATP -	Adenosina trifosfato
BAC -	“Bacterial Artificial Chromosome”
BC -	“BackCross”
BIBAC -	“Binary Bacterial Artificial Chromosome”
CAP -	“Cleavage Amplified Polimorphism”
CBCS -	“Clone-By-Clone Shotgun”
CBM -	Clonagem Baseada no Mapeamento
cDNA -	“Complementary DNA”
CENARGEN -	Centro Nacional de Recursos Genéticos e Biotecnologia
CIRAD -	“Centre de Coopération Internationale en Recherche Agronomique”
cM -	Centimorgan
CNPMF -	Centro Nacional de Pesquisa de Mandioca e Fruticultura Tropical
cpDNA -	“Chloroplast DNA”
cv. -	Cultivar
dATP -	Desoxiadenosina 5' - trifosfato
dCTP -	Desoxicitosina 5' - trifosfato
dGTP -	Desoxiguanina 5' - trifosfato
DH -	Duplo-Haplóide
DNA -	Ácido desoxirribonucleico
DTT -	Ditiotreitol
dTTP -	Desoxitimidina 5' - trifosfato

EDTA -	Ácido etilenodiaminotetracético
EM -	“Expectation-Maximization”
EMBRAPA -	Empresa Brasileira de Pesquisa Agropecuária
EST -	“Expressed Sequence Tag”
FAO -	“Food and Agriculture Organization”
FISH -	“Fluorescent In Situ Hybridization”
G6PDH -	Glucose-6-fosfato desidrogenase
HRR -	“Haplotype Relative Risk”
IDH -	Isocitrato desidrogenase
ISSR -	“Inter Simple Sequence Repeat”
LOD -	“Logarithm of ODDs”
MARDI -	“Malaysian Agricultural Research and Development Institute”
MDH -	Malato desidrogenases
MOC -	“Minimum Obligatory Crossingovers”
NAD -	Nicotinamida adenina dinucleotídeo
NADP -	Nicotinamida adenina dinucleotídeo fosfato
NBT-	Nitro azul de tetrazólio
OSKS -	“One Sample Kolmogorov-Smirnov”
PAC -	“Bacteriophage P1-derived Artificial Chromosome”
PARF -	“Minimum Product of Adjacent Recombination Fractions”
PBC -	“large-insert conventional Plasmid-Based Clone”
PCR -	“Polymerase Chain Reaction”
PDT -	“Pedigree Disequilibrium Test”
PFGE -	“Pulsed Field Gel Electrophoresis”
PGM -	Fosfo gluco mutases
PMS -	Fenazina metassulfato
PMSF -	Fluoreto de fenil metil sulfonil

PMWaV	“Pineapple Mealybug Wilt-associated Virus”
PPO -	Polifenol oxidase
PVP -	Polivinil pirrolidona
QTL -	“Quantitative Trait Loci”
RAPD -	“Random Amplified Polymorphic DNA”
RILs -	“Recombinant Inbred Lines”
SAM -	Seleção Assistida por Marcadores
SARF -	“Minimum Sum of Adjacent Recombination Fractions”
SCAR -	“Sequence Characterized Amplified Region”
SDS -	Dodecil sulfato de sódio
SNP -	“Single Nucleotide Polimorphism”
SSCP -	“Single-Stranded Conformational Polymorphism”
SSR -	“Simple Sequence Repeat”
STS -	“Sequence Tagged Site”
TBE -	Tris borato EDTA
TC7 -	Tris citrato pH 7.0
TDLM -	“Trans-Dominant Linked Markers”
TDT -	“Transmission Disequilibrium Test”
TE -	Tris-EDTA
Tris -	Tris(hidroximetil)aminometano
Var. -	Variedade botânica
VDA -	“Variant Detector Array”
WGS -	“Whole-Genome Shotgun”
YAC -	“Yeast Artificial Chromosome”

1- INTRODUÇÃO

O ananaseiro (*Ananas comosus* (L.) Merrill), cultivado em todos os países tropicais e subtropicais, produz um dos frutos de maior importância económica, o ananás. Nas estatísticas da “Food and Agriculture Organization” (FAO) (<http://apps.fao.org>) para o ano de 2004 este fruto aparece em 42º lugar a nível mundial, com uma produção de 15.287.413 toneladas. No entanto, considerando apenas as fruteiras ocupa o décimo lugar, sendo o terceiro entre os frutos tropicais. Cerca de 70% da produção mundial de ananás é destinada ao consumo como fruto fresco no país de origem, enquanto a restante produção se destina ao comércio internacional, principalmente em produtos processados como fatias enlatadas, pedaços frescos e congelados, sumo natural e sumo concentrado (Loeillet 1997, Rohrbach et al. 2003). No Brasil, por exemplo, o terceiro maior produtor do mundo (1,435,600 t – FAO 2004), o mercado interno consome cerca de 96% da produção: 65% como fruto fresco e os restantes 35% como sumo e outras formas processadas (Reinhardt et al. 2005).

Embora para espécies como a *Arabidopsis thaliana* e o arroz (*Oriza sativa*) o período em que o esforço científico se concentrou na sequenciação total dos genomas já tenha dado lugar a um período em que o esforço científico se concentra no estudo da função de cada gene, para muitas espécies, algumas das quais economicamente importantes como o ananaseiro, o estudo genómico está ainda no início.

Os mapas genéticos representam uma hipótese para a ordem de vários pontos (marcadores genéticos) ao longo dos cromossomas e são uma ferramenta importante na investigação genética, tanto fundamental como aplicada. Os mapas com elevada cobertura do genoma e elevada densidade de marcadores podem ser utilizados na ordenação de clones de bibliotecas de fragmentos de DNA genómico, constituindo uma estrutura de referência para a construção dos mapas físicos utilizados na sequenciação genómica (Murray et al. 1994, Dib et

al. 1996, Dietrich et al. 1996, Chen et al. 2002, Mueller et al. 2005). Os mapas genéticos são também utilizados para localizar genes relacionados com características de interesse, permitindo a identificação de marcadores que lhes estejam estreitamente ligados. Por um lado, estes marcadores podem servir para seleccionar indivíduos portadores dos genes a que estão ligados sem ser necessário avaliar a sua expressão fenotípica – Selecção Assistida por Marcadores (SAM), uma estratégia que tem sido utilizada com sucesso, acelerando a introgressão de genes por retrocruzamentos (“backcross”) (Charcosset e Moreau 2004, Francia et al. 2005). Por outro lado, os marcadores estreitamente ligados a genes podem servir para o isolamento e clonagem dos genes com base no seu posicionamento no mapa - Clonagem Baseada no Mapeamento (CBM). Esta estratégia tem permitido a clonagem e caracterização de vários genes identificados unicamente pela expressão do seu fenótipo, em particular genes de resistência a pragas e doenças (Martin et al. 1993, Cai et al. 1997, Büschges et al. 1997).

Um mapa genético do ananaseiro poderá ter um impacto importante na sua cultura ao possibilitar a aplicação da estratégia de selecção assistida por marcadores na introgressão de genes de resistência a doenças que a afectam, reduzindo a longa duração dos programas de melhoramento pelos métodos clássicos, ou a clonagem destes genes baseada no mapeamento, para posterior utilização dos mesmos no desenvolvimento de cultivares resistentes por transformação genética. Outros genes cujo mapeamento e clonagem tem particular importância são os relacionados com a floração do ananaseiro pois o estudo funcional dos mesmos é o ponto de partida para o desenvolvimento de formas de sincronizar o crescimento do fruto e a sua colheita.

Este trabalho teve como objectivo a construção do primeiro mapa genético do ananaseiro (*Ananas comosus*), um passo fundamental e de base no estudo do seu genoma e uma ferramenta importante para o seu melhoramento genético.

2- REVISÃO BIBLIOGRÁFICA

2.1- GENÓMICA ESTRUTURAL

O grande esforço científico das últimas décadas no estudo dos genomas de alguns animais e plantas, em particular a sequenciação do genoma humano, dos genomas da *Drosophila melanogaster*, *Arabidopsis thaliana* e *Oriza sativa* levaram à criação do campo interdisciplinar da genómica, uma área do conhecimento cujas fronteiras são difíceis de estabelecer pois o seu objecto de estudo, os genomas, são portadores de toda a informação que regula a vida dos organismos.

Encontram-se definições para genómica que abrangem o estudo da totalidade do genoma de um organismo, como: “the study of all of the nucleotide sequences, including structural genes, regulatory sequences, and noncoding DNA segments, in the chromosomes of an organism”, no dicionário “The American Heritage® Dictionary of the English Language” (<http://www.bartleby.com/61/>) ou “the study of the structure and function of the genome, including information about sequence, mapping, and expression, and how genes and their products work in organisms” no dicionário “Dorland's Illustrated Medical Dictionary” (<http://www.mercksource.com>). Encontram-se também definições para genómica que abrangem apenas o estudo dos genes e suas funções, como: “the large-scale investigation of the structure and function of genes” no glossário do “Drug Discovery and Development” (<http://www.dddmag.com/Glossary.aspx>) ou “the study of genes and their function” no glossário genómico do “Human Genome Project Information” (http://www.ornl.gov/sci/techresources/Human_Genome/glossary).

Aparentemente, em consequência das duas definições de genómica, também se encontram dois tipos de definições para o termo “genómica estrutural”. Um está de acordo com a definição mais abrangente de genómica, por exemplo: “structural genomics - the dissection of the architectural features of genes and chromosomes” no dicionário “MedTerms” do “MedicineNet[®]” (<http://www.medicinenet.com/script/main/hp.asp>). O outro parece corresponder à definição de genómica que se refere apenas aos genes, por exemplo: “structural genomics - the effort to determine the 3D structures of large numbers of proteins using both experimental techniques and computer simulation” no glossário genómico do “Human Genome Project Information” (http://www.ornl.gov/sci/techresources/Human_Genome/glossary).

No entanto, a definição de genómica no sentido de estudo de todo o genoma é mais concordante com o conceito inicial. Este foi expresso, pela primeira vez, há 19 anos da seguinte forma: "For the newly developing discipline of mapping/sequencing (including analysis of the information) we have adopted the term GENOMICS. We are indebted to T. H. Roderick of the Jackson Laboratory, Bar Harbor, Maine, for suggesting the term. The new discipline is born from a marriage of molecular and cell biology with classical genetics and is fostered by computational science." (McKusick e Ruddle 1987).

2.1.1- MAPEAMENTO FÍSICO

Geralmente designa-se mapa físico à descrição da ordem de um grupo de fragmentos de DNA clonados que representam parte ou a totalidade do genoma de um organismo em função das suas posições físicas na sequência nucleotídica. Neste tipo de mapa as distâncias são definidas em número de nucleótidos, mais precisamente em Kpb (Zhang e Wing 1997).

A introdução da técnica de “Pulsed Field Gel Electrophoresis” (PFGE) que permite separar fragmentos de DNA de grande dimensão (Schwartz e Cantor 1984), e de vectores que permitem clonar esses longos fragmentos de DNA, como os Cosmidos (Collins e Hohn 1978), os “Yeast Artificial Chromosomes” (YACs) (Murray e Szostak 1983), os Fosmidos (Kim et al. 1992), os “Bacterial Artificial Chromosomes” (BACs) (Shizuya et al. 1992), os “bacteriophage P1-derived Artificial Chromosomes” (PACs) (Iaonnou et al. 1994), os “Binary Bacterial Artificial Chromosomes” (BIBACs) (Hamilton et al. 1996) e os “large-insert conventional Plasmid-Based Clones” (PBCs) (Tao e Zhang 1998) foram passos essenciais para o que se pode considerar a primeira de duas fases da construção de um mapa físico: o desenvolvimento de bibliotecas de grandes fragmentos de DNA. A segunda fase consiste na determinação da ordem e da distância física em nucleótidos dos fragmentos de DNA clonados. Para atingir esse objectivo foram também desenvolvidos vários tipos de estratégias: “Interactive hybridisation”, “Mapped DNA marker-based chromosome landing”, “restriction fingerprinting” e “BAC end sequencing and fingerprint analysis” (para revisão, ver Zhang e Wu 2001).

Apesar de se designar por mapa genómico físico a descrição da ordem de fragmentos de DNA clonados, o mapa físico totalmente concluído de um genoma é em última análise a sua sequência de nucleótidos.

Vários avanços tecnológicos tornaram o processo de sequenciação de grandes genomas mais expedito. A melhoria da fidelidade e da eficiência do método de sequenciação foi crucial e tem na sua raiz o desenvolvimento de DNA polimerases termoestáveis, de terminadores de cadeia dideoxi-nucleotídeos marcados com fluorescência, de sequenciadores de leitura automática capazes de analisar múltiplas amostras em simultâneo, de sistemas robóticos capazes de realizar a maioria dos procedimentos e de métodos estatísticos e

programas informáticos de análise de dados que melhoraram e facilitaram o estabelecimento de rotinas para o alinhamento e encadeamento de sequências (para revisão, ver Green 2001).

A par da evolução tecnológica foram desenvolvidas várias estratégias para sequenciação de genomas. A estratégia de “shotgun sequencing” demonstrou ser mais eficiente, tendo-se tornado esta a mais utilizada (Zhang e Wu 2001). A base teórica desta estratégia, desenvolvida no início da década de 80 (Anderson 1981, Gardner et al. 1981, Sanger et al. 1982), consiste em que uma molécula de DNA muito longa pode ser sequenciada de forma rápida e eficiente se vários exemplares desta forem fragmentados em moléculas mais curtas, sendo estas sequenciadas e analisadas de forma a ser ordenadas em “contigs” até se revelar a sequência completa da molécula de DNA inicial. Esta estratégia foi aplicada para sequenciar grandes genomas segundo dois tipos de abordagem: “Clone-By-Clone Shotgun” (CBCS) e “Whole-Genome Shotgun” (WGS) (Venter et al. 1996, Zhang e Wu 2001, Green 2001). Na primeira abordagem (CBCS) são clonados longos fragmentos de DNA genómico e a partir de um mapa (físico) da sua ordem seleccionam-se para sequenciação os clones que cubram a totalidade do genoma com a mínima sobreposição. Na segunda (WGS) são clonados fragmentos de DNA genómico longos e curtos e as sequências das suas extremidades são utilizadas para aferir a sequência de “contigs” final.

A estratégia “Clone-By-Clone Shotgun” foi utilizada na sequenciação dos genomas de *Saccharomyces cerevisiae* (Goffeau et al. 1997), de *Caenorhabditis elegans* (The *C. elegans* Sequencing Consortium 1998) e de *Arabidopsis thaliana* (The *Arabidopsis* Genome Initiative 2000), respectivamente, o primeiro eucariota, o primeiro organismo multicelular e a primeira planta, cujos genomas foram sequenciados. Esta estratégia foi também escolhida para a sequenciação do genoma humano no âmbito do “Human Genome Project” (Bentley et al. 1998, Waterston e Sulston 1998), que em 2001 publicou uma sequência inicial do genoma humano (International Human Genome Sequencing Consortium 2001).

A estratégia “Whole-Genome Shotgun” foi utilizada na sequenciação do genoma do procarionta *Haemophilus influenzae* (Fleischmann et al. 1995) e desde então a sua utilização mais comum tem sido na sequenciação de pequenos genomas microbianos (Fraser et al. 2000). Esta foi, no entanto, a estratégia escolhida por uma empresa privada, Celera Genomics, para a sequenciação do genoma humano (Weber e Myers 1997). A principal crítica apontada a esta estratégia era a dificuldade de identificação de “contigs” causada pela elevada quantidade de zonas repetidas e pela grande dimensão dos genomas eucarióticos. No entanto, a utilização conjunta de CBCS e WGS na sequenciação do genoma da *Drosophila* (Adams et al. 2000, Myers et al. 2000), não só comprovou a utilidade de ambas as estratégias como evidenciou a vantagem de uma estratégia híbrida na sequenciação de grandes genomas. Em 2001 foi publicada a sequência inicial do genoma humano obtida pela Celera Genomics em conjugação com os dados obtidos pelo “Human Genome Project” (Venter et al. 2001).

Actualmente no “site” do NCBI (<http://www.ncbi.nlm.nih.gov/genomes/leuks.cgi>) encontra-se informação sobre os genomas totalmente sequenciados de 20 eucariotas. A estratégia de sequenciação utilizada foi a WGS em 10 espécies, a CBCS em 3 espécies e a combinação de WGS e CBCS em 3 espécies. Não existe, no “site”, informação sobre a estratégia utilizada para as restantes 4 espécies.

Os mapas citogenéticos também são enquadrados nas técnicas de mapeamento físico. No início da década de 80, com o mapeamento do gene da insulina por hibridação *in situ* (Harper et al. 1981), deu-se um passo determinante na citogenética: o desenvolvimento de métodos que permitem a localização de sequências específicas de DNA nos cromossomas. Foram desenvolvidos vários métodos de detecção (Carter 1994), no entanto, o mais utilizado é o “Fluorescent In Situ Hybridization” (FISH) (Lichter et al. 1990, Johnson et al. 1991, Boyle et al. 1992). Neste método, a hibridação de uma sonda de DNA marcada com um fluorocromo com os cromossomas intactos permite identificar a posição da sonda no

cromossoma. Apesar da baixa resolução deste tipo de mapeamento, a possibilidade de marcação com diferentes fluorocromos, que permite a análise de várias sondas em simultâneo, confere a esta técnica um enorme potencial de utilização.

2.1.2- MAPEAMENTO GENÉTICO

Um mapa genético é um modelo da ordem linear de um conjunto de genes e/ou marcadores, organizados em grupos de ligamento, que quando relativamente densos e completos deverão corresponder aos diferentes cromossomas da espécie.

Os princípios básicos do mapeamento genético mantêm-se desde os primeiros mapas publicados no início do século XX (Morgan 1912, Sturtevant 1913) e baseiam-se na recombinação de genes e/ou marcadores por “crossingover” durante a meiose. A probabilidade de um evento de recombinação - “crossingover” - ocorrer entre dois loci de um cromossoma depende da sua distância física: quanto mais distantes se encontrarem maior é a probabilidade de recombinarem e não se manterem geneticamente ligados. Assim, a ocorrência de “crossingover” e por conseguinte a recombinação pode ser utilizada como medida da distância física entre loci. A unidade de distância utilizada nos mapas genéticos é o centimorgan (cM), distância na qual ocorre em média um “crossingover” em cada 100 meioses. No entanto, a variação de frequência de recombinação ao longo dos cromossomas não é constante devido a factores como a redução de recombinação na heterocromatina, a possibilidade de ocorrência de “crossingover” duplos e a interferência genética *i.e.* efeito pelo qual um “crossingover” numa região reduz a probabilidade de ocorrência de “crossingovers” nas regiões adjacentes. Este facto levou ao desenvolvimento de equações para converter frequências de recombinação em centimorgans baseadas em diferentes modelos. A equação

de mapeamento mais simples é o modelo linear de Sturtevant (1913) e Morgan (1928) e baseia-se no pressuposto de existir interferência completa *i.e.* quando há um “crossingover” num cromossoma é inibida a ocorrência de outros “crossingovers” nesse cromossoma. As equações que têm sido mais utilizadas foram desenvolvidas por Haldane (1919) e Kosambi (1944) e baseiam-se no pressuposto de que os “crossingovers” ocorrem ao longo dos cromossomas segundo a distribuição de Poisson. A função de Haldane não tem em consideração a interferência genética *i.e.* parte do pressuposto que os “crossingovers” ocorrem aleatoriamente e independentemente. A função de Kosambi por outro lado é baseada num modelo que considera a existência de interferência genética (tabela 1).

Tabela 1. Funções de mapeamento: a mais simples (linear) e as duas mais utilizadas (Haldane e Kosambi).

Função linear	$x = \theta \cdot 100$
Função de Haldane	$x = -\frac{1}{2} \ln (1 - 2 \theta)$
Função de Kosambi	$x = \frac{1}{4} \ln [(1 + 2 \theta) / (1 - 2 \theta)]$

θ = Fração de recombinação (nº recombinantes / nº de meioses)
 x = Distância genética (cM)

Para construir um mapa genético é então necessário analisar a segregação alélica de vários loci numa população, e a partir das frequências de recombinação entre eles estimar as relações de ligamento genético. Ou seja, há três etapas principais na construção de um mapa genético: (1) obtenção da população de mapeamento, (2) análise da segregação dos alelos na população e (3) análise de ligamento entre os loci segregantes.

2.1.2.1- Populações e estratégias de mapeamento genético

Para agrupamento e ordenamento dos marcadores genéticos na construção de mapas genéticos foram propostas várias estratégias experimentais e de análise estatística que se baseiam no estudo da recombinação entre os marcadores analisados em populações obtidas por cruzamentos controlados ou em populações naturais (*p.e.* Reiter et al. 1992, George et al. 1999, Olson e Boehnke 1990, Mester et al. 2003, Vision et al. 2000).

2.1.2.1.1- Populações de cruzamentos controlados

Cruzamentos controlados para obtenção de populações úteis para mapeamento genético podem ser efectuados tanto em plantas como em animais. Em ambos os casos é essencial ter em conta o grau de polimorfismo entre os progenitores. Quanto mais divergentes geneticamente forem os progenitores maior será a probabilidade de identificação de marcadores polimórficos e a segregar na população em número suficiente para cobrir todo o genoma. É também necessário ter em linha de conta aspectos do ciclo de vida de cada espécie, como o número de descendentes que podem ser obtidos e o tempo necessário para os descendentes se tornarem férteis. Nas plantas é ainda importante saber se são autocompatíveis, o que possibilita a obtenção de linhas puras com elevado grau de homozigocidade por autofecundação sucessiva, ou autoincompatíveis, ou sensíveis à endogamia.

As populações de mapeamento geralmente utilizadas em plantas consistem em: populações F2, populações de retrocruzamento (“backcross” - BC), linhas endogâmicas recombinantes (“recombinant inbred lines” - RILs), linhas duplo-haplóides (linhas DH), e populações F1 de progenitores heterozigóticos.

As populações F2 são derivadas da autopolinização de uma planta da descendência F1 ou do cruzamento entre plantas F1 irmãs, sendo geralmente utilizadas duas linhas puras como progenitores iniciais. Os dados de segregação de marcadores codominantes em populações F2 são os que fornecem mais informação acerca do grau de ligamento genético entre marcadores (Liu 1998). Os marcadores dominantes, com uma razão de segregação esperada de 3:1 neste tipo de população, são pouco eficazes na detecção de ligação entre marcadores de cromossomas homólogos (ligação em repulsão) (Reiter et al. 1992, Liu 1998). Por isso, numa população F2, quando são analisados marcadores codominantes e marcadores dominantes é geralmente construído um mapa de referência com os primeiros, ao qual se ligam os segundos. Em *Medicago sativa* (n=8), por exemplo, Kaló et al. (2000) utilizaram uma população F2 de 137 plantas para construir um mapa genético com 868 marcadores distribuídos em 8 grupos de ligamento com 754 cM. Neste trabalho foi primeiro construído um mapa de referência com 121 marcadores codominantes (RFLPs), no qual foram depois integrados os restantes marcadores, na sua maior parte marcadores RAPD dominantes. Quando se analisam apenas marcadores dominantes em populações F2, alguns autores sugerem que se construam dois mapas independentes, um para cada conjunto haplóide de cromossomas, em vez de um mapa integrado (Knapp et al. 1995). Outros autores propõem que os pares de marcadores dominantes estreitamente ligados em repulsão (fracção de recombinação < 0.05) sejam designados por “Trans-Dominant Linked Markers” (TDLM) e recodificados e reanalisados como marcadores codominantes “megalocus”, para servirem de “âncoras” entre cromossomas homólogos (Plomion et al. 1996), uma estratégia que se pode considerar abusiva, pois nem sempre uma baixa recombinação corresponde a uma pequena distância física entre marcadores. Assim, a melhor forma de integrar os grupos de marcadores de cromossomas homólogos na construção de mapas genéticos com marcadores dominantes em populações F2 parece ser a análise de um grande número de loci, pois quantos mais

marcadores forem analisados maior é a probabilidade de serem detectados pares estreitamente ligados em repulsão. Em *Brassica oleracea* (n=9) Farinhó et al. (2004) construíram um mapa genético com 9 grupos de ligamento principais com base na análise de 402 marcadores dominantes e apenas 8 codominantes numa população F2 de 163 plantas.

Uma população de retrocruzamento (“backcross” - BC) é obtida quando um indivíduo F1 é cruzado com um dos seus progenitores (progenitor recorrente) e a proporção de genoma do outro progenitor é reduzida para metade. Neste caso, a razão de segregação esperada para cada marcador dominante ou codominante é de 1:1 (heterozigótico : homozigótico como o progenitor recorrente) e a detecção de ligamento entre marcadores dominantes é tão eficaz como entre marcadores codominantes (Reiter et al. 1992, Liu 1998). As populações BC são por isso ideais para construir mapas genéticos em espécies para as quais o segundo tipo de marcadores não foi ainda desenvolvido ou quando se pretende tirar partido do elevado número de marcadores obtido por reacção nas técnicas de marcadores dominantes baseadas na amplificação aleatória de fragmentos de DNA, como os AFLPs, os ISSRs e os RAPDs. Foi o que fizeram Wang et al. (1997 e 2000) quando utilizaram uma população BC de 60 plantas para construir um mapa genético de AFLPs para o melão (*Cucumis melo*), no qual identificaram marcadores estreitamente ligados ao gene de resistência ao fusarium (*Fom-2*) que se provou serem úteis na selecção assistida por marcadores.

A principal desvantagem das populações F2 e BC é a dificuldade em serem preservadas a longo prazo. A melhor solução para este problema é a utilização de linhas endogâmicas recombinantes (RILs) ou linhas duplo-haplóides (DH). As RILs são descendências homozigóticas de indivíduos de uma população F2 obtidas por autopolinização sucessiva ou por cruzamentos entre irmãos. Como no desenvolvimento de RILs são necessários vários ciclos meióticos até se atingir a homozigose, o nível de recombinação é mais elevado que nas populações F2 e consequentemente os mapas que geram têm maior

resolução que os gerados com estas últimas (Burr e Burr 1991). As linhas DH são produzidas a partir de plantas haplóides por duplicação cromossômica, pelo que são totalmente homozigóticas. Neste tipo de populações de mapeamento (RILs e linhas DH) a informação obtida com marcadores dominantes é igual à obtida com marcadores codominantes, pois praticamente todos os loci (nas RILs) ou todos (nas linhas DH) são homozigóticos (Reiter et al. 1992, Staub e Serquen 1996). Tanto umas como as outras (RILs e linhas DH) constituem populações de mapeamento permanentes que podem ser avaliadas em diferentes locais e em diferentes momentos, ou seja, podem ser mantidas e partilhadas por vários grupos de investigação sem que a sua constituição genética seja alterada e podem ser analisadas para novos marcadores possíveis de serem adicionados aos já mapeados (Reiter et al. 1992, Blanco et al. 1998). As linhas DH são também progenitores ideais para cruzamentos cujo objectivo é produzir populações para mapeamento genético uma vez que não possuem heterozigocidade residual.

Para plantas em que não é possível desenvolver linhas puras ou plantas com ciclos de vida plurianuais em que o processo de homozigotização se torna muito longo, podem ser utilizadas populações de mapeamento F1 obtidas a partir de progenitores heterozigóticos. Tal como referido por Ritter et al. (1990), podem ser utilizados marcadores presentes num progenitor no estado heterozigótico e ausentes no outro que segreguem 1:1 (presença:ausência) na descendência. O primeiro mapa genético de marcadores de DNA baseado neste tipo de segregação foi construído por Gebhardt et al. (1989) para a batateira (*Solanum tuberosum* ssp. *tuberosum*). Posteriormente, Grattapaglia e Sederoff (1994) designaram esta estratégia de mapeamento “pseudo-testcross”. No entanto, estes autores consideraram mais apropriado para o seu próprio trabalho a designação “two-way pseudo-testcross” por serem construídos simultaneamente dois mapas, um para cada progenitor (*Eucalyptus grandis* e *E. urophylla*). O termo “pseudo-testcross” é feliz e foi bem aceite pela

comunidade científica pois o tipo de segregação dos marcadores é igual ao das populações obtidas por “testcross” ou “backcross”. Os mapas construídos para os dois progenitores segundo esta estratégia podem ser integrados com base nos marcadores presentes em heterozigose nos dois progenitores que segreguem na população F1 na proporção 3:1 (presença:ausência). Por exemplo, Lespinasse et al. (2000) construíram dois mapas (um para cada progenitor) em *Hevea* spp. (árvore-da-borracha) segundo a estratégia “Two-way Pseudo-testcross” e integraram-nos num mapa único utilizando como “pontes” os loci heterozigóticos comuns. Pode também seguir-se uma estratégia “one-way pseudo-testcross”, em que se utiliza uma descendência F1 de um cruzamento entre um progenitor heterozigótico e um progenitor homozigótico para construir um mapa genético para o primeiro. Jones et al. (2002a) seguiram esta estratégia utilizando como progenitor homozigótico uma linha duplo-haplóide.

2.1.2.1.2- Populações naturais

Em muitas espécies é difícil, impossível ou eticamente impensável efectuar cruzamentos controlados. No entanto, por vezes é difícil fazer uma distinção clara entre população natural e população de cruzamento controlado, pois em alguns casos, como no Homem, podem seguir-se estratégias de mapeamento em que, apesar de não haver cruzamentos controlados propriamente ditos, é possível analisar marcadores genéticos nos progenitores e proceder à análise da sua segregação na descendência, ou seja, é possível proceder à análise de ligamento por co-segregação.

O avanço da biologia molecular e a possibilidade de analisar inúmeros marcadores no genoma humano levou a que fossem desenvolvidos modelos matemáticos de análise de múltiplos loci que possibilitaram mapear dezenas de marcadores de DNA em pedigrees de 3 gerações (Lander e Green 1987). Esta estratégia de mapeamento foi aplicada com sucesso,

não só na construção do primeiro mapa genético do Homem (Donis-Keller et al. 1987) como em vários outros mapas do genoma humano construídos posteriormente (*p.e.* NIH/CEPH Collaborative Mapping Group 1992, Broman et al. 1998). Em plantas, este tipo de populações foi utilizado, por exemplo, no mapeamento genético de uma espécie de pinheiro, *Pinus taeda*, (Devey et al. 1991) e do choupo, *Populus* spp., (Bradshaw e Stettler 1993), árvores para as quais o processo de obtenção de gerações sucessivas é muito longo. No entanto, nestes trabalhos foram utilizadas populações já existentes de 3 gerações obtidas por polinização livre.

Populações de meio irmãos (“half-sib families”) são outro exemplo em que o mapeamento genético se baseia no estudo de marcadores dos progenitores (neste caso apenas de um deles) e na análise da sua segregação na descendência. Em geral, as populações de meio irmãos utilizadas para mapeamento genético consistem em descendências F1 de um animal macho cujo sémen foi utilizado para fertilizar várias fêmeas (Crawford et al. 1994, Lindgren et al. 1998), ou de uma planta ou de um número limitado de plantas polinizadas por muitas outras desconhecidas ou parcialmente desconhecidas (Grattapaglia et al. 1996, Devey et al. 1995). O princípio desta estratégia de mapeamento genético pode também ser aplicado em populações de megagametófitos haplóides das coníferas, com a diferença que a análise de ligamento se faz directamente na “população” de gâmetas do progenitor em vez de ser na sua descendência (Tulsieram et al. 1992).

Foram também desenvolvidos métodos de mapeamento genético baseados na associação de genes e/ou marcadores em populações naturais de indivíduos não relacionados familiarmente. Neste caso o mapeamento baseia-se na associação ou desequilíbrio de ligamento (Bodmer 1986, Lander and Botstein 1986, Morton 2005). O princípio é simples: em vez de se procurar associação entre marcadores e características em populações de cruzamentos controlados ou indivíduos com parentesco, essa associação é procurada em

indivíduos sem relações de parentesco. Como os indivíduos não relacionados têm sempre alguma relação, por mais distante no tempo que seja, a expressão de uma característica num conjunto de indivíduos dever-se-á a estes terem herdado alelos comuns responsáveis por essa característica, muito provavelmente incluídos em segmentos genómicos também partilhados nos quais podem ser identificados marcadores (Aranzana et al. 2005). A detecção de associação não aleatória ou desequilíbrio de ligamento entre marcadores depois de muitas gerações de cruzamentos aleatórios sugere forte ligamento físico dos marcadores pois o desequilíbrio de ligamento decresce de geração para geração (Kaplan et al. 1995).

No entanto, fenómenos como a estratificação da população, a migração e a mistura de populações, podem levar a falsos resultados positivos ou negativos na detecção de ligamento por associação alélica em populações de indivíduos não relacionados (Gupta et al. 2005, Aranzana et al. 2005). Isto motivou o desenvolvimento de métodos de associação alélica em que, em vez de se estudarem indivíduos não relacionados para mapear genes responsáveis ou associados a determinada doença, se estudam descendências de famílias afectadas pela mesma. O “Transmission Disequilibrium Test” (TDT) é um dos mais utilizados e foi proposto para testar desequilíbrio de ligamento pela avaliação da segregação dos alelos em famílias constituídas pelos dois progenitores e a descendência afectada (Spielman et al. 1993), tendo sido depois alargado para ser aplicado em várias descendências (Martin et al. 1997). O “Haplotype Relative Risk” (HRR) é outro método baseado em estudos de família muito utilizado para mapear genes de doenças, no qual se compara a frequência dos alelos parentais transmitidos ao filho afectado com a frequência dos alelos parentais não transmitidos (Falk e Rubinstein 1987, Terwilliger 1992). Para doenças que se manifestam tarde, para as quais os progenitores não estão disponíveis, foram desenvolvidos vários métodos que usam descendências fenotipicamente discordantes (Curtis 1997, Boehnke e Langefeld 1998, Spielman e Ewens 1998, Horvath e Laird 1998). O PDT (“Pedigree Disequilibrium Test”)

utiliza todos os dados potencialmente informativos, isto é, de pedigrees de várias famílias nucleares e/ou de descendências fenotipicamente discordantes sem progenitores disponíveis (Martin et al. 2000).

2.1.2.2- Marcadores genéticos

A principal ferramenta da genómica é o marcador genético. Este representa variação num ponto particular do genoma e é hereditário, podendo ser seguido de geração em geração.

Desde as descobertas de Mendel até ao início da década de 70 a maior parte dos marcadores analisados e mapeados foram caracteres morfológicos controlados por um único gene, e por isso, apresentando segregação Mendeliana (Tanksley 1983). Nesse caso o fenótipo da característica morfológica pode ser utilizado como indicador alélico do respectivo gene, constituindo um marcador de referência no cromossoma fácil de analisar. No entanto, por vezes a sua expressão é alterada pelo ambiente ou por interacções epistáticas e frequentemente os seus alelos interagem de forma dominante/recessiva não sendo possível distinguir indivíduos heterozigóticos de homozigóticos (Tanksley et al. 1989, Staub e Serquen 1996).

Os marcadores proteicos são outro tipo de marcadores muito utilizados, principalmente no período anterior ao desenvolvimento de marcadores de DNA. O tipo de marcadores proteicos mais utilizado em melhoramento e genética aplicada em plantas são as isoenzimas. A expressão de diferentes alelos produz enzimas com o mesmo substrato que apresentam modificações na composição em aminoácidos e na carga eléctrica (Markert e Moller 1959, Staub e Serquen 1996) e que podem ser utilizadas como marcadores genéticos quando é possível detectar essas modificações. A forma geralmente utilizada para diferenciar os diferentes alelos é a sua separação por electroforese em géis de amido e a utilização da sua

actividade para gerar um produto visível nas zonas do gel para onde migram as diferentes formas alélicas da enzima (Staub e Serquen 1996). Desta forma é possível distinguir entre indivíduos heterozigóticos e homozigóticos, ou seja, as isoenzimas são marcadores codominantes. Um exemplo da aplicação destes marcadores foi o mapeamento de um alelo raro da enzima fosfatase ácida (*Aps-I*) ligado ao gene *Mi* de resistência a nemátodos em tomate - *Lycopersicon esculentum* (Rick e Fobes 1974, Medina-Filho 1980). Este alelo foi utilizado por muitas das principais companhias de sementes europeias e americanas na transferência do gene *Mi* para variedades de tomateiro (Tanksley 1983).

Apesar de bem estabelecidos como classes úteis de marcadores, os caracteres morfológicos e as isoenzimas são pouco numerosos para a construção de mapas genéticos que cubram a totalidade do genoma dos organismos. Este problema foi resolvido com o desenvolvimento das técnicas de marcadores moleculares que permitem seguir a herança de pequenos segmentos das moléculas de DNA que correspondem a partes ínfimas do genoma – os marcadores de DNA. Os primeiros a ser desenvolvidos, os RFLPs (“Restriction Fragment Length Polymorphism”) (Botstein et al. 1980), baseiam-se na restrição do DNA e são codominantes, permitindo a identificação de diferentes alelos num determinado *locus* do mesmo indivíduo. Embora esta técnica implique um longo trabalho prévio de selecção e clonagem de sequências que detectam polimorfismos para criar sondas de DNA e obrigue a procedimentos laboratoriais complexos, foi com ela que foram construídos os primeiros mapas genéticos completos de vários organismos, como por exemplo o Homem (Donis-Keller et al. 1987), o milho - *Zea mays* (Helentjaris 1987), o algodão - *Gossypium* spp. (Reinisch et al. 1994) e a soja - *Glycine max* (Keim et al. 1997).

Com o desenvolvimento da PCR (“Polymerase Chain Reaction”) (Mullis et al. 1986) foram potenciadas uma série de novas técnicas de marcadores de DNA de utilização muito mais simples e que requerem muito menos quantidades de DNA do que a técnica de RFLPs.

Um são baseadas na amplificação de fragmentos específicos com "primers" desenhados a partir de sequências conhecidas, como é o caso dos SSRs ("Simple Sequence Repeats"), enquanto outras, como os RAPDs ("Random Amplified Polymorphic DNAs"), os AFLPs ("Amplified Fragment Length Polymorphisms") e os ISSRs ("Inter Simple Sequence Repeats"), se baseiam na amplificação de vários fragmentos em simultâneo. Estas últimas têm a vantagem de gerar um elevado número de marcadores distribuídos pelo genoma, permitindo a construção de mapas genéticos de elevada densidade num curto período de tempo.

Os SSRs ou microssatélites são segmentos de DNA constituídos por sequências de 1 a 6 nucleótidos repetidos sequencialmente, cujas regiões adjacentes são geralmente conservadas entre génotipos da mesma espécie. Estas sequências podem ser usadas como marcadores codominantes pois numa amplificação de PCR com primers que as flanqueiam são gerados polimorfismos de tamanho para diferentes alelos em que o número de repetições do motivo de nucleótidos varia (Weber e May 1989, Litt e Luty 1989, Tantz 1989). A identificação de loci microssatélites em espécies para as quais não existem muitos dados genómicos requer um longo trabalho prévio de selecção e sequenciação. Apesar disso, o seu elevado conteúdo informativo, directamente relacionado com o número de alelos em cada *locus*, juntamente com o facto de existirem em grande número nos genomas dos eucariotas, com a facilidade de automação devido a serem obtidos por PCR, e com a fiabilidade e reprodutibilidade que lhes é reconhecida por serem amplificados com "primers" longos específicos, fizeram dos SSRs um dos tipos de marcadores mais utilizados em trabalhos de mapeamento genético. Por exemplo por Saal et al. (2001) em *Brassica* spp. e Chen et al. (1997) em *Oryza sativa*.

Os RAPDs, os primeiros marcadores baseados no PCR que não requerem informação de sequências de DNA a ser desenvolvidos, são geralmente marcadores dominantes, podendo no entanto ser analisados como codominantes quando fragmentos de DNA amplificados do mesmo *locus* e de diferentes tamanhos estão estritamente ligados em repulsão (Willians et al.

1990). Comparativamente às técnicas de RFLPs e SSRs, a técnica de RAPDs permite analisar um maior número de loci por ensaio e não necessita de trabalho prévio de selecção e clonagem de segmentos de DNA a sequenciar. Estes aspectos possibilitaram um significativo avanço na rapidez de construção de mapas genéticos, tendo os marcadores RAPD sido amplamente usados para esse fim. Exemplos disso são os trabalhos de Grattapaglia e Sederoff (1994) em *Eucalyptus* spp. e Chaparro et al. (1994) em *Prunus* spp.. No entanto, a estes marcadores foram apontadas desvantagens como sensibilidade a condições experimentais e consequentemente reduzida reprodutibilidade entre laboratórios (Jones et al. 1997), tendo-se mantido o interesse no desenvolvimento de novos tipos de marcadores de DNA.

A técnica de AFLPs de certa forma combina as estratégias de RFLPs e PCR pois baseia-se na amplificação selectiva de fragmentos obtidos pela restrição de DNA genómico total (Vos et al. 1995). No entanto, tal como nos RAPDs, é evitada a hibridação de sondas, não sendo necessário trabalho prévio de selecção e clonagem de sequências, aspectos que tornam os RFLPs bastante trabalhosos. Para além disso, por um lado, a utilização de "primers" longos e temperaturas de "annealing" mais elevadas tornam a técnica de AFLPs mais fiável que a técnica de RAPDs (Jones et al. 1997), e por outro, a utilização géis de poliacrilamida com maior poder de resolução permite a análise de um número muito maior de marcadores por reacção. Estas características têm levado a uma crescente utilização dos AFLPs pois tornam a técnica ideal para construir mapas genéticos de elevada densidade em relativamente pouco tempo (*p.e.* Qi et al. 1998). Os marcadores AFLP podem ser dominantes, devido à presença ou ausência de sítios de restrição entre genomas, ou codominantes devido a diferenças de tamanho de fragmentos amplificados. No entanto, as relações alélicas entre bandas polimórficas codominantes são raras e como sublinhado por Rafalski et al. (1996) podem não ser imediatamente óbvias.

Os ISSRs, cada vez mais utilizados na construção de mapas genéticos (*p.e.* Kojima et al. 1998, Casasoli et al. 2001), são outros marcadores que não requerem trabalho prévio de sequenciação. ISSRs são fragmentos amplificados por PCR com primers complementares de sequências repetidas, ou seja, são regiões entre microssatélites (Zietkiewicz et al. 1994). Apesar dos “primers” utilizados serem longos, o que confere elevada reprodutibilidade à técnica, como estes são constituídos por motivos microssatélites geralmente em elevado número nos genomas dos eucariotas, não amplificam marcadores de loci específicos, e pelo contrário, marcadores de vários loci. Esta técnica pode revelar polimorfismos resultantes da variação de microssatélites individuais cuja presença ou ausência difere entre genomas (loci dominantes) e resultantes da variação nas regiões entre microssatélites conservados (loci codominantes) (Rafalski et al. 1996). No entanto, tal como nas técnicas de RAPDs e de AFLPs, a identificação de marcadores ISSRs codominantes não é muito frequente.

A partir de marcadores cujas sequências são desconhecidas (amplificados por “primers” arbitrários) podem ser desenvolvidos marcadores específicos de loci únicos - SCARs (“Sequence Characterized Amplified Regions”) – amplificados com pares de “primers” longos (~20 pb). Estes marcadores, de fácil análise e elevada reprodutibilidade, permitem preencher a falha entre a relativa facilidade de obtenção de marcadores gerados com “primers” arbitrários e a eficiência de utilização dos mesmos. Originalmente derivados de marcadores RAPD clonados e sequenciados, foram desenvolvidos para resolver a necessidade de identificar de forma inequívoca um único *locus*, quando utilizados na clonagem de genes baseada no mapeamento, e para eliminar a sensibilidade a alterações nas condições de reacção dos mesmos (Paran e Michelmore 1993). Posteriormente foram também desenvolvidos SCARs a partir de marcadores AFLP (Negi et al. (2000) e a partir de ISSRs (Capparelli et al. 2004).

O crescente esforço em trabalhos de sequenciação e conseqüentemente a cada vez maior quantidade de seqüências disponíveis em bases de dados públicas e privadas permitiu o desenvolvimento de uma série de marcadores que podem ser utilizados em trabalhos de mapeamento genético, tanto derivados de STSs (“Sequence Tagged Sites”) como de ESTs (“Expressed Sequence Tags”). Os STSs, propostos por Olson et al. (1989) como referências nos cromossomas do genoma humano, são pequenos troços de DNA (~300 pb) usados no mapeamento físico para ligar grandes fragmentos de DNA clonados, em que clones contendo o mesmo STS devem sobrepor-se. Usando STSs como sondas, se forem detectados polimorfismos que permitam o seu mapeamento genético, podem estabelecer-se pontos de ancoragem entre mapas genéticos e mapas físicos (Weissenbach et al. 1992, Gyapay et al. 1994). Os ESTs são obtidos de clones de cDNA, sendo por isso derivados de mRNA e em geral representam genes funcionais. Tal como os STSs, podem ser utilizados como sondas em análises RFLP para mapeamento genético. Para além disso, foram desenvolvidos métodos mais eficientes para analisar STSs ou ESTs baseados na amplificação por PCR com “primers” específicos seguida de: (1) electroforese simples quando existe polimorfismo no tamanho do produto PCR ou (2) análise da restrição enzimática do fragmento amplificado - CAPs (“Cleavage Amplified Polimorphisms”), análise heteroduplex ou análise SSCP (“Single-Stranded Conformational Polymorphism”), quando o polimorfismo se resume a alterações na seqüência (Fischer e Lerman 1983, Orita et al. 1989, Konieczny e Ausubel 1993, Prior 1994, Harry et al. 1998, Schafer e Hawkins 1998, Sato e Nishio 2002, 2003). Foi ainda desenvolvido um método para converter ESTs em marcadores úteis para mapeamento baseado na amplificação por PCR das regiões imediatamente adjacentes (5’ ou 3’) às regiões codificantes dos genes (Cato et al. 2001). No trabalho de Gupta e Rustgi (2004) encontra-se uma boa revisão sobre o desenvolvimento de marcadores moleculares a partir de ESTs e a sua

aplicação em plantas. Por exemplo, Lan et al. (2000) utilizaram marcadores ESTs num trabalho de mapeamento comparativo entre *Brassica oleracea* e *Arabidopsis thaliana*.

Os SNPs (“Single Nucleotide Polimorphisms”), talvez o melhor exemplo de polimorfismos que não podem ser detectados pela análise directa dos produtos amplificados em electroforese simples, são qualquer diferença entre dois genomas baseada na substituição de uma base ou em pequenas inserções ou deleções (in-dels) (Brookes 1999), constituindo o tipo de variação mais frequente nos genomas, tanto de animais como de plantas. Na comparação de dois genomas humanos estima-se que há em média um SNP por cada 1000 nucleótidos (Wang et al. 1998b, Halushka et al. 1999). Em *Arabidopsis thaliana* ocorre em média um SNP num intervalo de menos de 300 pb de fragmento genómico aleatório entre qualquer entrada e o genoma da entrada Columbia (*Col-0*) (Schmid et al. 2003). As técnicas de RFLPs, SCARs CAPs, análise heteroduplex e SSCPs anteriormente referidas podem ser utilizadas na análise de SNPs. No entanto, o potencial de utilização dos SNPs é maximizado em técnicas que permitem analisar um elevado número de sequências em simultâneo como a VDAs (“Variant Detector Arrays”), baseada na hibridação de produtos de PCR com oligonucleótidos ordenados num “chip” (Wang et al. 1998b; Halushka et al. 1999). Para isso é necessário um extenso trabalho prévio de sequenciação genómica como acontece com o “The SNP Consortium” (<http://snp.cshl.org/>), um dos participantes no projecto “International HapMap Project” que visa desenvolver um poderoso mapa de SNPs do genoma humano que descreva os padrões comuns de variação da sequência de DNA (The International HapMap Consortium 2003). Em plantas, um dos primeiros mapas genómicos de SNPs foi construído para o arroz (*Oryza sativa*, *Oryza rufipogon*) por Nasu et al. (2002).

2.1.2.3- Análise de ligamento

Depois de analisada a co-segregação de marcadores numa população de mapeamento, a construção de mapas genéticos envolve três passos principais: (1) o agrupamento dos marcadores em vários grupos de ligamento, (2) a ordenação dos marcadores de cada grupo e (3) a estimativa das distâncias entre marcadores do mesmo grupo. A base de qualquer destes passos é o resultado da análise de ligamento entre todos os pares de marcadores.

O método de análise de ligamento com base na co-segregação de loci em populações de cruzamentos controlados mais utilizado é o da máxima probabilidade inicialmente desenvolvido por Mather (1938), em que o ligamento entre marcadores é calculado utilizando fracções de probabilidade de ligamento *versus* de não ligamento. Estas fracções são geralmente expressas na forma logarítmica e são conhecidas por valores LOD (“logarithm of odds”) (Morton 1955, Risch 1992). Por exemplo, um valor LOD de 3 entre dois marcadores indica que a probabilidade destes estarem ligados é 1000 vezes superior (1000:1) à probabilidade de não estarem ligados.

Depois de calculados os valores LOD para todos os pares de marcadores, o agrupamento faz-se segundo um processo experimental de definição de critérios como limites de LOD mínimo de ligamento e/ou de recombinação máxima entre marcadores, de forma a ser obtido um padrão de agrupamento com suporte estatístico aceitável. Estabelecidos os grupos de ligamento, vários métodos estatísticos podem ser utilizados para resolver o ordenamento de marcadores. Os métodos geralmente utilizados são o “minimum sum of adjacent recombination fractions (SARF) (Falk 1989), o “minimum product of adjacent recombination fractions (PARF) (Wilson 1988) e o “minimum obligatory crossovers” (MOC) (Thompson 1987). Finalmente, para calcular as distâncias genéticas entre marcadores também foram desenvolvidos vários métodos. Entre os mais utilizados estão o algoritmo EM

(“expectation-maximization”) (Lander e Green 1987) e o método dos mínimos quadrados (Jensen e Jorgensen 1975 a e b, Stam 1993).

Todos estes cálculos seriam uma tarefa interminável se fossem efectuados manualmente para um elevado número de marcadores. No entanto, a par da evolução das técnicas de marcadores moleculares, que permitiram a análise de um elevado número de loci, e beneficiando do “boom” da informática das últimas duas décadas, foram desenvolvidos programas informáticos específicos que facilitaram e aceleraram a análise de ligamento entre muitos marcadores. O MAPMAKER (Lander et al. 1987) e o JoinMap (Stam 1993, Van Ooijen e Voorrips 2001) são dois dos programas informáticos mais utilizados na construção de mapas genéticos.

2.1.2.4- Qualidade dos mapas genéticos

A qualidade de um mapa pode ser avaliada, por um lado, pela confiança no agrupamento e ordem dos marcadores, e por outro, pela percentagem de genoma que representa *i.e.* a cobertura do mapa. No primeiro há que considerar os critérios utilizados na construção dos mapas: obviamente que quanto mais elevados são os valores limites estatísticos LOD de agrupamento e LOD de ordenamento, maior é a confiança nos mapas obtidos. No segundo há que considerar dois aspectos: um é a proporção de genoma coberta pelo comprimento total do mapa, o outro é a densidade do mapa *i.e.* o tamanho dos intervalos entre os marcadores. Por outro lado, à medida que aumenta a cobertura de um mapa genético, o número de grupos de ligamento aproxima-se do número de cromossomas e o número de marcadores não ligados aproxima-se de zero.

No mapeamento genético de espécies com muitos cromossomas, como é o caso do ananaseiro ($n = 25$), é comum que o número de grupos de ligamento seja relativamente

superior ao número de cromossomas, sobretudo devido ao elevado número de pequenos grupos que se formam. Este fenómeno verificou-se em vários mapas construídos para diferentes espécies vegetais, com diferentes tipos de marcadores e com diferentes tipos de populações de mapeamento, como por exemplo: no algodão (n=26) em que foram obtidos 41, 43, 37 e 52 grupos de ligamento, respectivamente por Reinisch et al. (1994), Zhang et al. (2002), Lacape et al. (2003) e Waghmare et al. (2005) e na soja (n=20) onde foram estabelecidos 28 grupos de ligamento por Keim et al. (1997) e 35 por Ferreira et al. (2000) e por Matthews et al. (2001).

Nestes casos coloca-se o problema de como identificar os grupos de ligamento que representam diferentes cromossomas.

No algodão esse problema foi resolvido com mapas construídos a partir de cruzamentos interespecíficos entre linhas com o mínimo de introgressão recíproca e com a análise de marcadores dos mapas em diferentes linhas de plantas derivadas das linhas utilizadas como progenitores, em que, para cada uma, um dos pares de cromossomas homólogos de uma das espécies envolvidas fora substituído por um simples cromossoma ou apenas um braço de cromossoma do correspondente par da outra espécie (Reinisch et al. 1994, Liu et al. 2000, Zhang et al. 2002).

Outra solução é o alinhamento de mapas de diferentes indivíduos, através da análise de marcadores comuns. Por exemplo, no mapeamento genético de soja o alinhamento de mapas construídos com diferentes populações permitiu a identificação de 20 grupos ou conjuntos de grupos de ligamento que representam os 20 cromossomas da espécie (Cregan et al. 1999, Song et al. 2004). Outro exemplo é o trabalho de Yin et al. (2002) no choupo (n = 19). Estes autores, utilizando uma população “Pseudo testcross” de um cruzamento interespecífico construíram um mapa com 31 grupos de ligamento para um progenitor e com 34 para o outro,

e através do alinhamento dos dois mapas identificaram quais os grupos e conjuntos de grupos que representam os 19 cromossomas.

A construção dos mapas genéticos com base em populações de mapeamento muito numerosas parece ser, no entanto, a melhor forma de solucionar o problema. Estas permitem a detecção de ligamento entre marcadores fisicamente mais afastados e consequentemente com fracções de recombinação mais elevadas e também uma maior definição no cálculo das distâncias genéticas com base nas taxas de recombinação, comparativamente às populações pouco numerosas. Por exemplo, ainda no choupo, Yin et al. (2004) construíram um mapa genético com base numa população “backcross” de um cruzamento interespecífico, em que todos os marcadores analisados (439 AFLPs e 105 SSRs) ficaram ligados num número de grupos de ligamento igual ao número de cromossomas (19). Quando se compara o trabalho destes autores com outros trabalhos de mapeamento em espécies com muitos cromossomas e baseados também em populações “backcross”, ou populações F1 com o mesmo tipo de segregação, mas em que o número de grupos de ligamento obtido é superior ao número de cromossomas, a principal diferença é a utilização de uma população de mapeamento significativamente mais numerosa (180 plantas vs 58, 71, 80, 93, 94 e 120 plantas nos exemplos da tabela 2).

Tabela 2. Características de mapas construídos para algumas espécies de plantas com um elevado número de cromossomas.

Espécie	Número de cromossomas (n)	Tipo de população (número de plantas)	Número e tipo de marcadores mapeados	Número Total de grupos	Número de grupos com 2 ou 3 marcadores	Número de marcadores não ligados	Compr. do mapa (cM)	Estimativa de compr. do genoma (cM)	Referência	
Soja	<i>Glycine max</i> (cv. Noir x NIL BARC-2)	20	F2 (149)	207 loci (RFLPs, SSRs, AFLPs e RAPDs)	35	-	-	1400	-	Matthews et al. 2001
	<i>Glycine max</i> (BSR-101 x PI437,654)	20	RIL (300 para RFLPs) (42 para RAPDs e AFLPs)	165 RFLPs, 25 RAPDs e 650 AFLPs	28	-	-	3441	-	Keim et al. 1997
	<i>Glycine max</i>	20	RIL (330 para RFLPs) (? para RAPDs)	250 RFLPs e 106 RAPDs	35	-	-	-	3275	Ferreira et al. 2000
Kiwwi	<i>Actinidia chinensis</i> x <i>Actinidia callosa</i>	29	"Pseudo testcross" F1 (94)	80 SSRs e 123 AFLPs	38	14	22	1758,5	3849	Testolin et al. 2001
		29		31 SSRs e 111 AFLPs	30	7	25	1104,1	3799	
"Quinoa"	<i>Chenopodium quinoa</i> (Ku-2 x NL-6)	18	F2 (80)	19 SSRs, 230 AFLPs e 6 RAPDs	35	13	38	1020	1700	Maughan et al. 2004
Choupo	<i>Populus tomentosa</i> x <i>Populus bolleana</i>	19	"Pseudo testcross" F1 (120)	396 AFLPs	60	35	74	2682,8 *	3097	Zhang et al. 2004
		19		236 AFLPs	40	21	49	1956,3 *	2552	
	<i>Populus deltoides</i> x <i>Populus euramericana</i>	19	"Pseudo testcross" F1 (93)	85 RAPDs, 219 AFLPs e 6 ISSRs	31	8	27	3801	-	Yin et al. 2002
		19		68 RAPDs, 162 AFLPs e 11 ISSRs	34	8	18	3452	-	
	<i>Populus adenopoda</i> x <i>Populus alba</i>	19	"Pseudo testcross" F1 (80)	82 RAPDs	19	13	20	552,9	2104	Yin et al. 2001
		19		212 RAPDs	24	5	15	2299,7	2632	
<i>Populus deltoides</i> (I-69 x I-63)	19	"backcross" F1 (93)	259 AFLPs	43	24	69	2927 *	-	Wu et al. 2000	
Algodão	<i>Gossypium hirsutum</i> x <i>G. barbadense</i>	26	F2 (57)	683 RFLPs	41	6	22	4675	5125	Reinisch et al. 1994
	<i>Gossypium hirsutum</i> x <i>G. barbadense</i>	26	Haplóides e duplo-haplóides F1 (58)	489 (SSRs e RAPDs)	43	11	135	3314,5	-	Zhang et al. 2002
	<i>Gossypium hirsutum</i> x <i>G. barbadense</i>	26	"backcross" (71)	465 AFLPs, 229 SSRs, 192 RFLPs e 2 morfológicos	37	5	126	4400	5500	Lacape et al. 2003
	<i>Gossypium hirsutum</i> x <i>G. tomentosum</i>	26	F2 (82)	589 RFLPs com sondas de cDNA	52	8	26	4259,4	4979,4	Waghmare et al. 2005

* Excluindo grupos de dois e três marcadores.

2.1.2.5- Importância e aplicações dos mapas genéticos

Foram já construídos mapas genéticos tanto para espécies modelo, como o rato - *Mus musculus* (Dietrich et al. 1996) e a *Arabidopsis thaliana* (Alonso-Blanco et al. 1998), como para espécies importantes na alimentação do Homem, como o porco - *Sus scrofa* (Rohrer et al. 1996), a galinha - *Gallus gallus* (Groenen et al. 1998) e a ovelha - *Ovis aries* (Maddox et al. 2001), nos animais, ou o arroz - *Oriza sativa* (Mccough et al. 1988), a batata - *Solanum tuberosum* (Tanksley et al. 1992), a soja - *Glycine max* (Keim et al. 1997) e o trigo - *Triticum aestivum* (Messmer et al. 1999), nas plantas, e obviamente também para o Homem (Murray et al. 1994). No que diz respeito aos frutos, no grupo dos vinte mais produzidos no mundo segundo estatísticas da FAO no ano 2004 (<http://apps.fao.org>), só não foram publicados mapas genéticos para quatro (tabela 3).

Para além do interesse *per se* no estudo da estrutura dos genomas, os mapas genéticos podem ser utilizados na construção de mapas físicos *i.e.* na ordenação de clones de bibliotecas de fragmentos de DNA genómico. Esta estratégia foi inicialmente aplicada no desenvolvimento de mapas físicos de YACs, (Chumakov et al. 1995, Kurata et al. 1997, Canillieri et al. 1998) e posteriormente no mapeamento físico de BACs (Hoskins et al. 2000, Chen et al. 2002). A identificação de marcadores do mapa genético em troços longos de DNA clonados, por amplificação por PCR (Adam-Blondon et al. 2005, Morales et al. 2005) ou por hibridação em membranas (Aign et al. 2001, Wu et al. 2004), permite a sua ancoragem ao mapa genético e desta forma o estabelecimento de uma ordem linear de clones, o mapa físico. Desta forma os mapas genéticos podem contribuir para a selecção de clones com o mínimo de sobreposição em programas de sequenciação completa de genomas. Um bom exemplo foi a integração extensiva de um mapa físico de BACs com um mapa genético de elevada

densidade que constituiu uma referência para a sequenciação total do genoma do arroz (Chen et al. 2002).

Tabela 3. Referências de mapas genéticos das 20 fruteiras mais produzidas no mundo segundo estatísticas da FAO para o ano 2004.

Fruteiras		Referências	
1) Citrinos	<i>Citrus</i> spp.	Jarrell et al. (1992) Durham et al. (1992) Cai et al. (1994) Luro et al. (1996) Kijas et al. (1997)	Simone et al. (1998) Garcia et al. (1999) Cristofani et al. (1999) Roose et al. (2000) Ruiz e Asins (2003)
2) Bananeira	<i>Musa</i> spp.	Fauré et al. (1993)	
3) Melancia	<i>Citrullus lanatus</i> var. <i>lanatus</i>	Hashizume et al. (1996) Levi et al. (2002)	Hashizume et al. (2003)
4) Videira	<i>Vitis</i> spp.	Lodhi et al. (1995) Dalbó et al. (2000) Grando et al. (2003)	Riaz et al. (2004) Douceff et al. (2004) Adam-Blondon et al. (2004)
5) Macieira	<i>Malus</i> spp	Hemmat et al. (1994) Conner et al. (1997) Seglias e Gessler (1997)	Maliepaard et al. (1998) Liebhard et al. (2002) Liebhard et al. (2003)
6) Meloeiro	<i>Cucumis melo</i>	Baudracco-Arnas et al. (1996) Wang et al. (1997) Oliver et al. (2001)	Silberstein et al. (2003) Fazio et al. (2003) Gonzalo et al. (2005)
7) Manguieira	<i>Mangifera indica</i>	Chunwongse et al. (2000)	Kashkush et al. (2001)
8) Pereira	<i>Pyrus</i> spp.	Iketani et al. (2001)	Yamamoto et al. (2002)
9) Pessegueiro e nectarineiro	<i>Prunus</i> subg. <i>Amygdalus</i>	Chaparro et al. (1994) Rajapakse et al. (1995) Foolad et al. (1995) Dirlewanger et al. (1996) Dirlewanger et al. (1998) Lu et al. (1998) Joobeur et al. (1998)	Dettori et al. (2001) Bliss et al. (2002) Foulongne et al. (2003) Aranzana et al. (2003) Dirlewanger et al. (2004) Verde et al. (2005)
10) Ananaseiro	<i>Ananas comosus</i>	Carlier et al. (2004)	Carlier e Leitão (2006)
11) Ameixeira e damasqueiro	<i>Prunus</i> subg. <i>Prunus</i>	Hurtado et al. (2002) Salava et al. (2002)	Vilanova et al. (2003)
12) Tamareira	<i>Phoenix dactylifera</i>		-
13) Papaieira	<i>Carica papaya</i>	Ma et al. (2004)	Sondur et al. (1996)
14) Morangueiro	<i>Fragaria</i> spp.	Davis et al. (1997) Lerceteau-Köhler (2003)	Sargent et al. (2004)
15) Abacateiro	<i>Persea americana</i>	Sharon et al. (1997)	
16) Cerejeira	<i>Prunus</i> subg. <i>Cerasus</i>	Stockinger et al. (1996)	Wang et al. (1998a)
17) Diospireiro	<i>Diospyros</i> spp.		-
18) Cajueiro	<i>Anacardium occidentale</i>		-
19) Figueira	<i>Ficus</i> spp.		-
20) Kiwizeiro	<i>Actinidia</i> spp.	Testolin et al. (2001)	Fraser et al. (2004)

Desde meados da década de 80 que a selecção assistida por marcadores (SAM) é uma estratégia promissora no que diz respeito ao aumento da eficiência dos programas de melhoramento de plantas e animais. A utilidade desta estratégia não levanta dúvidas, principalmente nos casos em que a característica de interesse é de difícil avaliação, como quando são necessários testes laboratoriais ou ensaios biológicos longos, ou nos casos em que o programa de melhoramento envolve vários genes, genes recessivos ou genes de expressão tardia. Até finais da década de 90, apesar de terem sido identificados marcadores ligados a uma grande quantidade e variedade de genes, foram raras as publicações que descreveram casos de sucesso de aplicação prática da SAM (Young 1999). Um dos poucos exemplos foi a transferência de genes de tomateiros selvagens para a espécie cultivada visando o aumento de tamanho do fruto (Tanksley et al. 1996). No entanto, mais recentemente têm sido publicados vários casos de sucesso de aplicação da SAM, principalmente a descrever trabalhos de introgressão de genes ou regiões do genoma (QTLs) através de retrocruzamentos sucessivos (Charcosset e Moreau 2004; Francia et al. 2005). Por exemplo, Bouchez et al. (2002) e Concibido et al. (2003) descreveram a introgressão de segmentos de cromossoma portadores de QTLs favoráveis em linhas de elite através de retrocruzamentos com selecção assistida por marcadores, respectivamente, no milho (QTL para precocidade e rendimento em grão) e em soja (QTL para rendimento em grão).

A clonagem baseada em mapas genéticos (CBM) refere-se à utilização de marcadores estreitamente ligados a características de interesse para localizar os respectivos genes na sequência genómica e em seguida cloná-los e caracterizá-los. Com esse objectivo os marcadores são sequenciados e utilizados como ponto de partida para um processo de avanço contínuo (“chromosome walking”) convergente até ao gene alvo. Para tal é necessário construir uma biblioteca de fragmentos de DNA longos e, utilizando como referência os marcadores que flanqueiam o gene de interesse, fazer um mapa físico dessa região do

genoma. Após sequenciação total de um trecho longo de DNA clonado que contenha o gene alvo, este é identificado entre todos os genes que são codificados nesse trecho via transformação genética. O primeiro gene de plantas a ser isolado por clonagem baseada no mapeamento foi o gene *fad3* que controla a desnaturação do ácido gordo omega-3 na *Arabidopsis* (Arondel et al. 1992). Depois, utilizando esta estratégia, foram clonados genes de resistência em plantas, como são os casos do *Pto* (Martin et al. 1993) e *I2* (Simmons et al. 1996) no tomate - *Lycopersicon esculentum*, do *RPS2* (Bent et al. 1994) e *RPM1* (Grant et al. 1995) na *Arabidopsis*, de um gene de resistência a nemátodos na beterraba - *Beta vulgaris* (Cai et al. 1997), do *Mlo* na cevada - *Hordeum vulgare* (Büschges et al. 1997), e outros tipos de genes, como o *Rfl* (“fertility restorer locus”) no sorgo - *Sorghum bicolor* (Klein et al. 2005) e o *htd1* (“for high tillering and dwarf”) no arroz - *Oryza sativa* (Zou et al. 2005). Em teoria o marcador ideal para a clonagem baseada no mapeamento é um marcador cuja distância ao gene alvo é inferior ao tamanho médio dos fragmentos clonados da biblioteca de DNA genômico donde se espera isolar o gene. Nesse caso pode-se evitar o “chromosome walking” e fazer “chromosome landing”, identificando-se diretamente o clone que contém o gene alvo (Tanksley et al. 1995, Ballvora et al. 2001).

Os mapas genéticos têm ampla utilização em estudos de mapeamento comparativo. O estudo comparativo de diferentes genomas da mesma espécie ou de diferentes espécies do ponto de vista da homologia de genes e da conservação das distâncias e da ordem de genes e/ou marcadores nos cromossomas, contribui, por um lado, para o conhecimento sobre a evolução das espécies (Schmidt 2002), e por outro, para a obtenção de mapas comparativos de referência, pelo menos ao nível de famílias taxonômicas (Yu et al. 2004). Para além disso, as relações de sintenia podem ajudar a prever a posição de sequências ortólogas em espécies próximas. A identificação de loci ou haplótipos via genética comparativa é uma estratégia que já provou ser útil e cujo potencial aumenta à medida que aumentam os dados genômicos

disponíveis. Por exemplo, Hackauf e Wehling (2005), utilizando ESTs de cevada e de trigo e informação dos mapas genéticos e da sequência genómica do arroz, identificaram 12 novos marcadores STS a flanquear o *locus* de autoincompatibilidade (*locus Z*) no centeio.

2.1.3- GENÓMICA ESTRUTURAL NOS PRINCIPAIS FRUTOS TROPICAIS

A maior parte dos estudos na área da genómica estrutural foi efectuada em culturas de climas temperados. Das 26 espécies de plantas superiores para as quais há programas de sequenciação genómica em curso, apenas duas correspondem a culturas tropicais: *Musa acuminata* (bananeira) e *Carica papaya* (papaieira) (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj>). Estas correspondem, respectivamente, ao primeiro e quarto frutos tropicais mais produzidos no mundo (FAO 2004). Para a mangueira (*Mangifera indica*), a segunda fruteira tropical com maior produção mundial, com a construção de mapas genéticos foram dados os primeiros passos neste ramo da genómica. O mesmo se passa para o ananaseiro (*Ananas comosus*), a terceira fruteira tropical com mais produção mundial, com a construção dos mapas genéticos descrita neste trabalho.

A bananeira, a segunda fruteira com maior produção mundial e a primeira entre as tropicais (FAO 2004), pertence ao género *Musa*, secção *Eumusa* ($x = 11$), e inclui clones diplóides ($2n = 2x$) e poliplóides ($3x, 4x$). O tamanho do genoma haplóide (1C) da bananeira foi estimado por citometria de fluxo em cerca de 600 Mpb (Doležel et al. 1994). Na área da citogenética, estudos com a técnica de FISH permitiram a discriminação de cromossomas dos quatro genomas envolvidos nas variedades de bananeira (*Musa acuminata*, *M. balbisiana*, *M. schizocarpa* e *M. textilis*) (Osuji et al. 1997, D'Hont et al. 2000, D'Hont 2005) e a

localização física de genes ribossomais (Doleželová et al. 1998; Osuji et al. 1998), de sequências homologas a elementos genéticos móveis como retrotransposições (Balint-Kurti et al. 2000), “*Copia*-like elements” (Baurens et al. 1997), e de sequências de “baddnavirus” integradas no genoma da bananeira (Harper et al. 1999). Mais recentemente a técnica de FISH foi utilizada para localizar várias sequências repetitivas de DNA agrupadas nas regiões dos centrómeros (Valárik et al. 2002) e para demonstrar que as extremidades dos cromossomas de espécies de *Musa* consistem em motivos teloméricos, tipo-*Arabidopsis*, repetidos (Osuji et al. 1998). Embora tenha sido construído no início da década de noventa um mapa de baixa densidade com isoenzimas, RFLPs e RAPDs (Fauré et al. 1993), ainda não foi publicada nenhuma versão mais densa desse ou de qualquer outro mapa. No entanto, no âmbito do programa do “*Musa* Genomics Consortium” (MGC), formado em 2001 (<http://www.musagenomics.org>), está a ser utilizada uma população segregante para construir um mapa genético de elevada densidade com marcadores RFLPs, AFLPs e SSRs, no qual serão localizados genes de resistência ao fungo *Fusarium*. Por outro lado, foram já desenvolvidas bibliotecas de BACs para *M. acuminata* (Vilarinhos et al. 2003, Ortiz-Vázquez et al. 2005) e para *M. balbisiana* (Šafář et al. 2004), e no âmbito do MGC está em curso o desenvolvimento de um mapa físico para o genoma da bananeira e a sequenciação completa dos seus 11 cromossomas. Investigadores deste consórcio estão também a estabelecer várias populações segregantes adicionais para serem utilizadas na identificação de genes de resistência a outras doenças e pestes.

A papaieira (*Carica papaya*) é a décima terceira fruteira com maior produção mundial e a quarta entre as tropicais (FAO 2004). É uma espécie diplóide com $n = 9$ cromossomas e um genoma de 372 Mpb/1C (Arumuganathan e Earle 1991). Foram já construídos vários mapas genéticos e desenvolvida uma biblioteca de BACs, mas não há ainda publicações que refiram a localização física de marcadores ou genes nos cromossomas por técnicas de

hibridação *in situ*. Sondur et al. (1996), utilizando 62 marcadores RAPD, desenvolveram um mapa genético da papaieira e mapearam o gene que determina o sexo. Os mesmos autores mapearam QTLs para a altura da planta, para o diâmetro do caule e para o número do nó em que se desenvolve a primeira flor (Sondur et al. 1995). Mais recentemente, a partir de 54 plantas F2 de um cruzamento entre as cultivares Kapoho e SunUp, foi construído um mapa genético de elevada densidade com 1498 marcadores AFLP, um marcador do gene da proteína da cápsula do vírus “ringspot” (introduzido por engenharia genética) e dois marcadores morfológicos (tipo de sexo e cor da polpa do fruto), distribuídos em 12 grupos principais que cobrem 3294.2 cM com uma distância média entre marcadores de 2.2 cM (Ma et al. 2004). Com este mapa genético e com a biblioteca de BACs da papaieira, desenvolvida por Ming et al. (2001), foram criadas as bases para a sequenciação total do genoma desta fruteira. No ano 2004 um consórcio de várias instituições iniciou um programa que estima completar a sequência do genoma da papaieira em Março de 2006 (<http://cgpbr.mhpcc.hawaii.edu/tools/papaya/>).

A mangueira (*Mangifera indica*), com $n = 20$ cromossomas e um genoma estimado em 439 Mpb/1C (Arumuganathan e Earle, 1991), é a sétima fruteira com maior produção no mundo e a segunda com maior produção entre as fruteiras tropicais (FAO 2004). Para esta espécie não há publicações na área da citogenética molecular, *i.e.* que refiram a localização física de marcadores ou genes nos cromossomas, nem que relatem a construção de bibliotecas genómicas em YACs, BACs, etc. Contudo, Chunwongse et al. (2000) relataram a construção de mapas genéticos para as variedades Alphonso e Palmer, utilizando marcadores RFLPs e AFLPs e a estratégia “pseudo-testcross” com base em 31 plantas descendentes do seu cruzamento. O mapa da variedade Alphonso tem 365 marcadores distribuídos em 63 grupos que cobrem 1437.7 cM com uma distância média entre marcadores de 10.4 cM. O mapa da variedade Palmer tem 369 marcadores distribuídos em 59 grupos com uma distância média

entre marcadores de 7 cM. Foi ainda descrita a construção de um mapa genético preliminar baseado na descendência de um cruzamento entre as variedades Keitt e Tommy-Atkins. No entanto este mapa é muito incipiente, consistindo em apenas 34 marcadores AFLP distribuídos em 13 grupos de ligamento que cobrem 161.5 cM (Kashkush et al. 2001).

2.2- ANANASEIRO

Como referido acima, o ananaseiro (*Ananas comosus* (L.) Merrill) já entrou na era da genómica. A construção dos seus primeiros mapas genéticos (descrita neste trabalho) é um dos passos que marca a entrada nessa era. Neste contexto, não parece absurdo afirmar que não faltarão muitos anos para o início da sequenciação total do genoma de uma cultura tão importante como a de ananás.

2.2.1- IMPORTÂNCIA ECONÓMICA

O ananás é o terceiro fruto mais importante entre os frutos tropicais, depois da banana e da manga. Embora seja principalmente cultivado em países tropicais, entre 30°N e 30°S, podem ser encontradas pequenas plantações geralmente protegidas por resguardos em áreas com climas temperados para além destas latitudes (Nakasone e Paull 1998). É o caso da cultura de ananás em estufas nos Açores (37°N50'), com uma produção anual de cerca de 2000 toneladas (Tavares e Batista 2004, <http://www.angra.uac.pt/pessoais/docentes/jbatista/Ananas/>).

De acordo com estatísticas da “Food and Agriculture Organization” (FAO) (<http://apps.fao.org>), a produção mundial aumentou de 4,833,137 toneladas em 1961 para 15,287,413 toneladas em 2004. Cinco países, nomeadamente a Tailândia (1,700,000 t), as Filipinas (1,650,000 t), o Brasil (1,435,600 t), a China (1,475,000 t) e a Índia (1,300,000) contribuíram com cerca de metade da produção mundial em 2004. Um segundo grupo significativo de produtores, que inclui países tão díspares como a Nigéria, o México, a Costa Rica, a Indonésia, o Quênia, a Colômbia, a Costa do Marfim, a Venezuela, o Vietname, a Malásia, os Estados Unidos e a África do Sul, produziram cerca de um terço da produção mundial.

Aproximadamente 70% do ananás produzido é consumido como fruto fresco no país de origem enquanto os restantes 30% se destinam à exportação, principalmente em produtos processados como fatias enlatadas, pedaços frescos e congelados, sumos naturais e sumos concentrados (Loeillet 1997, Rohrbach e tal. 2003). Por exemplo, as exportações mundiais de sumo concentrado representam mais de 250 milhões de US dólares e as exportações de enlatados mais de 600 milhões de US \$. Do ananaseiro são também extraídas fibras de elevada qualidade utilizadas na produção de têxteis, assim como bromelaína, um derivado que contém, entre outros componentes, várias proteinases e que é utilizado para tornar a carne mais tenra, clarear a cerveja e tingir couro. Para além destas utilizações, a bromelaína tem um grande potencial para usos terapêuticos (Maurer 2001).

2.2.2- TAXONOMIA

O género *Ananas* (perene e monocotiledóneo) pertence à ordem *Bromeliales*, família *Bromeliaceae* e subfamília *Bromelioideae*. A família *Bromeliaceae* tem três subfamílias

(*Pitcairnioideae*, *Tillandsioideae* e *Bromelioideae*) que abrangem 56 géneros e 2921 espécies (Luther 2002). Em toda a subfamília *Bromelioideae* há tendência para a fusão das partes florais, no entanto é no género *Ananas* que esta característica é mais evidente. Este é o único género em que as flores e as brácteas se fundem completamente num fruto partenocárpico tipo sorose (sincarpo) formado por 50 a 200 bagas juntas (Coppens d'Eeckenbrugge et al. 1997a).

Até recentemente (2003) a taxonomia aceite para o ananaseiro era a de Smith e Downs (1979) com a modificação de Leal (1990) que invalidou a espécie *A. monstrosus* por a característica que a diferenciava (ausência de coroa no fruto) não ser consistente (tabela 4).

Tabela 4. Taxonomia do ananaseiro segundo Smith e Downs (1979) e Leal (1990).

Género	Espécie
<i>Pseudananas</i>	<i>Pseudananas sagenarius</i> (Arruda da Camara) Camargo
	<i>Ananas ananassoides</i> (Baker) L.B. Smith
	<i>Ananas nanus</i> (L.B. Smith) L.B. Smith
	<i>Ananas paraguayensis</i> Camargo & L.B. Smith
<i>Ananas</i>	<i>Ananas lucidus</i> Miller
	<i>Ananas bracteatus</i> (Lindley) Schultes
	<i>Ananas fritzmuelleri</i> Camargo
	<i>Ananas comosus</i> (L.) Merrill

Vários factos levaram, no entanto, a que fosse proposta uma simplificação desta classificação: (1) parte das características usadas na classificação serem quantitativas e fortemente influenciadas pelas condições ambientais, não sendo claramente discriminativas, (2) as características qualitativas usadas na classificação, como a presença ou ausência de espinhos nas folhas, serem determinadas por apenas um ou um número reduzido de loci que podem sofrer mutações e ter segregação tanto interespecífica como intraespecífica, (3) serem possíveis cruzamentos entre plantas de diferentes espécies do género *Ananas* e os híbridos resultantes serem férteis, (4) serem possíveis cruzamentos entre as espécies do género *Ananas*

e o tetraplóide *Pseudananas saganarius* em que são produzidos híbridos férteis, e (5) os dados disponíveis de diversidade bioquímica e molecular indicarem um processo de separação entre estas espécies muito incipiente (Leal e Coppens d'Eeckenbrugge 1996, Cabral et al. 1997, Coppens d'Eeckenbrugge et al. 1997a, Duval et al. 2001 e 2003, Coppens d'Eeckenbrugge e Leal 2003).

A classificação actualmente aceite para o ananaseiro foi proposta por Coppens d'Eeckenbrugge e Leal (2003). Esta considera apenas um género (*Ananas*) dividido em duas espécies: a espécie *A. macrodontes* Morren é restabelecida para substituir a única espécie do género *Pseudananas* da classificação anterior (*P. saganarius*) e as sete espécies do género *Ananas* são reclassificadas em cinco variedades botânicas de uma única espécie - *A. comosus* (tabela 5).

A espécie *A. macrodontes* é tetraplóide ($2n = 4x = 100$), autocompatível, vigorosa, propaga-se por longos talos basais ou subterrâneos, tem folhas de 2-3m de comprimento e 7cm de largura com espinhos, e não tem coroa de folhas no sincarpo. A espécie *A. comosus* é geralmente diplóide ($2n = 50$), autoincompatível e a propagação vegetativa é por rebentos que se desenvolvem no caule, no pedúnculo e na coroa do fruto. O ananaseiro cultivado para produção de fruto corresponde à variedade botânica *comosus*. As suas folhas são relativamente largas (mais de 5cm) e espinhosas, parcialmente espinhosas ou lisas, e o seu forte pedúnculo sustenta um fruto cujo tamanho pode atingir vários quilogramas. A variedade *ananassoides* corresponde à forma selvagem mais comum. Esta tem folhas espinhosas mais finas e um fruto muito menor suportado por um longo e relativamente fino pedúnculo. A variedade *parguazensis*, outra forma selvagem, tem folhas largas constringidas na base com espinhos antrorsos ou retrorsos e frutos globosos. A variedade *erectifolius* é muito semelhante à variedade *ananassoides*, excepto nas folhas lisas e fibrosas utilizadas pelos ameríndios para fazer camas de rede e linhas e redes de pesca. As plantas da variedade *bracteatus* são muito

vigorosas e espinhosas, e produzem frutos de tamanho médio com brácteas longas. Esta variedade é cultivada como cerca viva e o seu fruto é também utilizado para produzir sumo. Encontra-se ainda como planta espontânea em antigas povoações na América do Sul, e uma variante variegada é hoje uma planta ornamental comum nos jardins tropicais.

Tabela 5. Taxonomia do ananaseiro por Coppens d’Eeckenbrugge e Leal (2003).

Género	Espécie	Variedades
	<i>A. macrodontes</i> (Morren) Coppens & Leal Anteriormente <i>Pseudananas</i> <i>sagenarius</i>	-
<i>Ananas</i>		<i>comosus</i> ((L.) Merrill) Coppens & Leal Anteriormente <i>A. comosus</i>
		<i>ananassoides</i> (Baker) Coppens & Leal Anteriormente <i>A. ananassoides</i> e <i>A. nanus</i>
	<i>A. comosus</i>	<i>parguazensis</i> (Camargo & L.B. Smith) Coppens & Leal Anteriormente <i>A. parguazensis</i>
		<i>erectifolius</i> (L.B. Smith) Coppens & Leal Anteriormente <i>A. lucidus</i>
		<i>bracteatus</i> (Lindl.) Coppens & Leal Anteriormente <i>A. bracteatus</i> e <i>A. fritzmulleri</i>

2.2.3- HABITAT NATURAL E ORIGEM

Ambas as espécies do género *Ananas* (*A. comosus* e *A. macrodontes*) têm uma distribuição natural confinada ao subcontinente da América do Sul. No entanto quando Cristóvão Colombo chegou a Guadalupe em 1493 e os europeus viram pela primeira vez o ananaseiro, já os ameríndios o tinham domesticado e espalhado pela América do Sul, América Central, e pela região das Antilhas e Caraíbas (Leal e Coppens d’Eeckenbrugge 1996; Coppens d’Eeckenbrugge et al. 1997a). O cultivo do ananaseiro iniciou-se muito provavelmente com as variedades *comosus* e *erectifolius* a evoluírem da var. *ananassoides*

e/ou var. *paraguayensis* na região norte do rio Amazonas (nas bacias dos rios Orinoco e Negro, e na Guiana), onde há uma maior variabilidade morfológica e molecular tanto nas plantas selvagens como nas cultivadas (Duval et al. 2003). A var. *bracteatus* e a espécie *A. macrodontes* são originárias do sul do continente americano (Paraguai e sul do Brasil) (Leal e Coppens d'Eeckenbrugge 1996).

2.2.4- NÚMERO DE CROMOSSOMAS E TAMANHO DO GENOMA

O número de cromossomas mais comum na subfamília *Bromelioideae* é $2n = 50$, (Cotias-de-Oliveira et al. 2000). Este é também o número de cromossomas mais comum nas variedades e cultivares da espécie *A. comosus*, apesar de já terem sido identificados clones tetraplóides na var. *comosus* e na var. *ananassoides* (Capinpin e Rotor 1937, Sharma e Ghosh 1971, Lin et al. 1987, Dujardin 1991, Cotias-de-Oliveira et al. 2000). A dimensão dos cromossomas do ananaseiro varia entre 0.5 e 1.7 μm segundo Sharma e Ghosh (1971) e entre 0.67 e 0.89 μm segundo Cotias-de-Oliveira et al. (2000). No que se refere à dimensão do genoma, Arumuganathan e Earle (1991) baseados em citometria de fluxo estimaram-na em 444Mpb para a var. *bracteatus* e 526Mpb para a var. *comosus*.

A espécie *A. macrodontes* é tetraplóide ($2n = 100$) (Lin et al. 1987). O facto de já ter sido relatada a ocorrência de gâmetas não reduzidos em *A. comosus* (Collins 1960, Dujardin 1991), o facto de cruzamentos entre *A. comosus* e *A. macrodontes* produzirem alguns híbridos tetraplóides férteis (Collins 1960) e a relação próxima entre estas duas espécies indicada pelos dados moleculares, sugerem que a autoploidização da espécie *A. comosus* é uma forte hipótese para a origem da espécie *A. macrodontes* (Duval et al. 2003).

2.2.5- FLUXO GENÉTICO E PROPAGAÇÃO

Como referido anteriormente, dentro do género *Ananas* não há barreiras reprodutivas nem intraespecíficas nem interespecíficas. A polinização é fundamentalmente cruzada e os principais vectores de polinização natural são os beija-flores, enquanto as abelhas e formigas têm um papel secundário. No entanto, uma vez que as sementes retêm a capacidade de germinação durante apenas seis meses, germinam lentamente e/ou as novas plântulas raramente sobrevivem em condições naturais, a propagação vegetativa é o modo mais comum de multiplicação do ananaseiro no estado selvagem. Os estudos de diversidade genética em populações naturais confirmam este facto (Coppens d'Eeckenbrugge et al. 1997a). Na espécie *A. comosus* a propagação vegetativa faz-se por rebentos no caule, no pedúnculo e na coroa do fruto, e na espécie *A. macrodontes* por talos basais ou subterrâneos.

A espécie *A. comosus* possui um sistema de autoincompatibilidade gametófitica com uma reacção de autorejeição variável em intensidade. Esta é geralmente mais forte na variedade cultivada *comosus*, provavelmente como resultado do processo de domesticação e selecção de frutos sem sementes (Coppens d'Eeckenbrugge et al. 1993). Mesmo quando não se detecta qualquer depressão na formação ou germinação de sementes provindas de autopolinização observam-se severas depressões devido à endogamia em estados mais avançados de desenvolvimento das plantas, o que torna impossível qualquer programa de desenvolvimento de linhas puras para além da primeira ou segunda geração obtida por autopolinização (Collins 1960, Coppens d'Eeckenbrugge et al. 1993, Coppens d'Eeckenbrugge et al. 1997a, 1997b, Cabral et al. 2000). O *A. macrodontes*, pelo contrário, é autocompatível e as descendências de autopolinizações são muito homogéneas, o que indica que a espécie é altamente homozigótica e autogâmica (Collins 1960).

Apesar de não haver publicações ou relatos de cruzamentos na natureza entre o gênero *Ananas* e outros gêneros e de ser difícil a obtenção de híbridos intergenéricos, mesmo por polinização controlada, encontram-se referências a híbridos de plantas do gênero *Ananas* com plantas de outros gêneros como *Aechmea*, *Cryptanthus* e *Neoreglia*, obtidos com fins ornamentais (Grant 1998; <http://fcbs.org/articles/Bigenerics.htm>).

2.2.6- DIVERSIDADE MOLECULAR

No ananaseiro têm sido desenvolvidos estudos de diversidade molecular com recurso a várias técnicas: isoenzimas (Garcia 1988, Aradhya et al. 1994), RAPDs (Ruas et al. 2001), AFLPs (Kato et al. 2004), RFLPs (Noyer 1991, Noyer et al. 1997, Duval et al. 2001) e cpDNA PCR-RFLPs, fragmentos de DNA cloroplastídico amplificados e cortados com endonucleases (Duval et al. 2003). Todos estes trabalhos suportam uma separação clara entre *A. comosus* e *A. macrodontes*, apesar do nível baixo de diferenciação genética, e todos indicam a existência de fluxo genético tanto entre estas duas espécies como entre as cinco variedades botânicas da espécie *A. comosus*. Outras observações consistentes são, por um lado, uma maior variação genética dentro da variedade selvagem geograficamente mais dispersa e mais comum - var. *ananassoides* - do que dentro das outras variedades, e, por outro lado, uma forte proximidade genética entre esta variedade (*ananassoides*) e as outras variedades, particularmente as cultivadas (*comosus*, *erectifolius* e *bracteatus*), o que sugere estar a var. *ananassoides* na origem do processo de domesticação das variedades cultivadas. Para além disso, por exemplo, a existência na Guiana Francesa de clones morfológicamente situados entre a var. *comosus* e a var. *ananassoides*, provavelmente híbridos espontâneos,

com haplótipos presentes nestas duas variedades (Duval et al. 2001) sugere uma recente domesticação e uma introgressão contínua.

De acordo com os dados de análises por isoenzimas e RFLPs a variedade *comosus* revelou-se muito homogénea e a diversidade genética intracultivar revelou-se muito baixa independentemente do local de recolha dos diferentes clones, o que sugere uma origem única para as diferentes cultivares seguida de selecção clonal de mutantes com fenótipos com interesse agronómico. No entanto, a técnica de marcadores AFLP mostrou uma média de 20% de polimorfismos dentro da cv Smoth Cayenne, sugerindo alterações no genoma mais profundas do que as que seriam provocadas por selecção clonal (Kato et al. 2004). A instabilidade genómica provocada por transposição e a fertilização por cruzamento devido à autoincompatibilidade foram as causas prováveis apontadas pelos autores para a variação genética intracultivar. No que diz respeito à var. *erectifolius*, os dados de análises RFLP obtidos por Duval et al. (2001) indicam eventos de domesticação múltiplos, envolvendo processos de selecção convergente para folhas lisas (sem espinhos) e fibrosas a partir de diferentes genótipos. Quanto à var. *bracteatus*, todos os dados de análises de diversidade molecular indicam ser relativamente uniforme, melhor diferenciada das outras variedades e ter mais afinidade com a espécie *A. macrodontes*. Para além disso, Duval et al. (2003) sugerem que os dados que obtiveram de variação de DNA cloroplastídico, juntamente com outros dados genómicos de diversidade molecular publicados, apontam para uma origem híbrida da var. *bracteatus*.

2.2.7- MELHORAMENTO

Há mais de um século que a hibridação e a selecção clonal, envolvendo principalmente cultivares de elite, têm sido utilizadas no melhoramento do ananaseiro. No entanto, a natureza heterozigótica das cultivares de ananaseiro e a consequente forte segregação e recombinação têm limitado o sucesso na obtenção de híbridos melhorados. Até recentemente a selecção rigorosa entre milhões de plântulas resultou em cultivares de apenas qualidade mediana ou de importância local, tendo a produção mundial de ananás sido dominada por apenas uma cultivar - “Smooth Cayenne” (Coppens d’Eeckenbrugge et al. 1997a). Esta situação evoluiu no final da década de 90 com o sucesso comercial da cultivar MD-2 também conhecida por “Golden Ripe”, um novo híbrido que alterou o mercado mundial de ananás fresco e estimulou o interesse no desenvolvimento de novas cultivares (Loeillet 2003, Pineapple Working Group 2004).

Entre os objectivos dos programas de melhoramento está a introgressão de resistências a doenças como por exemplo as causadas pelos fungos *Phytophthora* e *Fusarium*, a prevenção de disfunções fisiológicas como o escurecimento interno (“internal browning”) também conhecido por “blackheart” e o controle de características específicas como a floração natural precoce, em cultivares de elite. Para evitar a aleatoriedade da segregação e recombinação, a engenharia genética é uma estratégia promissora uma vez que permite a transferência de apenas um gene, ou alguns genes, sem alterar substancialmente o genoma inicial. Para o ananaseiro foram já estabelecidos procedimentos de transformação genética, tanto por transferência com microprojecteis (Sripaoraya et al. 2001) como com *Agrobacterium* (Espinosa et al. 2002), e procedimentos de regeneração e propagação *in vitro* (Escalona et al. 1999, Firoozabady e Gutterson, 2003, Sripaoraya et al. 2003). Foram obtidos clones de ananaseiro geneticamente transformados com expressão reduzida de polifenol oxidase (PPO),

envolvida no escurecimento interno (“blackheart”) (Stewart et al. 2001) e com expressão reduzida de ACC (1-aminociclopropano-1-carboxilato) sintetase, envolvida na floração natural do ananaseiro (Botella 2005a). Apesar dos primeiros ensaios serem promissores para alguns clones transformados, ainda decorrem ensaios de campo para a qualidade dos frutos e performance agronómica que possibilitarão confirmar ou não o controlo do escurecimento interno e da floração natural em clones com interesse comercial (Smith et al. 2005). Foram também obtidos clones de ananaseiro geneticamente transformados que expressam o gene de um inibidor de protease de cisteína (PI) e o gene *bar* de resistência a bialafos (Rohrbach et al. 2000, Sripaoraya et al. 2001, Sipes 2003. Botella e Fairbairn 2005). Os inibidores de proteases de cisteína inibem a produção de enzimas digestivas dos nemátodos e, afectando a sua alimentação, reduzem a população dos mesmos (Urwin et al. 2000). No entanto, os ensaios de campo mostraram que as plantas transformadas que expressam esse gene crescem pior que as plantas não transformadas (Sipes 2003). Pelo contrário, o gene *bar* mostrou-se estável na sua expressão e sem interferir no crescimento das plantas transformadas, nem na produção ou na qualidade do fruto (Sripaoraya 2005).

2.2.8- RECURSOS GENÉTICOS E GENÓMICOS

As colecções de germoplasma de ananaseiro com maior diversidade genética são mantidas na Martinica pelo CIRAD, no Havai pelo “U.S. Pacific Basin Agricultural Research Center”, e no Brasil pelo EMBRAPA/CENARGEN em Brasília e pelo EMBRAPA/CNPMPF em Cruz das Almas. Há também colecções importantes mantidas por outros países produtores (*p.e.* Malásia, Austrália, Cuba e Japão) e por companhias privadas (Coppens d’Eeckenbrugge e Duval 1999). Para melhor gestão dos recursos genéticos, o CIRAD criou o TropGENE-DB

(<http://tropgenedb.cirad.fr>), um sistema de informação de culturas tropicais que brevemente incluirá o ananás.

Na área da genômica estrutural, para o ananaseiro foi isolada e caracterizada uma sequência tipo retrotransposição (“retroposon-like sequence”) repetidamente integrada no genoma em sequências múltiplas variadas e potencialmente capazes de transposar (Thomson et al. 1998) e está em curso a construção de um mapa genético de referência (Carlier et al. 2004, Carlier e Leitão 2006). No entanto, não foram ainda publicados trabalhos direccionados para o mapeamento físico ou para a sequenciação total do seu genoma nem é do conhecimento público a existência de qualquer grupo ou consórcio que tenha estes objectivos.

No que diz respeito a estudos na área da genômica funcional nesta espécie, são poucos os trabalhos publicados mas o esforço científico neste campo nos últimos anos leva a prever que o seu número aumente rapidamente. Até agora foram clonados e caracterizados funcionalmente genes que codificam: uma ACC (1-aminociclopropano-1-carboxilato) sintetase; uma ACC oxidase (Cazzonelli et al. 1998), uma malato desidrogenase (Cuevas e Podestá 2000), uma ananaína (Carter et al. 2000), uma Cu/Zn-superóxido dismutase (Lin et al. 2000), duas polifenol oxidases (Stewart et al. 2001) e dois tipos de inibidores de protease de cisteína (Sawano et al. 2002, Shyu et al. 2004). Para além disso, foi obtido um grande número de sequências ESTs em estudos recentes sobre genes envolvidos no desenvolvimento das raízes (Neuteboom et al. 2002) e do fruto (Moyle et al. 2005a). Na sequência destes trabalhos foram construídos “microarrays” contendo 10000 clones EST das bibliotecas da raiz e do fruto, os quais estão a ser utilizados para estudar as mudanças de expressão génica que ocorrem durante o amadurecimento do fruto e durante formação de inchaços na raiz provocados pela infecção com nemátodos (Botella 2005b). Foi também desenvolvida a base de dados PineappleDB que disponibiliza à comunidade científica dados anotados de ESTs de clones de cDNA isolados do fruto, da raiz e de inchaços na raiz provocados por infecção com

nemátodos (Moyle et al. 2005b). A PineappleDB tem actualmente mais de 5600 ESTs, e disponibiliza informação sobre os clones e “contigs”, incluindo: a sequência, a identificação da sequência, uma ligação ao homólogo mais próximo da base de dados “nr” do “National Center for Biotechnology Information” (NCBI) e o respectivo comprimento e percentagem de identidade, uma ligação ao homólogo mais próximo na *Arabidopsis*, classificações funcionais, ligações a todos os clones EST agrupados no mesmo “contig” e distribuição dos ESTs pelas bibliotecas de cada tipo de tecido de origem (“fruit, green mature fruit, yellow ripe fruit, root, root tips, root 1-4 days, root 5-10 days). O acesso a esta base de dados faz-se via <http://www.pgel.com.au>, estando os ESTs também disponíveis na base de dados do NCBI (<http://www.ncbi.nlm.nih.gov>).

Numa pesquisa de dados genómicos de ananaseiro no NCBI encontram-se 7 sequências de DNA cloroplastídico da var. *ananassoides* (4 delas registadas em *A. nanus* segundo a classificação taxonómica anterior), 50 loci microssatélites da var. *bracteatus* e 5722 ESTs da var. *comosus*, a maior parte dos quais do trabalho de Moyle et al. (2005a,b). Encontram-se também 10 sequências obtidas neste trabalho a partir de marcadores mapeados, com a particularidade de uma delas incluir um microssatélite.

No laboratório do CIRAD-BIOTROP em Montpellier, França, e no “Biotechnology Research Institute of Universiti Malaysia Sabah” foram identificados, respectivamente, 119 e 83 loci microssatélites em ananaseiro (Coppens d’Eeckenbrugge 2002, Kinsuat e Kumar 2005) que, apesar de não estarem publicados nem disponíveis nas bases de dados, poderão eventualmente ser utilizados em programas de cooperação com outros laboratórios. Por outro lado, foram identificados 22 loci microssatélites em outras espécies da família do ananaseiro (*Bromeliaceae*) mas pertencentes às outras duas subfamílias (*Pitcairnioideae* e *Tillandsioideae*): 7 loci em *Pitcairnia geyskesii* (Sarhou et al. 2003) e 15 loci em *Tillandsia fasciculata* e *Guzmania monostachya* (Boneh et al. 2003). Estes microssatélites poderão ser

úteis para o mapeamento genético do ananaseiro e para a análise comparativa dos genomas de toda a família *Bromeliaceae*, se as regiões adjacentes aos mesmos forem conservadas.

2.3- IMPORTÂNCIA DO MAPEAMENTO GENÉTICO NO ANANASEIRO

A elevada similaridade genética entre as variedades da espécie *A. comosus* assim como a possibilidade de cruzamentos com o *A. macrodontes*, a outra espécie do género, levam-nos a antecipar uma forte similaridade entre os seus genomas. Logo, a construção de um mapa genético em qualquer variedade da espécie *A. comosus* poderá constituir o ponto de partida para a construção de um mapa integrado do género *Ananas*. Tal mapa, por um lado, abriria caminho ao estudo genómico estrutural de toda a família taxonómica *Bromeliaceae*, e por outro lado, facilitaria a identificação de marcadores estreitamente ligados a características de interesse, os quais podem ser utilizados na selecção assistida por marcadores moleculares (SAM), assim como no isolamento de genes por clonagem baseada no mapeamento (CBM). A importância que a selecção assistida por marcadores moleculares pode ter no melhoramento do ananaseiro é evidenciada pelos longos períodos de tempo que os métodos clássicos implicam também nesta espécie. Por exemplo, no “Malaysian Agricultural Research and Development Institute” (MARDI) foram necessários 12 anos de melhoramento clássico desde o primeiro cruzamento até à obtenção de um clone comercial (Chan 2005). Por outro lado, a transferência de genes de interesse por engenharia genética pode ser particularmente útil no desenvolvimento de variedades melhoradas de ananaseiro, particularmente devido: (1) à dificuldade de aplicação das estratégias tradicionais de melhoramento nesta cultura, como o

prova o seu pouco sucesso até à actualidade, e (2) ao baixo nível de diversidade molecular entre as variedades desta espécie (Botella e Fairbain 2005).

A necessidade de desenvolvimento de novas variedades é premente dado o facto da produção mundial de ananás ser praticamente monovarietal, dominada pela cultivar Smooth Cayenne, e de apenas um pequeno número de outras cultivares, Queen, Red Spanish, Pernambuco ou Abacaxi, Motilona ou Perolera e mais recentemente o híbrido MD-2 ou Golden Ripe, terem importância comercial (Coppens d'Eeckenbrugge et al. 1997a, Nakasone e Paull 1998, Loeillet 2003, Pineapple Working Group 2004). Por outro lado, o facto do ananás ser maioritariamente produzido em monocultura em plantações relativamente grandes e com métodos de cultivo mais ou menos padronizados faz com que as doenças que afectam o ananaseiro, a maioria das quais estabelecidas no mundo inteiro e algumas confinadas a determinadas regiões, possam provocar grandes danos na produção (Nakasone e Paull 1998, Rohrbach e Schmitt 2003), o que também evidencia a necessidade de desenvolvimento de novas variedades, em particular variedades resistentes às doenças que mais afectam a cultura. De facto, apesar do esforço científico dos últimos anos com o objectivo de minimizar ou eliminar os efeitos negativos de doenças como: (1) as causadas pelos nemátodos *Meloidogyne javanica* e *M. Incognita*, (2) as causadas pelas cochonilhas *Dysmicoccus brevipes* e *D. neobrevipes* associadas aos vírus “Pineapple mealybug wilt-associated virus-1” (PMWaV-1) e “virus-2” (PMWaV-2) e a várias espécies de formigas *p.e. Pheidole megacephala*, *Iridomyrmex humilis* e *Solenopsis geminata* ou (3) a fusariose, causada pelo fungo *Fusarium subglutinans*, estas são ainda apontadas como uns dos principais problemas a resolver em futuros trabalhos de investigação (Sether et al. 2005, Matos et al. 2005, Petty et al. 2005, Sipes 2005, Galvão et al. 2005). Outras áreas de investigação também apontadas como importantes são a manutenção da qualidade do fruto durante o período de armazenamento, principalmente o controlo do escurecimento interno causado pelo frio, e a inibição da floração

natural, permitindo a sua sincronização tal como a do crescimento do fruto e da colheita (Lin et al. 2005, Rebolledo-Martinez et al. 2005, Trusov et al. 2005, Wijeratnam et al. 2005, Siriphanich e Chanjirakul 2005).

3- MATERIAIS E MÉTODOS

3.1- MATERIAL VEGETAL

Neste trabalho foram utilizadas duas populações de mapeamento: uma F1 e uma F2. A população F1 consiste em 46 plantas híbridas obtidas de um cruzamento entre o *A. comosus* var. *comosus* (L.) Merrill, cv. Rondon - clone BR 50 (progenitor fêmea) e *Ananas comosus* var. *bracteatus* (Lindley) Schultes, cv. Branco do mato - clone BR 20 (progenitor macho). A população F2 consiste em 142 plantas obtidas por autopolinização de uma das plantas da população F1 (planta n.º 5). O fruto ananás geralmente cultivado e comercializado corresponde à variedade botânica *comosus*, caracterizada por ter um fruto comestível grande. A var. *bracteatus*, cultivada como cerca viva e com fins ornamentais, é caracterizada por ter um fruto de tamanho médio com longas brácteas. Os cruzamentos e o crescimento das plantas realizaram-se no CIRAD-FLHOR na Martinica e as análises moleculares para mapeamento genético no Laboratório de Genética e Melhoramento de Plantas (LGMP) da Universidade do Algarve. A transferência de material vegetal da Martinica para Portugal fez-se via correio, tendo o LGMP recebido folhas dos progenitores e das duas descendências embrulhadas em papel e acondicionadas em caixas de cartão.

3.2- EXTRACÇÃO DE DNA

As folhas foram lavadas com água corrente e detergente (Teepol), imersas em água destilada e enxugadas com toalhetes de papel. Aproximadamente 1.5 g de material foliar foi

homogeneizado num almofariz com azoto líquido e o pó fino obtido ressuspensão em 3 mL de tampão A [200 mM de Tris-HCl (pH 8.0), 25 mM de EDTA, 250 mM de NaCl, 1% (p/v) de “sodium dodecyl sulfate” (SDS) e 2% (p/v) “polyvinylpyrrolidone” (PVP)]. Esta mistura foi então incubada a 65 °C durante 15 mins e em seguida tratada com RNase A (40 µg/mL) durante 30 mins a 37 °C. Depois, o DNA genómico total foi purificado por extracções sucessivas: uma com fenol, outra com fenol : clorofórmio : álcool isoamílico (25:24:1) e uma final com clorofórmio : álcool isoamílico (24:1), tal como descritas por Sambrook et al. (1989).

3.3- MARCADORES ANALISADOS

Para a construção do mapa foram utilizados: um marcador morfológico, isoenzimas e os marcadores - DNA “Random Amplified Polymorphic DNA” - RAPD (Williams et al. 1990), “Amplified Fragment length Polymorphism” - AFLP (Vos et al. 1995) e “Inter Simple Sequence Repeat” - ISSR (Zietkiewicz et al. 1994).

3.3.1- MARCADOR MORFOLÓGICO

O marcador morfológico analisado denomina-se “piping”. A sua expressão fenotípica é determinada pelo alelo dominante do gene *P* e caracteriza-se pelo enrolamento de uma prega de tecido da margem da folha que suprime os espinhos da mesma (Cabral et al. 1997, Leal e Coppens d’Eeckenbrugge 1996). A sua análise efectuou-se pela simples observação visual do fenótipo das (folhas) plantas.

3.3.2- ANÁLISES ISOENZIMÁTICAS

Extractos de folhas dos progenitores e das 46 plantas da população F1 foram analisados para as isoenzimas Fosfo-Gluco-Mutases (PGM) e Malato-Desidrogenases (MDH), de acordo com os procedimentos descritos por Lebrun e Chevalier (1988).

As folhas foram lavadas como descrito na extracção de DNA. Aproximadamente 0.5g de material foliar foi homogeneizado num almofariz com azoto líquido e o pó fino obtido ressuspenso em 1 mL de tampão B [200 mM de Tris-HCl (pH 7.2), 20 mM de “cystein”, 0.2% de Triton X-100, 5 mM de “Dithiotreitol” (DTT), 1 mM de “PhenylMethylSulfonylFluoride” (PMSF) e 2% (w/v) de “polyvinylpyrrolidone” (PVP)]. Esta mistura foi centrifugada a 13000 rpm durante 20 mins a 4 °C e os sobrenadantes imediatamente submetidos a electroforese em gel de amido.

Utilizaram-se géis de amido a 13% e como electrólito o tampão Tris-Citrato pH 7.0 (TC7) contendo 135 mM de Tris e 43 mM de ácido cítrico. Para preparar um gel dissolveram-se 57.2g de amido de batata hidrolisado (USB Corporation, Cleveland, OH USA) em 440 mL de TC7 : água destilada na razão de 1:15, cozeu-se o amido desta solução a 280 °C durante 20 mins sob agitação magnética, e fez-se uma desgasificação por vácuo.

Depois de ligeiramente arrefecido (até aproximadamente 60 °C) verteu-se o gel na tina de electroforese e deixou-se arrefecer à temperatura ambiente. Imediatamente antes se de proceder à electroforese, o gel foi arrefecido durante 30 mins num frigorífico a 4 °C, e em seguida colocou-se em cada poço 5 µL de solução de azul de bromofenol (30% de glicerol e 0,25% de azul de bromofenol) e um pequeno quadrado (8 mm x 8 mm) de papel 3 MM (Whatman, Maidstone, England) embebido com o extracto de uma amostra.

O aparelho de electroforese foi montado dentro de um frigorífico a 4 °C com uma placa de vidro arrefecido sobre o gel, e sobre esta um acumulador térmico congelado. A

tensão aplicada foi de 100 V durante os primeiros 30 mins e de 180 V durante as restantes 5 h 30 mins.

Terminada a electroforese, os géis foram cortados em fatias longitudinais com 1mm de altura, sendo estas imediatamente imersas em tampão Tris-HCl 50 mM (pH 8.5) em caixas acrílicas individuais. Em seguida, o tampão de cada caixa foi substituído pela solução de revelação de um sistema isoenzimático, preparada imediatamente antes de ser utilizada. A composição das soluções de revelação dos sistemas isoenzimáticos analisados está resumida na tabela 6. A revelação fez-se numa estufa a 37 °C durante pelo menos 20 mins e as fatias de gel, depois de lavadas com água corrente e fixadas com ácido acético a 13%, foram analisadas e fotografadas.

Tabela 6. Soluções de revelação dos sistemas isoenzimáticos analisados.

Sistema	Componentes	Quantidade
Fosfo-Gluco-Mutases (PGM)	Glucose-1-fosfato	0.1 g
	MgCl ₂ (1 M)	0.4 mL
	NADP (5 mg/mL)	2 mL
	PMS (1 mg/mL)	1 mL
	NBT (50 mg/mL)	200 µL
	Tris-HCl pH 8.5 (0.5 M)	20.8 mL
	G6PDH	25 µL
	Agar (solução de 2%)	25 mL
Malato-Desidrogenases (MDH)	Na Malato pH 7.0 (1 M)	5 mL
	NAD (10 mg/mL)	1 mL
	PMS (1 mg/mL)	1 mL
	NBT (50 mg/mL)	200 µL
	Tris-HCl pH 8.5 (0.5 M)	5 mL
	H ₂ O destilada	37.8 mL

3.3.3- ANÁLISES RAPD

Os marcadores RAPD foram amplificados com “primers” de 10 nucleótidos comercializados pela Operon Technologies (Alameda, CA, USA).

Os “primers” que revelaram bandas polimórficas entre os progenitores foram posteriormente testados numa amostra aleatória de 6 plantas da população F1. Desta forma, para marcadores presentes num progenitor no estado heterozigótico e ausentes no outro progenitor (marcadores polimórficos heterozigóticos) cuja razão de segregação esperada na população F1 é de 1:1, a probabilidade das 6 plantas amostradas terem o mesmo genótipo é $2(0.5)^6 = 0.03125$, pelo que marcadores polimórficos nos progenitores e presentes nas 6 plantas testadas foram assumidos como estando no estado homozigótico, *i.e.* a não segregar na população F1, não sendo por isso analisados nas restantes 40 plantas desta população. Inversamente, assumiu-se a ausência destes marcadores em pelo menos uma das 6 plantas testadas como indicação de heterozigocidade no *locus* do marcador, procedendo-se à sua análise nas restantes 40 plantas da população F1. Na população F2 foram utilizados maioritariamente “primers” que amplificam marcadores: (1) que fazem parte dos mapas construídos com base na população F1, e (2) que estão presentes na planta F1 autopolinizada para gerar a população F2. Foram também analisados seis novos “primers” que não tinham sido utilizados na população F1. A tabela 7 mostra os “primers” RAPD analisados nas populações de mapeamento F1 e F2.

Tabela 7. “Primers” RAPD analisados.

OPA02* ⁺	OPC05* ⁺	OPE04*	OPH13*	OPL12* ⁺	OPN11* ⁺	OPQ04*	OPU09 ⁺
OPA03* ⁺	OPC06*	OPE06*	OPH17*	OPL15* ⁺	OPN14*	OPQ18* ⁺	OPU10 ⁺
OPA04*	OPC08*	OPE12* ⁺	OPH18*	OPM03*	OPN15* ⁺	OPR04*	OPU12 ⁺
OPA18*	OPC09* ⁺	OPE16*	OPK04* ⁺	OPM06* ⁺	OPO02*	OPR10*	OPU16 ⁺
OPAA04*	OPC15*	OPF01*	OPK08*	OPM10* ⁺	OPO07*	OPR12*	OPU18 ⁺
OPAB09* ⁺	OPC18*	OPH02*	OPK10*	OPM12* ⁺	OPO09*	OPR13* ⁺	OPZ04*
OPAC07* ⁺	OPC19* ⁺	OPH04*	OPK16*	OPM16*	OPO18* ⁺	OPR15* ⁺	
OPB07*	OPD03*	OPH05*	OPL01*	OPN04* ⁺	OPP03* ⁺	OPT07*	
OPB11*	OPD15* ⁺	OPH07*	OPL03* ⁺	OPN06*	OPP11* ⁺	OPT08*	
OPB15* ⁺	OPE03* ⁺	OPH08* ⁺	OPL04*	OPN09*	OPQ01*	OPU06 ⁺	

* “Primers” analisados na população F1.

+ “Primers” analisados na população F2.

As amplificações RAPD fizeram-se num termociclador T-GRADIANT (Biometra, Göttingen, Germany) nas condições descritas na tabela 8. Depois da amplificação adicionou-se 3.75 µL de solução de azul de bromofenol a cada reacção e os produtos de amplificação foram separados por electroforese em gel de agarose a 2% durante cerca de 3 h a 100 v. Como tampão de gel e de electroforese foi utilizado 0.5xTBE (45 mM de Tris-Borate e 1 mM de EDTA).

Para revelação dos produtos amplificados os géis foram submersos em brometo de etídio e, após lavagem com água destilada durante 30 mins, foram fotografados e visualizados sob luz UV com um sistema Kodak EDAS 120. A imagem foi processada com o programa de computador 1D Image Analysis Software v. 2.0.1 (Eastman Kodak, Scientific Imaging Systems, New Haven, US).

Tabela 8. Condições de amplificação dos marcadores RAPD.

Mistura reactiva		
Componentes	Quantidade	
DNA genómico	15 ng	
10x "Taq buffer" (100 mM de Tris-HCl (pH 9.0), 500 mM de KCl e 15 mM de MgCl ₂)	1.5 µL	
Cada nucleótido (dATP, dGTP, dCTP e dTTP)	150 µM	
"Primer"	0.4 µM	
Taq DNA polymerase (Amersham Biosciences, Uppsala, Sweden)	0.6 U	
Volume total	15 µL	
Programa		
Número de ciclos	Temperatura ¹ (°C)	Tempo
1	94	1.5 mins
	94	30 s
35	36	30 s
	72	1 min
1	72	10 mins

¹ A variação de temperatura foi de 0.6 °C/s.

3.3.4- ANÁLISES AFLP

As análises de marcadores AFLP foram efectuadas utilizando o kit “AFLPTM Analysis System I” (Invitrogen, Paisley, UK) de acordo com as instruções do fabricante.

Tendo em conta o número relativamente alto de polimorfismos entre os progenitores revelado pela técnica RAPD e o elevado “multiplex ratio” esperado na técnica AFLP, as análises de marcadores AFLP realizaram-se directamente em toda a população F1. Quanto à população F2, foram analisadas maioritariamente combinações de “primers” AFLP que amplificam marcadores (1) que fazem parte dos mapas construídos com base na população F1, e (2) que estão presentes na planta F1 autopolinizada para gerar a população F2. Foi ainda analisada na população F2 uma combinação de “primers” que não foi previamente analisada na população F1. Na tabela 9 estão discriminadas as combinações de “primers” AFLP utilizadas e o número de código atribuído a cada uma.

O DNA genómico foi digerido com as endonucleases *EcoRI* e *MseI*. A reacção fez-se durante 2 h a 37 °C num volume de 12.5 µL com 125 ng de DNA genómico, 1.25 U de cada endonuclease, 10 mM de Tris-HCl (pH 7.5), 10 mM de acetato de magnésio e 50 mM de acetato de potássio. Após a reacção as endonucleases foram inactivadas por incubação a 70 °C durante 15 mins.

Aos fragmentos obtidos da restrição de DNA genómico ligaram-se os adaptadores *EcoRI* e *MseI* (tabela 10). Para tal, juntou-se aos 12.5 µL da solução de restrição igual volume de solução contendo 2.5 pMol de adaptador *EcoRI*, 25 pMol de adaptador *MseI*, 0.5 U de T4 DNA ligase, 0.4 mM de ATP, 10 mM de Tris-HCl (pH 7.5), 10 mM de Mg-acetato e 50 mM de K-acetato, e incubou-se durante 2 h a 20 °C.

Tabela 9. Combinações de “primers” AFLP analisados e respectivo número de código.

<i>MseI</i> +3/ <i>EcoRI</i> +3 AFLP primers	Nº de código	<i>MseI</i> +3/ <i>EcoRI</i> +3 AFLP primers	Nº de código	<i>MseI</i> +3/ <i>EcoRI</i> +3 AFLP primers	Nº de código
M-CTA/E-AAG* ⁺	01	M-CAC/E-ACT* ⁺	18	M-CTA/E-AAC* ⁺	35
M-CAT/E-ACG* ⁺	02	M-CAT/E-ACT* ⁺	19	M-CAC/E-AAG*	36
M-CTC/E-ACG* ⁺	03	M-CAG/E-AAC*	20	M-CTA/E-ACA* ⁺	37
M-CTT/E-AGC* ⁺	04	M-CTC/E-ACT* ⁺	21	M-CAG/E-ACG* ⁺	38
M-CAA/E-ACC* ⁺	05	M-CTC/E-ACA* ⁺	22	M-CAG/E-ACA* ⁺	39
M-CAT/E-ACC* ⁺	06	M-CTC/E-AAC*	23	M-CAG/E-ACT* ⁺	40
M-CTT/E-ACC* ⁺	07	M-CAT/E-AAC* ⁺	24	M-CTT/E-ACT* ⁺	41
M-CTT/E-AGG* ⁺	08	M-CTA/E-ACT* ⁺	25	M-CTT/E-AAC* ⁺	42
M-CAC/E-ACC*	09	M-CTC/E-AAG* ⁺	26	M-CTT/E-ACA* ⁺	43
M-CTG/E-AAC* ⁺	10	M-CTG/E-AAG* ⁺	27	M-CTT/E-ACG*	44
M-CAG/E-AGC* ⁺	11	M-CAA/E-AAG* ⁺	28	M-CAA/E-ACT* ⁺	45
M-CAA/E-AAC*	12	M-CAT/E-AAG* ⁺	29	M-CTG/E-ACA*	46
M-CAG/E-ACC*	13	M-CTT/E-AAG*	30	M-CAG/E-AGG* ⁺	47
M-CTC/E-ACC* ⁺	14	M-CAC/E-AGG*	31	M-CTG/E-ACG*	48
M-CTG/E-ACT* ⁺	15	M-CAA/E-AGG* ⁺	32	M-CTA/E-AT ⁺	49
M-CTG/E-ACC*	16	M-CAT/E-AGG* ⁺	33		
M-CTA/E-ACC* ⁺	17	M-CTA/E-ACG* ⁺	34		

* Combinações de “primers” analisadas na população F1.

⁺ Combinações de “primers” analisadas na população F2.**Tabela 10.** Adaptadores e “primers” utilizados nas análises AFLP.

“primers”/adaptadores	Sequência
Adaptador <i>EcoRI</i>	5'-CTCGTAGACTGCGTACC CATCTGACGCATGGTTAA-5'
Adaptador <i>MseI</i>	5'-GACGATGAGTCCTGAG TACTCAGGACTCAT-5'
“primer” <i>EcoRI</i> +A	5'-GACTGCGTACCAATTCA-3'
“primer” <i>MseI</i> +C	5'-GATGAGTCCTGAGTAA-3'

Após a ligação dos adaptadores aos fragmentos de DNA obtidos pela digestão do DNA genómico, a mistura reactiva da ligação de adaptadores (25 µL) foi diluída de 1:10, adicionando-se 225 µL de TE (0.1) (10 mM de Tris-HCl, pH 8.0 e 0.1 mM de EDTA). Esta diluição foi depois utilizada como amostra de DNA para uma primeira amplificação, a pré-amplificação. Nesta foram utilizados dois “primers” AFLP, cada um com 1 nucleótido selectivo na terminação 3': “primer” *EcoRI*+A e “primer” *MseI*+C (tabela 10). As pré-

amplificações foram realizadas segundo as condições descritas na tabela 11, num termociclador T-GRADIANT (Biometra, Göttingen, Germany).

Tabela 11. Condições da pré-amplificação dos marcadores AFLP.

Mistura reactiva		
Componentes	Quantidade	
Solução de DNA genómico (após restrição, ligação de adaptadores e diluição de 1:10)	2.5 µL	
10x “Taq buffer” (100 mM de Tris-HCl (pH 9.0), 500 mM de KCl e 15 mM de MgCl ₂)	2.5 µL	
Cada nucleótido (dATP, dGTP, dCTP e dTTP)	200 µM	
“primer” <i>EcoRI</i> +A	31.25 ng	
“primer” <i>MseI</i> +C	31.25 ng	
Taq DNA polymerase (Amersham Biosciences, Uppsala, Sweden)	0.5 U	
Volume total	25 µL	
Programa		
Número de ciclos	Temperatura ¹ (°C)	Tempo
	94	30 s
30	56	1 min
	72	1 min

¹ A variação de temperatura foi de 1 °C/s.

Após a pré-amplificação a mistura reactiva (25 µL) foi dividida em três partes: (1) para visualizar os produtos de amplificação submeteram-se 7 µL a electroforese durante 1 h a 100 v em géis de agarose a 1.5%, (2) depois de confirmada a presença de produtos de amplificação no gel, 3 µL foram diluídos de 1:50 adicionando-se 147 µL de TE (0.1), e (3) o restante foi guardado a -20 °C.

Na amplificação final de marcadores AFLP, a amplificação selectiva, utilizaram-se “primers” *EcoRI* marcados na extremidade 5’ com fosfato radioactivo [$\gamma^{33}\text{P}$] para análise dos produtos de amplificação por autoradiografia após electroforese em gel desnaturante de poliacrilamida. Para 48 amostras (progenitores e população F1) marcaram-se 130.1 ng de “primer” *EcoRI*. A marcação fez-se durante 1 h a 37 °C num volume de 13 µL com 5.2 U de T4 “polynucleotide kinase”, 2.6 µL de [$\gamma^{33}\text{P}$]ATP a 370 MBq/mL (10 mCi/mL) (Amersham

Biosciences, Uppsala, Sweden), 70 mM de Tris-HCl (pH 7.6), 1 mM de MgCl₂, 50 mM de KCl e 1 mM de 2-mercaptoethanol. Para a população F2 marcaram-se 390.3 ng de “primer” *EcoRI* numa reacção de 39 µL de volume total.

Na amplificação selectiva utilizaram-se 2 “primers” AFLP, cada um com 2 ou 3 nucleótidos selectivos na terminação 3’: *EcoRI*+AN ou *EcoRI*+ANN e *MseI*+CN ou *MseI*+CNN, sendo N qualquer um dos nucleótidos dATP, dGTP, dCTP ou dTTP. A amplificação selectiva fez-se nas condições descritas na tabela 12, num termociclador T-GRADIANT (Biometra, Göttingen, Germany).

Tabela 12. Condições da amplificação selectiva dos marcadores AFLP.

Mistura reactiva		
Componentes	Quantidade	
Solução de DNA genómico (após pré-amplificação e diluição de 1:50)	2.5 µL	
10x “Taq buffer” (100 mM de Tris-HCl (pH 9.0), 500 mM de KCl e 15 mM de MgCl ₂)	1 µL	
Cada nucleótido (dATP, dGTP, dCTP e dTTP)	200 µM	
“primer” <i>EcoRI</i> +AN(N) marcado	2.5 ng	
“primer” <i>MseI</i> +CN(N)	15.1 ng	
Taq DNA polymerase (Amersham Biosciences, Uppsala, Sweden)	0.25 U	
Volume total	10 µL	
Programa		
Número de ciclos	Temperatura ¹ (°C)	Tempo
	94	30 s
	65	30 s
1	72	1 min
	94	30 s
12	Inicia-se a 65 °C e decresce 0.7 °C por ciclo	1 min
	72	1 min
	94	30 s
23	56	30 s
	72	1 min

¹ A variação de temperatura foi de 1 °C/s.

Depois da amplificação juntou-se 10 µL de solução corada de formamida (98% de formamida, 10 mM de EDTA, pH 8.0, 0.025% de azul de bromofenol e 0.025% de xileno cianol) a cada reacção. Estas foram depois incubadas a 90 °C durante 3 mins para

desnaturação das cadeias de DNA e imediatamente a seguir colocadas em gelo. Em seguida os produtos de amplificação foram separados por electroforese em gel de poliacrilamida desnaturante (6% de poliacrilamida, Long Ranger Gel Solution, Cambrex Bio Science, Rockland inc., USA, e 7.5 M de ureia) durante cerca de 2 h 30 mins a 45 W. Utilizou-se como tampão de gel e de electroforese 1xTBE (90 mM de Tris-Borate e 2 mM de EDTA). Depois da electroforese, os géis foram secos em papel 3 MM (Whatman, Maidstone, England) e as autoradiografias foram obtidas em filmes HyperfilmTM MP (Amersham Biosciences, Uppsala, Sweden) expostos durante 24 a 168 h, dependendo da intensidade do sinal radioactivo.

3.3.5- ANÁLISES ISSR

Os marcadores ISSR foram amplificados utilizando “primers” de 17 a 19 pb, com ou sem nucleótidos aleatórios numa ou noutra extremidade (3’ ou 5’). Na tabela 13 estão discriminadas as sequências dos “primers” ISSR utilizados e seu número identificativo.

Na construção dos mapas com a população F1 os marcadores ISSR foram marcados com radioactividade, analisados em géis de poliacrilamida e revelados por autoradiografia. A análise fez-se directamente nas 46 plantas desta população e as reacções de amplificação realizaram-se num termociclador T-gradient (Biometra, Göttingen, Germany) nas condições descritas na tabela 14 e 15.

Na construção dos mapas com a população F2, os marcadores ISSR foram analisados em géis de agarose. Neste caso, a mistura reactiva foi semelhante à utilizada para os RAPDs (tabela 8) e o termociclador assim como os ciclos de amplificação foram os descritos para a análise de ISSRs na população F1 (tabela 15). Antes de analisados em toda a população F2, os marcadores ISSR foram testados em duas fases: primeiro foram seleccionados “primers” que

revelaram marcadores polimórficos entre os progenitores que se encontravam presentes na planta F1 autopolinizada para gerar a população F2; em segundo, estes “primers” foram analisados numa amostra aleatória de 5 plantas da população F2 para avaliação da intensidade da banda e da inexistência de outros marcadores moleculares de tamanho semelhante (com migração adjacente no gel de agarose) que dificultassem a análise.

Tabela 13. “Primers” ISSR analisados e respectivo número de código.

Primers ^a (5' a 3')	Nº de código	Primers ^a (5' a 3')	Nº de código	Primers ^a (5' a 3')	Nº de código
(CA)8ARG*	01	(GT)8YC ⁺	10	(CT)8RC ⁺	23
(CA)8ARY* ⁺	02	(AC)8YT ⁺	13	(TG)8RC ⁺	25
(GA)8YT* ⁺	03	(AG)8T ⁺	14	DBD(AC)7 ⁺	28
(GA)8YC*	04	(AG)8C ⁺	15	VHV(GT)7 ⁺	29
(GA)8AYC* ⁺	05	(GA)8T ⁺	17	HVH(TG)7 ⁺	30
(AG)8GYT* ⁺	06	(GA)8C ⁺	18	(AG)8VC ⁺	31
(AG)8YC*	07	(GA)8A ⁺	19	CCC(GT)7 ⁺	32
(AC)8YA*	08	(TC)8C ⁺	21		
(AC)8YG* ⁺	09	(TC)8G ⁺	22		

^a R=A+G; Y=C+T; B=T+C+G; D=A+T+G; H=A+T+C; V = A+C+G.

* “Primers” analisados na população F1.

⁺ “Primers” analisados na população F2.

Tabela 14. Mistura reactiva para amplificação de ISSRs marcados com [$\gamma^{33}\text{P}$].

Componentes	Quantidade
DNA genómico	10 ng
10x “Taq buffer” (100 mM de Tris-HCl (pH 9.0), 500 mM de KCl e 15 mM de MgCl ₂)	1 µL
Cada nucleótido (dATP, dGTP, dCTP e dTTP)	300 µM
“Primer”	1 µM
Taq DNA polymerase (Amersham Biosciences, Uppsala, Sweden)	1 U
[$\gamma^{33}\text{P}$]dCTP 370 MBq/mL (10 mCi/mL) (Amersham Bioscience, Uppsala, Sweden)	0.05 µL
Volume total	10 µL

Tabela 15. Ciclos de amplificação para os marcadores ISSR.

Programa		
Número de ciclos	Temperatura ¹ (°C)	Tempo
1	94	4 mins
40	94	30 s
	52	45 s
	72	2 mins
1	72	7 mins

¹ A variação de temperatura foi de 1 °C/s.

Na análise com a população F1, depois da amplificação adicionou-se 2.5 µL de solução de azul de bromofenol a cada reacção, e os produtos amplificados foram separados por electroforese em géis não desnaturantes de 6% de poliacrilamida (Long Ranger Gel Solution, Cambrex Bio Science, Rockland inc., USA) durante cerca de 2 h 30 mins a 45 W. Como tampão de gel e de electroforese utilizou-se 1xTBE. Depois da electroforese, os géis foram secos e as autoradiografias obtidas como acima descrito para os marcadores AFLP.

Na análise efectuada com a população F2, os marcadores ISSR foram revelados em géis de agarose. Depois da amplificação adicionou-se 3.75 µL de solução de azul de bromofenol a cada reacção e os produtos de amplificação foram separados por electroforese em gel de agarose de 2%, durante cerca de 3 h a 100 v. Utilizou-se 0.5xTBE como tampão de gel e de electroforese. Após electroforese, os marcadores ISSR foram visualizados como atrás descrito para os marcadores RAPD.

3.4- ANÁLISE DA SEGREGAÇÃO DOS MARCADORES

Os dados da segregação de marcadores foram registados numa matriz com códigos de presença ou ausência de bandas para os marcadores dominantes (RAPDs, AFLPs, ISSRs e a

característica morfológica “piping”) e códigos para heterozigóticos ou homozigóticos de um e/ou de outro alelo no caso dos marcadores codominantes (isoenzimas), de acordo com a nomenclatura utilizada pelo programa informático JoinMap 3.0 (Van Ooijen and Voorrips 2001). Para determinar a concordância da razão de segregação observada para cada marcador com a razão de segregação mendeliana esperada (1:1, 3:1 ou 1:2:1), utilizou-se o teste estatístico do qui quadrado (χ^2).

3.5- CONSTRUÇÃO DOS MAPAS

A análise de ligamento para distribuição dos marcadores por grupos, para a ordenação dos marcadores de cada grupo e para o cálculo das distâncias entre marcadores, foi efectuada com o programa informático JoinMap 3.0 (Van Ooijen e Voorrips 2001). Este programa utiliza um teste estatístico que não é afectado por marcadores com razão de segregação afastada da esperada, permite a utilização de todos os tipos de populações de mapeamento e a combinação de informação de diferentes populações para construir mapas integrados. Utilizou-se a função de Kosambi (1944) para calcular distâncias genéticas em centimorgans (cM) a partir de taxas de recombinação.

Na construção dos mapas com a população F1 utilizou-se um LOD mínimo de 4.0 para o agrupamento dos marcadores e um LOD mínimo de 1.0 para o ordenamento dos marcadores de cada grupo. Para identificar possíveis homologias entre grupos dos mapas independentes construídos para os dois progenitores foram utilizados os marcadores presentes nos dois progenitores no estado heterozigótico (marcadores heterozigóticos monomórficos) com segregação esperada de 3:1. A identificação de possíveis homologias fez-se apenas entre

grupos com pelo menos 3 marcadores polimórficos heterozigóticos (com segregação 1:1) nos dois mapas, uma restrição sugerida por Debener e Mattiesch (1999).

Na construção dos mapas com a população F2 o agrupamento dos marcadores fez-se com valores limite de LOD entre 3.0 e 7.0, em intervalos de 0.5. Os grupos finais foram estabelecidos de acordo com a estabilidade dos mesmos à medida que o LOD aumenta e tendo em conta a informação dos mapas construídos com a população F1. Só foram estabelecidos grupos de ligamento com LOD limite inferior a 4 quando a relação de ligamento dos marcadores constantes nesses grupos se confirma nos mapas construídos anteriormente com base na população F1. O ordenamento dos marcadores nos grupos fez-se com um LOD mínimo limite de 1.0.

Para a integração dos mapas construídos com as duas populações (F1 e F2) os grupos obtidos separadamente com ambas e com pelo menos 2 marcadores comuns foram combinados utilizando a função “combine groups for map integration” do menu “join” do JoinMap 3.0. Em seguida a ordem dos marcadores dos grupos combinados foi recalculada com base nos dados de segregação desses marcadores nas duas populações para um LOD mínimo de 1.0. Para os grupos calculados separadamente com as populações F1 e F2 com apenas 1 marcador comum é possível apenas estabelecer uma ligação, não sendo possível a integração e ordenação dos marcadores num só grupo de ligamento. Estes grupos foram dispostos lado a lado com uma linha a ligar o marcador comum, e para efeitos de cálculo de dimensão total do mapa tomou-se em consideração unicamente a dimensão do grupo maior.

3.6- IDENTIFICAÇÃO DOS GRUPOS E DOS MARCADORES

Os nomes dos grupos obtidos com base na População F1 começam com as letras Ab ou Ac, respectivamente para a variedade *bracteatus* ou para a var. *comosus*, seguidas de um número atribuído segundo a ordem decrescente da dimensão dos grupos na primeira versão dos mapas publicada (Carlier et al. 2004). Os grupos obtidos com a População F2, depois de divididos em 3 tipos (com marcadores dos dois progenitores, com marcadores só de um progenitor e com marcadores só do outro progenitor) foram identificados por um número atribuído segundo a ordem decrescente dos seus comprimentos. Os grupos que integram marcadores mapeados com as duas populações estão identificados pela designação do grupo F2 seguida da designação do(s) grupo(s) F1 integrado(s), *p.e.* 12_Ab4_Ac6 para o grupo 12 obtido com a população F2, integrado com os grupos Ab4 e Ac6 dos mapas obtidos com a população F1.

Os marcadores RAPD estão identificados pela(s) letra(s) que identificam o “Kit” de “primers”, seguidos do número de código do “primer” (de 01 a 20) e do tamanho estimado do fragmento (pb), *p.e.* OPA03700 para o marcador de 700 pb gerado pelo “primer” 03 do “kit” A da Operon Technologies Inc. (Alameda, CA, USA). Os marcadores AFLP e ISSR foram identificados, respectivamente, pelas letras maiúsculas AFLP ou ISSR, dois dígitos correspondentes a cada combinação de “primers” AFLP (Tabela 9) ou “primer” ISSR (Tabela 13), seguidos do tamanho estimado do fragmento, *p.e.* AFLP01200 para um marcador AFLP de 200 pb gerado pela combinação de “primers” 01 e ISSR06300 para um marcador ISSR de 300 pb gerado pelo “primer” 06. Os marcadores que apresentam um pequeno desvio da segregação esperada ($\chi^2_{\alpha=0.05} < \chi^2 \leq \chi^2_{\alpha=0.01}$) estão identificados com um asterisco (*). Os que têm um desvio mais acentuado ($\chi^2 > \chi^2_{\alpha=0.01}$) estão identificados com dois asteriscos (**). Os

marcadores analisados na população F2 que não foram analisados na população F1 (novos marcadores) estão identificados com “n_” no início do nome do marcador. As letras Ab ou Ac no fim do nome do marcador identificam a origem do mesmo, respectivamente, var. *bracteatus* ou var. *comosus*, e as letras cb identificam os marcadores heterozigóticos monomórficos entre os progenitores. O símbolo # identifica marcadores heterozigóticos monomórficos entre os progenitores utilizados para estabelecer homologias: ligados nos dois mapas (var. *bracteatus* e var. *comosus*) a pelo menos 3 marcadores heterozigóticos polimórficos.

3.7- COMPRIMENTO E COBERTURA DO MAPA

O comprimento total do mapa foi calculado somando o comprimento (em cM) de todos os grupos, aos quais foi adicionada a cada extremidade a distância média entre marcadores, como sugerido por Marques et al. (1998).

A cobertura do mapa foi estimada como a fração do comprimento total do mapa pelo tamanho do genoma estimado segundo o método 3 de Chakravarti et al. (1991):

$$G = (n \times (n-1) \times X) / k$$

onde G é a estimativa do tamanho do genoma, n o número de marcadores, k o número de pares de marcadores observado para LOD de agrupamento maior ou igual a Z e X é a distância entre o par de marcadores com o maior valor de frequência de recombinação entre os pares k.

3.8- DISTRIBUIÇÃO DOS MARCADORES NO MAPA

Para estudar a distribuição dos marcadores no mapa fez-se o teste “Kolmogorov-Smirnov one sample test” para uma distribuição normal (Sokal e Rohlf 1981), utilizando o programa informático SPSS 13.0 (2004).

3.9- CLONAGEM E SEQUENCIAÇÃO DE MARCADORES

Os marcadores a sequenciar foram amplificados e as respectivas bandas cortadas do gel de agarose. Os fragmentos de DNA contidos nessas porções de gel foram purificados por ligação a uma matriz de fibra de vidro e eluição em TE com o kit “GFX PCR DNA and gel Band purification” (Amersham Biosciences, Uppsala, Sweden). Uma alíquota do DNA isolado foi analisado por electroforese em gel de agarose para confirmar a dimensão do fragmento e estimar a quantidade de DNA purificado.

Para a clonagem dos fragmentos de DNA utilizou-se o “Topo TA cloning[®] Kit for sequencing” (Invitrogen, Paisley, UK) e a transformação bacteriana fez-se por electroporação.

A mistura reactiva de ligação ao plasmídeo consistiu em 0.5 a 4µl do produto de PCR (quantidade estimada de acordo com a intensidade da banda no gel de agarose), 50 mM de NaCl, 2.5 mM de MgCl₂ e 10 ng de vector pCR[®]4- Topo[®] (Invitrogen, Paisley, UK), num volume final de 6µl. A reacção fez-se durante 30 mins a 22.5 °C.

Para a electroporação adicionou-se 2 µl da reacção de ligação a 50 µl de células electrocompetentes TOP10 Electrocomp[™] (Invitrogen, Paisley, UK) e misturou-se

suavemente. Esta solução foi transferida cuidadosamente, evitando a formação de bolhas, para uma cuvette de electroporação de 0.1 cm (Invitrogen, Paisley, UK) previamente arrefecida em gelo e procedeu-se à electroporação de acordo com o manual do electroporador (Electroporator II, Invitrogen, Paisley, UK). A tabela 16 resume as condições de electroporação.

Tabela 16. Condições da electroporação bacteriana.

Condições do “Power supply”		Condições do Electroporador	
Voltagem:	1500 V	Capacidade:	50 μ F
Corrente:	25 mA	Resistência:	150 Ω
Potência:	25 W		
“Pulse Length” (com cuvette de 0.1 cm) =		7,5 ms	
Campo de força nestas condições =		15000 V/cm	

Imediatamente depois da inserção do vector nas bactérias por electroporação adicionou-se a estas 250 μ l de meio SOC à temperatura ambiente e foram transferidas para um tubo de cultura bacteriana que foi colocado sob agitação orbital durante uma hora a 37°C para permitir a expressão dos genes de resistência ao antibiótico antes da selecção via resistência vs susceptibilidade ao mesmo.

Para a selecção de colónias transformadas plaquearam-se 10 a 50 μ l da suspensão bacteriana obtida no passo anterior em caixas de Petri contendo meio LB com antibiótico (75 μ g/mL de ampicilina) e incubaram-se durante a noite a 37 °C. Depois, recolheram-se colónias transformadas e amplificou-se o DNA bacteriano por PCR (1) com os “primers” universais M13 e (2) com os “primers” RAPD específicos para cada marcador, a fim de confirmar a presença e a dimensão dos fragmentos clonados. Finalmente, para cada marcador foram seleccionadas para sequenciação três colónias bacterianas cujo plasmídeo revelou por PCR ter o “insert” de DNA de tamanho esperado.

Para multiplicação dos plasmídeos contendo os marcadores de DNA, as respectivas colónias bacterianas foram transferidas para 5 mL de meio LB líquido com antibiótico (75 µg/mL de ampicilina) e incubadas durante a noite a 37 °C num agitador orbital. Na manhã seguinte as culturas foram renovadas a partir de um inóculo de 10% da cultura anterior e incubadas a 37 °C num agitador orbital até atingirem uma densidade de 1 a 1.5 x 10⁹ células/mL.

A purificação dos plasmídeos fez-se com o Kit “S.N.A.PTM Miniprep” (Invitrogen, Paisley, UK). Após a purificação o DNA plasmídico foi analisado em gel de agarose com o objectivo de se confirmar a presença das bandas correspondentes ao plasmídeo e a ausência de bandas correspondentes a RNA. Depois fez-se a quantificação do DNA plasmídico por espectrofotometria UV.

A sequenciação dos marcadores foi encomendada à empresa Macrogen (Seoul, Korea), e o marcador foi considerado sequenciado quando pelo menos duas das três sequências eram iguais.

3.10- ANÁLISE DE SEQUÊNCIAS

A procura de similaridades entre os marcadores sequenciados e sequências de bases de dados mundiais fez-se através da página da Internet do “National Center for Biotechnology Information – NCBI” (<http://www.ncbi.nlm.nih.gov>), com o programa Blastn (Altschul et al. 1997). Procuraram-se similaridades em duas bases de dados: (1) a generalizada “nr” para sequências nucleotídicas (<http://www.ncbi.nlm.nih.gov/BLAST/producttable.shtml#db>) e, (2) como sugerido por McGinnis e Madden (2004) para aumentar a eficiência da procura, numa base de dados mais restrita, neste caso a “plant genomes mapped DNA sequences” do recurso

de procura do NCBI orientado para plantas “Customized plant genome blast” (<http://www.ncbi.nlm.nih.gov/BLAST/Genome/PlantBlast.shtml?2>).

Os marcadores para os quais foram encontradas sequências com “expect value” $\leq 10^{-8}$, como sugerido por Schloss et al. (2002) e Jones et al. (2002b), foram seleccionados para posterior avaliação. Com estes marcadores fez-se uma pesquisa com o programa Blastx (Altschul et al. 1997) de procura de proteínas similares a partir da tradução de uma sequência de nucleótidos. Esta pesquisa foi efectuada na base de dados generalizada “nr” de sequências proteicas (<http://www.ncbi.nlm.nih.gov/BLAST/producttable.shtml#db>).

4- RESULTADOS

4.1- MAPEAMENTO COM A POPULAÇÃO F1

4.1.1- MARCADORES ANALISADOS

Dos 2 sistemas de isoenzimas analisados, apenas o PGM foi útil para mapeamento. Este sistema revelou 2 bandas (monomórficas) em cada progenitor com uma segregação na população de mapeamento típica de 2 alelos do mesmo *locus*. Denominou-se este *locus* Isoz_PGM. O sistema MDH revelou 5 bandas monomórficas em toda a população de mapeamento (figura 1).

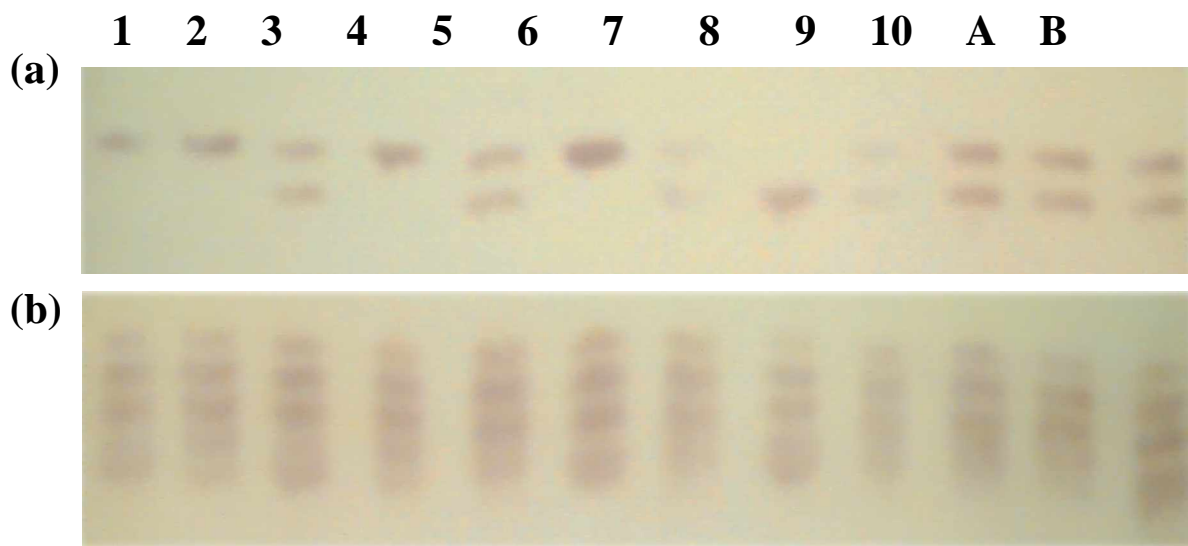


Figura 1. Zimogramas dos sistemas isoenzimáticos PGM (a) e MDH (b) de 10 plantas da população F1 (1-10) e dos progenitores var. *bracteatus* (A) e var. *comosus* (B).

Dos 300 “primers” RAPD testados inicialmente nos progenitores, 70 (23.3%) (tabela 7) amplificaram 134 marcadores polimórficos heterozigóticos (83 da var. *bracteatus* e 51 da var. *comosus*) e 9 marcadores monomórficos heterozigóticos apropriados para mapeamento na população F1.

As 48 combinações de “primers” AFLP analisadas na população de mapeamento F1 (tabela 9) revelaram 492 marcadores polimórficos heterozigóticos (339 da var. *bracteatus* e 153 da var. *comosus*) e 46 marcadores monomórficos heterozigóticos.

Na população F1 os marcadores ISSR foram analisados em géis de poliacrilamida e revelados por autoradiografia. Os nove “primers” ISSR analisados nesta população (tabela 13) revelaram 49 marcadores polimórficos heterozigóticos (26 da var. *bracteatus* e 23 da var. *comosus*) e 8 marcadores monomórficos heterozigóticos, cuja segregação foi possível analisar.

Apesar do modo de hereditariedade dominante do tipo de marcadores-DNA utilizados foi possível identificar 2 loci (1 RAPD e 1 AFLP) que apresentam dois alelos em segregação codominante (figura 2).

O número de marcadores polimórficos heterozigóticos identificados na var. *bracteatus* foi quase o dobro dos identificados na var. *comosus*, o que sugere um nível de heterozigocidade mais elevado para a primeira. Uma estimativa com base em 150 marcadores polimórficos revelou um nível de heterozigocidade de 93.8% para a var. *bracteatus* e de 62.9% para a var. *comosus*. Conseqüentemente a construção do mapa genético foi praticamente duas vezes mais eficiente para a var. *bracteatus* do que para a var. *comosus*.

Uma maior eficiência na construção do mapa do progenitor com um nível de heterozigocidade mais elevado foi também descrita por Lespinasse et al. (2000) na árvore-da-borracha (*Hevea* spp.) e por Hurtado et al. (2002) em damasqueiro (*Prunus armeniaca*). Em ambos os casos o progenitor mais heterozigótico era de origem híbrida. No que se refere à

variedade mais heterozigótica utilizada neste trabalho, a var. *bracteatus*, os estudos de diversidade molecular indicam que é a mais diferenciada das outras variedades e a que tem mais afinidade com outra espécie do género *Ananas* (*A. macrodontes*), o que segundo Duval et al. (2003) aponta para uma origem híbrida da var. *bracteatus*. O facto desta variedade revelar mais marcadores específicos que não estão presentes no outro progenitor (var. *comosus*) também contribuiu para uma maior eficiência na construção do seu mapa genético.

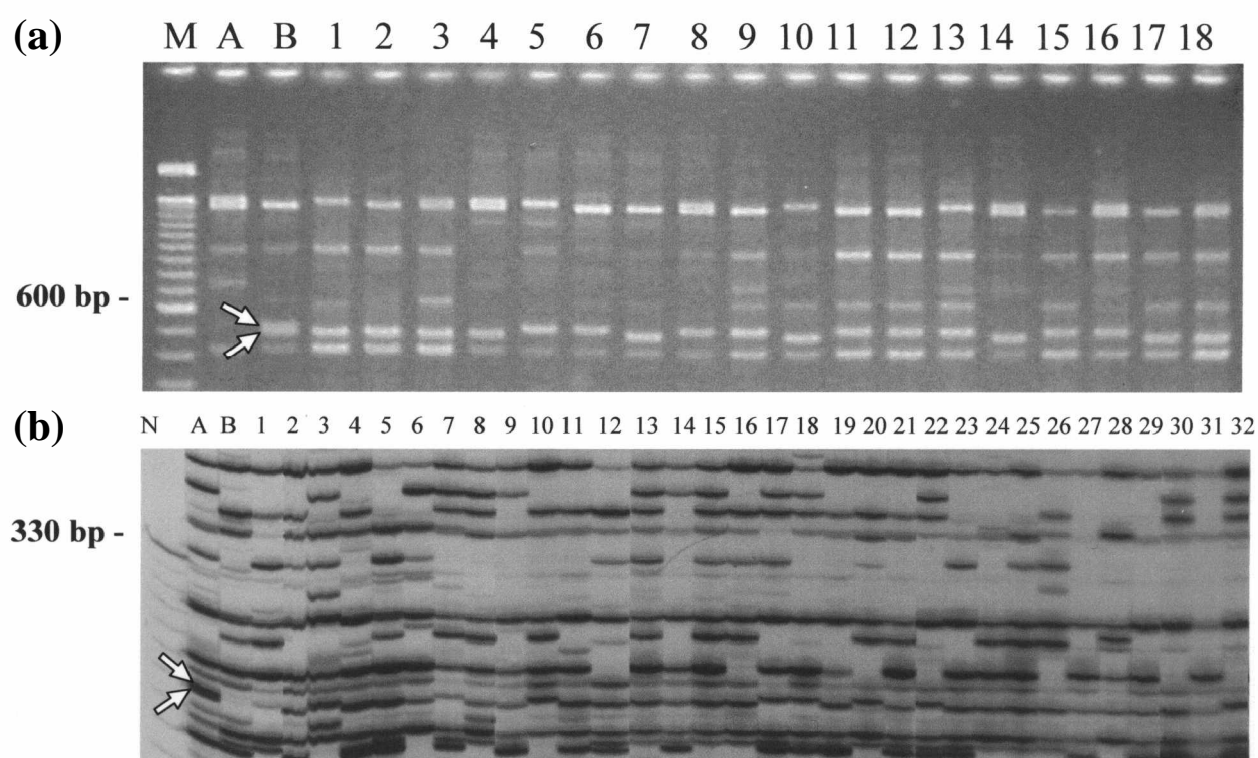


Figura 2. (a) Padrões de amplificação RAPD de 18 plantas da população F1 (1-18) e dos progenitores var. *bracteatus* (A) e var. *comosus* (B), obtidos com o “primer” OPB15. M-“escada” de marcadores padrão de 100 em 100 pb. (b) Marcadores AFLP em 32 plantas da população F1 (1-32) e nos progenitores var. *bracteatus* (A) e var. *comosus* (B), obtidos com a combinação de “primers” M-CTA/E-AAG. Tanto em (a) como em (b), as setas indicam pares de marcadores cuja segregação em alternância sugere serem alelos do mesmo *locus*. N-“escada” de marcadores padrão de 10 em 10 pb.

Dos 448 marcadores de DNA polimórficos heterozigóticos da var. *bracteatus*, 34 (7.6%) evidenciaram uma razão de segregação muito desviada ($\chi^2 > \chi^2_{\alpha=0.01}$) da razão de segregação esperada para uma população F1 (1:1), e 37 (8.3%) apresentaram uma razão de segregação com um desvio menos acentuado ($\chi^2_{\alpha=0.05} < \chi^2 \leq \chi^2_{\alpha=0.01}$). Dos 227 marcadores de DNA polimórficos heterozigóticos da var. *comosus*, 15 (6.6%) evidenciaram uma razão de segregação desviada da esperada com um $\chi^2 > \chi^2_{\alpha=0.01}$, e 18 (7.9%) com $\chi^2_{\alpha=0.05} < \chi^2 \leq \chi^2_{\alpha=0.01}$. Quanto aos marcadores monomórficos heterozigóticos, apenas 3 marcadores de DNA apresentaram uma razão de segregação desviada da esperada (3:1), e os três com $\chi^2_{\alpha=0.05} < \chi^2 \leq \chi^2_{\alpha=0.01}$.

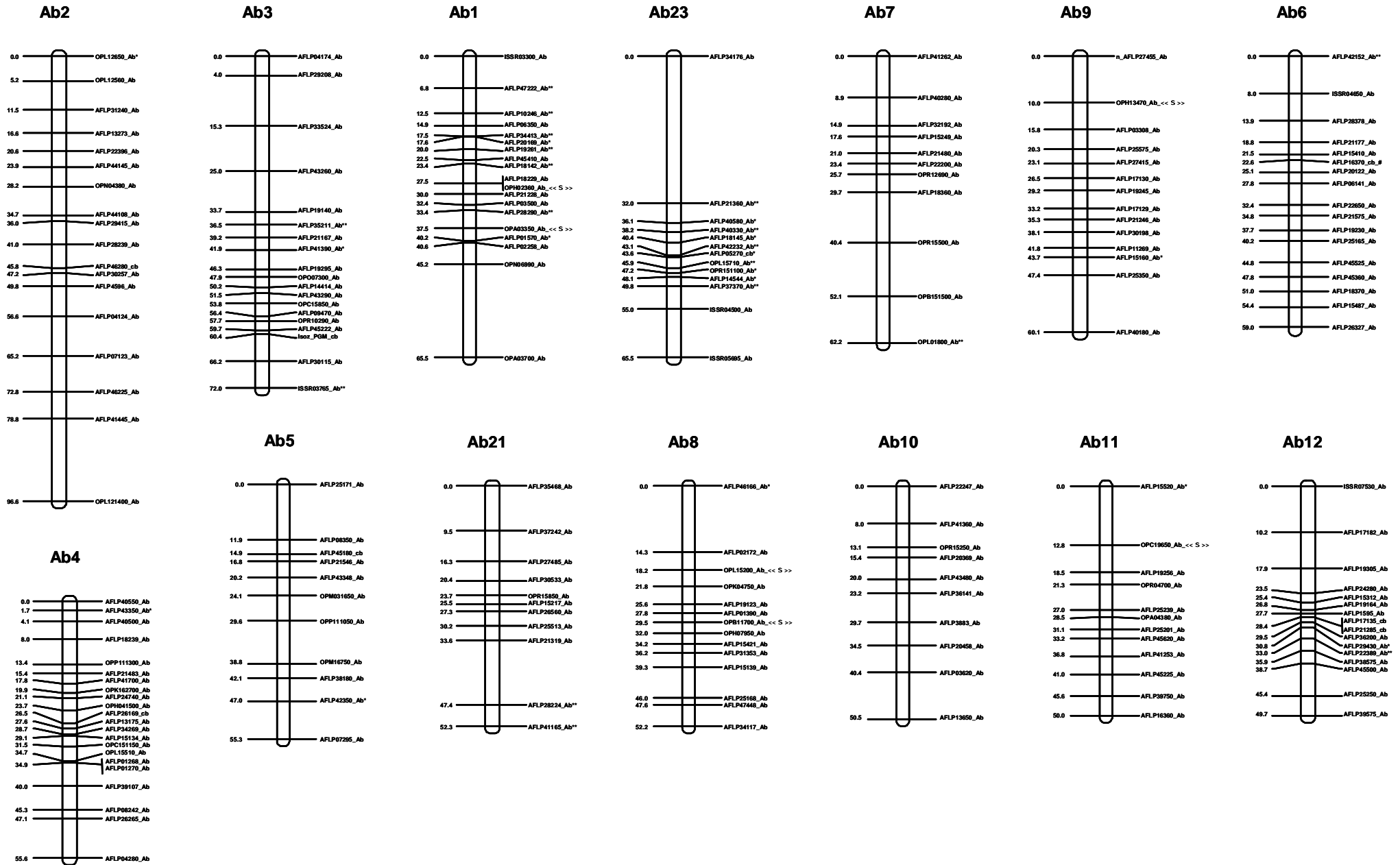
Devido ao acaso são esperados 5% de marcadores que apresentam uma segregação cujo desvio da segregação esperada revela um valor de $\chi^2 > \chi^2_{\alpha=0.05}$. Na população F1 a proporção de marcadores com segregação desviada da esperada ($\chi^2 > \chi^2_{\alpha=0.05}$) foi de 15.9% para a var. *bracteatus* e 14.5% para a var. *comosus*. Estes valores estão entre os geralmente estimados para outras espécies de plantas e podem ter como causas interações alélicas “intra-locus” ou “inter-loci” desfavoráveis ou o emparelhamento ectópico dos cromossomas na meiose levando à formação de cromossomas aberrantes depois da recombinação e consequentemente a distorções na segregação (Jenczewski et al. 1997).

O fenótipo “piping” apresentou na população F1 uma razão de segregação concordante ($\chi^2 = 0.4$) com a razão de segregação mendeliana esperada de 1:1, o que corrobora o modelo de hereditariedade monogénica proposto para esta característica. Collins e Kerns (1946) mostraram que o tipo de margem da folha é controlado por dois genes (*S* e *P*), tendo os seus resultados sido posteriormente corroborados por Cabral et al. (1997). Segundo estes autores, o fenótipo “Spiny”, com espinhos ao longo de toda a folha, corresponde ao alelo recessivo *s*. O fenótipo “spiny tip” característico da cultivar Smooth Cayenne, com espinhos apenas na ponta da folha, corresponde à presença do alelo *S*, dominante sobre o alelo *s*. O

fenótipo “spineless”, as folhas lisas do *A. comosus* var. *erectifolius* (antes *A. lucidus*), corresponde à presença do *Se*, um terceiro alelo que é dominante sobre os alelos *S* e *s*. A característica “piping”, enrolamento e fusão da margem da folha suprimindo os espinhos da mesma, deve-se à presença do alelo dominante do gene *P*, o qual é epistático sobre o gene *S*.

4.1.2- MAPA DA VAR. *BRACTEATUS*

Testou-se o agrupamento dos marcadores para os valores limite de LOD 3.0, 4.0 e 5.0. A utilização de um LOD mínimo de agrupamento de 3.0 permite reduzir significativamente o número de marcadores não ligados. No entanto a incorporação desses marcadores no mapa é acompanhada pela formação de grupos anormalmente grandes, cuja ordem de marcadores não é possível determinar. Inversamente, um LOD mínimo de agrupamento de 5.0 faz aumentar significativamente o número de marcadores não ligados. Por isso escolheu-se 4.0 como valor limite de LOD de agrupamento. Com este valor foram encontradas relações de ligamento em 341 marcadores polimórficos heterozigóticos (62 RAPDs, 267 AFLPs e 12 ISSRs) e 22 marcadores monomórficos heterozigóticos (2 RAPDs, 18 AFLPs, 1 ISSR e o locus *Isoz_PGM*), distribuídos em 51 grupos. Vinte e sete grupos de ligamento têm pelo menos 4 marcadores, seis grupos têm 3 marcadores e 18 são pares de marcadores (figura 3). Cento e sete marcadores polimórficos heterozigóticos (23.9%) e 42 monomórficos heterozigóticos (65.6%) não ficaram ligados.



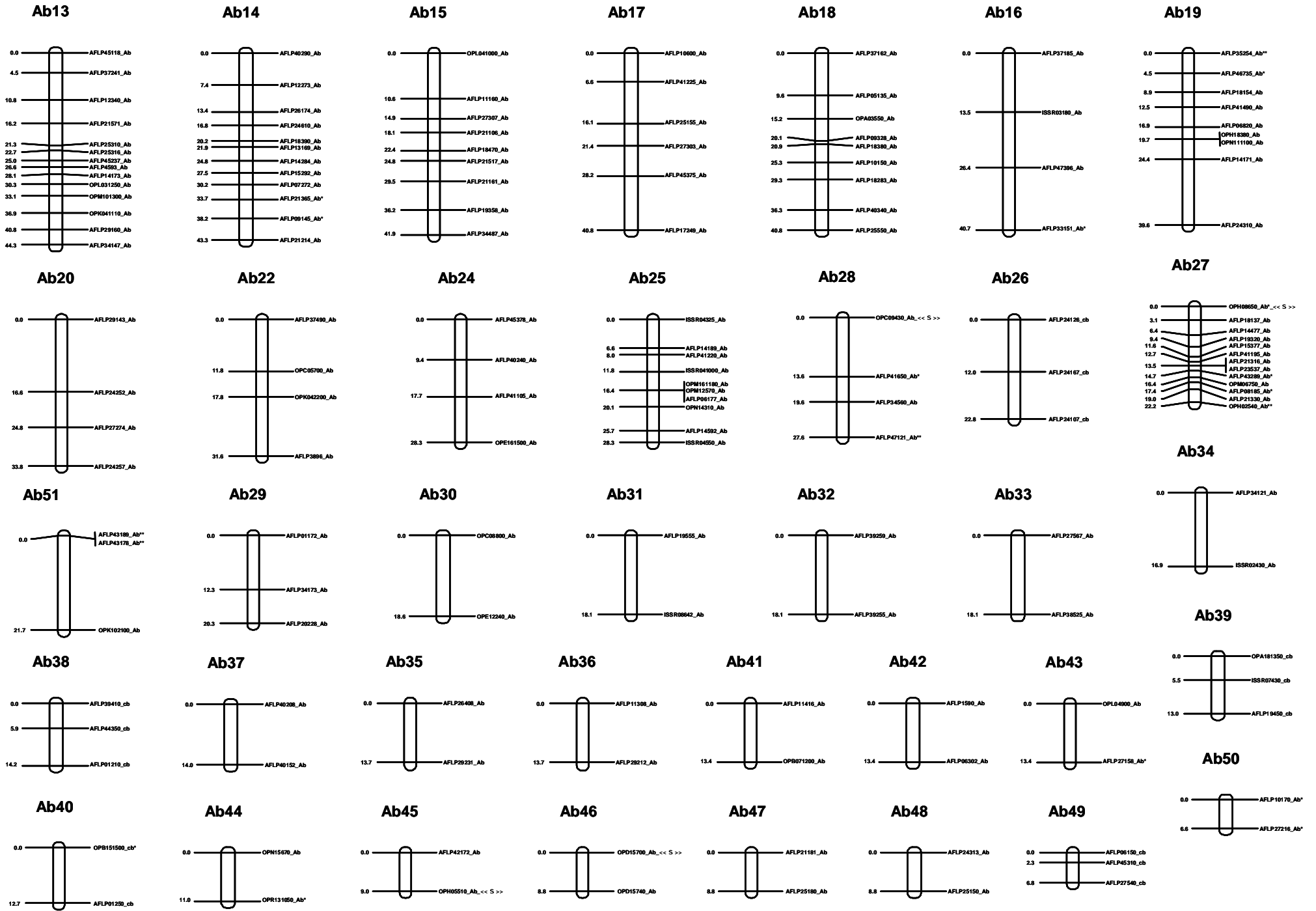
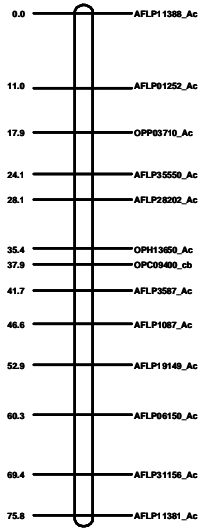
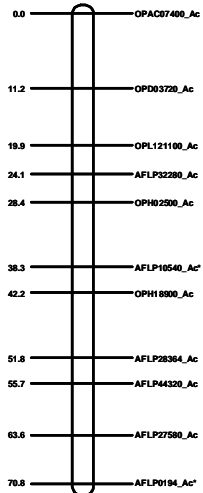
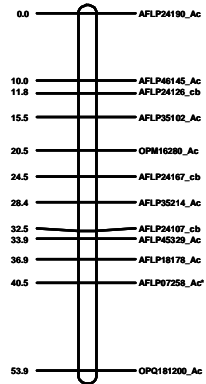
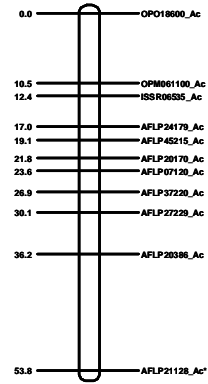
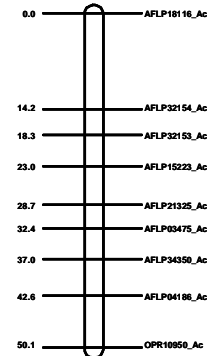
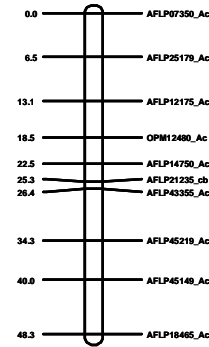
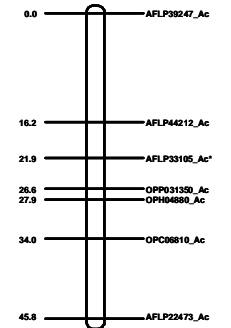
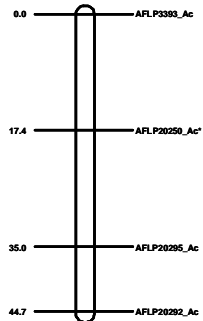
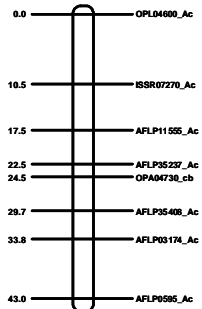
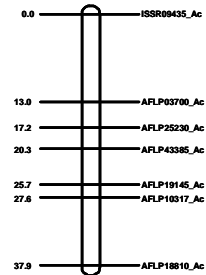
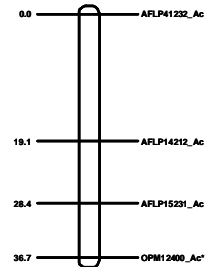
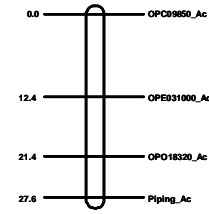
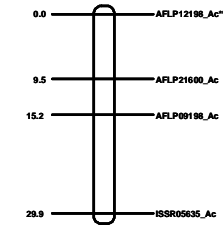
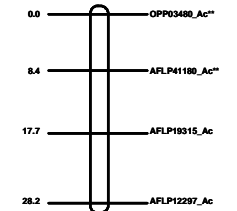


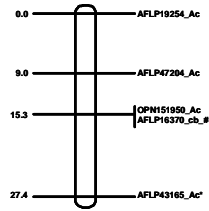
Figura 3. Mapa genético obtido com a população F1 para a var. *bracteatus*. Os números à esquerda indicam a distância em cM ao topo do grupo. As letras Ab no fim do nome do marcador identificam a origem do mesmo – var. *bracteatus*; as letras cb identificam os marcadores heterozigóticos monomórficos entre os progenitores *i.e.* a segregar para ambos. Os marcadores cuja segregação têm um desvio relativamente pouco afastado da segregação esperada ($\chi^2_{\alpha=0.05} < \chi^2 \leq \chi^2_{\alpha=0.01}$) estão identificados com um asterisco (*), os marcadores com segregação com um desvio mais acentuado ($\chi^2 > \chi^2_{\alpha=0.01}$) estão identificados com dois asteriscos (**). O símbolo # identifica um marcador a segregar para os dois progenitores que permitiu identificar uma possível homologia entre um grupo do mapa da var. *bracteatus* e um grupo do mapa da var. *comosus*. O símbolo «s» indica marcadores que foram sequenciados. Os nomes dos grupos começam com as letras Ab, seguidas de um número atribuído segundo a ordem decrescente da dimensão dos grupos na primeira versão dos mapas publicada (Carlier et al. 2004).

4.1.3- MAPA DA VAR. *COMOSUS*

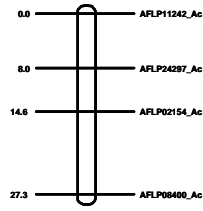
Por razões de análise e comparação dos mapas obtidos para os dois progenitores o agrupamento dos marcadores da var. *comosus* fez-se, com o valor de LOD de agrupamento mínimo igual ao escolhido para a var. *bracteatus* (4.0). Com este valor, foram encontradas relações de ligamento em 140 marcadores polimórficos heterozigóticos (31 RAPDs, 101 AFLPs, 7 ISSRs e o marcador morfológico “piping”) e 22 marcadores monomórficos heterozigóticos (4 RAPDs, 17 AFLPs e 1 ISSR), os quais ficaram distribuídos em 31 grupos. Vinte grupos têm pelo menos 4 marcadores, três grupos têm 3 marcadores e 8 grupos são pares de marcadores (Figura 4). Oitenta e oito marcadores polimórficos heterozigóticos (38.6%) e 42 monomórficos heterozigóticos (65.6%), incluindo o marcador Isoz_PGM, não ficaram ligados. Baixando o LOD de agrupamento para 2.5 é possível estabelecer ligamento entre o locus Isoz_PGM e o grupo Ac9, no entanto, devido ao fraco suporte estatístico desta ligação não é possível determinar a sua posição no grupo.

Ac1**Ac2****Ac3****Ac4****Ac5****Ac6****Ac7****Ac8****Ac9****Ac10****Ac11****Ac12****Ac17****Ac25**

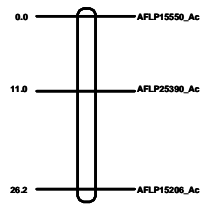
Ac13



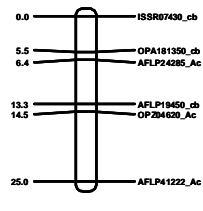
Ac14



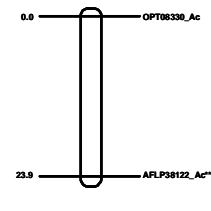
Ac15



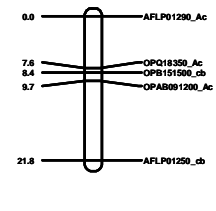
Ac16



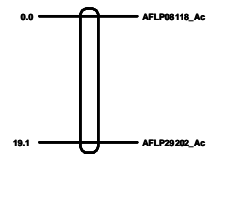
Ac31



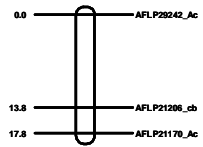
Ac18



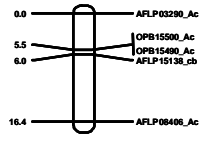
Ac19



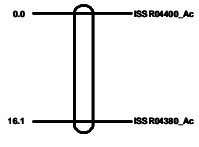
Ac20



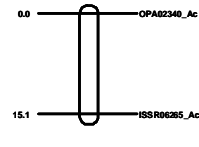
Ac21



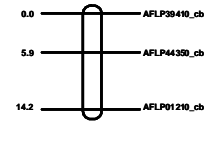
Ac22



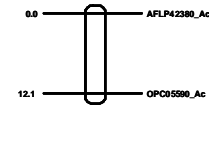
Ac23



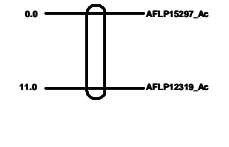
Ac24



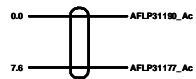
Ac26



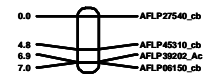
Ac27



Ac28



Ac29



Ac30



Figura 4. Mapa genético obtido com a população F1 para a var. *comosus*. Os números à esquerda indicam a distância em cM ao topo do grupo. As letras Ac no fim do nome do marcador identificam a origem do mesmo - var. *comosus*; as letras cb identificam os marcadores heterozigóticos monomórficos entre os progenitores. Os marcadores com segregação desviada da esperada com $\chi^2_{\alpha=0.05} < \chi^2 \leq \chi^2_{\alpha=0.01}$ estão identificados com um asterisco (*), os marcadores com segregação desviada da esperada com $\chi^2 > \chi^2_{\alpha=0.01}$ estão identificados com dois asteriscos (**). O símbolo # identifica um marcador que permitiu identificar uma possível homologia entre grupos dos dois mapas (var. *bracteatus* e var. *comosus*). Os nomes dos grupos começam com as letras Ac, seguidas de um número atribuído segundo a ordem decrescente da dimensão dos grupos na primeira versão dos mapas publicada (Carlier et al. 2004).

4.1.4- INTEGRAÇÃO DOS MAPAS DA VAR. *BRACTEATUS* E DA VAR. *COMOSUS*

Neste trabalho considerou-se condição mínima para identificar grupos possivelmente representantes de cromossomas homólogos a ligação dos marcadores a segregar para os dois progenitores (heterozigóticos e monomórficos entre os progenitores) a pelo menos 3 marcadores heterozigóticos polimórficos de cada mapa *i.e.* 3 marcadores a segregar só para um progenitor e 3 a segregar só para o outro). O marcador AFLP16370_cb_# foi o único que preencheu este requisito, permitindo identificar uma possível homologia entre o grupo de ligamento Ab6 do mapa da var. *bracteatus* e o grupo Ac13 do mapa da var. *comosus*. No entanto, devido à falta de mais marcadores ligados a ambos os grupos não foi possível a sua integração nem a determinação da sua orientação relativa.

4.1.5- COMPRIMENTO E COBERTURA DOS MAPAS

Utilizando para o cálculo o método 3 de Chakravarti et al. (1991), obteve-se uma estimativa para o comprimento do genoma do *A. comosus* var. *bracteatus* de 3693 cM. O comprimento total do mapa construído para esta variedade botânica, calculado somando a distância média entre marcadores (5.27 cM) a ambas as extremidades de cada grupo, é de 2184 cM e representa 59.1% de cobertura do genoma. Utilizando a mesma metodologia de cálculo, o mapa da variedade *comosus*, com uma distância média entre marcadores de 7.13 cM e um comprimento total de 1377 cM, cobre 33.2% do genoma estimado em 4146 cM.

A estimativa da dimensão genética dos genomas de 15 espécies de plantas superiores aponta para comprimentos médios por cromossoma de 90 a 217 cM, e um valor médio para as 15 espécies de 143 cM (tabela 17). Os valores obtidos neste trabalho de 148 cM e 166 cM por cromossoma, respectivamente para as variedades *bracteatus* e *comosus*, estão dentro deste intervalo e próximos da respectiva média.

Tendo em conta as estimativas de tamanho físico dos genomas haplóides da var. *bracteatus* e var. *comosus*, respectivamente, de 444 e 526 Mpb, baseadas em dados de citometria de fluxo (Arumuganathan e Earle 1991), pode estabelecer-se uma relação média entre distância física e distância genética de 120 Kpb/cM para a var. *bracteatus* e 127 Kpb/cM para a var. *comosus*. Estes valores são aproximadamente metade do valor estimado para *Arabidopsis thaliana*, o valor mais baixo da tabela 17, o que está de acordo com o facto do ananaseiro ter um número elevado de cromossomas ($n = 25$) e de a dimensão do seu genoma ser relativamente pequena.

Tabela 17. Características genômicas de 15 espécies de plantas (adaptado de Chagné et al. 2002).

Espécies	Dimensão física (Mpb)	Comprimento genético (cM)	Número de cromossomas	Dimensão genética média por cromossoma (cM)	Dimensão física / Dimensão genética (Mpb/cM)
<i>Arabidopsis thaliana</i>	150	675	5	135	0.22
<i>Prunus pérsica</i>	300	712	8	90	0.42
<i>Oriza sativa</i>	450	1490	12	125	0.3
<i>Populus deltoides</i>	550	2300	19	121	0.23
<i>Eucalyptus grandis</i>	600	1370	11	125	0.43
<i>Brassica rapa</i>	650	1850	10	185	0.35
<i>Quercus robur</i>	900	1200	12	100	0.75
<i>Lycopersicon esculentum</i>	980	1280	12	107	0.76
<i>Solanum tuberosum</i>	1540	1120	12	90	1.37
<i>Zea mays</i>	2500	1860	10	186	1.34
<i>Lactuca sativa</i>	2730	1950	9	217	1.4
<i>Triticum tauschii</i>	4200	1330	7	190	3.15
<i>Hordeum vulgare</i>	5500	1250	7	178	4.4
<i>Pinus taeda</i>	21000	1700	12	141	12.35
<i>Pinus pinaster</i>	25700	1850	12	154	13.78

4.1.6- DISTRIBUIÇÃO DOS MARCADORES NOS MAPAS

O teste estatístico “One Sample Kolmogorov-Smirnov” (OSKS) revelou que no mapa construído para a variedade botânica *bracteatus* a distribuição dos intervalos entre marcadores AFLP se desvia significativamente de uma distribuição normal ($P < 0.001$). Isto sugere que a distribuição destes marcadores não ocorreu de forma aleatória e que houve tendência para a sua agregação em determinadas zonas. No mapa construído para a variedade botânica *comosus* o teste OSKS revelou uma distribuição normal ($P = 0.017$) dos intervalos entre marcadores AFLP.

Os intervalos entre RAPDs no mapa da var. *bracteatus* apresentam uma distribuição normal ($P = 0.113$, teste OSKS). No entanto, o pequeno número (35) de marcadores RAPD mapeados para a var. *comosus* não permitiu a análise da sua distribuição com o teste OSKS.

Quanto aos ISSRs, parece haver uma certa tendência na sua distribuição pois em ambos os mapas estes marcadores estão localizados nas extremidades de grupos maiores ou em grupos pequenos. No entanto é prematuro tirar qualquer ilação desse facto pois o número deste tipo de marcadores mapeados foi muito baixo (13 na var. *bracteatus* e 8 na var. *comosus*).

A concentração de marcadores em áreas específicas de grupos de ligamento é um fenómeno comum na construção de mapas genéticos que foi descrito anteriormente em várias espécies e para diferentes tipos de marcadores. Para AFLPs este fenómeno foi descrito por exemplo em batata (*Solanum tuberosum*) por Rouppe van der Voort et al. (1997), em soja (*Glycine max*) por Keim et al. (1997) e em rosa (*Rosa* spp.) por Crespel et al. (2002). Distribuições não aleatórias de marcadores RAPD foram descritas por Krutovskii et al. (1998) em *Pseudotsuga menziessi* (uma conífera), e por Nilson et al. (1997) em *Beta vulgaris* (beterraba sacarina). Para explicar esta aglomeração de marcadores foram apontados factores biológicos como a baixa taxa de recombinação nas regiões dos centrómeros e nos telómeros, a distribuição descontínua dos níveis de polimorfismo e a variação do número de cópias de determinadas sequências ao longo do genoma (Tanksley et al. 1992, Sherman e Stack 1995, van Eck et al. 1995, Alonso-Blanco et al. 1998, Keim et al. 1997, Vuylsteke et al. 1999, Young et al. 1999, Castiglioni et al. 1999, Miklas et al. 2001, Peters et al. 2001, Qi et al. 2004).

Qualquer que seja o factor que levou à aglomeração de marcadores AFLP no mapa da var. *bracteatus*, não há evidências de razões biológicas para o mesmo não acontecer no mapa da var. *comosus*. Assim, a justificação que parece mais provável para que isso não se tenha verificado é o facto de na primeira estarem mapeados mais do dobro (285) do número de marcadores AFLP mapeados na segunda (118). Isto sugere que o resultado do teste OSKS que indica uma distribuição normal dos marcadores RAPD no mapa da var. *bracteatus* tem de ser

analisado com muitas reservas pois o número relativamente baixo de marcadores mapeados (64) pode também não ser suficiente para que seja detectada uma eventual tendência de aglomeração deste tipo de marcadores.

4.2- MAPEAMENTO COM A POPULAÇÃO F2

4.2.1- MARCADORES ANALISADOS

O número de primers e marcadores analisados na população F2 é apresentado na tabela 18. As listas de “primers” RAPDs, AFLPs e ISSRs utilizados nesta população encontram-se, respectivamente, nas tabelas 7, 9 e 13. Na população F2 os marcadores ISSR foram analisados em géis de agarose (figura 5).

Tabela 18. Número de “primers” e de marcadores analisados na população F2.

Tipo de marcadores	Número de “primers” ^a	Número de marcadores			
		<i>var. bracteatus</i>		<i>var. comosus</i>	
		Analisados na F1 e F2	Analisados só na F2	Analisados na F1 e F2	Analisados só na F2
RAPDs	34	21	21	14	15
ISSRs	21	0	16	0	22
AFLPs	37	119	95	46	81
Morfológico	-	0	0	1	0
Sub-total	-	140	132	61	118
Total	-	272		179	

^a Combinações de “primers” AFLPs.

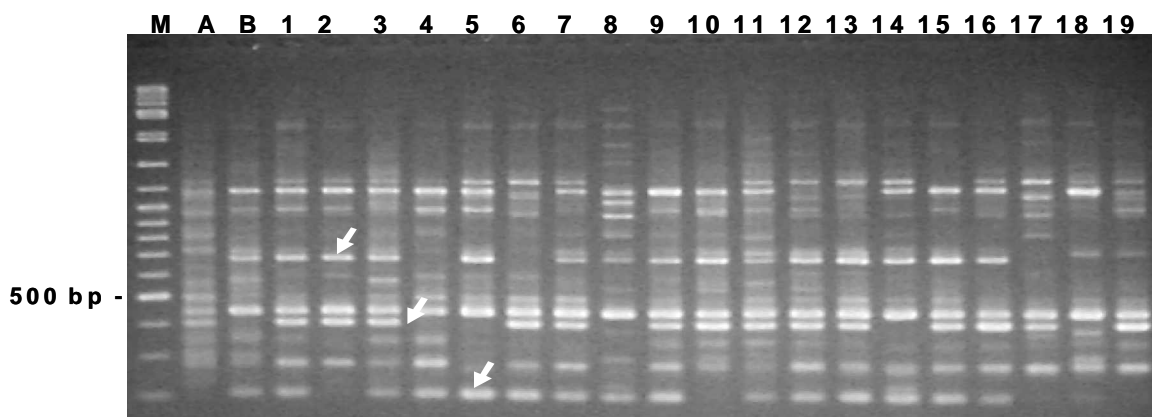


Figura 5. Padrões de amplificação ISSR de 19 plantas da população F2 (1-19) e dos progenitores var. *bracteatus* (A) e var. *comosus* (B), obtidos com o “primer” (GA)₈YT. As setas indicam marcadores polimórficos heterozigóticos analisados em toda a população. M- “escada” de marcadores padrão.

Dos 272 marcadores polimórficos da var. *bracteatus*, 30 (11%) revelaram uma razão de segregação muito desviada ($\chi^2 > \chi^2_{\alpha=0.01}$) da razão de segregação esperada (3:1) e trinta e seis (13.2%) apresentaram uma razão de segregação com um desvio menos acentuado ($\chi^2_{\alpha=0.05} < \chi^2 \leq \chi^2_{\alpha=0.01}$). Dos 178 marcadores-DNA polimórficos da var. *comosus*, 17 (9.6%) evidenciaram uma razão segregação com um desvio acentuado da razão esperada de 3:1 ($\chi^2 > \chi^2_{\alpha=0.01}$), e 12 (6.7%) uma de segregação menos desviada da esperada ($\chi^2_{\alpha=0.05} < \chi^2 \leq \chi^2_{\alpha=0.01}$). Tal como na população F1, também na população F2 a proporção de marcadores com segregação desviada da esperada para $\chi^2 > \chi^2_{\alpha=0.05}$ (24.2% para a var. *bracteatus* e 16.3% para a var. *comosus*) está entre os valores geralmente estimados para outras espécies de plantas (Jenczewski et al. 1997) e é maior que a esperada devido ao acaso (5%).

A hereditariedade monogénica da característica “piping” (Collins e Kerns 1946, Cabral et al. 1997) confirmou-se novamente na população F2 com uma razão de segregação aproximada da razão mendeliana esperada de 3:1 ($\chi^2 = 0.5$).

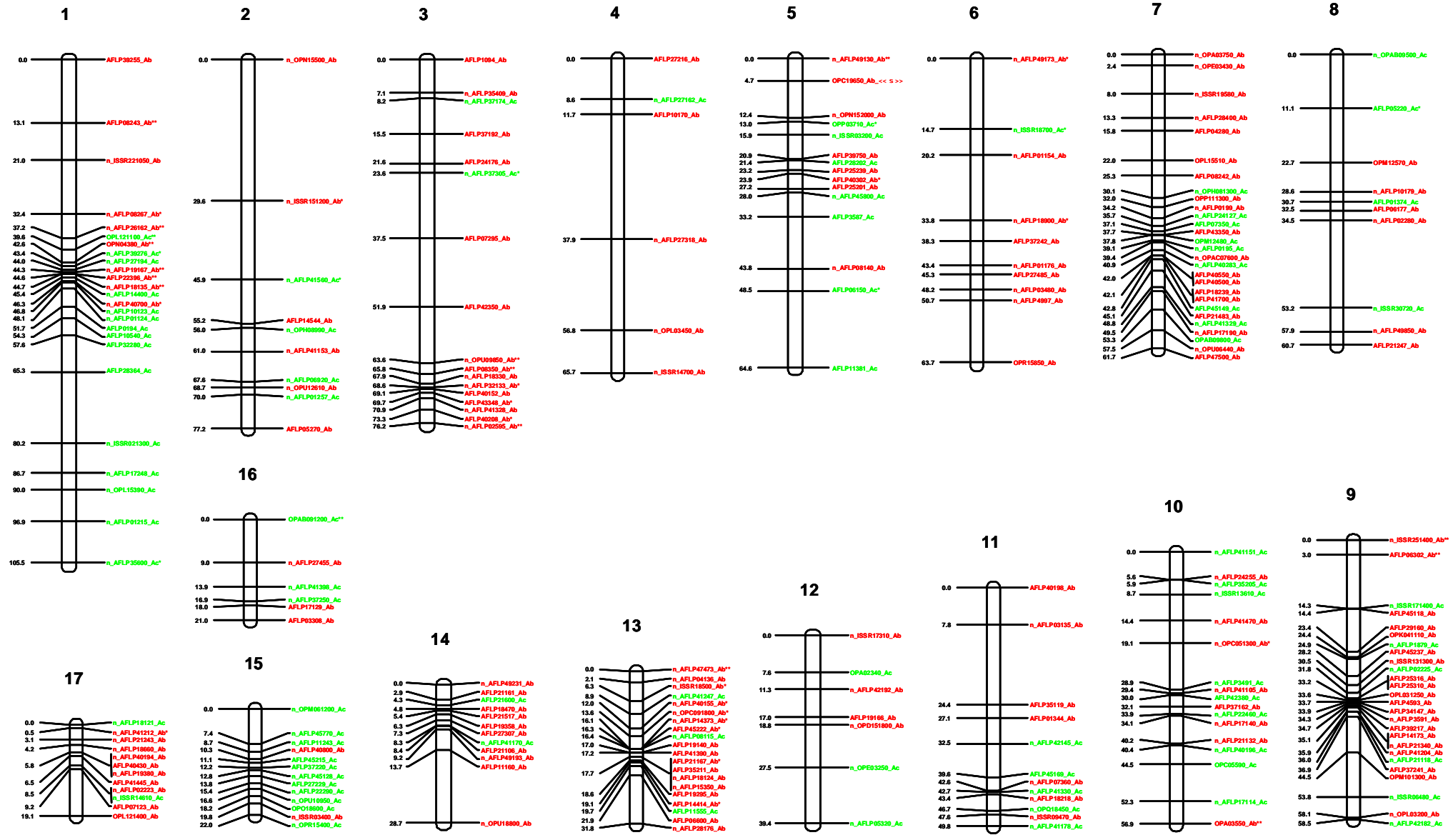
4.2.2- MAPA

O agrupamento dos marcadores em grupos de ligamento fez-se com LODs de 3.0 a 7.0 em intervalos de 0.5. Foram escolhidos 49 grupos de acordo com a sua estabilidade à medida que o LOD aumenta e com a informação dos mapas obtidos com a população F1. Nesses 49 grupos ficaram distribuídos 410 marcadores (65 RAPDs, 311 AFLPs, 33 ISSRs e o marcador morfológico “piping”). Foram obtidos três tipos de grupos: (1) 17 grupos com marcadores dos dois progenitores, (2) 15 grupos só com marcadores de um progenitor (var. *bracteatus*), e (3) 17 grupos só com marcadores do outro progenitor (var. *comosus*) (Figura 6). Todos os 17 grupos com marcadores dos dois progenitores (var. *bracteatus* e var. *comosus*) têm pelo menos 4 marcadores. Dos 15 grupos só com marcadores da var. *bracteatus*, 8 têm pelo menos 4 marcadores, dois têm 3 marcadores e cinco são pares de marcadores. Dos 17 grupos só com marcadores da var. *comosus*, 7 têm pelo menos 4 marcadores, quatro têm 3 marcadores e seis são pares de marcadores.

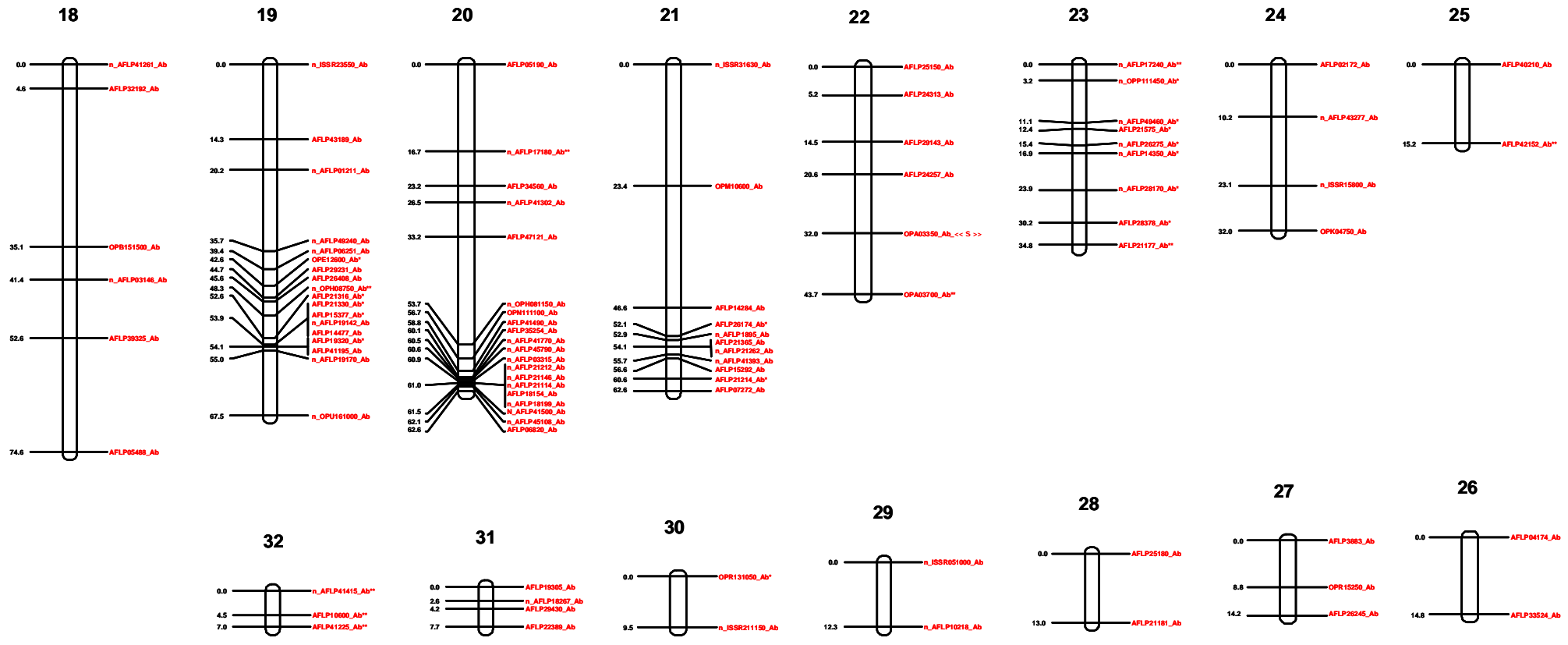
A percentagem de marcadores que não ficaram ligados neste mapa (9.1%) é significativamente mais baixa que a verificada no mapeamento com a população F1 (23.9% para a var. *bracteatus* e 38.6% para a var. *comosus*). Para esta diferença contribuíram vários factores. Um dos factores foi a selecção de “primers”: (1) que geraram maior número de marcadores que fazem parte dos mapas-F1, e (2) que estão presentes na planta F1 autopolinizada para gerar a população F2. Ou seja, eram já conhecidas relações de ligamento genético para uma grande parte (44.6%) dos marcadores analisados na população F2, apesar dos “primers” seleccionados terem permitido analisar novos marcadores distribuídos ao acaso no genoma. Outro factor foi o estabelecimento de alguns grupos de ligamento com LOD limite de agrupamento 3.0 na construção do mapa com a população F2, ao contrário dos mapas construídos com base na população F1, em que se utilizou um LOD limite de

agrupamento 4.0, igual para todos os grupos. O tamanho da população utilizada também contribuiu para a diferença na percentagem de número de marcadores não ligados. Uma população F2 de 142 descendentes permite detectar ligamento entre marcadores em acoplamento (no mesmo cromossoma físico) até uma fracção de recombinação mais elevada do que uma população F1 de 46 descendentes (Liu 1998).

a



b



C

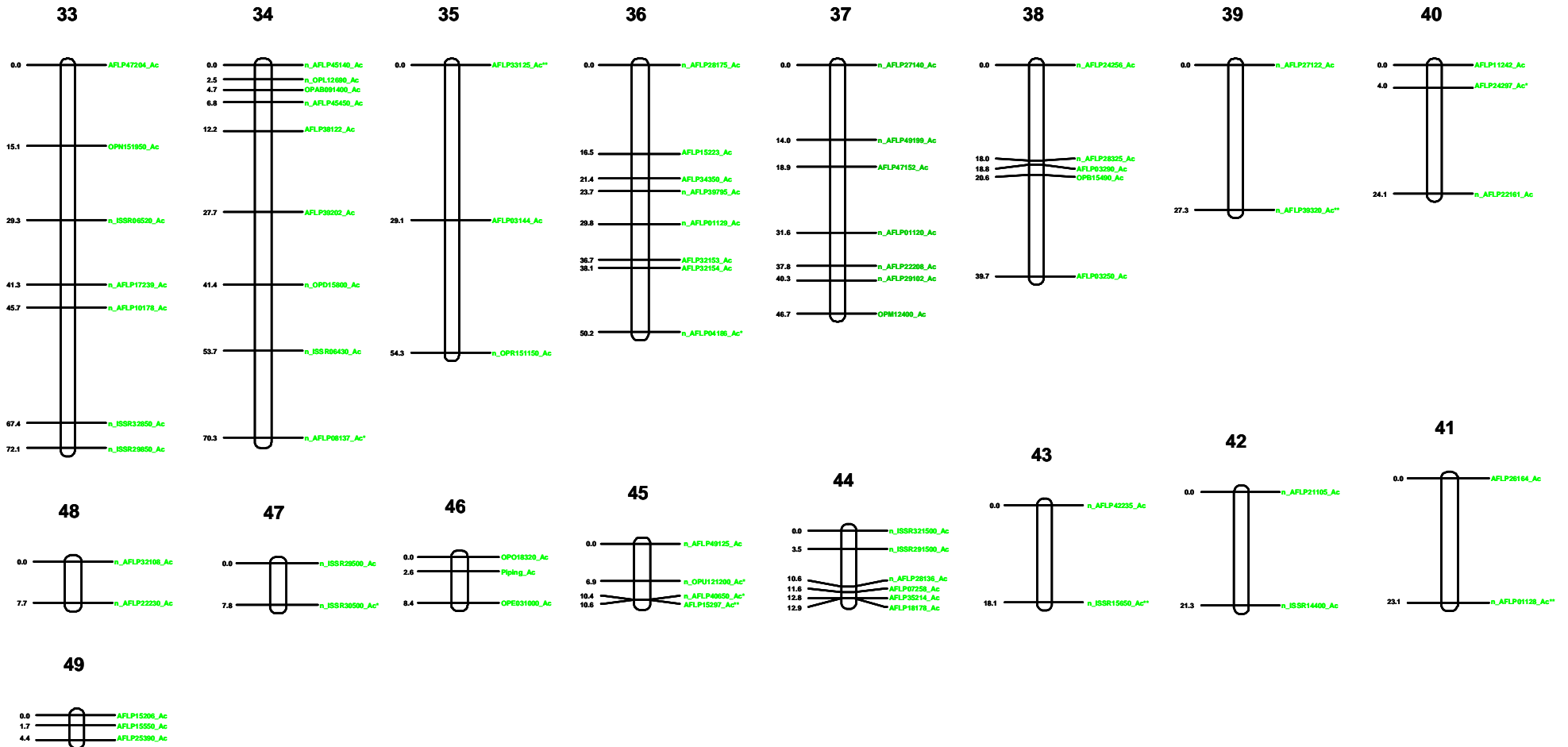


Figura 6. Mapa genético obtido com a população F2. (a) 17 grupos com marcadores dos dois progenitores, (b) 15 grupos só com marcadores de um progenitor (var. *bracteatus*), e (c) 17 grupos só com marcadores do outro progenitor (var. *comosus*). Os números à esquerda indicam a posição em cM. As cores vermelho ou verde e as letras Ab ou Ac no fim do nome do marcador identificam a origem do mesmo, respectivamente, var. *bracteatus* ou var. *comosus*. Os marcadores com segregação desviada da esperada para $\chi^2_{\alpha=0.05} < \chi^2 \leq \chi^2_{\alpha=0.01}$ estão identificados com um asterisco (*), os marcadores com segregação desviada da esperada para $\chi^2 > \chi^2_{\alpha=0.01}$ estão identificados com dois asteriscos (**). Os marcadores analisados na população F2 que não foram analisados na população F1 (novos marcadores) estão identificados com “n_” no início do nome do marcador. O símbolo «>>» indica marcadores sequenciados.

4.2.3- COMPRIMENTO E COBERTURA DO MAPA

O comprimento total do mapa construído com a população F2, calculado adicionando a distância média entre marcadores (5.19 cM) a ambas as extremidades de cada grupo de ligamento é 2381 cM. A estimativa da dimensão do genoma da var. *bracteatus* e do genoma da var. *comosus*, segundo o método 3 de Chakravarti et al. (1991) com os dados de mapeamento da população F2 é de, respectivamente, 2872 cM e 2962 cM. Com estas estimativas e com as estimativas em distância física, respectivamente de 444 e 526 Mpb (Arumuganathan e Earle 1991), estabelecem-se as relações de 155 kpb/cM para a var. *bracteatus* e 177 kpb/cM para a var. *comosus*.

Se a média das estimativas de tamanho dos genomas da var. *bracteatus* e da var. *comosus* calculada com base nos dados da população F2 (2917cM) for utilizada como valor do comprimento genético do genoma do ananaseiro, a cobertura do mapa F2 é de 81.6%. No entanto, é necessário atentar que, como referido anteriormente, o facto dos marcadores analisados na população F2 não terem sido escolhidos de forma totalmente aleatória contribuiu para uma menor percentagem de marcadores não ligados, o que faz diminuir a estimativa do tamanho do genoma calculada segundo o método 3 de Chakravarti et al. (1991). Se for utilizada como denominador a média das estimativas das dimensões dos genomas obtidas com base nos dados do mapeamento com a população F1, 3920cM, o mapa construído com a população F2 cobre 60.7% do genoma do ananaseiro.

4.2.4- DISTRIBUIÇÃO DOS MARCADORES NO MAPA

O teste “One sample Kolmogorov-Smirnov” (OSKS) revelou que a distribuição dos intervalos entre marcadores AFLP no mapa construído com a população F2 se desvia significativamente de uma distribuição normal ($P < 0.001$), indicando agregação de marcadores em determinadas zonas. Para as distâncias entre marcadores RAPD, o teste OSKS revelou uma distribuição normal ($P = 0.194$). Quanto aos ISSRs, mais uma vez tendem a localizar-se nas extremidades dos grupos de ligamento de maior dimensão ou em pequenos grupos. No entanto, também na população F2 o pequeno número deste tipo de marcadores mapeados (33) não permitiu a análise da sua distribuição com o teste OSKS.

4.3- INTEGRAÇÃO DOS MAPAS F1 E F2

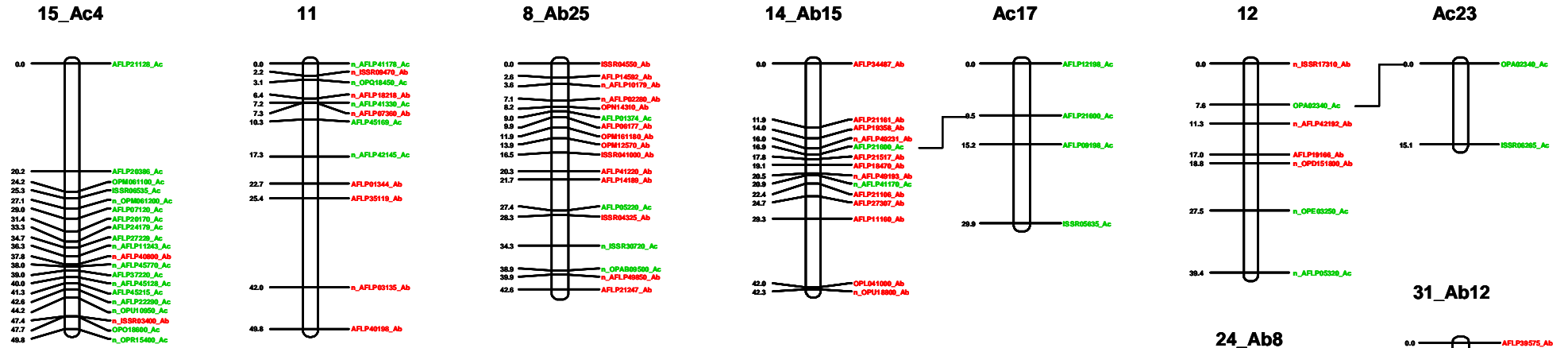
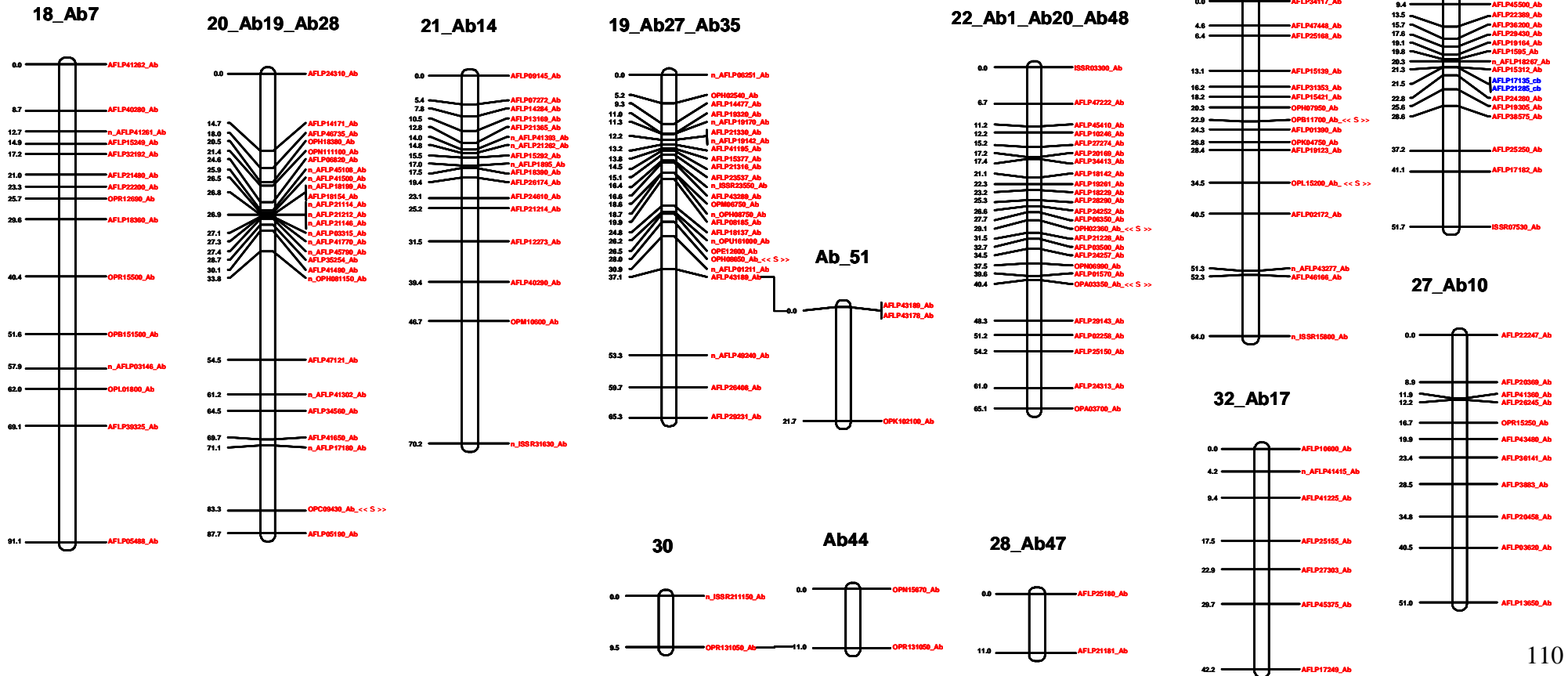
Os marcadores cuja segregação foi analisada na população F1 e na população F2 permitiram a integração de grupos dos mapas-F1 com grupos dos mapas-F2. Foram integrados 40 grupos do mapa F2 com 32 grupos do mapa F1 da var. *bracteatus* e 19 grupos do mapa F1 da var. *comosus*. Alguns grupos do mapa F2 ficaram integrados com grupos de ambos os mapas F1 (da var. *bracteatus* e da var. *comosus*) e outros com grupos de apenas um dos mapas F1 (da var. *bracteatus* ou da var. *comosus*).

A análise conjunta de todos os marcadores mapeados permite estabelecer um mapa integrado com 77 grupos (figura 7). Este número é muito diferente do número haplóide de cromossomas do ananaseiro ($n=25$). No entanto, considerando apenas os grupos com mais de 50 cM quando adicionada a distância média entre marcadores nas duas extremidades podemos

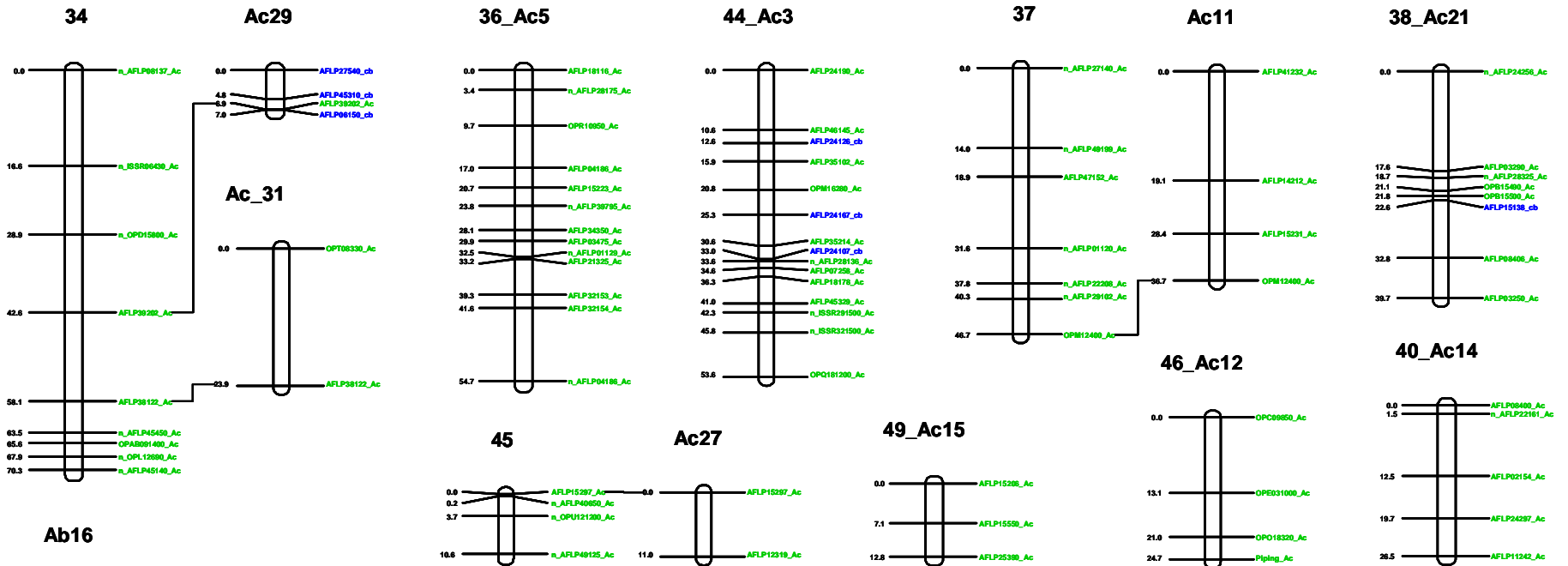
contabilizar 31 grupos maiores: 16 que integram marcadores das duas variedades, 9 só com marcadores da var. *bracteatus* e 6 só com marcadores da var. *comosus*.

O mapa integrado reúne 782 marcadores-DNA (588 AFLPs, 137 RAPDs, 57 ISSRs), o locus *Isoz_PGM* e o locus determinante da característica morfológica “piping”, distribuídos em: (1) 17 grupos de ligamento com marcadores da var. *bracteatus* e da var. *comosus*; (2) 11 grupos que integram marcadores analisados nas duas populações, mas só com marcadores identificados na var. *bracteatus*; (3) 9 grupos de ligamento com marcadores analisados nas duas populações, mas contendo só marcadores da var. *comosus*; e (4) 40 grupos constituídos por marcadores que segregam apenas numa das populações de mapeamento e somente de um dos progenitores (20 dos quais só têm marcadores da var. *bracteatus* e outros 20 só marcadores da var. *comosus*).

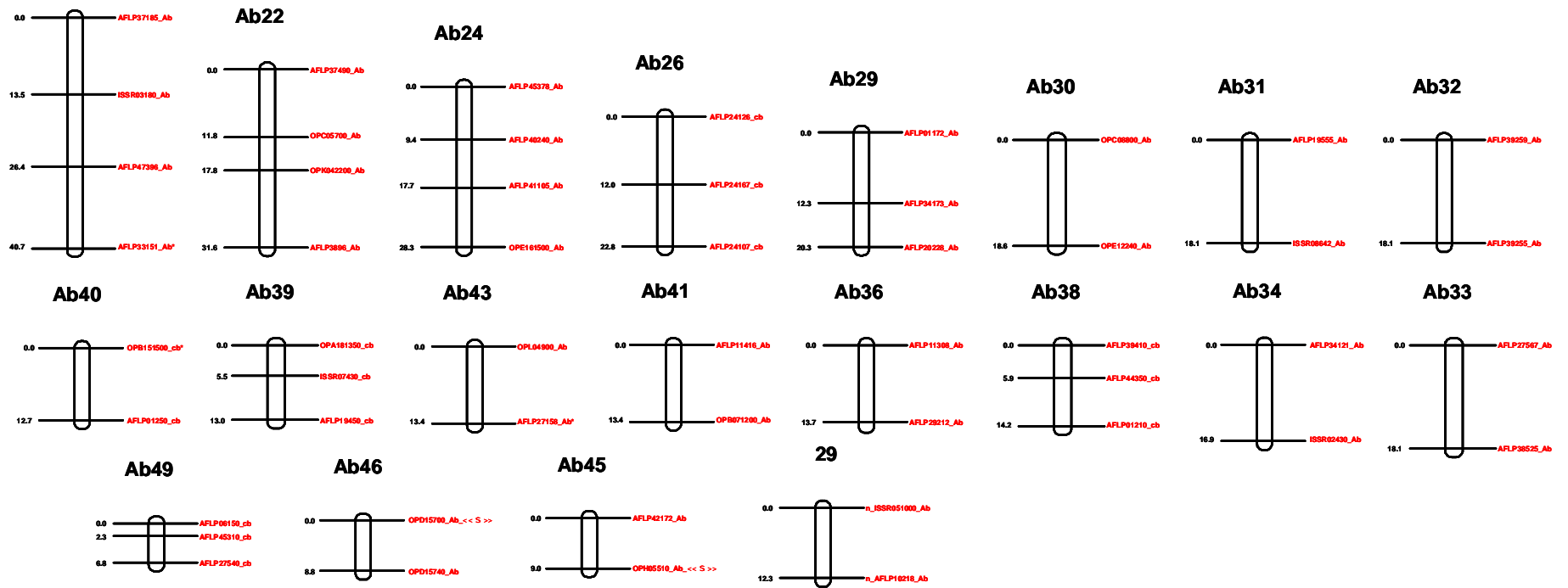
Com a metodologia de cálculo utilizada nos mapas construídos com as populações F1 e F2, este mapa tem uma distância média entre marcadores de 4.16 cM e um comprimento de 3582 cM, o que corresponde a 91.4% da dimensão do genoma do ananaseiro de 3920 cM (média das estimativas da var. *bracteatus* e da var. *comosus* com base no mapeamento com a população F1). No entanto, este valor para o comprimento total do mapa está provavelmente inflacionado pois há grupos correspondentes a cromossomas homólogos que não foi possível integrar e os grupos só com marcadores a segregar numa população que podem corresponder a regiões genómicas representadas também por grupos só com marcadores a segregar na outra população.

a**b**

c



d



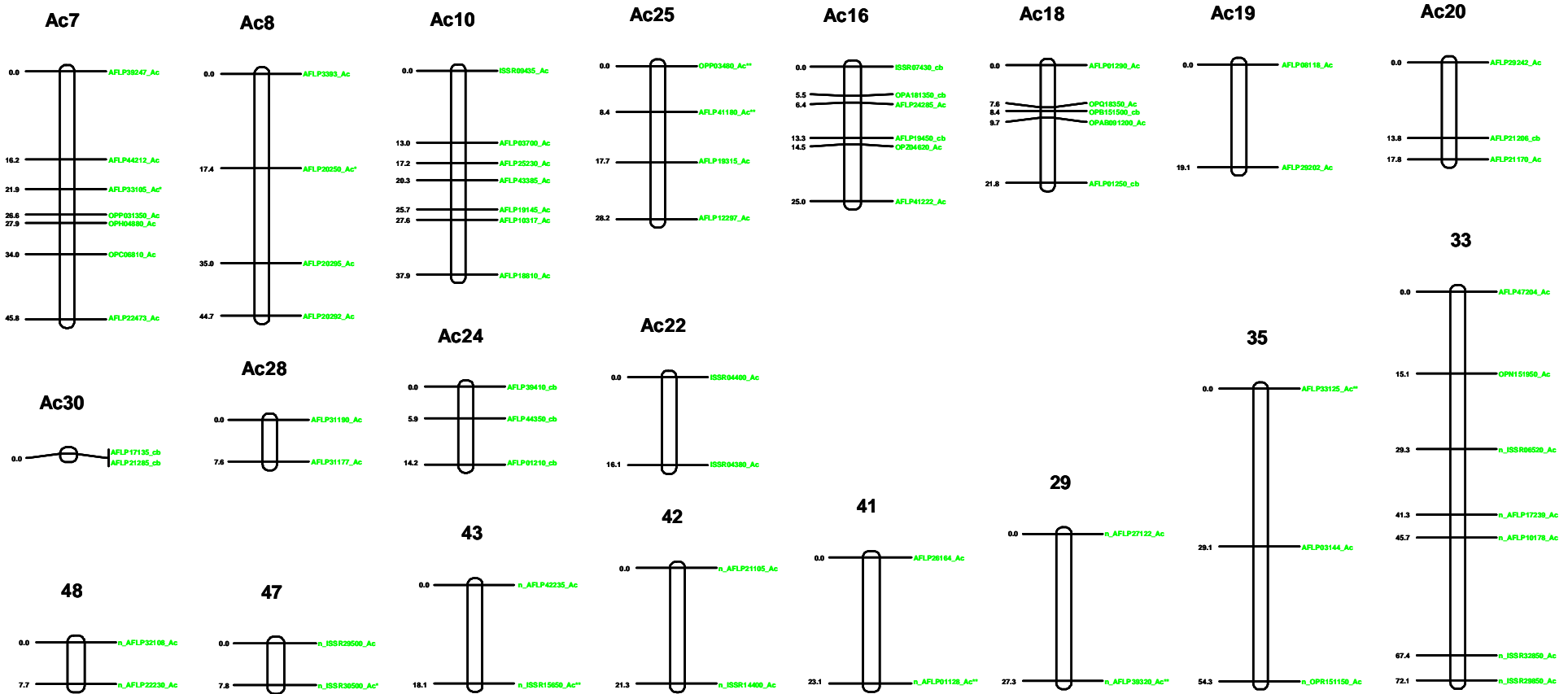
d

Figura 7. Mapa integrado. (a) 17 grupos de ligamento com marcadores da var. *bracteatus* e da var. *comosus*, (b) 11 grupos que integram marcadores analisados nas duas populações, mas só com marcadores identificados na var. *bracteatus*, (c) 9 grupos de ligamento com marcadores analisados nas duas populações, mas só com marcadores da var. *comosus*; e (d) 40 grupos com marcadores que segregam apenas numa das populações de mapeamento e só de um dos progenitores. Os números à esquerda indicam a posição em cM. As cores vermelho ou verde e as letras Ab ou Ac no fim do nome do marcador identificam a origem do mesmo, respectivamente, var. *bracteatus* ou var. *comosus*; a cor azul e as letras cb identificam os marcadores heterozigóticos monomórficos entre os progenitores. Os marcadores com segregação desviada da esperada para $\chi^2_{\alpha=0.05} < \chi^2 \leq \chi^2_{\alpha=0.01}$ estão identificados com um asterisco (*), os marcadores com segregação desviada da esperada para $\chi^2 > \chi^2_{\alpha=0.01}$ estão identificados com dois asteriscos (**). Os marcadores analisados na população F2 que não foram analisados na população F1 (novos marcadores) estão identificados com “n_” no início do nome do marcador. O símbolo «s» indica marcadores que foram sequenciados.

4.4- MARCADORES SEQUENCIADOS

Com o objectivo de no futuro desenvolver marcadores específicos de fácil análise e elevada reprodutibilidade SCARs (Sequence Characterized Amplified Regions) que funcionem como marcos de referência no genoma do ananaseiro, tornando a utilização dos mapas mais eficiente, iniciou-se a clonagem e sequenciação de marcadores mapeados. Na tabela 19 são apresentadas as sequências dos 10 marcadores sequenciados, os grupos de ligamento a que pertencem nos mapas construídos com base nas populações F1 e F2, e o respectivo número de entrada na base de dados genómicos “GenBank” do NCBI.

As sequências destes 10 marcadores foram comparadas com as sequências das bases de dados do NCBI. Na procura por Blastn nas bases de dados de sequências nucleotídicas “nr” e “plant genomes mapped DNA sequences” só foram identificadas sequências com elevada similaridade (“expect value” $\leq 10^{-8}$) para o marcador OPD15-700. Para os restantes marcadores apenas foram encontradas similaridades em pequenos segmentos e com “expect value” $> 10^{-3}$ em ambas as bases de dados. A sequência com elevada similaridade com o marcador OPD15-700 corresponde a DNA genómico de arroz - *Oryza sativa* (tabela 20).

O facto de só um dos 10 marcadores sequenciados ter homologia significativa com sequências das bases de dados internacionais revela um certo grau de originalidade do genoma do ananaseiro em comparação com os genomas parcial ou totalmente sequenciados. No entanto, é necessário um número maior de sequências para se poder fazer uma análise comparativa mais profunda entre o genoma do ananaseiro e outros, pois as sequências destes 10 marcadores perfazem um total de 4868 pb, representando apenas uma pequena fracção (1.0964^{-5}) do genoma a que pertencem (*var. bracteatus*), o qual segundo (Arumuganathan e Earle 1991) tem um comprimento estimado em 444 Mpb.

Na procura por Blastx na base de dados “nr” de sequências proteicas com o marcador OPD15-700 não foram encontrados alinhamentos com “expect value” $\leq 10^{-8}$. No entanto, das 100 primeiras sequências encontradas (“expect values” entre $1e^{-13}$ e $8e^{-6}$), 10 correspondem a produtos proteicos desconhecidos e as restantes 90 correspondem a partes de sequências de Isocitrato-Desidrogenase (IDH), semelhantes a IDH ou hipotéticas IDH, de várias espécies: *p.e. Oryza sativa, Eucalyptus globulus, Citrus limon, Arabidopsis thaliana, Daucus carota, Lupinus albus, Prunus persica, Apium graveolens, Nicotiana tabacum, Pisum sativum*. Apesar deste facto sugerir a hipótese do marcador OPD15-700_Ab fazer parte da sequência de um gene de uma IDH, isso só pode ser confirmado com uma caracterização mais profunda da mesma, assim como das sequências adjacentes a este marcador, de forma a identificar um eventual gene candidato e finalmente fazer um estudo da expressão do mesmo. Na tabela 21 encontra-se referenciada a entrada com similaridade mais elevada encontrada por Blastx para o marcador OPD15-700.

Tabela 19. Sequências de marcadores.

Nome do marcador	Grupo de ligamento ^a Mapa F1	Grupo de ligamento Mapas F2	Número de entrada (Nº ent) e definição (Def) no GenBank. Sequência ^b
OPA03350	Ab1	22	Nº ent: DQ386596 / Def: Ananas bracteatus RAPD marker OPA03_332 sequence. 1 <u>AGTCAGCCAC</u> CCGTACAACA GACAAGATTT GTTCAAATGG CCGTGGACTA CCAACCACCT GTTTACAGTC 70 71 AAATCCCAAT ACCAATTTCT CACAGACGGC GGTTGTAGAG ATTGGTTGAA CCAGCACATC TGGAGCCTGA 140 141 AAATACCATT TAAAATCAAA ATTTTCGTTT GTTTGCTGCT TCGAAAGAGA CTACTTACCA CTGACAGACT 210 211 GCTCTCCAGA GGCATGCC TG TAGACCCGTA CTGCGCACTC TGTGGGGTAC TGCCGGAAAG CTGTGATCAC 280 281 CTTTTCTCTA AGTGCCTCTT TGTTCGCTTC CTCCTACTCT CTGTGGCTGA CT 332
OPH02360	Ab1	-	Nº ent: DQ386597 / Def: Ananas bracteatus RAPD marker OPH02_349 sequence 1 <u>TCCGACGTGA</u> GGGCGTGGAC AACCCGGAGG GTAGCCGGAA GATTTCATAGG AGGTTAAGAG AGGTTGAGTC 70 71 CGTGGAGTCG GTAATCACCT TATAATCTTT TCTCTTAGTG AATTGTTTTTT TCGCGTGTG GTCCCGTGGG 140 141 TTTTTTCTCC AAGTTGGAGT TTTCCCACGT AAATCTCGGT GTGCTTTTTTA TTTTCCGCAT TTATTTTTAT 210 211 TGCATCGAGA TTTGTGGTTT TTGGCCGTTA CACCTATTCA CCCCCCTCTA GGTTAGTAGAT ACAATCTAGA 280 281 TAGATCCTTG GGCTCCTCAA AACTTTCACT GATTGTGCC GACCTCCCCA GTGTGTGTGT CACGTCCGA 349
OPB11700	Ab8	-	Nº ent: DQ386598 / Def: Ananas bracteatus RAPD marker OPB11_733 sequence 1 <u>GTAGACCCGT</u> GATGTATATT TACTGTTGGG ACCAGATACT ATAGCTCTGA TACTTTCTCC TTTTTGCTGA 70 71 GATTTGGGAT TGTATCCTCT TTTGTCCCTT TATGGTTAGT ATATCGACTT GTACTTTGCT CTGATGTCAG 140 141 GATAGGATTT TACTATACTT TTCTCCTATC GACTTGTGTT CTTGTAGACG CTTTATATAC TGCAGGTGTA 210 211 TGGCGGGTCT GTGCACGTGC CCGATATGCT TCCGCTGGTA CCCGGGCGTG ACACATCATG AGTCTATTTAT 280 281 TTCTTAGGGT TAATGTCTCT ATGTTGGCCT TATGTGGTTT CTAGTGTAGA GCTCCAAAAA TGTAGGTTGA 350 351 CTTGGTGGTC GCACTGTTCA TTACTCTATA TTAATGGAAT CAATAATGGT GTAATTGAGA AAGTGGTTAT 420 421 TGTTAGATGT ATGCCCTAGA AGCCAACCAG GCCGACACAT GTATTTATTC TAGGACATAA ATTTGTATTT 490 491 GACTTTAAGA TATTATGAAT AATTTGGGTT CTTATTTTCA TTCATGTTGT GTATGTGTCC ATGAATCGTC 560 561 CAAGAAATTA ATAAAAATGGT TACATATATT CTCAATAGTT GAGAATTTGA GACATGTATC ATTGGTGATT 630 631 AATTCCATAA TGCTCCCGAT CAATGGATCA TCAC TGGGGA CGGTAATTGA TCCGATGAGA TTGGTGCATA 700 701 GGTCGCTTCC CTTTTTTGGG TAGACGGGTC TAC 733
OPL15200	Ab8	-	Nº ent: DQ386599 / Def: Ananas bracteatus RAPD marker OPL15_185 sequence 1 <u>AAGAGAGGGG</u> CAGGCACAAT TGAGAGGCCA TATCAGCCTG TGTGTAAGG AAGAAAAATG AGATTCGAAT 70 71 AAATAATAAT AATAATAATA AAAGTGAAGA GAGTCTACAA AATGTTAAAG AAATGTCAGT GAAGATAAAA 140 141 CTACAAAACCT CTACAGTGTT AGAGATATAT ATGTACCCCT CTCTT 185

OPH13470	Ab9	-	N° ent: DQ386600 / Def: Ananas bracteatus RAPD marker OPH13_431 sequence					
1	<u>GACGCCACAC</u>	CCACACCCAT	ATCCATTAGG	GTGGATTAAT	AGCAACATCG	AGATGAAGAT	CGATCGCCAG	70
71	TGTAAGGTGA	AGTTTGTAGT	GACCAGCGTC	TACATCGACA	AAATGCTATA	AGAGTTTGTG	CCGTTGAATA	140
141	TTTGCAACTT	TATCTTCGGC	AGTCCCTACC	AATAGGACCG	TGACGC TACT	TTTTTTTCGT	GTCCTCAGCA	210
211	ATACTAATTT	ATTGAGGATG	GACAATGTGA	CACCC TAATA	GTCCACATC	GGATGGGAAT	GGGGTTATTA	280
281	TTGGGTTTAT	AAGAGACTTA	GACACTAGTA	ATAATAACTG	GGCTTAAGCA	TTTTGGGCTG	TTGACTGGGC	350
351	CCAACAAGTT	ATTATTGCTA	GTGGGCTGGG	TCGTTACATT	TGGTATTAGA	GCCGGCTCAC	TAGCTGGACA	420
421	TGTGTGGCGT	C	431					
OPC19650	Ab11	5	N° ent: DQ386601 / Def: Ananas bracteatus RAPD marker OPC19_621 sequence					
1	<u>GTTGCCAGCC</u>	CAATGTTCAA	TACAATCAAA	GAGAAAATTG	TTCCAAGCCG	TAAGGAGACC	CCCGCGAGTC	70
71	CCTGTAGCAT	CCAAGGTACG	AAATTCCTGG	AGATAAAAAC	CACAAAACGA	TTGAAATTTA	TCCAGAGAAA	140
141	TAGATGAGAG	CTTGGTTTCT	TGAAGACACA	CCACCGAGAC	TTTACTAGAG	CGGATAAAAG	ACCGTACAAT	210
211	GCAACATTTG	GCCCTGTTGT	TCAGGCC TCT	CACATTCAC	GACAAAAAGC	AAAAGATTTT	ATTTCTAGCC	280
281	ATAAAAGAAC	AAAATGCACT	CAACGCAGGA	AAGAAGCCGA	GGGTTAGACC	CCTTCC TGT	CAGGTCCATC	350
351	ACCCCTCGCG	CCCGCTAGCA	GAGCGGCAAG	AACCTCCAAA	GAAGGCCCA	TTTTCGTACT	GGTGATTCCA	420
421	CAGGATAATA	AGATCTTCTG	GATCTCCTGA	CTGGTAAGTG	GGTGAGCAGG	ATCTCTT GAG	GAGGATGACA	490
491	CCTTATCAAT	CGAAGGAAGG	GCCTCAGTAG	GAGGATCTGG	GCTAACATCC	TCTACCGCTG	CTTGCGCATT	560
561	TGGGGCCTTG	TTCAGAAAGT	CATCGAGAGT	CGGGGTGATT	GAGGAGTGGT	CGGCTGGCAA	C	621
OPH08650	Ab27	-	N° ent: DQ386605 / Def: Ananas bracteatus microsatellite SSR_H8 sequence					
1	<u>GAAACACCCC</u>	TTCC TAACCT	TTCCGCAAGA	TGGCATCTTT	ATCAC TTTAC	CAGCACTATT	GCGTGATTAC	70
71	CGTTATTATT	ATTATTATTA	TTTATACATT	TTAAACCTTA	ATAAACATCT	TTTCC TGGAT	CTGCTCAGCA	140
141	ATTTGGTTTT	CTATTTTTGC	CCAGTTTTGT	CACACCCAAG	CCCAGTTTGG	CAACTCATT	CGAAGCCGCG	210
211	CACAATCTAG	GCCGGATCTC	GACATTTTTGG	ATGAAATTTT	AGCACCTATT	GTTATTGGGG	TTTTGGATCG	280
281	TCCGATCGAG	ACCCGAAC TA	GATACAATTG	GACCTGGTCA	CCTTAATTAA	CAAGCAACTG	CGTTCGGTGA	350
351	GTGAGAAAAA	AAATGTTAAA	TTTTCTGTCTG	CTCAATGCTC	TTCC TCTATC	TCTTTCTCCT	GAGCTCTAAC	420
421	CATGAAC TTT	CATAATCAAT	TCCTGAGTCC	CAAAAATGGT	GCCACAAC TC	AACTCAAAGA	AAAGGTGAAA	490
491	CTGGTTTTTT	TTTCAAAAAA	AA CACACACA	CACACACA	CACGCACACA	CACACACA	CACACACA	560
561	CAACCTACTA	GTAGACCAGC	TAAGAACTGC	GACAAAAGCT	AAAGGGAAGG	GCAATTAAGG	AAATGAAAAG	630
631	ATTGCTCTTT	GGGGGTGTTT	C	651				
OPC09430	Ab28	-	N° ent: DQ386602 / Def: Ananas bracteatus RAPD marker OPC09_387 sequence					
1	<u>CTCACCGTCC</u>	ATCAAAATAT	AAAAT TGCAA	AAACATATGG	TGGGTAAAGG	TGATTGATTG	CGTTTAGTTC	70
71	TATAGGACTT	AATTTATGGT	GCAAAC TACT	TTGGGACACG	TGGTTTGACT	TGTTATTTGA	GAGGGCACAG	140
141	GATTCGAGGG	AAAAAAGAAG	GTTGACTTCG	CTGGCAAGAA	AAT TATCGTT	GACTTG TAGT	TATAACCGAT	210
211	TAAGGATTAA	GGTTGTGTTT	TTGTTTCTTG	TACATTAATC	TCAATTTTTT	TTTAATTAAT	TGAAAAAAGA	280
281	TATTCCTTGT	TTCGTGGCTG	GTA AATTTTT	TTTTTTTTCAG	ATTTCAAAGC	AAAACGTAGC	CTTACGTTGC	350
351	TGCAGCGAGC	TTTTCACTCT	ACCTGGTGG A	CGGTGAG	387			

OPH05510	Ab45	-	N° ent: DQ386603 / Def: Ananas bracteatus RAPD marker OPH05_483 sequence								
			1	<u>AGTCGTC</u> CCCC	CACCC <u>TAAGC</u>	AAATC <u>ACCCA</u>	TAATTT <u>TACGC</u>	A <u>TAGCATA</u> AAA	GCAAGCAGCA	TGAAAA <u>CAGT</u>	70
			71	<u>ATCCTCT</u> AGA	CAAGCAA <u>ATA</u>	TGCAACA <u>AGA</u>	AACTG <u>TCACT</u>	ACTAGA <u>AATTG</u>	AGCTC <u>CTATT</u>	CAGTAAG <u>ACA</u>	140
			141	<u>CATCCTG</u> TGA	C <u>TT</u> CAGGCAC	TAACAA <u>AGCA</u>	ACCAGG <u>GCAA</u>	GAAAC <u>TACTG</u>	AGTTA <u>CTAAG</u>	CACACA <u>AACT</u>	210
			211	<u>CAGTACCA</u> AT	G <u>TTTTCT</u> TTTT	T <u>TTTTT</u> TCCT	TTTT <u>CTTG</u> TT	T <u>CTTCT</u> TTTTG	TTTT <u>CCCT</u> TTT	TTG <u>TCT</u> TTTTT	280
			281	<u>TACAACT</u> G <u>T</u>	C <u>ATTACC</u> AGG	G <u>TGGTC</u> ATAT	A <u>ATCTG</u> G <u>TAG</u>	G <u>CTATAT</u> GCC	TG <u>CTTCC</u> ACG	A <u>CTCAT</u> ATGA	350
			351	<u>TGAAAT</u> TTGA	GATGATG <u>ACA</u>	GAGTATAT <u>TG</u>	GGAAAA <u>AGTA</u>	GCTGGA <u>ACAA</u>	GCTAGAT <u>C</u> TT	TCCAA <u>AGCTA</u>	420
			421	<u>TATGAT</u> AGTT	TGTACATG <u>TA</u>	TTCCAGAT <u>AA</u>	TATTT <u>CTT</u> G <u>T</u>	GTACATG <u>C</u> TT	CATGGG <u>GACG</u>	<u>ACT</u>	483
OPD15700	Ab46	-	N° ent: DQ386604 / Def: Ananas bracteatus RAPD marker OPD15_696 sequence								
			1	<u>CATCCG</u> TGCT	GGTAACGGCT	ACTT <u>CTTC</u> CT	TCTTTT <u>TGCT</u> T	TCCAGGATTTT	CTAATATTT <u>C</u>	TGTTGAGAAC	70
			71	<u>CACATA</u> AAAAT	TCATTTTTTTTT	TTTTCTTTG <u>CA</u>	GTTATCTT <u>CA</u>	GATGGAAAAA	CATTAGAAGC	TGAAGCAGCT	140
			141	<u>CATGGT</u> ACTG	TTACTAG <u>ACA</u>	TTTCAGG <u>CTA</u>	CACCAA <u>AGG</u>	GAACAAGAGA	CCAGCACC <u>AA</u>	TAGTATCGCT	210
			211	<u>TCCATC</u> TTTG	CATGGACTCG	CGGT <u>CTTG</u> CA	CACAGG <u>TACG</u>	TTTTAT <u>TGTT</u>	CATTTGGAG <u>A</u>	TATCTTTACT	280
			281	<u>TCTTGG</u> ATTA	TGTGATTTTA	C <u>TATTTG</u> TGA	GTTGGATTTG	GTACATCATA	TAACAATGTT	TTTGCCTAGT	350
			351	<u>TCTCACT</u> TAC	ACTGCATGGT	C <u>TTTTTTTT</u> T	C <u>TTTTTTTT</u> CT	TTTTTTTTTTTT	TAACAGAGCA	AAGCTTGATA	420
			421	<u>AAAATG</u> ATAG	GCTGCTAGAA	TTTGTGCAG <u>A</u>	ACCTTGAG <u>TC</u>	TGCATGCATT	GAAACAGT <u>CG</u>	AGTCCGGTAA	490
			491	<u>AATGAC</u> AAAG	GATCTTGCGC	TTCTTAT <u>CCA</u>	CGGTCCCGAG	TCAGTATGCT	CACTCTGTTT	ATTCCAATTA	560
			561	<u>GAAATA</u> CTTC	TATAATTGAT	TATTTT <u>C</u> AAA	TCAAGAAGAA	TACAGAGCAG	AAATCTATTT	TCTCCCTCCC	630
			631	<u>TTTATG</u> ACAA	ACTGA <u>ACTAT</u>	TCTGTTTTG <u>T</u>	TTAGAGTATC	AAGGGAACAT	TACTTGAGCA	CGGATG	696

^a Ab = Grupos do mapa da var. *bracteatus*.

^b Os nucleótidos sublinhados correspondem à sequência do “primer” utilizado para amplificar o marcador. Os nucleótidos em itálico no marcador OPH08650 identificam um microsatélite (imperfeito) com 24 repetições CA (com uma repetição CG intercalar).

Tabela 20. Entradas das sequências com similaridades mais elevadas com o marcador OPD15-700 identificadas por Blastn no NCBI.

Base de dados “nr”	
Marcador	Entrada
OPD15-700 (696 nucleótidos):	gi 32489509 emb AL662969.3 OSJN00169 Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNBa0043L24:
- Nucleótidos 97 a 216	- Nucleótidos 94066 a 94198 [Score = 115 bits (58), Expect = $2e^{-22}$, Identities = 116/134 (86%), Gaps = 1/134 (0%)]
- Nucleótidos 404 a 520	- Nucleótidos 94364 a 94480 [: Score = 89.7 bits (45), Expect = $9e^{-15}$, Identities = 99/117 (84%)]
Base de dados “plant genomes mapped DNA sequences”	
OPD15-700 (696 nucleótidos):	ref NT_079864.2 Oryza sativa (japonica cultivar-group):
- Nucleótidos 97 a 216	- Nucleótidos 315028 a 315146 [Score = 115 bits (58), Expect = $6e^{-24}$, Identities = 116/134 (86%), Gaps = 1/134 (0%)]
- Nucleótidos 404 a 520	- Nucleótidos 315326 a 315442 [Score = 89.7 bits (45), Expect = $3e^{-16}$, Identities = 95/117 (81%)]

Tabela 21. Entrada da sequência com maior similaridade ao marcador OPD15-700 identificada por Blastx na base de dados “nr” do NCBI.

Marcador	Entrada
OPD15-700 (696 nucleótidos):	gi 34904654 ref NP_913674.1 putative NADP-isocitrate dehydrogenase [Oryza sativa (japonica cultivar-group)]:
- Nucleótidos 111 a 527	- Aminoácidos 293 a 377 [Score = 79.0 bits (193), Expect = $1e^{-13}$, Identities = 54/139 (38%), Positives = 62/139 (44%), Frame = +3]

5- DISCUSSÃO

O elevado número de cromossomas do ananaseiro foi o principal desafio na tarefa da construção de um mapa genético para o seu genoma. Tal como em outros trabalhos de mapeamento genético em espécies com um elevado número de cromossomas (*p.e.* Reinisch et al. 1994, Keim et al. 1997, Ferreira et al. 2000, Matthews et al. 2001, Wu et al. 2000, Yin et al. 2002, Zhang et al. 2002, Lacape et al. 2003, Zhang et al. 2004, Waghmare et al. 2005), o número de cromossomas envolvidos no cruzamento que gerou as populações de mapeamento neste trabalho ($2n = 50$ cromossomas em cada progenitor) levou à formação de um elevado número de pequenos grupos de ligamento. Este fenómeno fez com que o número total de grupos se afastasse do número haplóide de cromossomas. Nos mapas construídos com a população F1 formaram-se 35 grupos só com 2 ou 3 marcadores (24 para a var. *bracteatus* e 11 para a var. *comosus*) e no mapa construído com a população F2 formaram-se 17 destes grupos. Apesar disso, e de não haver uma distinção clara entre um conjunto de grupos de tamanhos maiores e outro de tamanhos mais pequenos, no mapa construído com a população F1 para a var. *bracteatus* há 21 grupos que abarcam mais de 50 cM quando adicionada a distância média entre marcadores (5.27cM) a ambas as extremidades que provavelmente correspondem a 21 dos 25 pares de cromossomas desta variedade botânica de ananaseiro. Este mapa tem uma cobertura do comprimento do respectivo genoma (estimado em 3693 cM pelo método 3 de Chakravarti et al. 1991) que se pode considerar média (59.1%) e que está entre valores de cobertura de mapas publicados para outras espécies (tabela 22). No entanto, as estimativas de comprimento genético dos genomas devem ser analisadas com alguma reserva pois há grandes variações quando os cruzamentos utilizados para o seu cálculo são diferentes. Para além disso, podem obter-se valores significativamente diferentes para estimativas de

comprimento do genoma quando se utilizam diferentes subconjuntos de dados de marcadores a segregar na mesma população (Liebhard et al. 2003).

Tabela 22. Exemplos de cobertura de mapas genéticos construídos para diferentes espécies.

Espécies	Marcadores	Cobertura do genoma	Referências
<i>Brassica napus</i>	RFLPs	61%	Landry et al. (1991)
	RFLPs e RAPDs	63%	Uzunova et al. (1995)
	RFLPs e RAPDs	71%	Foisset et al. (1996)
<i>Actinidia callosa</i>	SSRs e AFLPs	29%	Testolin et al. (2001)
<i>Actinidia chinensis</i>	SSRs e AFLPs	46%	
<i>Chenopodium quinoa</i>	SSRs, AFLPs e RAPDs	60%	Maugham et al. (2004)
<i>Citrus</i> spp.	Isoenzimas e RFLPs	23%	Jarrel et al. (1992)
	Isoenzimas e RFLPs	37%	Durham et al. (1992)
	AFLPs, RAPDs e RFLPs	40%	Simone et al. (1998)
	AFLPs, RAPDs e RFLPs	67%	
	AFLPs, RAPDs e SSRs	62%	
<i>Malus</i> spp.	AFLPs, RAPDs e SSRs	62%	Liebhard et al. (2003)
	AFLPs, RAPDs e SSRs	69%	

Expressões como “mapa saturado” ou “mapa de densidade elevada” têm sido aplicadas em mapas com distâncias médias entre marcadores da ordem dos 3 cM ou inferiores, mesmo quando nesses mapas há zonas mais densas e zonas menos densas em marcadores. Num trabalho de mapeamento genético na árvore-da-borracha (*Hevea* spp.) Lespinasse et al. (2000) descreveram como saturado um mapa com uma distância média entre marcadores adjacentes de 3 cM, com aglomeração de marcadores (principalmente AFLPs) em algumas zonas e com segmentos longos sem qualquer marcador noutras, sendo o maior de 26.4 cM. Noutro exemplo, Hori et al. (2003) utilizaram o termo “densidade elevada” na descrição de um mapa da cevada (*Hordeum vulgare*) com uma densidade média de marcadores de 1.4 cM e também com aglomeração de marcadores em algumas zonas e dispersão noutras. De facto há muitos exemplos em que foram utilizados os termos “mapa saturado” ou “mapa de elevada densidade” para valores de distância média entre marcadores

inferiores a 3 cM. Por exemplo, 2.6 cM em *Brassica oleracea* (Sebastian et al. 2000), 2.1 cM em *Theobroma cacao* - cacauero (Risterucci et al. 2000), 2 cM em *Prunus* (Joobeur et al. 1998) e menos de 1 cM em *Zea mays* - milho (Vuylsteke et al. 1999) e *Lycopersicon esculentum* - tomateiro (Haanstra et al. 1999). A descrição dos mapas genéticos do meloeiro (*Cucumis melo*), com distâncias médias entre marcadores entre 3 e 3.8 cM, como tendo uma densidade moderadamente elevada (Oliver et al. 2001) também corrobora a ideia do limite de 3 cM para a classificação de um mapa como tendo uma densidade elevada. Por outro lado, quando as distâncias se aproximam de 10 cM ou ultrapassam este valor, os mapas são considerados de baixa densidade (King et al. 1998, Baudracco-Arnas e Pitrat 1996, Wang et al. 1997, Oliver et al. 2001). Pode então considerar-se que o mapa construído com a população F1 para a var. *bracteatus*, com uma distância média entre marcadores adjacentes de 5.27 cM, tem uma densidade média/elevada. No entanto, o teste “One Sample Kolmogorov-Smirnov” (OSKS) confirmou que houve aglomeração do principal tipo de marcadores mapeados (AFLPs) em algumas zonas e dispersão noutras ($P < 0.001$). O valor da maior distância entre marcadores adjacentes (32.0 cM) é revelador da baixa densidade em algumas zonas do mapa.

Segundo Liu (1998) a aglomeração de marcadores em certas zonas dos mapas faz com que para se atingir a saturação dos últimos seja necessário analisar um número de marcadores maior do que seria necessário se estes se distribuíssem de forma aleatória. Utilizando a equação de Lange e Boehnke (1982), para um valor máximo de distância entre marcadores adjacentes de 32.0 cM, se os marcadores tivessem uma distribuição normal, seria necessário mapear apenas 171 marcadores para haver uma probabilidade de 95% de ficar coberta a totalidade do genoma da var. *bracteatus* (estimado em 3693 cM). O facto de se ter obtido para a var. *bracteatus* um mapa com 363 marcadores, no qual o maior valor de distância entre

marcadores adjacentes é 32.0 cM e que cobre apenas 59.1% do genoma, é revelador de como a aglomeração de marcadores impediu uma construção mais eficiente do mapa.

No mapa construído com a população F1 para a var. *comosus* há apenas 11 grupos que depois de adicionada a distância média entre marcadores (7.13 cM) a ambas as extremidades ficam com mais de 50cM e que provavelmente representam 11 dos 25 pares de cromossomas desta variedade de ananaseiro. Para este mapa (F1 - var. *comosus*) os valores de 33.2% de representação do genoma (estimado em 4146 cM), de 7.13 cM de distância média entre marcadores adjacentes e de 23.9 cM de distância máxima entre marcadores adjacentes, indicam uma cobertura baixa e uma densidade média-baixa.

Quanto ao mapa obtido com a população F2, tal como nos mapas obtidos com a população F1, também há um número excessivo de grupos e não há uma distinção clara entre um conjunto de grupos de tamanhos maiores e outro de tamanhos relativamente mais pequenos, o que impossibilita a identificação de 25 grupos que representem os 25 pares de cromossomas do ananaseiro. No entanto há 23 grupos com mais de 50 cM depois de adicionada a ambas as suas extremidades a distância média entre marcadores (5.19 cM). Destes, há 12 que integram marcadores dos 2 progenitores e que provavelmente representam 12 dos 25 pares de cromossomas do ananaseiro, há 5 que só têm marcadores da var. *bracteatus* e 6 que só têm marcadores da var. *comosus*. O facto de não ter sido detectado ligamento entre os grupos com marcadores só de um, ou só de outro, progenitor não exclui a hipótese de poderem corresponder a cromossomas homólogos. Com uma população F2 de 142 plantas, para se detectar ligamento entre marcadores dominantes ligados em repulsão é necessário que a recombinação entre eles seja menor que 0.1 (Liu 1998). Logo, se em 2 grupos de cromossomas homólogos não estiverem mapeados loci suficientemente próximos para que seja detectada ligação em repulsão, a homologia entre os grupos não é identificada.

As estimativas de tamanho dos genomas calculadas com os dados da população F2 podem estar deturpadas devido a terem sido seleccionados para a construção dos mapas F2 os “primers” que amplificam maior número de marcadores mapeados anteriormente com a população F1. Por isso, e para se ter uma referência comum na análise comparativa dos mapas F1 e F2, a cobertura deste último foi calculada em relação à média dos comprimentos genómicos estimados com os dados da população F1 (3920 cM). Deste modo, o mapa obtido com a população F2 tem uma cobertura (60.7%) similar à do obtido com a população F1 para a var. *bracteatus* (59.1%), tal como a distância média entre marcadores adjacentes, respectivamente de 5.19 e 5.27 cM e a distância máxima entre marcadores adjacentes, respectivamente de 30.5 e 32.0 cM, revelam semelhanças na qualidade destes dois mapas.

A vantagem do mapa F2 é que integra marcadores de ambas as variedades, ao contrário dos mapas F1. Para além disso o mapa F2 inclui marcadores dos mapas F1, o que permitiu a integração dos mapas construídos com ambas as populações (figura 7). O agrupamento de marcadores estabelecido com a população F1 confirmou-se na população F2 com os marcadores analisados em ambas as populações, não havendo casos de ligamentos conflituosos, o que demonstra uma boa análise genotípica dos marcadores e revela segurança na qualidade dos mapas obtidos.

No mapa integrado a distância média entre marcadores (4.16 cM) é inferior à dos mapas construídos com cada uma das populações (mapas F1: 5.27 cM – var. *bracteatus*, 7.13 cM – var. *comosus*; mapa F2 - 5.19 cM), o que indica uma densidade média de marcadores superior à destes últimos. No entanto, a densidade de marcadores não é igual em todo o mapa integrado: os grupos que integram marcadores dos dois progenitores apresentam, como esperado, uma distância média entre marcadores mais reduzida (2.80 cM) do que os que têm apenas marcadores de um progenitor (5.98 cM).

No que respeita à cobertura do mapa integrado, o valor estimado (91.4%) é significativamente mais elevado que os valores estimados para os mapas F1 (59.1% – var. *bracteatus*, 33.2% – var. *comosus*) e para o mapa F2 (60.7%). Apesar desta estimativa apontar para uma representação do genoma do ananaseiro próxima de 100%, o seu valor deverá estar inflacionado, pois não há ainda uma integração completa de todos os grupos de cromossomas homólogos e há grupos só com marcadores analisados numa população e grupos só com marcadores analisados na outra população que podem corresponder à mesma região genómica. A cobertura real deste mapa deverá estar entre o valor do mapa com maior cobertura construído com base nas populações F1 e F2 e o valor estimado para o próprio, ou seja, entre 60.7 e 91.4%.

Apesar de não ser possível estabelecer uma relação grupo/cromossoma definitiva por haver relações de ligamento ainda por resolver, o mapa integrado apresentado disponibiliza à comunidade científica um conjunto de marcadores, principalmente os marcadores dos 31 grupos cuja cobertura é superior a 50 cM e os quais correspondem a 64% da cobertura do mapa, que podem ser utilizados como marcos de referência no genoma do ananaseiro (*A. comosus*) para o mapeamento de genes de características de interesse, monogénicas ou poligénicas (QTLs), em qualquer que seja a variedade botânica desta espécie. No Laboratório de Genética e Melhoramento de Plantas da FERN, Universidade do Algarve, onde foram construídos os mapas do ananaseiro aqui apresentados, foi já explorada a utilidade dos mapas genéticos neste tipo de aplicação: marcadores RAPD mapeados ao longo de vários grupos de ligamento de um mapa de *Brassica oleracea* (Farinhó et al. 2004) foram utilizados como pontos de referência para “percorrer” o genoma da couve algarvia - *Brassica oleracea* var. *tronchuda* – e mapear um segundo gene de resistência ao míldio (Leitão et al. 2006). Por outro lado, trabalhos como os de Waugh et al. (1997) e Rouppe van der Voort et al. (1997), respectivamente em cevada (*Hordeum vulgare*) e batata (*Solanum tuberosum*), em que foram

mapeados marcadores AFLP co-migrantes em diferentes genótipos destas espécies e foi confirmada a sua posição no mesmo *locus*, sugerem que este tipo de marcadores, o qual é o que está em maior número nos mapas aqui apresentados, também poderá ser utilizado em todas as variedades de ananaseiro.

O alelo dominante do gene *P* responsável pela característica morfológica “piping”, a segregar nas duas populações utilizadas neste trabalho, foi mapeado num grupo de 24.7 cM com quatro marcadores (figura 7, grupo 45_Ac12). A identificação de um marcador - OPO18320_Ac - a uma distância genética do *locus* “piping” de apenas 3.7 cM poderá ser o ponto de partida para a identificação de mais marcadores nesta zona do genoma (mapeamento “fino”) para posterior clonagem e identificação de genes candidatos a gene *P* por “chromosome walking”. A caracterização deste gene e a possibilidade de ser utilizado para transformação genética de variedades de ananaseiro é importante pois a sua expressão elimina os espinhos das folhas que dificultam o manuseamento da cultura de ananás e a sua colheita.

Para tornar mais eficiente a utilização do mapa iniciou-se o processo de sequenciação de marcadores distribuídos ao longo do mesmo. Oito dos 10 marcadores sequenciados estão em seis grupos maiores (cobertura > 50 cM) provavelmente a representar 6 cromossomas diferentes: 2 grupos têm dois marcadores sequenciados e 4 grupos têm apenas um. Com estas sequências, ou partindo destas para as sequências adjacentes, poderão ser desenvolvidos marcadores específicos polimórficos em populações que segreguem determinadas características, e o(s) gene(s) por elas responsáveis serão rapidamente mapeados se estiverem num dos cromossomas de origem das sequências. As sequências podem ser testadas como sondas para desenvolver marcadores RFLPs ou servir de base para desenhar “primers” para marcadores SCARs ou CAPs. Para além disso, um dos marcadores tem uma sequência tipo microsatélite. Um conjunto de marcadores específicos de referência para o genoma do ananaseiro poderá ser estabelecido pela sequenciação de pelo menos 3 marcadores de cada um

dos grupos de ligamento maiores (cobertura > 50 cM) do mapa integrado apresentado. Por outro lado, quer os 50 microssatélites disponíveis para o ananaseiro nas bases públicas de dados genómicos, quer alguns microssatélites que possam ser encontrados nos ESTs também já publicados para esta espécie, poderão ser testados nas populações de mapeamento utilizadas neste trabalho e eventualmente contribuir para melhorar o actual mapa.

6- CONSIDERAÇÕES FINAIS

O objectivo central do trabalho foi alcançado com a construção do primeiro mapa genético do ananaseiro (*Ananas comosus*). O elevado grau de cobertura e de densidade em marcadores deste mapa permitirá que venha a constituir o ponto de partida para a construção de um mapa genético de referência desta espécie. Este mapa é também o primeiro de toda a família *Bromeliaceae*, abrindo caminho para o estudo genómico, tanto estrutural como funcional, da mesma.

O facto de cruzamentos entre as cinco variedades botânicas da espécie *Ananas comosus* e de cruzamentos entre esta espécie e o *Ananas macrodontes* gerarem descendentes férteis faz com que este mapa possa ser utilizado transversalmente em todo o género *Ananas* para o desenvolvimento de novas variedades melhoradas, quer por selecção assistida por marcadores, quer através da clonagem de genes baseada no mapeamento e posterior utilização desses genes na transformação genética.

Apesar da densidade de marcadores do mapa não ser ainda suficiente para permitir a identificação de todos os clones de uma biblioteca genómica de largos fragmentos de DNA (em BACs por exemplo) relativamente completa, este permite desde já posicionar sequências chave ao longo do genoma onde as restantes sequências poderão facilmente ser ancoradas, permitindo a construção de mapas genómicos físicos. Estes mapas, por seu turno, poderão constituir uma base sólida para qualquer projecto de sequenciação genómica total do Ananaseiro.

7- REFERÊNCIAS BIBLIOGRÁFICAS

- Adam-Blondon A-F, Bernole A, Faes G, Lamoureux D, Pateyron S, Grando MS, Caboche M, Velasco R, Chalhoub B (2005) Construction and characterization of BAC libraries from major grapevine cultivars. *Theor Appl Genet* 110:1363–1371
- Adam-Blondon A-F, Roux ,C Claux D, Butterlin G, Merdinoglu D, This P (2004) Mapping 245 SSR markers on the *Vitis vinifera* genome: a tool for grape genetics *Theor Appl Genet* 109:1017–1027
- Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yandell MD, Zhang Q, Chen LX, Brandon RC, Rogers YH, Blazej RG, Champe M, Pfeiffer BD, Wan KH, Doyle C, Baxter EG, Helt G, Nelson CR, Gabor GL, Abril JF, Agbayani A, An HJ, Andrews-Pfannkoch C, Baldwin D, Ballew RM, Basu A, Baxendale J, et al. (2000) The genome sequence of *Drosophila melanogaster*. *Science* 287:2185-2195
- Aign V, Schulte U, Hoheisel JD (2001) Hybridization-Based Mapping of *Neurospora crassa* Linkage Groups II and V. *Genetics* 157:1015–1020
- Alonso-Blanco C, Peters AJM, Korneef M, Lister C, Dean C, van den Bosch N, Pot J, Kuiper MTR (1998) Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. *Plant J* 14:259-271
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389-3402
- Anderson S (1981) Shotgun DNA sequencing using cloned DNase I-generated fragments. *Nucleic Acids Res* 9:3015-3027
- Aradhya MK, Zee F, Manshardt RM (1994) Isozyme variation in cultivated and wild pineapple. *Euphytica* 79:87-89
- Aranzana MJ, Kim S, Zhao K, Bakker E, Horton M, Jakob K, Lister C, Molitor J, Shindo C, Tang C, Toomajian C, Traw B, Zheng H, Bergelson J, Dean C, Marjoram P, Nordborg M (2005) Genome-Wide Association Mapping in *Arabidopsis* Identifies Previously Known Flowering Time and Pathogen Resistance Genes. *PLoS Genet* 1(5):e60

- Aranzana MJ, Pineda A, Cosson P, Dirlewanger E, Ascasibar J, Cipriani G, Ryder CD, Testolin R, Abbott AG, King GJ, Iezzoni AF, Arús P (2003) A set of simple-sequence repeat (SSR) markers covering the *Prunus* genome. *Theor Appl Genet* 87:805–815
- Arondel V, Lemieux B, Hwang I, Gibson S, Goodman HM, Somerville CR (1992) Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. *Science* 258:1353-1355
- Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. *Plant Mol Biol Rep* 9:208-218
- Balint-Kurti PJ, Clendennen SK, Doleželová M, Valárik M, Doležel J, Beetham PR, May GD (2000) Identification and chromosomal localization of the monkey retrotransposon in *Musa* sp. *Mol Gen Genet* 263:908–15
- Ballvora A, Schornack S, Baker BJ, Ganai M, Bonas U, Lahaye T (2001) Chromosome landing at the tomato *Bs4* locus. *Mol Genet Genomics* 266:639-645
- Baudracco-Arnas S, Pitrat M (1996) A genetic map of melon (*Cucumis melo* L.) with RFLP, RAPD, isozyme, disease resistance and morphological markers. *Theor Appl Genet* 93: 57-64
- Baurens FC, Noyer JL, Lanaud C, Lagoda PJJ (1997) A repetitive sequence family of banana (*Musa* sp.) shows homology to Copia-like elements. *J Genet Breed* 51:135-142
- Bent AF, Kunkel BN, Dahlbeck D, Brown KL, Schmidt R, Giraudat J, Leung J, Staskawicz BJ (1994) *RPS2* of *A. thaliana*, a leucine rich repeat class of plant disease resistance genes. *Science* 265:1856-1860
- Bentley DR, Pruitt KD, Deloukas P, Schuler GD, Ostell J (1998) Coordination of human genome sequencing via a consensus framework map. *Trends Genet* 14:381-384
- Blanco A, Bellomo MP, Lotti C, Maniglio T, Pasqualone A, Simeone R, Troccoli A, di Fonzo N (1998) Genetic mapping of sedimentation volume across environments using recombinant inbred lines of durum wheat. *Plant Breed* 117:413-417
- Bliss FA, Arulsekhar S, Foolad MR, Becerra V, Gillen AM, Warburton ML, Dandekar AM, Kocsisne GM, Mydin KK (2002) An expanded genetic map of *Prunus* based on an interspecific cross between almond and peach. *Genome* 45:520–529
- Bodmer WF (1986) Human genetics: the molecular challenge. *Cold Spring Harb Symp Quant Biol* 51:1–13
- Boehnke M, Langefeld CD (1998) Genetic association mapping based on discordant sib pairs: the discordant alleles test (DAT). *Am J Hum Genet* 62:950-961
- Boneh L, Kuperus P, Van Tienderen PH (2003) Microsatellites in the bromeliads *Tillandsia fasciculata* and *Guzmania monostachya*. *Mol Ecol Notes* 3:302-303

- Botella J (2005a) Control of flowering in pineapple (*Ananas comosus*) by genetic engineering. 5th International Pineapple Symposium. South Africa. Oral presentation
- Botella J (2005b) The pineapple EST sequencing and microarray project. 5th International Pineapple Symposium. South Africa. Oral presentation
- Botella JR, Fairbairn DJ (2005) Present and future potential of pineapple biotechnology. *Acta Hort* 666:23-28
- Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a linkage map in man using Restriction Fragment Length Polymorphisms. *Am J Hum Genet* 32: 314-331
- Bouchez A, Hospital F, Causse M, Gallais A, Charcosset A (2002) Marker-assisted introgression of favorable alleles at Quantitative Trait Loci between maize elite lines. *Genetics* 162:1945–1959
- Boyle AL, Feltquite DM, Dracopoli NC, Housman DE, Ward DC (1992) Rapid physical mapping of cloned DNA on banded mouse chromosomes by fluorescence in situ hybridization. *Genomics* 12:106-115
- Bradshaw HD, Stettler RF (1993) Molecular genetics of growth and development in *Populus*. I. Triploidy in hybrid poplars. *Theor Appl Genet* 86: 301-307
- Broman KW, Murray JC, Sheffield VC, White RL, Weber JL (1998) Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am J Hum Genet* 63:861-869
- Brookes AJ (1999) The essence of SNPs. *Gene* 234:177-186
- Burr B, Burr FA (1991) Recombinant inbreds for molecular mapping in maize: theoretical and practical considerations. *Trends Genet* 7:55-60
- Büschges R, Hollricher K, Panstruga R, Simons G, Wolters M, Frijters A, van Daelen R, van der Lee T, Diergaarde P, Groenendijk J, Toopsch S, Vos P, Salamini F, Schultze-Lefert P (1997) The barley *Mlo* gene: a novel control element of plant pathogen resistance. *Cell* 88:695–705
- Cabral JRS, Coppens d'Eeckenbrugge G, de Matos AP (2000) Introduction of selfing in pineapple breeding. *Acta Hort* 529:165-168
- Cabral JRS, de Matos AP, Coppens d'Eeckenbrugge, G (1997) Segregation for resistance to fusariose, leaf margin type and leaf colour from the EMBRAPA Pineapple Hybridization Programme. *Acta Hort* 425:193-200
- Cai DG, Kleine M, Kifle S, Harloff HJ, Sandal NN, Marcker KA, Klein-Lankhorst R, Salentijn EMJ, Lange W, Stiekema WL, Wyss U, Grundler FMW, Jung C (1997) Positional cloning of a gene for nematode resistance in sugar beet. *Science* 275:832–834

- Cai Q, Guy CL, Moore GA (1994) Extension of the linkage map in *Citrus* using random amplified polymorphic DNA (RAPD) markers and RFLP mapping of cold-acclimation-responsive loci. *Theor Appl Genet* 89:606-614
- Canillieri C, Lafleurriel J, Macadre C, Varoquaux F, Parmentier Y, Picard G, Caboche M, Bouchez D (1998) A YAC contig map of *Arabidopsis thaliana* chromosome 3. *Plant J* 15:633-642
- Capinpin JM, Rotor GB (1937) A cytological and morphogenetic study of some pineapple varieties and their mutant and hybrid derivatives. *The Philippine Agriculturist* 26, 139-15
- Capparelli R, Viscardi M, Amoroso MG, Blaiotta G, Bianco M (2004) Inter-simple sequence repeat markers and flow cytometry for the characterization of closely related *Citrus limon* germplasms. *Biotechnol Lett* 26:1295-1299
- Carlier J, Leitão J (2006) Constructing An Integrated Map Of DNA Markers In Pineapple. *Plant & Animal Genomes XIV Conference*, January 14-18, San Diego, CA
- Carlier JD, Reis A, Duval MF, Coppens d'Eeckenbrugge G, Leitão JM (2004) Genetic maps of RAPD, AFLP and ISSR markers in *Ananas bracteatus* and *A. comosus* using the pseudo-testcross strategy. *Plant Breed* 123:186-192
- Carter CE, Marriage H, Goodenough PW (2000) Mutagenesis and kinetic studies of a plant cysteine proteinase with an unusual arrangement of acidic amino acids in and around the active site. *Biochemistry* 39:11005-11013
- Carter NP (1994) Cytogenetic Analysis by Chromosome Painting. *Cytometry* 18:2-10
- Casasoli M, Mattioni C, Cherubini M, Villani F (2001) A genetic linkage map of European chestnut (*Castanea sativa* Mill.) based on RAPD, ISSR and isozyme markers. *Theor Appl Genet* 102:1190-1199
- Castiglioni P, Ajmone-Marsan P, van Wijk R, Motto M (1999) AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution. *Theor Appl Genet* 99:425-431
- Cato SA, Gardner RC, Kent J, Richardson TE (2001) A rapid PCR-based method for genetically mapping ESTs. *Theor Appl Genet* 102:296-306
- Cazzonelli CI, Cavallaro AS, Botella JR (1998) Cloning and characterisation of ripening-induced ethylene biosynthetic genes from non-climacteric pineapple (*Ananas comosus*) fruits. *Aust J Plant Physiol* 25:513-518
- Chagné D, Lalanne C, Madur D, Kumar S, Frigerio J-M, Krier C, Decroocq S, Savouré A, Bou-Dagher-Kharratc M, Bracha J, Plomion C (2002) A high density genetic map of maritime pine based on AFLPs. *Ann For Sci* 59:627-636

- Chakravarti A, Lasher LA, Reefer JE (1991) A maximum likelihood method for estimating genome length using genetic linkage data. *Genetics* 128:175-182
- Chan YK (2005) Hybridization and selection in pineapple improvement: the experience in Malaysia. 5th International Pineapple Symposium 2005. South Africa. Oral presentation
- Chaparro JX, Werner DJ, Malley DO, Sederoff RR (1994) Targeted mapping and linkage analysis of morphological, isozyme and RAPD markers in peach. *Theor Appl Genet* 87: 805-815
- Charcosset A, Moreau L (2004) Use of molecular markers for the development of new cultivars and the evaluation of genetic diversity. *Euphytica* 137:81-94
- Chen M, Presting G, Barbazuk WB, Goicoechea JL, Blackmon B, Fang G, Kim H, Frisch D, Yu Y, Sun S, Higingbottom S, Phimpilai J, Phimpilai D, Thurmond S, Gaudette B, Li P, Liu J, Hatfield J, Main D, Farrar K, Henderson C, Barnett L, Costa R, Williams B, Walser S, Atkins M, Hall C, Budiman MA, Tomkins JP, Luo M, Bancroft I, Salse J, Regad F, Mohapatra T, Singh NK, Tyagi AK, Soderlund C, Dean RA, Wing RA (2002) An Integrated Physical and Genetic Map of the Rice Genome. *Plant Cell* 14:537-545
- Chen X, Temnykh S, Xu Y, Chu YG, McCouch SR (1997) Development of a microsatellite framework map providing genomewide coverage in rice (*Oryza sativa* L.). *Theor Appl Genet* 95: 553-567
- Chumakov IM, Rigault P, Le Gall I, Bellanne-Chantleot C, Billault A, Guillou S, Soularue P, Guasconi G, Poullier E, Gros I, Belova M, Sambucy I-L, Susini L, Gervy P, Glibert F, Beaufils S, Bui H, Massart C, De Tand M-F, Dukasz F, Lecoulant S, Ougen P, Perrot V, Saumier M, Soravito C, Bahouyila R, Cohen-Akenine A, Barillot E, Bertrand S, Codani J-J, Caterina D, Georges I, Lacroix B, Lucotte G, Sahbatou M, Schmit C, Sangouard M, Tubacher E, Dib C, et al. (1995) A YAC contig map of the human genome. *Nature* 377:175-297
- Chunwongse J, Phumichai C, Barbrasert C, Chunwongse C, Sukonsawan S, Boonreungrawd R (2000) Molecular mapping of mango cultivars 'Alphonso' and 'Palmer'. *Acta Horti* 509:193-206
- Collins J, Hohn B (1978) Cosmids: a type of plasmid gene-coding vector that is packageable in vitro in bacteriophage lambda heads. *Proc Natl Acad Sci USA* 83:4242-4246
- Collins JL (1960) The pineapple, botany, utilisation, cultivation. Leonard Hill Ltd, London
- Collins JL, Kerns KR (1946) Inheritance of three leaf types in the pineapple. *J Hered* 32:123-128

- Concibido VC, Vallee BL, Mclaird P, Pineda N, Meyer J, Hummel L, Yang J, Wu K & Delannay X (2003) Introgression of a quantitative trait locus for yield from *Glycine soja* into commercial soybean cultivars. *Theor Appl Genet* 106:575–582
- Conner JP, Brown SK, Weeden NF (1997) Randomly amplified polymorphic DNA-based genetic linkage maps of three apple cultivars. *J Am Soc Hortic Sci* 122:350–359
- Coppens d'Eeckenbrugge (2002) Evaluation and utilization of pineapple genetic resources from the Amazon to breed resistant varieties. EU INCO-DC Project, Final Report
- Coppens d'Eeckenbrugge G, Duval MF, Van Miegroet F (1993) Fertility and self-incompatibility in the genus *Ananas*. *Acta Hortic* 334:45-51
- Coppens d'Eeckenbrugge G, Leal F (2003). Morphology, Anatomy and Taxonomy. In: Bartholomew, DP, Paull, RE and Rohrbach, KG (eds) *The Pineapple: Botany, Production and Uses*. CABI Publishing, Oxon
- Coppens d'Eeckenbrugge G, Bernasconi B, Messiaen B, Duval MF (1997b). Using incompatibility alleles as genetic markers to identify pineapple varieties. *Acta Hortic* 425:161-170
- Coppens d'Eeckenbrugge G, Duval MF (1999) Pineapple germplasm conservation: experiences from the Martinique field collection. In: F. Engelmann (ed.): *Management of field and in vitro genebanks*. International Plant Genetic Resources Institute, Rome
- Coppens d'Eeckenbrugge G, Leal F, Duval MF (1997a) Germplasm resources of pineapple. *Hortic Rev* 21:133-175
- Cotias-de-Oliveira ALP, Assis JGA, Bellintani MC, Andrade JCS, Guedes MLS (2000) Chromosome numbers in Bromeliaceae. *Genet Mol Biol* 23:173-177
- Crawford AM, Montgomery GW, Pierson CA, Brown T, Dodds KG, Sunden SLF, Henry HM, Ede AJ, Swarbrick PA, Berryman T, Penty JM, Hill DF (1994) Sheep Linkage Mapping: Nineteen Linkage Groups Derived From the Analysis of Paternal Half-sib Families. *Genetics* 137:573-579
- Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J, Specht JE (1999) An integrated genetic linkage map of the soybean. *Crop Sci* 39:1464 –1490
- Crespel L, Chirollet M, Durel CE, Zhang D, Meynet J, Gudin S (2002) Mapping of qualitative and quantitative phenotypic traits in *Rosa* using AFLP markers. *Theor Appl Genet* 105:1207-1214

- Cristofani M, Machado MA, Grattapaglia D (1999) Genetic linkage maps of *Citrus sunki* Hort. ex. Tan. and *Poncirus trifoliata* (L.) Raf. and mapping of citrus tristeza virus resistance gene. *Euphytica* 109:25-32
- Cuevas IC, Podestá FE (2000) Purification and physical and kinetic characterization of an NAD⁺-dependent malate dehydrogenase from leaves of pineapple (*Ananas comosus*). *Physiol Plant* 108:240-248
- Curtis D (1997) Use of siblings as controls in case-control association studies. *Ann Hum Genet* 61:319-333
- D'Hont A (2005) Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. *Cytogenet Genome Res* 109:27-33
- D'hont A, Paget-Goy A, Escoute J, Carreel F (2000) The interspecific genome structure of cultivated banana, *Musa* spp. revealed by genomic DNA in situ hybridization, *Theor Appl Genet* 100:177-183
- Dalbó MA, Ye GN, Weeden NF, Steinkellner H, Sefc KM, Reisch BI (2000) A gene controlling sex in grapevines placed on a molecular marker-based genetic map. *Genome* 43:333–340
- Davis TM, Yu H (1997) A linkage map of the diploid strawberry, *Fragaria vesca*. *J Hered* 88:215-221
- Debener T, Mattiesch L (1999) Construction of a genetic linkage map for roses using RAPD and AFLP markers. *Theor Appl Genet* 99:891-899
- Dettori MT, Quarta R, Verde I (2001) A peach linkage map integrating RFLPs, SSRs, RAPDs and morphological markers. *Genome* 44:783–790
- Devey ME, Delfino-Mix A, KINLOCH Jr BB, Neale DB (1995) Random amplified polymorphic DNA markers tightly linked to a gene for resistance to white pine blister rust in sugar pine. *Genetics* 92:2066-2070
- Devey ME, Jermstad KD, Tauer CG, Neale DB (1991) Inheritance of RFLP loci in a loblolly pine three-generation pedigree. *Theor Appl Genet* 83:238-242
- Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 380:152-154
- Dietrich WF, Miller J, Steen R, Merchant MA, Damron-Boles D, Husain Z, Dredge R, Daly MJ, Ingalls KA, O'Connor TJ, Evans CA, Deangelis MM, Levinson DM, Kruglyak L, Goodman N, Copeland NG, Jenkins NA, Hawkins TL, Stein L, PAGE DC, LANDER ES (1996) A comprehensive genetic map of the mouse genome. *Nature* 380:149-152

- Dirlewanger E, Graziano E, Joobeur T, Garriga-Calderé F, Cosson P, Howad W, Arús P (2004) Comparative mapping and marker-assisted selection in *Rosaceae* fruit crop. *Proc Natl Acad Sci USA* 101:9891–9896
- Dirlewanger E, Pascal T, Zuger C, Kervella J (1996) Analysis of molecular markers associated with powdery mildew resistance genes in peach (*Prunus persica* (L.) Batsch) x *Prunus davidiana* hybrids. *Theor Appl Genet* 93:909–919
- Dirlewanger E, Pronier V, Parvery C, Rothan C, Guye A, Monet R (1998) Genetic linkage map of peach [*Prunus persica* (L.) Batsch] using morphological and molecular markers. *Theor Appl Genet* 97:888–895
- Doležel J, Doleželová M, Novák FJ (1994) Flow cytometric estimation of nuclear DNA amount in diploid bananas (*Musa acuminata* and *M. balbisiana*). *Biol Plant* 36:351–357
- Doleželová M, Valárik M, Swennen R, Horry JP, Doležel J (1998) Physical mapping of the 18S–25S and 5S ribosomal RNA genes in diploid bananas. *Biol Plant* 41:497–505
- Donis-Keller H, Green P, Helms C, Cartinhour S, Weiffenbach B, Stephens K, Keith TP, Bowden DW, Smith DR, Lander ES, Botstein D, Akots G, Rediker KS, Gravius T, Brown VA, Rising MB, Parker C, Powers JA, Watt DE, Kauffman ER, Bricker A, Phipps P, Muller-Kahle H, Fulton TR, Siu Ng, Schumm JW, Braman JC, Knowlton RG, Barker DF, Crooks SM, Lincoln SE, Daly MJ, Abrahamson J (1987) A genetic linkage map of the human genome. *Cell* 51:319–337
- Doucleff M, Jin Y, Gao F, Riaz S, Krivanek AF, Walker MA (2004) A genetic linkage map of grape, utilizing *Vitis rupestris* and *Vitis arizonica*. *Theor Appl Genet* 109:1178–1187
- Dujardin M (1991) Cytogénétique de l'ananas. *Fruits* 46:376–379
- Durham RE, Liou PC, Gmitter FG Jr, Moore GA (1992) Linkage of restriction fragment length polymorphisms and isozymes in Citrus. *Theor Appl Genet* 84:39–48
- Duval MF, Buso GSC, Ferreira FR, Noyer JL, Coppens d'Eeckenbrugge G, Hamon P, Ferreira ME (2003) Relationships in Ananas and other related genera using chloroplast DNA restriction site variation. *Genome* 46:990–1004
- Duval MF, Noyer JL, Perrier X, Coppens d'Eeckenbrugge G, Hamon P (2001) Molecular diversity in pineapple assessed by RFLP markers. *Theor Appl Genet* 102:83–90
- Escalona M, Lorenzo JC, González B, Daquinta M, González JL, Desjardins Y, Borroto CG (1999) Pineapple (*Ananas comosus* (L.) Merr) micropropagation in temporary immersion systems. *Plant Cell Rep* 18:743–748

- Espinosa P, Lorenzo JC, Iglesias A, Yabor L, Menéndez E, Borroto J, Hernández L, Arencibia AD (2002) Production of pineapple transgenic plants assisted by temporary immersion bioreactors. *Plant Cell Rep* 21:136–140
- Falk CT (1989) A simple scheme for preliminary ordering of multiple loci: application to 45 CF families. In: Elston, Spence, Hodge and MacCluer (ed.), *Multipoint mapping and linkage based upon affected pedigree members*. Genetic Workshop 6. Liss, New York
- Falk CT, Rubinstein P (1987) Haplotype relative risks: An easy reliable way to construct a proper control sample for risk calculations. *Am J Hum Genet* 51:227-233
- Farinhó M, Carlier J, Svetleva D, Coelho P, Monteiro A, Leitão J (2004) Mapping adult plant resistance to downy mildew in *Brassica oleracea*. *Theor Appl Genet* 109:1392-1398
- Fauré S, Noyer JL, Horry JP, Bakry F, Lanaud C, González de León D (1993). A molecular-based linkage map of diploid bananas (*Musa acuminata*). *Theor Appl Genet* 87:517-526
- Fazio G, Staub JE, Stevens MR (2003) Genetic mapping and QTL analysis of horticultural traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. *Theor Appl Genet* 107:864–874
- Ferreira AR, Foutz KR, Keim P (2000) Soybean genetic map of RAPD markers assigned to an existing scaffold RFLP map. *J Hered* 91:392-396
- Firoozabady E, Gutterson N (2003) Cost-effective in vitro propagation methods for pineapple. *Plant Cell Rep* 21:844–850
- Fischer SG, Lerman LS (1983) DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: correspondence with melting theory. *Proc Natl Acad Sci USA* 80:1579-1583
- Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb J-F, Dougherty BA, Merrick JM, McKenney K, Sutton G, Fitzhugh W, Fields C, Gocayne JD, Scott J, Shirley R, Liu L-I, Glodek A, Kelley JM, Weidman JF, Phillips CA, Spriggs THE, Cotton MD, Utterback TR, Hanna MC, Nguyen DT, Saudek DM, Brandon RC, Fine LD, Fritchman JL, Fuhrmann JL, Geoghagen NSM, Gnehm CL, McDonald LA, Small KV, Fraser CM, Smith HO, Venter JC (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269: 496-512
- Foisset N, Delourme R, Barret P, Hubert N, Landry BS, Renard M (1996) Molecular-mapping analysis in *Brassica napus* using isozyme, RAPD and RFLP markers on a double haploid progeny. *Theor Appl Genet* 93:1017-1025
- Foolad MR, Arulsekar S, Becerra V, Bliss FA (1995) A genetic map of *Prunus* based on an interspecific cross between peach and almond. *Theor Appl Genet* 91:262–269

- Foulongne M, Pascal T, Arús P, Kervella J (2003) The potential of *Prunus davidiana* for introgression into peach [*Prunus persica* (L.) Batsch] assessed by comparative mapping. *Theor Appl Genet* 107:227–238
- Francia E, Tacconi G, Crosatti C, Barabaschi D, Bulgarelli D, Dall’Aglia E, Vale G (2005) Marker assisted selection in crop plants. *Plant Cell Tissue Organ Cult* 82:317-342
- Fraser CM, Eisen JA, Salzberg SL (2000) Microbial genome sequencing. *Nature* 406:799–803
- Fraser LG, Harvey CF, Crowhurst RN, De Silva HN (2004) EST-derived microsatellites from *Actinidia* species and their potential for mapping. *Theor Appl Genet* 108:1010-1016
- Galvão OP, Palhano FL, Santos RB, Orlando MTD, Ventura JA, Fernandes PMB (2005) 5th International Pineapple Symposium. South Africa. Poster
- García ML (1988) Etude taxonomique du genre *Ananas*. Utilisation de la variabilité enzymatique. These de Doctorat. Université des Sciences et Techniques du Languedoc. Montpellier
- Garcia R, Asins MJ, Forner J, Carbonell EA (1999) Genetic analysis of apomixis in *Citrus* and *Poncirus* by molecular markers. *Theor Appl Genet* 99:511-518
- Gardner RC, Howarth AJ, Hahn P, Brown-Luedi M, Shepherd RJ, Messing J (1981) The complete nucleotide sequence of an infectious clone of cauliflower mosaic virus by M13mp7 shotgun sequencing. *Nucleic Acids Res* 9:2871-2888
- Gebhardt C, Ritter E, Debener T, Schachtschabel U, Walkemeier B, Uhrig H, Salamini F (1989) RFLP-analysis and linkage mapping in *Solanum tuberosum*. *Theor Appl Genet* 78:65-75
- George AW, Mengersen KL, Davis GP (1999) A Bayesian Approach to Ordering Gene Markers. *Biometrics* 55:419-429
- Goffeau A, Aert R, Agostini-Carbone ML, Ahmed A, Aigle M, Alberghina L, Albermann K, Albers M, Aldea M, Alexandraki D, Aljinovic G, Allen E, Altmann R, Alt-Mörbe J, André B, Andrews S, Ansorge W, Antoine G, Anwar R, Aparicio A, Araujo R, Arino J, Arnold W, Arroyo J, Aviles E, Backes U, Baclet MC, Badcock K, Bahr A, Baladron V, Ballesta JPG, Bankier AT, Banrevi A, Bargues M, Baron L, Barreiros T, Barrell BG, Barthe C, Barton AB, Baur A, Bécam A-M, et al. (1997) The yeast genome directory. *Nature* 387:S1–S105
- Gonzalo MJ, Oliver M, Garcia-Mas J, Monfort A, Dolcet-Sanjuan R, Katzir N, Arús P, Monforte AJ (2005) Simple-sequence repeat markers used in merging linkage maps of melon (*Cucumis melo* L.) *Theor Appl Genet* 110:802–811
- Grando MS, Bellin D, Edwards KJ, Pozzi C, Stefanini M, Velasco R (2003) Molecular linkage maps of *Vitis vinifera* L. and *Vitis riparia* Mchx. *Theor Appl Genet* 106:1213-1224

- Grant JR (1998) An annotated catalogue of the generic names of the Bromeliaceae. *Selbyana* 19:91-121
- Grant MR, Godiard L, Straube E, Ashfield T, Lewald J, Sattler A, Innes RW, Dangl JL (1995) Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance. *Science* 269:843-846
- Grattapaglia D, Bertolucci FLG, Penchel R, Sederoff RR (1996) Genetic Mapping of Quantitative Trait Loci Controlling Growth and Wood Quality Traits in *Eucalyptus grandis* Using a Maternal Half-sib Family and RAPD Markers. *Genetics* 144:1205-1214
- Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics* 137:1121-1137
- Green ED (2001) Strategies for the systematic sequencing of complex genomes. *Nat Rev Genet* 2:573-583
- Groenen MAM, Crooijmans RPMA, Veenendaal A, Cheng HH, Siwek M, van der Poel JJ (1998) A Comprehensive Microsatellite Linkage Map of the Chicken Genome. *Genomics* 49:265-274
- Gupta PK, Rustgi S (2004) Molecular markers from the transcribed/expressed region of the genome in higher plants. *Funct Integr Genomics* 4:139-162
- Gupta PK, Rustgi S, Kulwal PL (2005) Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant Mol Biol* 57:461-485
- Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, Bernardi G, Lathrop M, Weissenbach J (1994) The 1993-94 Genethon human linkage map. *Nat Genet* 7:246-249
- Haanstra JPW, Wye C, Verbankel H, Meijer-Dekens F, van den Berg P, Odinet P, van Heusden AW, Tanksley S, Lindhout P, Peleman J (1999) An integrated high-density RFLP-AFLP map of tomato based on two *Lycopersicon esculentum* x *L. pennellii* F2 populations. *Theor Appl Genet* 99:254-271
- Hackauf B, Wehling P (2005) Approaching the self-incompatibility locus Zin rye (*Secale cereale*L.) via comparative genetics. *Theor Appl Genet* 110: 832-845
- Haldane JBS (1919) The combination of linkage values and the calculation of distances between the loci of linked factors. *J Genet* 8:299-309
- Halushka MK, Fan JB, Bentley K, Hsie L, Shen N, Weder A, Cooper R, Lipshutz R, Chakravarti A (1999) Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat Genet* 22:239-247

- Hamilton CM, Frary A, Lewis C, Tanksley SD (1996) Stable transfer of intact high molecular weight DNA into plant chromosomes. *Proc Natl Acad Sci USA* 93:9975-9979
- Harper G, Osuji JO, Heslop-Harrison JS, Hull R (1999) Integration of Banana streak badnavirus into the *Musa* genome: molecular and cytogenetic evidence. *Virology* 255:207–213
- Harper ME, Ullrich A, Saunders GF (1981) Localization of the human insulin gene to the distal end of the short arm of chromosome 11. *Proc Natl Acad Sci USA* 78:4458-4460
- Harry DE, Temesgen B, Neale DB (1998) Codominant PCR-based markers for *Pinus taeda* developed from mapped cDNA clones. *Theor Appl Genet* 97:327-336
- Hashizume T, Shimamoto I, Harusima, Yui M, Sato T, Imai T, Hirai M (1996) Construction of a linkage map for watermelon (*Citrullus lanatus*) using random amplified polymorphic DNA (RAPD). *Euphytica* 90:265–273
- Hashizume T, Shimamoto I, Hirai M (2003) Construction of a linkage map and QTL analysis of horticultural traits for watermelon [*Citrullus lanatus* (Thunb.) Matsum & Nakai] using RAPD, RFLP and ISSR markers. *Theor Appl Genet* 106:779–785
- Helentjaris T, Slocum M, Wright S, Schaefer A, Nienhuis J (1986) Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphism. *Theor Appl Genet* 72:761-769
- Hemmat M, Weeden NF, Manganaris AG, Lawson DM (1994) Molecular marker linkage map for apple. *J Hered* 85:4–11
- Hori K, Kobayashi T, Shimizu A, Sato K, Takeda K, Kawasaki S (2003) Efficient construction of high-density linkage map and its application to QTL analysis in barley. *Theor Appl Genet* 107:806–813
- Horvath SM, Laird NM (1998) A discordant-sibship test for disequilibrium and linkage: no need for parental data. *Am J Hum Genet* 63:1886-1897
- Hoskins RA, Nelson CR, Berman BP, Laverty TR, George RA, Ciesiolka L, Naeemuddin M, Arenson AD, Durbin J, David RG, Tabor PE, Bailey MR, DeShazo DR, Catanese J, Mammoser A, Osoegawa K, de Jong PJ, Celniker SE, Gibbs RA, Rubin GM, Scherer SE (2000) A BAC-based physical map of the major autosomes of *Drosophila melanogaster*. *Science* 287:2271–2274
- <http://apps.fao.org>
- <http://cgpbr.mhpcc.hawaii.edu/tools/papaya/>
- <http://fcbs.org/articles/Bigenerics.htm>
- <http://snp.cshl.org/>
- <http://tropgenedb.cirad.fr>

<http://www.angra.uac.pt/pessoais/docentes/jbatista/Ananas/>
<http://www.bartleby.com/61/>
<http://www.dddmag.com/Glossary.aspx>
<http://www.medicinenet.com/script/main/hp.asp>
<http://www.mercksource.com>
<http://www.musagenomics.org>
<http://www.ncbi.nlm.nih.gov>
<http://www.ncbi.nlm.nih.gov/BLAST/Genome/PlantBlast.shtml?2>
<http://www.ncbi.nlm.nih.gov/BLAST/producttable.shtml#db>
<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj>
<http://www.ncbi.nlm.nih.gov/genomes/leuks.cgi>
http://www.ornl.gov/sci/techresources/Human_Genome/glossary
<http://www.pgel.com.au>

Hurtado MA, Romero C, Vilanova S, Abbott AG, Ll acer G, Badenes ML (2002) Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.), and mapping of PPV (sharka) resistance. *Theor Appl Genet* 105:182–191

Iaonnou I, Amemiya CT, Garnes J, Kroisel PM, Shizuya H, Chen C, Batzer MA, de Jong PJ, (1994) A new bacteriophage P1-derived vector for propagation of large human DNA fragments. *Nat Genet* 6:84-89

Iketani H, Abe K, Yamamoto T, Kotobuki K, Sato Y, Saito T, Terai O, Matsuta N, Hayashi T (2001) Mapping of disease-related genes in Japanese pear using a molecular linkage map with RAPD markers. *Breed Sci* 51:179-184

International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860-921

Jarrell DC, Roose ML, Traugh SN, Kupper RS (1992) A genetic map of citrus based on segregation of isozymes and RFLPs in an intergeneric cross. *Theor Appl Genet* 84:49-56

Jenczewski E, Gherardi M, Bonnin I, Prospero JM, Olivieri I, Huguet T (1997) Insight on segregation distortions in two intraspecific crosses between annual species of *Medicago* (*Leguminosae*). *Theor Appl Genet* 94:682-691

Jensen J, Jorgensen JH (1975a) The barley chromosome 5 linkage map. *Hereditas* 80:5-16

Jensen J, Jorgensen JH (1975b) The barley chromosome 5 linkage map II. *Hereditas* 80:17-26

Johnson CV, Singer RH, Lawrence JB (1991) Fluorescent detection of nuclear RNA and DNA: implications for genome organization. *Methods Cell Biol* 35:73-98

- Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, van deWiel C, Bredemeijer G, Vosman B, Matthes M, Daly A, Brettschneider R, Bettini P, Buiatti M, Maestri E, Malcevski A, Marmiroli N, Aert R, Volckaert G, Rueda J, Linacero R, Vazquez A, Karp A (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Mol Breed* 3: 381–390
- Jones ES, Dupal MP, Dumsday JL, Hughes LJ, Forster JW (2002a) An SSR-based genetic linkage map for perennial ryegrass (*Lolium perenne* L.). *Theor Appl Genet* 105:577-584
- Jones ES, Mahoney NL, Hayward MD, Armstead IP, Jones JG, Humphreys MO, King IP, Kishida T, Yamada T, Balfourier F, Charmet G, Forster JW (2002b) An enhanced molecular marker based genetic map of perennial ryegrass (*Lolium perenne*) reveals comparative relationships with other Poaceae genomes. *Genome* 45:282-295
- Joobeur T, Viruel MA, de Vicente MC, Jáuregui B, Ballester J, Dettori MT, Verde I, Truco MJ, Messeguer R, Batlle I, Quarta R, Dirlwanger E, Arus P (1998) Construction of a saturated linkage map of *Prunus* using an almond x peach F2 progeny. *Theor Appl Genet* 97:1034-1041
- Kaló P, Endre G, Zimányi L, Csanádi G, Kiss GB (2000) Construction of an improved linkage map of diploid alfalfa (*Medicago sativa*). *Theor Appl Genet* 100:641-657
- Kaplan NL, Hill WG, Weir BS (1995) Likelihood methods for locating disease genes in nonequilibrium populations. *Am J Hum Genet* 56:18–32
- Kashkush K, Jinggui F, Tomer E, Hillel J, Lavi U (2001) Cultivar identification and genetic map of mango (*Mangifera indica*). *Euphytica* 122:129–136
- Kato CY, Nagai C, Moore PH, Zee F, Kim MS, Steiger DL, Ming R (2004) Intra-specific DNA polymorphism in pineapple (*Ananas comosus* (L.) Merr.) assessed by AFLP markers. *Genet Resour Crop Evol* 51:815-825
- Keim P, Schupp JM, Travis SE, Clayton K, Zhu T, Shi L, Ferreira A, Webb DM (1997) A high density genetic map of soybean based upon AFLP markers. *Crop Sci* 37:537-543
- Kijas JMH, Thomas MR, Fowler JCS, Roose ML (1997) Integration of trinucleotide microsatellites into a linkage map of *Citrus*. *Theor Appl Genet* 94:701-706
- Kim UJ, Shizuya H, de Jong PJ, Birren B, Simon MI (1992) Stable propagation of cosmid sized human DNA inserts in an F factor based vector. *Nucleic Acids Res* 20:1083-1085
- King JJ, Bradeen JM, Bark O, McCallum JA, Havey MJ (1998) A low-density genetic map of onion reveals a role for tandem duplication in the evolution of an extremely large diploid genome. *Theor Appl Genet* 96:52-62

- Kinsuat MJ, Kumar SV (2005) Isolation of Di-, Tri- and Tetranucleotide Microsatellite Loci from Pineapple (*Ananas comosus* var *comosus*) using the 5'- Anchored PCR Technique. Proceedings of the National genetics Congress of Malaysia.
- Klein RR, Klein PE, Mullet JE, Minx P, Rooney WL, Schertz KF (2005) Fertility restorer locus *Rf1* of sorghum (*Sorghum bicolor* L.) encodes a pentatricopeptide repeat protein not present in the colinear region of rice chromosome 12. *Theor Appl Genet* 111:994-1012
- Knapp SJ, Holloway JL, Bridges WC, Liu BH (1995) Mapping dominant markers using F₂ matings. *Theor Appl Genet* 91:74-81
- Kojima T, Nagaoka T, Noda K, Ogihara Y (1998) Genetic linkage map of ISSR and RAPD markers in Einkorn wheat in relation to that of RFLP markers. *Theor Appl Genet* 96:37-45
- Konieczny A, Ausubel FM (1993) A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. *Plant J* 4: 403-410
- Kosambi DD (1944) The estimation of map distances from recombination values. *Ann Eugen* 12:172-175
- Krutovskii KV, Vollmer SS, Sorensen FC, Adams WT, Knapp SJ, Strauss SH (1998) RAPD Genome Maps of Douglas-Fir. *J Hered* 89:197-205
- Kurata N, Umehara Y, Tanoue H, Sasaki T (1997) Physical mapping of the rice genome with YAC clones. *Plant Mol Biol* 35:101-113
- Lacape J-M, Nguyen T-B, Thibivilliers S, Bojinov B, Courtois B, Cantrell RG, Burr B, Hau B (2003) A combined RFLP-SSR-AFLP map of tetraploid cotton based on a *Gossypium hirsutum* × *Gossypium barbadense* backcross population. *Genome* 46:612-626
- Lan T-H, DelMonte TA, Reischmann KP, Hyman J, Kowalski SP, McFerson J, Kresovich S, Paterson AH (2000) An EST-enriched Comparative Map of *Brassica oleracea* and *Arabidopsis thaliana*. Cold Spring Harbour Lab. Press. *Genome Res* 10: 776-788
- Lander ES, Botstein D (1986) Mapping complex genetic traits in humans: new methods using a complete RFLP linkage map. *Cold Spring Harb Symp Quant Biol* 51:49-62
- Lander ES, Green P (1987) Construction of multilocus genetic linkage maps in humans. *Proc Nat Acad Sci USA* 84:2363-2367
- Lander RS, Green P, Abrahamson J, Baarlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174-181.
- Landry BS, Hubert N, Etoh T (1991) A genetic map for *Brassica napus* based on restriction fragment length polymorphisms detected with expressed DNA sequences. *Genome* 34: 543-552

- Lange K, Boehnke M (1982) How many polymorphic genes will it take to span the human genome? *Am J Hum Genet* 34:842–845
- Leal F (1990) On the validity of *Ananas monstrosus*. *J Bromel Soc* 40:246-249
- Leal F, Coppens d'Eeckenbrugge G (1996) Pineapple. In: Janick J and Moore JN (eds): *Fruit Breeding. I. Tree and tropical fruits*. Wileys and Sons, New York
- Lebrun P, Chevalier MH (1988) Starch and polyacrylamide gel electrophoresis of *Hevea brasiliensis*: A laboratory manual. IRCA-CIRAD, Paris
- Leitão JM, Farinhó MJ, Leitão JV, Sousa NF, Coelho PS, Monteiro AA, Bowers JE, Estill JC, Paterson AH (2006) Tracking Down Pp523 And Pp1A, Two *Brassica oleracea* L Downy Mildew Resistance Genes. Plant & Animal Genomes XIV Conference, January 14-18, San Diego, CA
- Lerceteau-Köhler E, Guerin G, Laigret F, Denoyes-Rothan B (2003) Characterization of mixed disomic and polysomic inheritance in the octoploid strawberry (*Fragaria x ananassa*) using AFLP mapping. *Theor Appl Genet* 107:619–628
- Lespinasse D, Rodier-Goud M, Grivet L, Leconte A, Legnate BH, Seguin CM (2000) A saturated genetic linkage map of rubber tree (*Hevea* spp.) based on RFLP, AFLP, microsatellite, and isozyme markers. *Theor Appl Genet* 100:127-138
- Levi A, Thomas CE, Joobeur T, Zhang X, Davis A (2002) A genetic linkage map for watermelon derived from a testcross population: (*Citrullus lanatus* var. *citroides* x *C. lanatus* var. *lanatus*) x *Citrullus colocynthis*. *Theor Appl Genet* 105:555–563
- Lichter P, Tang CJC, II K, Hermanson G, Evans GA, Housman D, Ward DC (1990) High-resolution mapping of human chromosome 11 by in situ hybridization with cosmid clones. *Science* 247:64-69
- Liebhart R, Gianfranceschi L, Koller B, Ryder CD, Tarchini R, Van de Weg E, Gessler C (2002) Development and characterization of 140 new microsatellites in apple (*Malus domestica* Borkh.) *Mol Breed* 10:217–241
- Liebhart R, Koller B, Gianfranceschi L, Gessler C (2003) Creating a saturated reference map for the apple (*Malus domestica* Borkh.) genome. *Theor Appl Genet* 106:1497-1508
- Lin BY, Ritschel PS, Ferreira FR (1987) Número cromossômico de exemplares da família *Bromeliaceae*. *Rev Bras Fruticult* 9:49-55
- Lin C-H, Kuan C-S, Hsu Y-M, Bartholomew (2005) 5th International Pineapple Symposium. South Africa. Oral presentation
- Lin MT, Fu MT, Ken CF, Lin CT (2000) Cloning and characterization of a cDNA encoding for Cu/Zn-superoxide dismutase from pineapple. *Plant Physiol* 122:619-619

- Lindgren G, Sandberg K, Persson H, Marklund S, Breen M, Sandgren B, Carlstén J, Ellegren H (1998) A Primary Male Autosomal Linkage Map of the Horse Genome. *Genome Res* 8:951–966
- Litt M, Luty JA (1989) A Hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle action gene. *Am J Genet* 44:397-401
- Liu BH (1998) *Statistical Genomics: Linkage, Mapping and QTL Analysis*. CRC Press, Boca Raton
- Liu S, Saha S, Stelly DM, Burr B, Cantrell RG (2000) Chromosomal assignment of microsatellite loci in cotton. *J Hered* 91:326–332
- Lodhi MA, Ye GN, Weeden NF, Reisch BI (1995) A molecular marker-based linkage map of *Vitis*. *Genome* 38:786–794
- Loeillet D (1997) The world pineapple market: the importance of Europe. *Acta Hort* 425:37-48
- Loeillet D (2003) The world pineapple market. When growth goes hand in hand with diversity. *FruiTrop* 100:9-11
- Lu ZX, Sosinski B, Reighard GL, Baird WV, Abbott AG (1998) Construction of a genetic linkage map and identification of AFLP markers for resistance to root-knot nematodes in peach rootstocks. *Genome* 41:199–207
- Luro F, Laigret F, Lorieux M, Ollitrault P (1996) *Citrus* genome mapping with molecular markers: two maps obtained by segregation analysis of progeny of one intergeneric cross. *Proceedings of the International Society of Citriculture* 2:862-866
- Luther HE (2002) An alphabetical list of bromeliad binomials, 8th edition. The Bromeliad Society International, Newberg, Oregon, 82 pp
- Ma H, Moore PH, Liu Z, Kim MS, Yu Q, Fitch MMM, Sekioka T, Paterson AH, Ming R (2004) High-Density Linkage Mapping Revealed Suppression of Recombination at the Sex Determination Locus in Papaya. *Genetics* 166:419–436
- Maddox JF, Davies KP, Crawford AM, Hulme DJ, Vaiman D, Cribiu EP, Freking BA, Beh KJ, Cockett NE, Kang N, Riffkin CD, Drinkwater R, Moore SS, Dodds KG, Lumsden JM, van Stijn TC, Phua SH, Adelson DL, Burkin HR, Broom JE, Buitkamp J, Cambridge L, Cushwa WT, Gerard E, Galloway SM, Harrison B, Hawken RJ, Hiendleder S, Henry HM, Medrano JF, Paterson KA, Schibler L, Stone RT, van Hest B (2001) An Enhanced Linkage Map of the Sheep Genome Comprising More Than 1000 Loci. *Genome Res* 11:1275 –1289
- Maliepaard C, Alston FH, Van Arkel G, Brown LM, Chevreau E, Dunemann F, Evans KM, Gardiner S, Guilford P, van Heusden AW, Janse J, Laurens F, Lynn JR, Manganaris AG, Den Nijs APM, Periam N, Rikkerink E, Roche P, Ryder C, Sansavini S, Schmidt H, Tartarini S,

- Verhaegh JJ, Vrieling-Van Ginkel M, King GJ (1998) Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. *Theor Appl Genet* 97:60–73
- Markert CL, Moller F (1959) Multiple forms of enzymes: tissues, ontogenetic, and species specific patterns. *Proc Natl Acad Sci USA* 45:753-763
- Marques CM, Araújo JA, Ferreira JG, Whetten R, O'Malley DM, Liu BH, Sederoff R (1998): AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. *Theo Appl Genet* 96, 727-737
- Martin ER, Kaplan NL, Weir BS (1997) Tests for linkage and association in nuclear families. *Am J Hum Genet* 61:439-448
- Martin ER, Monks SA, Warren LL, Kaplan NL (2000) A Test for Linkage and Association in General Pedigrees: The Pedigree Disequilibrium Test. *Am J Hum Genet* 67:146-154
- Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganai MW, Spivey R, Wu T, Earle ED, Tanksley SD (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* 262:1432-1435.
- Mather K (1938) Measurement of linkage in heredity. Methuen, London
- Matos AP, Cabral JRS, Teixeira Júnior T (2005) 5th International Pineapple Symposium. South Africa. Oral presentation
- Matthews BF, Devine TE, Weisemann JM, Beard HS, Lewers KS, MacDonald MH, Park Y-B, Maiti R, Lin J-J, Kuo J, Pedroni MJ, Cregan PB, Saunders JA (2001) Incorporation of Sequenced cDNA and Genomic Markers into the Soybean Genetic Map. *Crop Sci* 41:516–521
- Maughan PJ, Bonifacio A, Jellen EN, Stevens MR, Coleman CE, Ricks M, Mason SL, Jarvis DE, Gardunia BW, Fairbanks DJ (2004) A genetic linkage map of quinoa (*Chenopodium quinoa*) based on AFLP, RAPD and SSR markers. *Theor Appl Genet* 109:1188-1195
- Maurer HR (2001) Bromelain: biochemistry, pharmacology and medical use. *Cell Mol Life Sci* 58:1234-1245
- Mccough SR, Kochert G, Yu ZH, Wang ZY, Khush GS, Coffman WR, Tanksley SD (1988) Molecular mapping of rice chromosomes. *Theor Appl Genet* 76:815-829
- McGinnis S, Madden TM (2004) BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res* 32: (Web Server issue) W20-W25
- McKusick VA, Ruddle FH (1987) A new discipline, a new name, a new journal. *Genomics* 1:1-2.
- Medina-Filho HP (1980) linkage of *Aps-1* , *Mi* and other markers on chromosome 6. *Tomato Genet Coop* 30:26-28

- Messmer MM, Keller M, Zanetti S, Keller B (1999) Genetic linkage map of a wheat X spelt cross. *Theor Appl Genet* 98:1163-1170
- Mester D, Ronin Y, Minkov D, Nevo E, Korol A (2003) Constructing Large-Scale Genetic Maps Using an Evolutionary Strategy Algorithm. *Genetics* 165:2269-2282
- Miklas PN, Johnson WC, Delorme R, Gepts P (2001) QTL conditioning physiological resistance and avoidance to white mold in dry bean. *Crop Sci* 41:309-315
- Ming R, Moore PH, Zee F, Abbey CA, Ma H, Paterson AH (2001) Construction and characterization of a papaya BAC library as a foundation for molecular dissection of a tree-fruit genome. *Theor Appl Genet* 102:892-899
- Morales M, Orjeda G, Nieto C, van Leeuwen H, Monfort A, Charpentier M, Caboche M, Arús P, Puigdomènech P, Aranda MA, Dogimont C, Bendahmane A, Garcia-Mas J (2005) A physical map covering the nsv locus that confers resistance to Melon necrotic spot virus in melon (*Cucumis melo* L.) *Theor Appl Genet* 111:914-922
- Morgan TH (1912) Complete linkage in the second chromosome of the male of *Drosophila*. *Science* 36:719-720
- Morgan TH (1928) *The theory of genes*. Yale University Press, New Haven
- Morton NE (1955) Sequential tests for the detection of linkage. *Am J Hum Genet* 7:277-318
- Morton NE (2005) Linkage disequilibrium maps and association mapping. *J Clin Invest* 115:1425-1430
- Moyle R, Fairbairn DJ, Ripi J, Crowe M, Botella JR (2005a) Developing pineapple fruit has a small transcriptome dominated by metallothionein. *J Exp Bot* 56:101-112
- Moyle RL, Crowe ML, Ripi-Koia J, Fairbairn DJ, Botella JR (2005b) PineappleDB: An online pineapple bioinformatics resource. *BMC Plant Biol.* 5:21 (<http://www.biomedcentral.com/1471-2229/5/21>)
- Mueller LA, Tanksley SD, Giovannoni JJ, van Eck J, Stack S, Choi D, Kim BD, Chen M, Cheng Z, Li C, Ling H, Xue Y, Seymour G, Bishop G, Bryan G, Sharma R, Khurana J, Tyagi A, Chattopadhyay D, Singh NK, Stiekema W, Lindhout P, Jesse T, Lankhorst RK, Bouzayen M, Shibata D, Tabata S, Granell A, Botella MA, Giuliano G, Frusciante L, Causse M, Zamir D (2005) The Tomato Sequencing Project, the first cornerstone of the International Solanaceae Project (SOL). *Comp Funct Genom* 6:153-158
- Mullis KB, Faloona FA, Scharf S, Saiki R, Hoen G, Erlich H (1986) Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* 51:263-273

- Murray AW, Szostak JW (1983) Construction of artificial chromosomes in yeast. *Nature* 305:189-193
- Murray JC, Buetow KH, Weber JL, Ludwigsen S, Scherpbier-Heddema T, Manion F, Quillen J, Sheffield VC, Sunden S, Duyk GM, Weissenbach J, Gyapay G, DIB C, Morissette J, Lathrop GM, Vignal A, White R, Matsunami N, Gerken S, Melis R, Albertsen H, Plaetke R, Odelberg S, Ward D, Dausset J, Cohen D, CANN H (1994) A comprehensive human linkage map with centimorgan density. *Science* 265:2049-2054
- Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA, Mobarry CM, Reinert KH, Remington KA, Anson EL, Bolanos RA, Chou HH, Jordan CM, Halpern AL, Lonardi S, Beasley EM, Brandon RC, Chen L, Dunn PJ, Lai Z, Liang Y, Nusskern DR, Zhan M, Zhang Q, Zheng X, Rubin GM, Adams MD, Venter JC. (2000) A whole-genome assembly of *Drosophila*. *Science* 287:2196-2204
- Nakasone HY, Paull RE (1998) Pineapple. In: *Tropical Fruits (Crop production science in horticulture, No 7)*. CAB International, Wallingford
- Nasu S, Suzuki J, Ohta R, Hasegawa K, Yui R, Kitazawa N, Monna L, Minobe Y (2002) Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (*Oryza sativa*, *Oryza rufipogon*) and establishment of SNP markers. *DNA Res* 9:163-171
- Negi MS, Devic M, Delseny M, Lakshmikumaram M (2000) Identification of AFLP fragments to seed colour in *Brassica juncea* and conversion to a SCAR marker for rapid selection. *Theor Appl Genet* 101:146-152
- Neuteboom LW, Kunimitsu WY, Webb D, Christopher DA (2002) Characterization and tissue-regulated expression of genes involved in pineapple (*Ananas comosus* L.) root development. *Plant Sci* 163:1021-1035
- NIH/CEPH Collaborative Mapping Group (1992) A Comprehensive Genetic Linkage Map of the Human Genome. *Science* 258:67-86
- Nilsson NO, Hallden C, Hansen M, Hjerdin A, Säll T (1997) Comparing the distribution of RAPD and RFLP markers in a high density linkage map of sugar beet. *Genome* 40:644-651
- Noyer JL (1991) Etude préliminaire de la diversité génétique du genre *Ananas* par les RFLPs. *Fruits* 46:372-375
- Noyer JL, Lanaud C, Duval MF, Coppens d'Eeckenbrugge G (1997) RFLP study on rDNA variability in *Ananas* genus. *Acta Hort* 425:153-160
- Oliver M, Garcia-Mas J, Cardús M, Pueyo N, López-Sesé AI, Arroyo M, Gómez-Paniagua H, Arús P, De Vicente MC (2001) Construction of a reference linkage map for melon. *Genome* 44:836-845

- Olson JM, Boehnke M (1990) Monte Carlo comparison of preliminary methods for ordering multiple genetic loci. *Am J Hum Genet* 47:470-82
- Olson JM, Hood L, Cantor C, Botstein D (1989) A common language for physical mapping of the human genome. *Science* 245:1434-1435
- Orita M, Suzuki Y, Sekiya T, Hayashi K (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* 5:874-879
- Ortiz-Vázquez E, Kaemmer D, Zhang H-B, Muth J, Rodríguez-Mendiola M, Arias-Castro C, Andrew James (2005) Construction and characterization of a plant transformation-competent BIBAC library of the black Sigatoka-resistant banana *Musa acuminata* cv. Tuu Gia (AA). *Theor Appl Genet* 110:706-713
- Osuji JO, Crouch J, Harrison G, Heslop-Harrison JS (1998) Molecular cytogenetics of *Musa* species, cultivars and hybrids: location of 18S-5.8S-25S and 5S rDNA and telomere-like sequences. *Ann Bot* 82:243–248
- Osuji JO, Harrison G, Crouch JH, Heslop-Harrison JS (1997) Identification of the genomic constitution of *Musa* L. lines (banana, plantains and hybrids) using molecular cytogenetics. *Ann Bot* 80:787-793
- Paran I, Michelmore RW (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theor Appl Genet* 85:985-993
- Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature* 335:721-726
- Peters JL, Constandt H, Neyt P, Cnops G, Zethof J, Zabeau M, Gerats T (2001) A Physical Amplified Fragment-Length Polymorphism Map of *Arabidopsis*. *Plant Physiol* 127:1579-1589
- Petty GJ, Tustin HA, Dicks HM (2005) 5th International Pineapple Symposium. South Africa. Oral presentation
- Pineapple Working Group (2004) Changes in the Pineapple Industry. In: Newsletter of the Pineapple Working Group, International Society for Horticultural Science, Issue No. 11:2-3
- Plomion C, Liu BH, O'Malley DM (1996) Genetic analysis using trans-dominant linked markers in an F₂ family. *Theor Appl Genet* 93:1083-1089
- Prior TW, Papp AC, Snyder PJ, Sedra MS, Western LM, Bartolo C, Moxley RT, Mendell JR. (1994) Heteroduplex analysis of the dystrophin gene: application to point mutation and carrier detection. *Am J Med Genet.* 50:68-73

- Qi X, Pittaway TS, Lindup S, Liu H, Waterman E, Padi FK, Hash CT, Zhu J, Gale MD, Devos KM (2004) An integrated genetic map and a new set of simple sequence repeat markers for pearl millet, *Pennisetum glaucum*. *Theor Appl Genet* 109: 1485-1493
- Qi X, Stam P, Lindhout P (1998) Use of locus-specific AFLP markers to construct a high-density molecular map in barley. *Theor Appl Genet* 96:376-384
- Rafalski JÁ, Vagel JM, Margante M, Powell W, (1996) Generating and using DNA markers in plants. In: Birren B, Lai (eds). *Nonmammalian genomic analysis: a practical guide*. Academic press Inc, San Diego
- Rajapakse S, Beltho. LE, He G, Estager AE, Scorza R, Verde I, Ballard RE, Baird WV, Callahan A, Monet R, Abbott AG (1995) Genetic mapping in peach using morphological, RFLP and RAPD markers. *Theor Appl Genet* 90:503–510
- Rebolledo-Martinez A, Rebolledo Martinez L, Uriza-Avila D, Del Ángel-Pérez A (2005) 5th International Pineapple Symposium. South Africa. Oral presentation
- Reinhardt DH, de Almeida CO, Vilar L da C (2005) Characteristics and status of the brazilian pineapple industry. 5th International Pineapple Symposium. South Africa. Oral presentation
- Reinisch A, Dong JM, Brubaker CL, Stelly DM, Wendel JF, Paterson AH (1994) A detailed RFLP map of cotton, *Gossypium hirsutum* × *Gossypium barbadense*: chromosome organization and evolution in a disomic polyploid genome. *Genetics* 138: 829–847
- Reiter RS, Williams JGK, Feldmann KA, Rafalski ATJ, Tingeyt SV, Scolnik PA (1992) Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. *Genetics* 89:1477-1481
- Riaz S, Dangl GS, Edwards KJ, Meredith CP (2004) A microsatellite marker based framework linkage map of *Vitis vinifera* L. *Theor Appl Genet* 108:864–872
- Rick CM, Fobes JF (1974) Association of an allozyme with nematode resistance. *Rep Tomato Genet Coop* 24:25
- Risch N (1992) Genetic linkage: Interpreting LOD scores. *Science* 255:803804
- Risterucci AM, Grivet L, N’Goran JAK, Pieretti I, Flament MH, Lanaud C (2000) A high-density linkage map of *Theobroma cacao* L. *Theor Appl Genet* 101:948-955
- Ritter E, Gebhardt C, Salamini F (1990) Estimation of recombination frequencies and construction of RFLP linkage maps in plants from crosses between heterozygous parents. *Genetics* 125:645-654
- Rohrbach KG, Christopher D, Hu J, Paull R, Sipes B, Nagai C, Moore P, McPherson M, Atkinson H, Levesley A, Oda C, Fleisch H, McLean M (2000) Management of a multiple goal pineapple genetic engineering program. *Acta Hort* 529:111-113

- Rohrbach KG, Leal F, Coppens d'Veckenbrugge G (2003) History, distribution and world production. In: Bartholomew, DP, Paull, RE and Rohrbach, KG (eds) *The Pineapple: Botany, Production and Uses*. CABI Publishing, Oxon
- Rohrbach KG, Schmitt D (2003) Diseases of pineapple. In: R.C. Ploetz (ed.): *Diseases of tropical fruit crops*. CAB International, Oxon
- Rohrer GA, Alexander LJ, Hu Z, Smith KPL, Keele JW, Beattie CW (1996) A comprehensive map of the porcine genome. *Genome Res* 6:371–391
- Roose ML, Fang D, Cheng FS, Tayyar RI, Federici CT, Kupper RS (2000) Mapping the citrus genome. In: *Proc 1st Int Citrus Biotechnol Symp*. Israel
- Roupe van der Voort JNAM, van Zandvoort P, Eck HJvan, Folkertsma RT, Hutten RCB, Draaistra J, Gommers FJ, Jacobsen E, Helder J, Bakker J (1997) Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. *Mol Gen Genet* 255:438-447
- Ruas CD, Ruas PM, Cabral JRS (2001) Assessment of genetic relatedness of the genera *Ananas* and *Pseudananas* confirmed by RAPD markers. *Euphytica* 119:245-252
- Ruiz C, Asins MJ (2003) Comparison between Poncirus and Citrus genetic linkage maps. *Theor Appl Genet* 106:826-836
- Saal B, Plieske QJ, Hu J, Quiros CF, Struss D (2001) Microsatellite markers for genome analysis in *Brassica* II. Assignment of rapessed microsatellites to the A and C genomes and genetic mapping in *Brassica oleracea* L. *Theor Appl Genet* 102:695-699
- Šafář J, Noa-Carranza JC, Vrána J, Bartoš J, Alkhimova O, Sabau X, Šimková H, Lheureux F, Caruana M-L, Dolezel J, Piffanelli P (2004) Creation of a BAC resource to study the structure and evolution of the banana (*Musa balbisiana*) genome. *Genome* 47:1182–1191
- Salava J, Krska B, Wang E, Polak J, Kominek P, Reighard GL, Miller W, Dowler W, Abbott AG (2002) Bulk segregant analysis and construction of a linkage map in apricot (*Prunus armeniaca* L.) crosses segregating for resistance to plum pox virus. *Plant, Animal and Microbe Genomes Xth Conference*, San Diego, California, January 12-16
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, New York
- Sanger F, Coulson AR, Hong GF, Hill DF, Petersen GB (1982) Nucleotide sequence of the bacteriophage lambda DNA. *J Mol Biol* 162:729-773
- Sargent DJ, Davis TM, Tobutt KR, Wilkinson MJ, Battey NH, Simpson DW (2004) A genetic linkage map of microsatellite, gene-specific and morphological markers in diploid *Fragaria*. *Theor Appl Genet* 109: 1385-1391

- Sarthou C, Boisselier-Dubayle MC, Lambourdiere J, Samadi S (2003) Polymorphic microsatellites for the study of fragmented populations of *Pitcairnia geyskesii* L. B. Smith (Bromeliaceae), a specific saxicolous species of inselbergs in French Guiana. *Mol Ecol Notes* 3:221-223
- Sato Y, Nishio T (2002) Efficient detection of DNA polymorphism in cabbage and rice cultivars by PCR-RF-SSCP (PRS). *Plant Cell Rep* 21:276-281
- Sato Y, Nishio T (2003) Mutation detection in rice waxy mutants by PCR_RF_SSCP. *Theor Appl Genet* 107:560-567
- Sawano Y, Muramatsu T, Hatano K, Nagata K, Tanokura M (2002) Characterization of genomic sequence coding for Bromelain inhibitors in Pineapple and expression of its recombinant isoform. *J Biol Chem* 277:28222-28227
- Schafer AJ, Hawkins JR (1998) DNA variation and the future of human genetics. *Nat Biotechnol* 16:33-39
- Schloss SJ, Mitchell SE, White GM, Kukatla R, Bowers JE, Paterson AH, Kresovich S (2002) Characterization of RFLP probe sequences for gene discovery and SSR development in *Sorghum bicolor*(L.) Moench. *Theor Appl Genet* 105:912-920
- Schmid KJ, Sørensen RT, Stracke R, Törjek O, Altmann T, Mitchell-Olds T, Weisshaar B (2003) Large-Scale identification and analysis of genome-wide single-nucleotide polymorphisms for mapping in *Arabidopsis thaliana*. *Genome Res* 13:1250-1257
- Schmidt R (2002) Plant genome evolution: lessons from comparative genomics at the DNA level. *Plant Mol Biol* 48:21-37
- Schwartz DC, Cantor CR (1984) Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell* 37:67-75
- Sebastian RL, Howell EC, King GJ, Marshall DF, Kearsey MJ (2000) An integrated AFLP and RFLP *Brassica oleracea* linkage map from two morphologically distinct doubled-haploid mapping populations. *Theor Appl Genet* 100:75–81
- Seglias NP, Gessler C (1997) Genetics of apple powdery mildew resistance from *Malus zumi* (PI2). *Integrated control of pome fruit diseases. IOBC/WPRS Bull* 20:195–208
- Sether DS, Perez E, Melzer MJ, Subere CV, Busto JJ, Hu JS (2005) 5th International Pineapple Symposium. South Africa. Oral presentation
- Sharma KA, Ghosh I (1971) Cytotaxonomy of the family Bromeliaceae. *Cytologia* 36:237-247
- Sharon D, Cregan PB, Mhameed S, Kusharska M, Hillel J, Lahav E, Lavi U (1997) An integrated genetic linkage map of avocado *Theor Appl Genet* 95:911-921

- Sherman JD, Stack SM (1995) Two-dimensional spreads of synaptonemal complexes from solanaceous plants. VI. High-resolution recombination nodule map for tomato (*Lycopersicon esculentum*). *Genetics* 141:683-708
- Shizuya H, Birren B, Kim UJ, Mancino V, Slepak T, Tachiiri Y, Simon M (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci USA* 89:8794-8797
- Shyu DJH, Chyan CL, Tzen JTC, Chou WM (2004) Molecular cloning, expression, and functional characterization of a cystatin from pineapple stem. *Biosci Biotechnol Biochem* 68:1681-1689
- Silberstein L, Kovalski I, Brotman Y, Perin C, Dogimont C, Pitrat M, Klingler J, Thompson G, Portnoy V, Katzir N, Perl-Treves R (2003) Linkage map of *Cucumis melo* including phenotypic traits and sequence-characterized genes. *Genome* 46:761-773
- Simmons G, Vos P, Groenendijk J, Wijbrandi J, Diergaard P, Groenen J, van der Lee T, Bleeker M, Reijans M, Onstenk J, de Both M, Zabeau M (1996) Isolation and characterisation of the I2 *Fusarium oxysporum* resistance locus from tomato. In: *Plant Genome, Proceedings of the fourth International Conference on the status of Plant Genome, Research*. San Diego, CA. pp. 14-18
- Simone M de, Russo MP, Puelo G, Marsan PA, Lorenzoni C, Marocco A, Recupero GR (1998) Construction of genetic maps for *Citrus aurantium* and *C. latipes* based on AFLP, RAPD and RFLP markers. *Fruits* 53:383-390
- Sipes B (2003) Nematode Control Research in Hawaii. In: *Newsletter of the Pineapple Working Group, International Society for Horticultural Science*, Issue No. 10:14-15
- Sipes BS (2005) 5th International Pineapple Symposium. South Africa. Oral presentation
- Siriphanich J, Chanjirakul K (2005) 5th International Pineapple Symposium. South Africa. Poster
- Smith LB, Downs RJ (1979) Bromelioideae (*Bromeliaceae*). Monograph 14, Pt 3, *Flora Neotropica*. NYBG, New York
- Smith M, Lien Ko, Karen Kretschmann, Paul Campbell, Tim Holton (2005) Genetic Engineering of Pineapple for Blackheart Resistance and Control of Natural Flowering. In: *Newsletter of the Pineapple Working Group, International Society for Horticultural Science*, Issue No. 12:10-12
- Sokal RR, Rohlf FJ (1981) *Biometry*, 2nd ed. WH Freeman, San Francisco

- Sondur SN, Manshardt RM, Stiles JI (1995) Genetics of growth rate and flowering time in papaya (*Carica papaya* L.). J Quant Trait Loci 1: article 4 [now J. Agric. Genomics]. <http://www.cabi-publishing.org/gateways/jag/papers95/paper495/indexp495.html>
- Sondur SN, Manshardt RM, Stiles JI (1996) A genetic linkage map of papaya based on randomly amplified polymorphic DNA markers. Theor Appl Genet 93:547–553
- Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, Cregan PB (2004) A new integrated genetic linkage map of the soybean. Theor Appl Genet 109:122–128
- Spielman R, Ewens W (1998) A sibship test for linkage in the presence of association: The sib transmission/disequilibrium test. Am J Hum Genet 62:450-458
- Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 52:506-516
- SPSS 13.0 (2004) <http://www.spss.com/spss/>
- Sripaoraya S (2005) Transgene stability expression and herbicide tolerance of genetically manipulated pineapple (*Ananas comosus*) under field conditions. 5th International Pineapple Symposium 2005. South Africa. Oral presentation
- Sripaoraya S, Marchant R, Power JB, Davey MR (2001) Herbicide-tolerant transgenic pineapple (*Ananas comosus*) produced by microprojectile bombardment. Ann Bot 88:597-603
- Sripaoraya S, Marchant R, Power JB, Davey MR (2003) Plant regeneration by somatic embryogenesis and organogenesis in commercial pineapple (*Ananas comosus* L.). In Vitro Cell Dev Biol Plant 39:450-454
- Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J 3:739-744
- Staub JE, Serquen FC (1996) Genetic markers, map construction, and their application in plant breeding. HortScience 31:729-741
- Stewart RJ, Sawyer BJB, Bucheli CS, Robinson SP (2001) Polyphenol oxidase is induced by chilling and wounding in pineapple. Aust J Plant Physiol 28: 181-191
- Stockinger EJ, Brettin TS, Iezzoni AF, Mulinix CA, Long CM (1996) A linkage map of sweet cherry based on RAPD analysis of a microspore-derived callus culture population. J Hered 87:214-221
- Sturtevant AH (1913) The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. J Exp Zool 14:43-59
- Tanksley SD (1983) Molecular markers in plant breeding. Plant Mol Biol Rep 1: 3-8

- Tanksley SD, Ganai MW, Martin GB (1995) Chromosome landing, a paradigm for map-based gene cloning in plants with large genomes. *Trends Genet* 11:63-68
- Tanksley SD, Ganai MW, Prince JW, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin O, Röder MS, Wing RA, Wu W, Young ND (1992) High density molecular linkage maps of the tomato and potato genomes. *Genetics* 132:1141-1160
- Tanksley SD, Grandillo S, Fulton TM, Zamir D, Eshed Y, Petiard V, Lopez J, Beckbunn T (1996) Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. *Theoretical and Applied Genetics* 92:213-224
- Tanksley SD, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding: new tools for an old science. *Biotechnology* 7:257-264
- Tantz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* 17:6463-6471
- Tao QZ, Zhang HB (1998) Cloning and stable maintenance of DNA fragments over 300 kb in *Escherichia coli* with conventional plasmid-based vectors. *Nucleic Acids Res* 26:4901-4909
- Tavares JFP, Batista JAB (2004) Cultura de ananás em estufa Ilha S. Miguel – Açores (Portugal). Edição: Profrutos - Cooperativa de Produtores de Frutas, Produtos Hortícolas e Florícolas de S. Miguel. Ponta Delgada
- Terwilliger J (1992) A haplotype-based “haplotype relative risk” approach to detecting allelic associations. *Hum Hered* 42:337-346
- Testolin R, Huang WG, Lain O, Messina R, Vecchione A, Cipriani G (2001) A kiwifruit (*Actinidia* spp.) linkage map based on microsatellites and integrated with AFLP markers. *Theor Appl Genet* 103:30-36
- The *Arabidopsis* Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796-815
- The *C. elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282:2012-2018
- The International HapMap Consortium (2003) The International HapMap Project. *Nature* 426:789-796
- Thompson EA (1987) Crossingover counts and likelihood in multipoint linkage analysis. *IMA J Math Appl Med Biol* 4:93-108
- Thomson KG, Thomas JE, Dietzgen RG (1998) Retrotransposon-like sequences integrated into the genome of pineapple, *Ananas comosus*. *Plant Mol Biol* 38:461-465
- Trusov Y, Firoozabady E, Gutterson N, Botella (2005) 5th International Pineapple Symposium. South Africa. Oral presentation

- Tulsieram LK, Glaubitz JC, Kiss G, Carlson JE (1992) Single tree genetic linkage mapping in conifers using haploid DNA from megagametophytes. *Biotechnology* 10:686–690
- Urwin PE, Aurora Levesley, Michael J. McPherson and Howard J. Atkinson (2000) Transgenic resistance to the nematode *Rotylenchulus reniformis* conferred by *Arabidopsis thaliana* plants expressing proteinase inhibitors. *Mol Breed* 6:257-264
- Uzunova M, Ecke W, Weissleder K, Röbbelen G (1995) Mapping the genome of rapeseed (*Brassica napus* L.) I. Construction of an RFLP linkage map and localization of QTLs for glucosinolate content. *Theor Appl Genet* 90:194-204
- Valárik M, Šimková H, Hřibová E, Šafář J, Doleželová M, Doležel J (2002) Isolation, characterization and chromosome localization of repetitive DNA sequences in bananas (*Musa* spp.). *Chromosome Res* 10:89–100
- Van Eck HJ, Rouppe van der Voort JNAM, Draaistra J, Van Zandvoort P, Van Enckevort E, Segers B, Peleman B, Jacobsen E, Helder J, Bakker J (1995) The inheritance and chromosomal localization of AFLP markers in a non-inbred potato offspring. *Mol Breed* 1:397-410
- Van Ooijen JW, Voorrips RE (2001) JoinMap version 3.0, software for the calculation of genetic linkage maps. Plant Research International, Wageningen
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, et al. (2001) The sequence of the human genome. *Science* 291:1304-1351
- Venter JC, Smith HO, Hood L (1996) A new strategy for genome sequencing. *Nature* 381:364-366
- Verde I, Lauria M, Dettori MT, Vendramin E, Balconi C, Micali S, Wang Y, Marrazzo MT, Cipriani G, Hartings H, Testolin R, Abbott AG, Motto M, Quarta R (2005) Microsatellite and AFLP markers in the *Prunus persica* [L. (Batsch)]· *P. ferganensis* BC1 linkage map: saturation and coverage improvement. *Theor Appl Genet* 111:1013–1021
- Vilanova S, Romero C, Abbott AG, Lacer G, Badenes ML (2003) An apricot (*Prunus armeniaca* L.) F2 progeny linkage map based on SSR and AFLP markers, mapping plum pox virus resistance and self-incompatibility traits. *Theor Appl Genet* 107:239–247

- Vilarinhos AD, Piffanelli P, Lagoda P, Thibivilliers S, Sabau X, Carreel F, D'Hont A (2003) Construction and characterization of a bacterial artificial chromosome library of banana (*Musa acuminata* Colla). *Theor Appl Genet* 106:1102–1106
- Vision TJ, Brown DG, Shmoys DB, Durrett RT, Tanksley ST (2000) Selective Mapping: A Strategy for Optimizing the Construction of High-Density Linkage Maps. *Genetics* 155:407-420
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* 23:4407-4414
- Vuylsteke M, Mank R, Antonise R, Bastiaans E, Senior ML, Stuber CW, Melchinger AE, L_bberstedt T, Xia XC, Stam P, Zabeau M, Kuiper M (1999) Two high-density AFLP linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. *Theor Appl Genet* 99:921-935
- Waghmare VN, Rong J, Rogers CJ, Pierce GJ, Wendel JF, Paterson AH (2005) Genetic mapping of a cross between *Gossypium hirsutum* (cotton) and the Hawaiian endemic, *Gossypium tomentosum*. *Theor Appl Genet* 111:665–676
- Wang D, Karle R, Brettin TS, Iezzoni AF (1998a) Genetic linkage map in sour cherry using RFLP markers. *Theor Appl Genet* 98:1217–1224
- Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen N, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lipshutz R, Chee M, Lander ES (1998b) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 280:1077-1082
- Wang YH, Thomas CE, Dean RA (1997) A genetic map of melon (*Cucumis melo* L.) based on amplified fragment polymorphism (AFLP) markers. *Theor Appl Genet* 95:791-798
- Wang YH, Thomas CE, Dean RA (2000) Genetic mapping of a fusarium wilt resistance gene (*Fom-2*) in melon (*Cucumis melo* L.) *Mol Breed* 6:379-389
- Waterston R, Sulston JE (1998) The Human Genome Project: reaching the finish line. *Science* 282:53-54
- Wagh R, Bonar N, Baird E, Thomas B, Graner A, Hayes P, Powell W (1997) Homology of AFLP products in three mapping populations of barley. *Mol Gen Genet* 255:311-21
- Weber JL, May PE, (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* 44:388-396

- Weber JL, Myers EW (1997) Human whole-genome shotgun sequencing. *Genome Res* 7:401-409
- Weissenbach J, Gyapay G, Dib C, Vignal A, Moressette J, Millasseau P, Vaysseix G, Lathrop M (1992) A second generation linkage map of the human genome based on Highly informative microsatellite loci. *Nature* 359:794-802
- Wijeratnam W, Hewajulige IGN, Perera RS (2005) 5th International Pineapple Symposium. South Africa. Oral presentation
- Williams JGK, Kubelik AR, Livak JK, Rafalski JÁ, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 18:6531-6535
- Wilson SR (1988) A major simplification in the preliminary ordering of linked loci. *Genet Epidemiol* 5:75-80
- Wu C, Sun S, Nimmakayala P, Santos FA, Meksem K, Springman R, Ding K, Lightfoot DA, Zhang H-B (2004) A BAC- and BIBAC-Based Physical Map of the Soybean Genome. *Genome Res* 14:319–326
- Wu RL, Han YF, Hu JJ, Fang JJ, Li L, Li ML, Zeng ZB (2000) An integrated genetic map of *Populus deltoids* based on amplified fragment length polymorphisms. *Theor Appl Genet* 100:1249–1256
- Yamamoto T, Kimura T, Shoda M, Imai T, Saito T, Sawamura Y, Kotobuki K, Hayashi T, Matsuta N (2002) Genetic linkage maps constructed by using an interspecific cross between Japanese and European pears. *Theor Appl Genet* 106:9-18
- Yin T, Zhang X, Huang M, Wang M, Zhuge Q, Tu S, Zhu L-H, Wu R (2002) Molecular linkage maps of the *Populus* genome. *Genome* 45: 541–555
- Yin TM, DiFazio SP, Gunter LE, Riemenschneider D, Tuskan GA (2004) Large-scale heterospecific segregation distortion in *Populus* revealed by a dense genetic map *Theor Appl Genet* 109:451–463
- Yin TM, Huang MR, Wang Mx, Zhu LH, Zeng ZB, Wu RL (2001) Preliminary interspecific genetic maps of the *Populus* genome constructed from RAPD markers. *Genome* 44:602–609
- Young ND (1999) A cautiously optimistic vision for marker-assisted breeding. *Mol Breed* 5:505-510
- Young WP, Schupp JM, Keim P (1999) DNA methylation and AFLP marker distribution in the soybean genome. *Theor Appl Genet* 99:785-790
- Yu JK, La Rota M, Kantety RV, Sorrells ME (2004) EST derived SSR markers for comparative mapping in wheat and rice. *Mol Genet Genomics* 271:742-751

- Zhang D, Zhang Z, Yang K, Li B (2004) Genetic mapping in (*Populus tomentosa* x *Populus bolleana*) and *P. tomentosa* Carr. using AFLP markers. *Theor Appl Genet* 108:657–662
- Zhang H-B, Wing RA (1997) Physical mapping of the rice genome with BACs. *Plant Mol Biol* 35:115-127
- Zhang HB, Wu C (2001) BAC as tools for genome sequencing. *Plant Physiol Biochem* 39:195-209
- Zhang J, Guo W, Zhang T (2002) Molecular linkage map of allotetraploid cotton (*Gossypium hirsutum* L. x *Gossypium barbadense* L.) with a haploid population. *Theor Appl Genet* 105:1166–1174
- Zietkiewicz E, Rafaski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* 20:176-183
- Zou J, Chen Z, Zhang S, Zhang W, Jiang G, Zhao X, Zhai W, Pan X, Zhu L (2005) Characterizations and fine mapping of a mutant gene for high tillering and dwarf in rice (*Oryza sativa* L.) *Planta* 222:604-612