

WILLIAN SILVA BARROS

**GENOTIPAGEM SELETIVA E OUTRAS ESTRATÉGIAS DE
AMOSTRAGEM NO MAPEAMENTO GENÉTICO E NA
DETECÇÃO DE QTL EM POPULAÇÕES F₂ SIMULADAS**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS - BRASIL
2007

**Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV**

T

B277g
2007

Barros, Willian Silva, 1978-

Genotipagem seletiva e outras estratégias de amostragem no mapeamento genético e na detecção de QTL em populações F₂ simuladas / William Silva Barros. – Viçosa, MG, 2007.

xiv, 158f. : il. ; 29cm.

Orientador: Cosme Damiano Cruz.

Tese (doutorado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. Genética molecular - Métodos de simulação.
2. Genética quantitativa. 3. Locos de caracteres quantitativos. 4. Amostragem (Estatística). I. Universidade Federal de Viçosa. II. Título.

CDD 22.ed. 572.8

WILLIAN SILVA BARROS

**GENOTIPAGEM SELETIVA E OUTRAS ESTRATÉGIAS DE
AMOSTRAGEM NO MAPEAMENTO GENÉTICO E NA
DETECÇÃO DE QTL EM POPULAÇÕES F₂ SIMULADAS**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.

APROVADA: 28 de março de 2007.

Prof. José Marcelo Soriano Viana
(Co-orientador)

Prof. Pedro Crescêncio Souza Carneiro
(Co-orientador)

Prof. Fabyano Fonseca e Silva

Dra. Eveline Teixeira Caixeta

Prof. Cosme Damião Cruz
(Orientador)

AGRADECIMENTOS

A Deus, pelo dom da vida, da persistência e da vitória.

À Giselda Maria Pereira, pela tolerância, pela paciência e pelo incentivo.

À Universidade Federal de Viçosa, em especial ao Programa de Genética e Melhoramento pela oportunidade de realizar e concluir este curso.

Ao meu orientador, prof. Cosme Damião Cruz, um exemplo a ser seguido, pela paciência, pela dedicação e amizade sempre presentes.

Ao co-orientador prof. José Marcelo Soriano Viana, pelos conselhos e amizade.

Ao co-orientador e amigo, prof. Pedro Crescêncio Souza Carneiro, por ter-me estimulado a ingressar no Programa de Genética e Melhoramento.

Aos professores Marcos Ribeiro Furtado, Pedro Crescêncio Souza Carneiro, Derly José Henriques da Silva, Adair José Regazzi, Jorge Abdala Dergam dos Santos, Lúcio Antônio de Oliveira Campos e Cosme Damião Cruz, pelos ensinamentos em suas disciplinas.

A todos os amigos e colegas do Programa de Genética e Melhoramento e do laboratório de Bioinformática, pelo convívio agradável durante a realização deste curso.

Aos amigos da República dos Desqualificados, Fábio Medeiros Ferreira Fabiano Ricardo Brunele Caliman e Renato Brandão de Oliveira pelo agradável convívio.

Aos professores Fabyando Fonseca e Silva, Antônio Policarpo Souza

Carneiro e Marcos Ribeiro Furtado, pela amizade.

Aos amigos Fábio Medeiros Ferreira e Rodrigo Barros Rocha, pela convivência, pelas palavras de incentivo, pelo companheirismo e pelos obstáculos vencidos juntos.

Às funcionárias Rita de Cássia Rosado Cruz, Rosemary Inácio Tomás e Maria Salvadora Saraiva Torres, pela amizade, pela dedicação e apoio.

Ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), pela concessão da bolsa de estudos.

À Pró-Reitoria de Ensino, pela concessão da bolsa de monitor nível II no Departamento de Informática - UFV.

BIOGRAFIA

WILLIAN SILVA BARROS, filho de Vera Lúcia Silva Barros e Wilson Júlio de Barros, nasceu em 26 de outubro de 1978, em Belo Horizonte - Minas Gerais.

Em 1996, formou-se Técnico em Contabilidade pela Escola Estadual Padre José Maria de Man, em Coronel Fabriciano, MG. Foi homenageado com diploma de HONRA AO MÉRITO concedido pelo Conselho Regional de Contabilidade de Minas Gerais (CRCMG).

Em 1997, formou-se em Técnico em Eletrônica pelo Instituto Municipal de Educação Técnica de Timóteo (IMETT), Timóteo, MG.

Em março de 1998, ingressou no Curso de Agronomia, na Universidade Federal de Viçosa (UFV), Viçosa - MG, concluindo-o em setembro de 2002. Recebeu do Centro de Ciências Agrárias (CCA) da UFRV o certificado de VOTOS DE LOUVOR, pelo destaque entre seus pares. Nesse período, desenvolveu atividades de monitoria nas disciplinas Genética Básica, no Departamento de Biologia Geral - UFRV e Hidráulica, Irrigação e Drenagem, no Departamento de Engenharia Agrícola - UFRV. Além disso, desenvolveu atividade de pesquisa como bolsista de Iniciação Científica na área de Genética da Conservação da poaia (*Psychotria ipecacuanha*), no Departamento de Bioquímica e Biologia Molecular - UFRV.

Em setembro de 2002, iniciou o curso de Mestrado em Genética e Melhoramento na UFRV, vindo a concluí-lo em março de 2004. Desenvolveu atividade de pesquisa na área de Genética Molecular, com ênfase em Filogeografia Molecular de poaia, na UFRV.

Iniciou o curso de Doutorado em Genética e Melhoramento, na UFRV, em

março de 2004. Atuou como representante dos estudantes do programa de Genética e Melhoramento, coordenador financeiro da associação dos pós-graduandos (APG) e representou os estudantes de pós-graduação no conselho Universitário (CONSU), durante o período de agosto 2004 a julho 2005. No período de março a julho de 2005, foi monitor II em Iniciação à Estatística e Estatística I, no Departamento de Informática - UFV. Durante o doutorado, desenvolveu atividade de pesquisa na área de Estatística Genômica com ênfase em Biometria, Simulação e Genética molecular. No dia 28 de março de 2007, submeteu-se aos exames finais de defesa de tese.

Em dezembro de 2006, obteve aprovação do CNPq para uma bolsa de pós-doutorado júnior (PDJ) junto ao departamento de Biologia Geral - UFV, sob a supervisão do professor Cosme Damião Cruz, com início de trabalho em abril de 2007.

À minha filha Luana Fonseca Barros

À minha amada Giselda Maria Pereira

Aos meus pais, Vera Lúcia Silva Barros e Wilson Júlio de Barros

À minha querida avó Odete Torres da Silva (*in memoriam*)

Aos meus irmãos, Wallace Silva Barros e Vanessa Silva Barros

À minha tia e madrinha Laecy Silva Souza

Ao meu tio e padrinho Olindo Souza (*in memoriam*)

À minha tia Maria das Graças Torres Miranda

À toda minha família

À minha grande amiga Lucina Soares Moraes e toda sua família

Dedico

SUMÁRIO

RESUMO	xi
ABSTRACT	xii
1. INTRODUÇÃO GERAL	1
2. REVISÃO DE LITERATURA	4
2.1. Mapeamento genético	4
2.1.1. Marcadores genéticos	4
2.1.2. Ligação fatorial e recombinação	5
2.1.3. Populações segregantes para mapeamento genético	11
2.2. Detecção de QTL	13
2.2.1. Caráter quantitativo	13
2.2.2. Mapeamento por marcas simples	15
2.2.3. Mapeamento por intervalo simples	16
2.3. Genotipagem seletiva	19
2.4. Efeito da genotipagem seletiva sobre mapeamento	21
3. REFERÊNCIAS BIBLIOGRÁFICAS	23
Capítulo I: Mapeamento genético e detecção de QTL em populações F₂	
simuladas	33
1. INTRODUÇÃO	34
2. OBJETIVOS	36
3. MATERIAL E MÉTODOS	37
3.1. Simulação de dados	37

3.1.1. Simulação do genoma, níveis de saturação e tipos de marcas	37
3.1.2. Simulação de genitores	39
3.1.3. Procedimento de simulação dos indivíduos da população	40
3.1.4. Simulação das populações e da característica quantitativa	41
3.2. Análises genômicas - mapeamento genético	42
3.2.1. Análise de segregação de locos individuais - teste de qui- quadrado	42
3.2.2. Análise de pares de marcas - estimação da porcentagem de recombinação	42
3.2.3. Análise de pares de marcas - determinação dos grupos de ligação e ordenamento das marcas	43
3.2.4. Comparação de genomas	43
3.3. Análises genômicas - detecção de QTL	44
3.3.1. Método de mapeamento de QTL por marca simples	44
3.3.2. Método de mapeamento de QTL por intervalo simples	47
3.3.3. Obtenção do limite crítico	48
4. RESULTADOS E DISCUSSÃO	50
4.1. Mapeamento genético em populações F_2	50
4.1.1. Análise de segregação de locos individuais	51
4.1.2. Número de grupos de ligação e ordenamento das marcas	51
4.1.3. Construção do mapa de ligação em populações F_2	52
4.1.4. Comparação dos genomas	54
4.2. Detecção de QTL em populações F_2	57
4.2.1. Detecção da presença do QTL	59
4.2.2. Detecção da presença do QTL no grupo de ligação	63

4.2.3. Detecção do QTL pelo método de marca simples e sua contribuição para a variação da característica	65
4.2.4. Detecção e localização do QTL pelo método de intervalo simples e sua contribuição para a variação da característica	69
4.2.5. Estimação dos efeitos dos marcadores sobre a característica	72
5. CONCLUSÕES	77
6. REFERÊNCIAS BIBLIOGRÁFICAS	78
Capítulo II: Genotipagem seletiva em populações F ₂ simuladas para fins de mapeamento genético e detecção de QTL	86
1. INTRODUÇÃO	87
2. OBJETIVOS	89
3. MATERIAL E MÉTODOS	90
3.1. Populações amostradas	90
3.2. Genotipagem seletiva	91
3.3. Comparação das estratégias experimentais de amostragem	91
3.4. Análises genômicas - mapeamento genético em populações amostradas	92
3.5. Análises genômicas - detecção e mapeamento de QTL em populações amostradas	92
4. RESULTADOS E DISCUSSÃO	93
4.1. Mapeamento genético em populações estabelecidas por genotipagem seletiva	93
4.1.1. Análise de segregação de locos individuais	93
4.1.2. Número de grupos de ligação e ordenamento das marcas	96
4.1.3. Construção do mapa de ligação	96

4.1.4. Comparação dos genomas	99
4.2. Detecção de QTL em populações amostradas por genotipagem seletiva	100
4.2.1. Detecção do QTL pelo método de marca simples e sua contribuição para a variação da característica	102
4.2.2. Detecção e localização do QTL pelo método de intervalo simples e sua contribuição para a variação da característica	109
4.2.3. Estimação do efeito do QTL sobre a característica em populações amostradas	116
4.3. Outras estratégias de amostragem	126
4.3.1 Mapeamento genético em populações obtidas a partir de diferentes estratégias de amostragem	126
4.3.1.1. Análise de segregação de locos individuais	127
4.3.1.2. Número de grupos de ligação e ordenamento das marcas ..	127
4.3.1.3. Construção do mapa de ligação	129
4.3.1.4. Comparação dos genomas	129
4.3.2. Detecção de QTL em populações amostradas por meio de diferentes estratégias de amostragem	131
4.3.2.1. Detecção do QTL pelo método de marca simples e sua contribuição para a variação da característica	131
4.3.2.2. Detecção e localização QTL pelo método de intervalo simples e sua contribuição para a variação da característica	138
4.3.2.3. Estimação do efeito do QTL sobre a característica	14
5. CONCLUSÕES	153
6. REFERÊNCIAS BIBLIOGRÁFICAS	155

RESUMO

BARROS, Willian Silva, DSc., Universidade Federal de Viçosa, março de 2007.
Genotipagem seletiva e outras estratégias de amostragem no mapeamento genético e na detecção de QTL em populações F₂ simuladas. Orientador: Cosme Damião Cruz. Co-Orientadores: José Marcelo Soriano Viana e Pedro Crescêncio Souza Carneiro.

A análise genética ao nível de DNA está se tornando cada vez mais eficiente. Grande quantidade de informações sobre genomas de plantas cultivadas está sendo gerada. No entanto, o principal desafio hoje refere-se à utilização deste conhecimento de forma integrada na prática efetiva do melhoramento genético de plantas. O mapeamento genético e de QTL contém um número considerável de estudos de natureza analítica. Porém, genotipagem de grande número de indivíduos para vários marcadores é uma limitação, mesmo com a revolução causada pela introdução de técnicas baseadas em PCR “Polymerase Chain Reaction”. Dessa forma, são necessários estudos que envolvem estratégias experimentais como a genotipagem seletiva, que permitem reduzir o número de indivíduos genotipados, mantendo-se o poder de detecção do QTL. A simulação em computador tem sido utilizada para obtenção de propriedades e verificação do desempenho de modelos, métodos, tamanho e tipos de população no mapeamento genético e detecção de QTL. Portanto, o objetivo do trabalho foi investigar, por meio de simulação, tomando como referência populações F₂ com 1.000 indivíduos em 36 cenários para características fenotípicas, a influência dos níveis de herdabilidade, da ação gênica e da saturação dos grupos de ligação no mapeamento genético e detecção do QTL. Outro objetivo foi investigar a influência da intensidade de seleção em amostras obtidas por genotipagem seletiva das populações F₂ simuladas no mapeamento genético e

detecção do QTL. Para intensidade de seleção de 20%, os resultados obtidos pela genotipagem seletiva foram comparados com outras estratégias experimentais de amostragem sistemática, mista e aleatória. Verificou-se no primeiro capítulo que o mapeamento genético em todas as 3.600 populações F_2 foi de acordo com o genoma simulado. Em relação à detecção de QTL, pode-se observar que níveis de herdabilidade de 20% e níveis de ação gênica de 5 e 10%, independente da saturação do grupo de ligação, são cenários que não permitem detectar facilmente o QTL. No segundo capítulo, quando a intensidade de seleção da genotipagem seletiva foi de 10%, em muitos cenários, o mapa de ligação obtido foi diferente do esperado. Nas outras intensidades de seleção foi possível a recuperação do mapa de ligação como esperado, mesmo com distorções das marcas próxima ao QTL simulado. Quanto à detecção do QTL, os valores médios dos “LOD scores” nas populações de 500, 300 e 200 indivíduos amostrados por genotipagem seletiva foram ligeiramente menores que os encontrados nas populações com 1.000 indivíduos. Os resultados obtidos das análises genômica em estratégias experimentais de amostragem mista e genotipagem seletiva foram satisfatórios. Eles mantiveram, em média mapas, de ligação próximos ao esperado, com alto poder de detecção de QTL em comparação com resultados obtidos em amostragem aleatória e sistemática.

ABSTRACT

BARROS, Willian Silva, DSc., Universidade Federal de Viçosa, March of 2007.
Selective genotyping and other sampling strategies for the genetic mapping and QTL detection in F₂ populations simulated. Adviser: Cosme Damião Cruz.
Co-Advisers: José Marcelo Soriano Viana and Pedro Crescêncio Souza Carneiro.

The genetic analysis at DNA level has becoming more and more efficient. Much information has been generated on the genomes of plants. Though, the main challenge refers to the integrally application of this knowledge on the effective practice of the genetic plant breeding. A considerable number of analytical studies on the mapping of the genes and QTL have been conducted. However, genotyping a great number of individuals for several markers is a limitation, in spite of the revolution caused by the introduction of techniques based on Polymerase Chain Reaction (PCR). So, there is a need for studies involving experimental strategies such as the selective genotyping, that would allow to reduce the number of genotyped individuals, but keeping the detection power of the QTL. The computer simulation has been used for either obtainment of the properties and verification of the performance by the models, methods, size and population types in the genetic mapping and QTL detection. So, a study was carried out to investigate the influence of the heritability levels, gene action, and the linkage group saturations in the genetic mapping and QTL detection. By using simulation, some F₂ populations with 1,000 individuals in 36 sceneries were analyzed for phenotypic traits. Another objective was to investigate the influence of the selection intensity in samples obtained by selective genotyping of the F₂ populations simulated in the genetic mapping and QTL

detection. For 20% selection intensity, the results from the selective genotyping were compared to other experimental strategies for the samplings of the systematic, mixed and random types. In the first chapter, it was found that the genetic mapping was according to the simulated genome in all 3,600 F_2 populations. Concerning to the QTL detection, it is observed that 20% heritability levels and the gene levels of 5 and 10% are sceneries that do not allow for an easy detection of QTL, independent from the saturation of the linkage group. In the second chapter, the linkage map differed from the expected one, when the selection intensity of the selective genotyping was 10% in many sceneries. At the other selection intensities, the recovery of the linkage map was possible as expected, even with distorted marks near to the simulated QTL. Concerning to QTL detection, the average values of the "LOD scores" in the populations with 500, 300 and 200 individuals sampled for selective genotyping were slightly lower than those found in the populations with 1,000 individuals. The results obtained from the genomic analyses in experimental strategies of the mixed sampling and selective genotyping were satisfactory. On average, they maintained linkage maps close to the expected, with high power to detecting QTL in comparison to the results obtained in both random and systematic samplings.

1. INTRODUÇÃO GERAL

Nos primórdios da agricultura, quando iniciou-se a “domesticação” das espécies, utilizava-se a seleção artificial, baseada no fenótipo, conscientemente, ou não. Os resultados desses esforços primitivos contribuíram, decisivamente, para o processo evolucionário das espécies cultivadas. Com a descoberta da reprodução sexual no reino vegetal, a hibridação de indivíduos foi incorporada às técnicas de melhoramento. Com o experimento clássico desenvolvido por Gregor Johann Mendel, foram obtidas as bases para o entendimento e manipulação da hereditariedade, possibilitando o melhoramento e o desenvolvimento de novas variedades (MENDEL, 1866).

Em 1902, George Udny Yule propôs que a variação contínua de caracteres seria explicável em função do efeito acumulado de vários genes, segregando segundo as leis mendelianas (TANKSLEY, 1993). Essa teoria poligênica da variação contínua permitia explicar a distribuição gaussiana de muitos caracteres. Em 1918, Sir Ronald Aylmer Fisher generalizou o modelo poligênico de variação contínua, incorporando-lhe as frequências dos genes e extraindo todas as conseqüências, em termos de parentescos genotípicos e respostas fenotípicas à seleção (FISHER, 1941). Complementando a pressuposição de que características quantitativas são influenciadas por muitos genes de pequeno efeito, LANDE (1981) sugeriu que poucos genes poderiam explicar uma proporção, relativamente, grande da variação genética para

características quantitativas. Estes locos são conhecidos como QTL “Quantitative trait loci” ou locos controladores de características quantitativas.

Com o desenvolvimento da biologia molecular, surgiram ferramentas novas que possibilitaram a construção de mapas genéticos. Os marcadores de DNA têm permitido a construção de mapas genéticos para várias espécies vegetais e animais. Tais mapas podem atingir alto grau de saturação devido à disponibilidade de grande número de marcas genéticas, que têm a vantagem de não ser influenciadas pelo ambiente, além de serem altamente polimórficas.

No estudo genético de determinada característica, tem interesse em conhecer o número de genes e alelos envolvidos no controle da sua expressão, a localização e posição relativa desses genes nos cromossomos, assim com a sua relação com outros genes. Neste contexto, os mapas genéticos são de fundamental importância, uma vez que permitem a visualização, mesmo que de forma relativa, da organização dos genes nos cromossomos.

No mapeamento genético podem ser usados, dependendo da espécie de interesse, vários tipos de populações segregantes, tais como populações F_2 , retrocruzamentos, F_1 “pseudo-testcross”, duplo-haplóides e linhas endogâmicas recombinantes “RILs - Recombinat Inbred Lines”.

Na construção de mapas genéticos e detecção de QTL, são usados modelos estatísticos para descrever sistemas genéticos e biológicos reais. No entanto, estes sistemas são complexos, impossibilitando a inclusão de todas as variáveis nos modelos utilizados. Portanto, existem diversos modelos estatísticos para estudos de populações segregantes e para o mapeamento de QTL. Alguns destes são muito complexos, enquanto outros são bastante simples.

O conhecimento das propriedades dos estimadores dos parâmetros desses modelos é de fundamental importância. Essas propriedades podem ser obtidas parametricamente, quando a distribuição do amostral do estimador é conhecida e bem caracterizada. Contudo, na maioria dos

modelos de mapeamento e detecção de QTL, é difícil obter, parametricamente, as propriedades do estimador. Portanto, a simulação em computadores tem sido utilizada para obtenção de propriedades e verificação do desempenho de modelos, métodos, tamanhos e tipo de população. Não é possível examinar o desempenho de um modelo por meio de dados experimentais, devido ao desconhecimento dos verdadeiros valores dos parâmetros. No entanto, na simulação de dados via computador, os parâmetros são conhecidos e podem ser usados para comparação da eficiência dos diversos modelos.

Este trabalho foi conduzido para avaliar a eficácia da genotipagem seletiva, via simulação, sobre a detecção de QTL e suas consequências no mapeamento de marcadores moleculares. Como a detecção de QTL depende de vários fatores, como a herdabilidade da característica, o controle da ação gênica e o tamanho e saturação do genoma, avaliou-se a influência destes fatores no mapeamento e detecção do QTL em F_2 simuladas sem restrição do tamanho.

2. REVISÃO DE LITERATURA

2.1. Mapeamento genético

2.1.1. Marcadores genéticos

O termo marcador tem sido utilizado para designação dos fatores morfológicos, fisiológicos, bioquímicos ou genéticos, passíveis de serem identificados e que permitam o estudo comparativo de genótipos e de suas progênes (SAKIYAMA, 1993). O marcador genético corresponde a uma característica do organismo, que pode ser facilmente detectada, visualmente, ou com auxílio de algum instrumento tecnológico e que co-segrega mendelianamente com genes de interesse (BORÉM & CAIXETA, 2006). Para ser útil como marcador genético, a característica deve apresentar polimorfismos entre os indivíduos analisados (LIU, 1998).

Até meados da década de 60, em estudos de genética e melhoramento, eram utilizados basicamente marcadores morfológicos, em geral fenótipos de fácil identificação visual, como cor do hipocótilo, de pétalas, de brotações, de sementes e morfologia foliar, entre outros. Esses marcadores contribuíram bastante para o desenvolvimento teórico da análise de ligação gênica e para a construção das primeiras versões de mapas genéticos (BORÉM & CAIXEITA, 2006).

Apesar de sua extrema importância na elucidação de uma série de eventos genéticos, os marcadores morfológicos e citológicos apresentam desvantagem, pois, ocorrerem em número

reduzido, o que limita sua aplicação em análise genética e conseqüentemente, a probabilidade de encontrar associações entre esses marcadores e características de interesse (FERREIRA & GRATTAPAGLIA, 1998). Esta limitação foi resolvida a partir do desenvolvimento de outra classe de marcadores, denominados como marcadores moleculares, que incluem os protéicos e os de DNA. A utilização de marcadores moleculares no estudo teórico e aplicado a espécies vegetais vem sendo cada vez maior. Estes marcadores possuem várias características desejáveis, como herança mendeliana simples, ausência de efeitos epistáticos, abundância e alto polimorfismo (FERREIRA & GRATTAPAGLIA, 1998; BORÉM & CAIXETA, 2006).

Atualmente, diversos tipos de marcadores moleculares estão disponíveis e são facilmente conhecidos e identificados por meio de suas siglas. Geneticamente, pode-se separá-los em dois grupos principais: (a) marcadores moleculares loco-específicos codominantes, como o RFLP - “Restriction Fragment Length Polymorphism” (GRODZICKER et al., 1974; BOTSTEIN et al., 1980), os SSR - “Simple Sequence Repeats” (LITT & LUTY, 1989; SCHLOTTERER, 2000) e marcadores bioquímicos isoenzimáticos; e (b) marcadores loco não específicos dominantes, como o RAPD - “Random Amplified Polimorphic DNA” (WELSH & MCCLELLAND, 1990; WILLIAMS et al., 1990) e o AFLP - “Amplified Fragment Length Polymorphism” (VOS et al., 1995).

Com desenvolvimento de variedade e quantidade de marcadores moleculares é possível obter mapas genéticos, cada vez mais saturados, sendo que vários já foram construídos para plantas e animais (BENTO, 2006).

2.1.2. Ligação fatorial e recombinação

Ligação fatorial é definida como a associação entre genes localizados num mesmo cromossomo, sendo que, para estes genes, a segunda Lei de Mendel (Lei da Segregação Independente) não se aplica, exceto quando eles estão separados por uma grande distância.

Portanto, genes ligados tendem a ser herdados em conjunto (CRUZ, et al., 2001; VIANA et al., 2003). Assim, a ligação fatorial possibilita a construção de mapas de ligação. A base genética para a construção de um mapa de ligação é a recombinação genética, resultante do “crossing over” entre cromossomos homólogos durante a meiose. A recombinação é medida pela fração de recombinação, que é a razão de gametas recombinantes pelo total de gametas (LIU, 1998).

Na construção de um mapa de ligação, geralmente, envolve a obtenção de uma população segregante a partir de duas linhagens progenitoras, a identificação de vários marcadores no genoma e a utilização de diversas metodologias de análises estatística e computacional, para estimar a ligação e distância entre marcadores (LYNCH & WALSH , 1998; SCHUSTER & CRUZ, 2004).

Para gerar um mapa de ligação, é necessário avaliar o padrão de segregação dos alelos dos marcadores, individualmente, em uma população, detectar o desequilíbrio de ligação entre marcadores, juntamente com a determinação da distância entre locos e o ordenamento dos mesmos em grupos de forma linear (LIU, 1998). Desequilíbrio de ligação é qualquer desvio das frequências alélicas dos locos em relação às frequências esperadas, sob a hipótese destes não estarem ligados e segregando independentemente, indicando ligação entre os locos (FALCONER & MACKAY, 1996).

A distância entre locos pode ser expressa em porcentagem de recombinação ou traduzida em centiMorgans (cM), utilizando-se diferentes funções de mapeamento (SCHUSTER & CRUZ, 2004). Estas funções são utilizadas para a correção de distâncias calculadas em porcentagem de recombinação para distâncias calculadas em cM, levando-se em conta, ou não, o fenômeno da interferência, que ocorre quando um “crossing over” em uma região interfere na ocorrência de mesmo fenômeno em uma região adjacente (LIU, 1998).

Um centiMorgan equivale a 1% de recombinação, aproximadamente, quando os marcadores estão bem próximos, mas pode diferir consideravelmente da porcentagem de

recombinação, quando os marcadores estão mais distantes. Isto ocorre devido ao aumento da probabilidade de “crossing over” duplos e triplos, que podem influenciar a estimativa da proporção de genótipos recombinantes em relação a genótipos parentais (FERREIRA & GRATTAPAGLIA, 1998).

A relação universal entre distância de mapeamento (cM) e tamanho físico do fragmento de DNA (em pares de bases) não existe, pois um cM pode variar entre 10.000 e 1.000.000 pares de bases, dependendo da espécie, podendo também existir grandes diferenças entre segmentos de um mesmo cromossomo (LYNCH & WALSH, 1998).

No mapeamento genético, é necessário estimar a probabilidade de recombinação entre cada par de marcas. Por exemplo, a geração F_2 , pode ser avaliada a partir das informações considerando-se, para cada par de marcas, marcadores apenas codominantes, dominantes ou marcadores codominantes e dominantes, simultaneamente.

No processamento de dados, para estimação da frequência de recombinação, é fundamental a codificação adequada das classes fenotípicas, de modo a possibilitar a associação do número de ocorrências de indivíduos na população e a frequência esperada em cada situação, estabelecida pelo tipo de marcador e pela fase de ligação. Para a população F_2 derivada do cruzamento entre indivíduos A_1A_1 e A_2A_2 (SCHUSTER & CRUZ, 2004, CRUZ & SILVA, 2006), pode ser utilizada a escala, apresentada a seguir.

Código	Genótipo	Referência
0	A ₁ A ₁	Homozigoto para P1 ¹
1	A ₁ A ₂	Heterozigoto
2	A ₂ A ₂	Homozigoto para P2 ²
3	A ₂ —	Não homozigoto para P1
4	A ₁ —	Não homozigoto para P2
9	— —	Valor perdido

¹\ P1 corresponde ao parental 1

²\ P2 corresponde ao parental 2

Considerando-se uma população F₂, derivada de F₁ em fase de aproximação, quanto marcadores codominantes, as frequências esperadas dos nove genótipos, observados na população, são apresentadas na Tabela 1, com a codificação numérica de cada indivíduo, para os dois locos.

Tabela 1 - Frequências genotípicas esperadas na população F₂, utilizando marcadores codominantes, derivada do cruzamento entre indivíduos com genótipo A₁A₁B₁B₁ x A₂A₂B₂B₂

Genótipo	Código	Número observado	Frequência esperada	P _i (R G)
A ₁ A ₁ B ₁ B ₁	0-0	n ₁	$p_1 = \frac{1}{4} (1-r)^2$	0,0
A ₁ A ₁ B ₁ B ₂	0-1	n ₂	$p_2 = \frac{1}{2} r (1-r)$	0,5
A ₁ A ₁ B ₂ B ₂	0-2	n ₃	$p_3 = \frac{1}{4} r^2$	1,0
A ₁ A ₂ B ₁ B ₁	1-0	n ₄	$p_4 = \frac{1}{2} r (1-r)$	0,5
A ₁ A ₂ B ₁ B ₂	1-1	n ₅	$p_5 = \frac{1}{2} (1-r)^2 + \frac{1}{2} r^2$	$r^2/[(1-r)^2+r^2]$
A ₁ A ₂ B ₂ B ₂	1-2	n ₆	$p_6 = \frac{1}{2} r(1-r)$	0,5
A ₂ A ₂ B ₁ B ₁	2-0	n ₇	$p_7 = \frac{1}{4} r^2$	1,0
A ₂ A ₂ B ₁ B ₂	2-1	n ₈	$p_8 = \frac{1}{2} r (1-r)$	0,5
A ₂ A ₂ B ₂ B ₂	2-2	n ₉	$p_9 = \frac{1}{4} (1-r)^2$	0,0

P_i(R|G) é a probabilidade de que um gameta seja recombinante dado o genótipo dos dois locos.

Na Tabela 1, verifica-se que

$$p_1 = p_9 = \alpha_1 = \frac{1}{4}(1-r)^2$$

$$p_2 = p_4 = p_6 = p_8 = \alpha_2 = \frac{1}{2}r(1-r)$$

$$p_3 = p_7 = \alpha_3 = \frac{1}{4}r^2$$

e

$$p_5 = \alpha_4 = \frac{1}{2}(1-r)^2 + \frac{1}{2}r^2$$

Para facilidade de cálculos, pode-se utilizar:

$$n_a = n_1 + n_9$$

$$n_b = n_2 + n_4 + n_6 + n_8$$

$$n_c = n_3 + n_7$$

$$n_d = n_5$$

e
$$N = n_a + n_b + n_c + n_d$$

Nos trabalhos de mapeamento, o método da máxima verossimilhança para estimação das frequências de recombinação, conforme descrito a seguir tem sido utilizado (SCHUSTER & CRUZ, 2004). Para a utilização deste método, é necessário o conhecimento da função densidade de probabilidade (f.d.p.) da distribuição de uma variável aleatória X. Para populações F_2 , esta função segue uma distribuição multinomial. Portanto, a função de máxima verossimilhança para uma população de tamanho N, com n possíveis classes, assumindo distribuição multinomial, é dada por

$$L(p_i | n_i) = \lambda p_1^{n_1} p_2^{n_2} \dots p_9^{n_9}$$

sendo

$$\lambda = \frac{N!}{n_1! n_2! \dots n_9!}$$

$$N = \sum_{i=1}^9 n_i$$

em que,

n_i é o número de ocorrência do evento x_i com probabilidade p_i .

ou, de forma equivalente,

$$L(p_i | n_i) = \lambda \alpha_1^{n_a} \alpha_2^{n_b} \alpha_3^{n_c} \alpha_4^{n_d}$$

ou ainda,

$$L(r | n_i) = \lambda \left[\frac{1}{4}(1-r)^2 \right]^{n_a} \left[\frac{1}{2}r(1-r) \right]^{n_b} \left[\frac{1}{4}r^2 \right]^{n_c} \left[\frac{1}{2}(1-r)^2 + \frac{1}{2}r^2 \right]^{n_d}$$

Tem-se, portanto, a função suporte, aplicando o logaritmo natural,

$$\ell(p_i | n_i) = \ln(\lambda \alpha_1^{n_a} \alpha_2^{n_b} \alpha_3^{n_c} \alpha_4^{n_d})$$

ou ainda

$$\ell(p_i | n_i) = \ln(\lambda) + n_a \ln(\alpha_1) + n_b \ln(\alpha_2) + n_c \ln(\alpha_3) + n_d \ln(\alpha_4)$$

O último passo do processo de estimação, pelo método da máxima verossimilhança, é a obtenção da função escore, que é dada pela primeira derivada da função suporte, em relação ao parâmetro que se deseja estimar.

$$f'[\ell(p_i | n_i)] = \frac{\partial \ell(p | n)}{\partial p}$$

Como o único parâmetro a ser estimado, neste caso, é a frequência de recombinação entre duas marcas, o estimador de máxima verossimilhança é obtido, igualando-se a equação anterior a zero.

$$f'[\ell(p_i | n_i)] = \frac{\partial \ell(p | n)}{\partial p} = 0$$

Tendo demonstrado o processo de obtenção dos estimadores de máxima verossimilhança, a estimação *per se* fica dependente, apenas, do conhecimento dos valores de n_i (número de ocorrência de cada classe genotípica) e da probabilidade p_i associada a cada uma dessas classes. A obtenção destes valores é apresentada na Tabela 1.

Uma vez obtidos os valores de n_i e as probabilidades p_i associadas, realizam-se as devidas substituições na equação de máxima verossimilhança, de modo a obter

$$L(r|n_i) = \lambda \left[\frac{1}{4}(1-r)^2 \right]^{n_a} \left[\frac{1}{2}r(1-r) \right]^{n_b} \left[\frac{1}{4}r^2 \right]^{n_c} \left[\frac{1}{2}(1-r)^2 + \frac{1}{2}r^2 \right]^{n_d}$$

Tem-se, portanto, a função suporte

$$\ell(r|n_i) = \ln(\lambda) + n_a \ln(\alpha_1) + n_b \ln(\alpha_2) + n_c \ln(\alpha_3) + n_d \ln(\alpha_4)$$

Obtém-se, por derivação da função suporte em relação a r , a função escore dada a seguir:

$$\frac{\partial \ell(r|n_i)}{\partial r} = \frac{n_a}{(1-r)^2} 2(1-r)(-1) + \frac{n_b}{r(1-r)} (1-2r) + \frac{n_c}{r^2} 2r + \frac{n_d}{(1-r)^2 + r^2} [2(1-r)(-1) + 2r]$$

de forma que

$$\frac{\partial \ell(r|n_i)}{\partial r} = \frac{(n_b + 2n_c) - 2r(n_a + n_b + n_c)}{r(1-r)} + \frac{2(2r-1)n_d}{(1-r)^2 + r^2}$$

igualando a zero a função escore, tem-se

$$(n_b + 2n_c) - 2(n_a + 2n_b + 3n_c + n_d)\hat{r} + 2(2n_a + 3n_b + 4n_c + 3n_d)\hat{r}^2 - 4N\hat{r}^3 = 0$$

Assim, a solução de máxima verossimilhança é obtida, utilizando-se uma das raízes da equação polinomial, de grau 3, descrita anteriormente.

2.1.3. Populações segregantes para mapeamento genético

Para o mapeamento genético, diferentes tipos de populações segregantes podem ser empregados. Tradicionalmente, são utilizadas populações derivadas do cruzamento de linhas

puras, o que origina uma geração F_1 , a que é autofecundada ou retrocruzada com um dos pais para a produção de uma geração F_2 ou RC_1 , respectivamente (FERREIRA & GRATTAPAGLIA, 1998; SCHUSTER & CRUZ, 2004). Alternativamente, são utilizadas populações F_n ($n= 3, 4, \dots, \infty$), duplo-haplóides (DH) e linhagens endogâmicas recombinantes ou RILs “Recombinat Inbred Lines” (BURR et al., 1988; KNAPP, 1991). Na escolha da população, deve-se levar em conta os objetivos, bem como o tempo e os recursos disponíveis para a execução do trabalho. Nas plantas F_1 , o desequilíbrio de ligação é máximo. Assim, estudos das populações derivadas a partir destas plantas F_1 são conduzidos no sentido de explorar este desequilíbrio (SCHUSTER & CRUZ, 2004).

Nas populações F_2 derivadas de cruzamentos de progenitores homozigotos contrastantes, os marcadores codominantes segregam na proporção 1:2:1 e os dominantes na proporção 3:1. Estas populações apresentam as vantagens de maior rapidez em sua obtenção e maior precisão no mapeamento genético e na detecção de QTL, devido à disponibilidade de informações dos três genótipos (A_1A_1 , A_1A_2 e A_2A_2). No entanto, há perda de precisão na mensuração de características quantitativas, podendo ocorrer a impossibilidade do mapeamento de genes de resistência a doenças, em que seja necessária a inoculação da população segregante com diferentes raças fisiológicas ou diferentes patógenos, uma vez que não é possível a replicação dos indivíduos da população (LIU, 1998; SCHUSTER & CRUZ, 2004).

Estas desvantagens podem ser contornadas por meio da propagação vegetativa ou com a utilização de populações $F_{2:n}$, geralmente $F_{2:3}$. Neste caso, os genótipos das plantas são determinados na geração F_2 , enquanto as características fenotípicas são medidas em plantas F_3 , com repetições. Os dados das médias das famílias F_3 são, geralmente utilizados como o valor fenotípico das plantas F_2 , que originaram estas famílias. Entretanto, isto só é verdadeiro na ausência de dominância. Para a utilização de dados de famílias F_3 , na análise de plantas F_2 é necessário levar em consideração a composição genotípica das plantas na geração F_3 (SCHUSTER & CRUZ, 2004; ZHANG & XU, 2004).

2.2. Detecção de QTL

2.2.1. Carácter quantitativo

A maioria dos caracteres de importância econômica são governados por muitos genes de pequeno efeito, que são influenciados pelo ambiente e são denominados como caracteres quantitativos ou métricos (MATHER & JINKS, 1984; VENCovsky & BARRIGA, 1992; FALCONER & MACKAY, 1996; CRUZ, 2005). Os caracteres quantitativos possuem natureza complexa e devido a sua importância são muito estudados, porém pouco compreendidos, tanto do ponto de vista evolutivo quanto do melhoramento genético (COELHO, 2000). Embora os princípios genéticos de segregação das características quantitativas sejam mendelianos, a identificação da segregação dos genes, individualmente, não é uma tarefa fácil. Por este motivo, a média e a variância da distribuição dos valores das características quantitativas são essenciais ao estudo da herança e variação destas características (TANKSLEY, 1993; MACKAY, 2001).

No estudo da herança de caracteres quantitativos, os geneticistas e melhoristas utilizam uma abordagem biométrica, em que um carácter é função da sua constituição genética, do efeito ambiental e da interação de genótipos com ambientes (HALLAUER & MIRANDA FILHO, 1988; RAMALHO et al., 1993; FALCONER & MACKAY, 1996; LYNCH & WALSH, 1998; CRUZ, 2005). Vários delineamentos genéticos foram propostos com o intuito de estimar as variâncias associadas aos efeitos genéticos (aditivo e dominância) dos genes, que controlam os caracteres quantitativos (CRUZ et al., 2004).

A aplicação de técnicas de marcadores moleculares, no estudo de características quantitativas permite identificar maior fração da variância genética total, relacionada à expressão de determinada característica, possibilitando, ainda detectar as regiões dos cromossomos, ou grupos de ligação, que são mais determinantes para a expressão da característica. O termo QTL “Quantitative trait loci” foi criado por GELDERMANN (1975) e baseia-se no princípio de existência de locos de maior importância, relacionados à expressão de determinada característica

quantitativa. Além da determinação da posição do QTL ao longo do grupo de ligação, a utilização de técnicas de marcadores moleculares permite a obtenção de estimativas de ação gênica (efeito aditivo) e de interação entre os alelos (efeito de dominância) e não alelos (efeito epistático) (BENTO, 2006).

A detecção e caracterização dos QTL dependem, principalmente, do mecanismo de herança da característica, da população avaliada, do número de indivíduos genotipados e do tipo de marcador molecular utilizado (DEKKERS & HOSPITAL, 2002; MACKAY, 2001; SCHUSTER & CRUZ, 2004). O efeito do QTL na expressão da característica é associado à precisão experimental, à herdabilidade da característica avaliada e aos efeitos aditivos e de dominância (LIU, 1998; LYNCH & WALSH, 1998; SCHUSTER & CRUZ, 2004). Além destas três características, também destacam-se o efeito do QTL na expressão da característica, a precisão experimental e a herdabilidade.

Portanto, utilizando-se informação indireta dos marcadores, é possível construir um mapa de QTL. Entende-se por mapeamento de QTL não apenas sua localização física no genoma, mediante alguma escala apropriada, mas também a quantificação e caracterização de seus efeitos, como interações alélicas (grau de dominância) de cada um, e mesmo interações epistáticas ou pleiotrópicas (ROCHA, 2004; MACHADO, 2004). Estudos mais detalhados possibilitam, ainda, a detecção e a decomposição da interação genótipos x ambientes em nível de cada QTL (SILVA, 2001; BENTO, 2006).

O mapeamento de QTL pode ser, também, útil no melhoramento genético. Embora a seleção assistida por marcadores (SAM) seja possível sem o mapeamento de QTL, este pode ser útil no direcionamento de cruzamentos específicos, maximizando tanto o número de QTL com alelos favoráveis, como o valor genotípico global, incluindo os efeitos epistáticos (FERREIRA & GRATTAPAGLIA, 1998).

2.2.2. Mapeamento por marcas simples

A primeira metodologia desenvolvida para detecção de QTL foi o mapeamento de marcas simples, que consiste na associação da expressão do QTL à presença de um marcador, sendo realizadas análises para cada marcador separadamente. O método baseia-se no mapeamento de QTL, sendo o procedimento adotado desde as primeiras tentativas de se associar um caráter quantitativo a um marcador genético de qualquer natureza, o qual poderia corresponder inclusive a algum outro caráter, cujo controle genético fosse mais simples, responsável pela expressão da coloração ou morfologia de uma determinada estrutura da planta (SAX, 1923; RASMUSSEN, 1933; McMILLAN & ROBERTSON, 1974; SOLLER et al., 1976). A análise é feita, verificando-se a diferença entre as médias fenotípicas do caráter para cada uma das classes genotípicas de um dado marcador, que devem possuir distribuição de frequência correspondente ao tipo de população utilizada, como 1:2:1 no caso de populações F_2 e marcador codominante ou 1:1 no retrocruzamentos (KNAPP, 1991). Se as diferenças entre as médias fenotípicas das classes forem estatisticamente significativas, pode-se inferir que existe um QTL ligado à marca analisada (LIU, 1998; LYNCH & WALSH, 1998; SCHUSTER & CRUZ, 2004). Vários procedimentos estatísticos podem ser utilizados como o teste t, análise de variância, regressão linear simples ou método da máxima verossimilhança (KEARSEY & HYNE, 1994; ZENG, 1994; DOERGE et al., 1997).

O método possui a vantagem de ser simples, mas apresenta limitações. Como cada marcador é analisado, individualmente, não é possível identificar se uma marca está ligada a um ou mais QTL simultaneamente e, assim, determinar o número de QTL em questão (HYNE & KEARSEY, 1995). O método também confunde efeito e posição de QTL, não havendo distinção entre a ocorrência de um QTL de pequeno efeito, situado próximo ao marcador e de um QTL de grande efeito mais distante do mesmo marcador. Além disso, os efeitos genéticos do QTL são subestimados, pois, seus estimadores são viesados pela fração de recombinação entre o marcador

e o QTL. Este viés, por sua vez, gera a necessidade de se utilizar amostras de tamanho grande na análise (LANDER & BOTSTEIN, 1989). Deve-se notar, entretanto, que se a abordagem for realizada por meio do método da máxima verossimilhança é possível a estimação dos parâmetros relativos aos efeitos e à distância entre o QTL e o marcador, uma vez que os estimadores de máxima verossimilhança são não viesados (BEARZOTI, 2000). Não se pode, entretanto, determinar se o QTL está posicionado à direita ou à esquerda da marca.

2.2.3. Mapeamento por intervalo simples

Para contornar os problemas da análise de marcas simples, LANDER & BOTSTEIN (1989) propuseram o mapeamento por intervalo, que passou a ser o método padrão de mapeamento de QTL, nos anos subsequentes (BOHN et al., 1997). Neste método, ao invés de marcas individuais, são considerados na análise pares de marcas adjacentes. A detecção da presença e a estimação dos efeitos dos QTL são realizados, dentro de cada intervalo entre marcas, analisando-se cada intervalo, separadamente e utilizando o método da máxima verossimilhança.

Por exemplo, a construção de um mapa de ligação, utilizando-se um conjunto de “n” marcadores, M_1, M_2, \dots, M_n . Como princípio, o método assume a existência de um QTL “Q” no intervalo entre as marcas adjacentes $[M_k, M_{k+1}]$, sendo realizado o teste da razão de verossimilhança (TRV) para a presença do QTL, em cada posição ao longo do intervalo $[M_x, M_{x+1}]$, obtendo-se as estimativas de verossimilhança, tanto para os efeitos do possível QTL quanto para sua posição. Em seguida, o procedimento é repetido no intervalo seguinte $[M_{k+1}, M_{k+2}]$ e, assim, sucessivamente até que todo o genoma coberto pelos marcadores seja mapeado.

Dentre as vantagens do mapeamento por intervalo, tem-se a representação mais clara da presença dos QTL ao longo do genoma, mostrando, para cada posição o valor do TRV correspondente, sendo apresentado mais facilmente o indício da presença de um ou mais QTL controlando o caráter. As estimativas dos efeitos genéticos são não-viesadas, o que é garantido pelo método da máxima verossimilhança. Outra vantagem é o menor número de progênies, requeridas para a análise (LANDER & BOTSTEIN, 1989). É possível, ainda, o estabelecimento de intervalos de confiança para as posições estimadas dos QTL, utilizando-se o valor do TRV expresso na forma de “LOD score”, de acordo com o critério denominado “one LOD support interval” (LANDER & BOTSTEIN, 1989; LYNCH & WALSH, 1998).

O mapeamento por intervalo apresentou considerável avanço, em relação ao de marca simples, mas possui também restrições. A principal restrição refere-se às situações em que múltiplos QTL estão presentes em um mesmo cromossomo, pois, nesses casos, os efeitos de QTL situados fora do intervalo sob análise influenciarão o mapeamento do QTL em questão, cujo efeito não pode ser individualizado pelo modelo empregado. O TRV para a presença do QTL, por sua vez, perde confiabilidade e, mesmo que não haja nenhum QTL presente, o mesmo pode ser significativo devido à influência do efeito de QTL presentes em outros intervalos, no mesmo cromossomo. Os QTL inexistentes, detectados nessas situações, são denominados “QTL fantasmas”. Devido a esta deficiência, o método não possibilita a estimação precisa do número e da posição dos QTL. Além disso, o uso de apenas duas marcas por vez limita a eficiência do método, que não aproveita a informação de marcas fora do intervalo mapeado e que porventura possuam associação com o caráter (LANDER & BOTSTEIN, 1989; ZENG, 1994).

Uma questão de grande importância para o mapeamento de QTL é a determinação do limite crítico ou “threshold” a ser utilizado na análise, que corresponde ao valor do TRV acima do qual o QTL será declarado presente. Conforme detalhadamente discutido por ZENG (1994), de acordo com uma das propriedades da regressão múltipla, os coeficientes de regressão parciais

associados a quaisquer dois marcadores não são correlacionados, a menos que esses marcadores sejam adjacentes. Por consequência, os TRVs correspondentes a dois intervalos adjacentes não são independentes, apresentando pequena correlação entre si. O autor afirma que, sendo esta correlação de pequena magnitude e tratando-se de uma situação de testes múltiplos, em que a presença de QTL é verificada para todo o genoma, a razão de verossimilhança em um intervalo qualquer pode ser aproximada pela distribuição de $\chi^2_{\alpha/M, GL}$, sendo α o nível de significância global e M o número de intervalos, com GL graus de liberdade. Essa aproximação supõe os intervalos como independentes entre si, o que nem sempre é considerado pelos autores, sendo adequada para situações em que a amostra utilizada é de grande tamanho e o número de marcadores presentes no modelo não é excessivo. O limite crítico pode ser obtido, também, a partir de permutações ou da estimativa do número de testes independentes (JIN et al., 2007).

O valor do TRV pode ser expresso na forma de “LOD score”, o que facilita a compreensão da magnitude do teste pelo uso de escala logarítmica. A relação entre os valores do TRV e do “LOD score” é dada por (LANDER & BOTSTEIN, 1989; LYNCH & WALSH, 1998), como

$$\text{LOD}_{\text{score}} = \frac{\text{TRV}}{2\ln 10} \cong \frac{\text{TRV}}{4,6052}$$

Assim, um valor de “LOD score” igual a 3, em uma dada posição no intervalo mapeado indica que a hipótese de ocorrência do QTL é 1.000 vezes mais provável que a hipótese da não ocorrência.

2.3. Genotipagem seletiva

A genotipagem seletiva é uma estratégia, que pode reduzir significativamente o número de indivíduos genotipados para um dado poder de detecção de QTL, a partir de um grande número de indivíduos fenotipados (LANDER & BOTSTEIN, 1989; DARVASI & SOLLER, 1992). O termo genotipagem seletiva é usado, quando se determina a ligação entre os locos dos marcadores e locos que controlam a característica quantitativa, genotipando somente os indivíduos presentes no extremo superior e inferior da distribuição normal dos fenótipos de uma grande população. Esta metodologia dispensa a genotipagem de todos os indivíduos para todos os marcadores e permite uma análise preliminar, indicando marcadores que, possivelmente, estejam associados à QTL (DARVASI, 1997).

Indivíduos de uma parte da progênie contribuem mais para a informação de ligação do marcador com o QTL que outros. Os indivíduos, que fornecem maior informação, são aqueles cujo genótipo pode ser claramente inferido do seu fenótipo, ou seja, os organismos com fenótipos extremos contribuem com maior informação de ligação que os organismos concentrados próximos à média fenotípica da população (VAN OOIJEN, 1992). Progênies com valores fenotípicos de mais de um desvio-padrão da média compreendem cerca de 33% de toda a população e contribuem com cerca de 81% de toda a informação de ligação (VAN GESTEL et al., 2000). Quando a população é grande e genotipam-se apenas os extremos, pode-se reduzir em até 5,5 vezes o número de indivíduos genotipados com a mesma informação de ligação (LANDER & BOTSTEIN, 1989).

A partir de uma população de 500 indivíduos e genotipando-se a metade ($p=0,5$), DARVASI & SOLLER (1992) afirmaram que o poder de detecção de QTL é reduzido de 0,80 para 0,77. Isto significa que metade da população pouco contribui com informação para a análise e,

quando o tamanho da população é grande, uma proporção muito menor da população pode ser selecionada.

A proporção ótima de indivíduos a serem genotipados, do ponto de vista de minimizar custos para uma dada precisão experimental, depende da relação entre o custo da genotipagem completa de um indivíduo para todos os marcadores incluídos no experimento e o custo da fenotipagem. O número de indivíduos, selecionados para a genotipagem seletiva, nunca deverá ser maior que 50% da população (25% de cada extremo), segundo LANDER & BOTSTEIN (1989). Por outro lado, quando o efeito do QTL é grande, foi constatado um viés de grande magnitude na posição do QTL, mesmo selecionando 50% da população (LIN & RITLAND, 1996).

Ao aplicar a estratégia de genotipagem seletiva, LANDER & BOTSTEIN (1989), DARVASI & SOLLER (1992), MURANTY et al. (1997) e BOVENHUIS & SPELMAN (2000) afirmaram que é possível aumentar, significativamente, o poder de detecção do QTL para a característica pelo quais os indivíduos foram selecionados.

A genotipagem seletiva é, em essência, uma amostragem. A amostragem é um campo da estatística bastante sofisticado, que estuda técnicas de planejamento de pesquisa para possibilitar inferências sobre um universo, a partir do estudo de uma pequena parte de seus componentes.

O termo amostragem refere-se ao processo (probabilístico ou não probabilístico), pelo qual se obtém uma amostra, devendo ser realizado com técnicas adequadas (amostra probabilística) a fim de garantir a representatividade da população em estudo. Cabe ainda ressaltar que, geralmente, sempre que possível cada elemento da população deve ter igual probabilidade de participar da amostra, evitando, assim, um viés de amostragem. As pesquisas por amostragem proporcionam algumas vantagens, como menor custo, resultado em menor tempo, objetivos mais amplos e dados fidedignos.

A amostragem pode ser aleatória em que se toma ao acaso, na população, indivíduos que constituirão a amostra a ser genotipada, ou sistemática quando os elementos que constituirão a

amostra são coletados, segundo um padrão estabelecido pelo pesquisador. Neste caso, por exemplo, amostraram-se indivíduos representativos das várias classes da distribuição da característica analisada. A amostragem por conglomerado, que é o caso da genotipagem seletiva, ocorre quando a população é subdividida em estratos sobre o qual se pratica a amostragem.

2.4. Efeito da genotipagem seletiva sobre mapeamento

Ao realizar a análise genômica, dentre os vários objetivos, há o interesse de estabelecer mapas de ligação e a detecção de QTL para fins de melhoramento. Em termos de redução nos custos e aumento relativo no poder de detecção de QTL, as vantagens têm sido ressaltadas quando para a genotipagem seletiva (RUY et al., 2005). Por outro lado, o processo de amostragem pode conduzir à perda da estrutura genética da população, causando problemas de distorção na segregação dos marcadores e comprometendo o mapeamento genético.

Em populações segregantes, a distorção da razão de segregação é o desvio da proporção mendeliana esperada nos indivíduos, em dada classe genotípica (SANDLER & GOLIC, 1985). Resulta na falha ou violação de pressupostos da teoria genética convencional e suas análises (LU et al., 2002). Esta ocorrência tem sido relatada em diversos organismos, incluindo plantas, nas quais espécies ou raças híbridas exibem, preferencialmente, disfunção de gametas (XU et al., 1997) ou problemas relativos à seleção genotípica. As causas podem ser biológicas ou atribuídas à amostragem, incluindo os processos de amostragem não aleatório.

As estimativas de recombinação entre pares de locos, com distorção de segregação em pelo menos um deles são viesadas (LORIEUX et al. 1995). Assim, os mapas construídos com distorção de segregação seriam mapas poucos acurados (HACKETT & BROADFOOT, 2003).

Para o mapeamento, a utilização ou não de marcadores que apresentam distorção de segregação é discutível. De acordo com KAO et al. (1999), recomenda-se, preferencialmente, o

descarte dos locos que apresentam distorção de segregação, para que a qualidade do mapa não seja comprometida. São raros os casos em que conjuntos de dados, com a presença de marcadores com distorção de segregação, são utilizados para o mapeamento (HARUSHIMA et al., 1996).

Em muitas situações, após a construção de mapas de ligação, segue-se a análise de QTL. Assim, se frações de recombinação ou a ordem dos marcadores forem inferidas incorretamente, as pressuposições básicas da análise de QTL serão violadas e os resultados poderão ser imprecisos (Vogl & Xu, 2000).

O tamanho da população é fundamental para a detecção da distorção de segregação. Tamanho mínimo para detecção de segregação ambígua deve ser observado quando há suspeita de marcadores sob seleção. Distorções proporcionadas por inviabilidade gamética ou genotípica, que ocorre após fertilização, afetam de forma diferencial o mapeamento genético. Os efeitos da distorção genotípica sobre o mapeamento genético são mais drásticos e erráticos. Mapas genéticos estabelecidos a partir de funções de verossimilhança que negligenciam a distorção de segregação de marcadores, proporcionam estimativas viesadas da distância entre marcas (FERREIRA, 2006).

3. REFERÊNCIAS BIBLIOGRÁFICAS

BEARZOTI, E. Mapeamento de QTL. In: PINHEIRO, J.B.; CARNEIRO, I.F. **Análise de QTL no melhoramento de plantas**: Segunda jornada em genética e melhoramento de plantas. Goiânia: FUNAPE, 2000. p.63-209.

BENTO, D.A.V. **Mapeamento de QTLs para produção de grãos e seus componentes em uma população de milho tropical**. 2006. 133p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba.

BOHN, M.; KHAIRALLAH, M.M.; JIANG, C.; GONZÁLEZ-DE-LEÓN, D.; HOISINGTON, D.A.; UTZ, H.F.; DEUTSCH, J.A.; JEWELL, D.C.; MIHM, J.A.; MELCHINGER, A.E. QTL mapping in tropical maize: II. Comparison of genomic regions for resistance to *Diatraea* spp. **Crop Science**, Madison, v.37, n.6, p.1892-1902, 1997.

BORÉM, A.; CAIXETA, E.T. **Marcadores moleculares**. Viçosa: Jard, 2006. 374p.

BOTSTEIN, D.; WHITE, R.L.; SKOLNICK, M.; DAVIS, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetics**, Chicago, v.32, n.3, p.314-331, 1980.

BOVENHUIS, H.; SPELMAN, R.J. Selective genotyping to detect quantitative trait loci for multiple traits in outbred populations. **Journal of Dairy Science**, Savoy, v.83, n.1, p.173-180, 2000.

BURR, B.; BURR, F.A.; THOMPSON, K.H.; ALBERTSON, M.; STUBER, C.W. Gene mapping with recombinant inbreds in maize. **Genetics**, Bethesda, v.118, n.3, p.519-526, 1988.

COELHO, A.S.G. Considerações gerais sobre a análise de QTLs. In: PINHEIRO, J.B.; CARNEIRO, I.F. **Análise de QTL no melhoramento de plantas**: Segunda jornada em genética e melhoramento de plantas. Goiânia: FUNAPE, 2000. p.1-36.

CRUZ, C.D. **Princípios de genética quantitativa**. Viçosa: UFV, 2005. 394p.

CRUZ, C.D.; REGAZZI, A.J.; CARNEIRO, P.C.S. 3.ed. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, 2004. v.1, 480p.

CRUZ, C.D.; SILVA, L.C. Análise de marcadores moleculares. In: BORÉM, A.; CAIXETA, E.T. (Eds.). **Marcadores Moleculares**. Viçosa: Jard, 2006. p.307-374.

CRUZ, C.D.; VIANA, J.M.S.; Carneiro, P.C.S. **Genética: GBOL - Software para ensino e aprendizagem de genética**. Viçosa: UFV, 2001. v.1, 476p.

DARVASI, A. The effect of selective genotyping on QTL mapping accuracy. **Mammalian Genome**, New York, v.8, n.1, p.67-68, 1997.

DARVASI, A.; SOLLER, M. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. **Theoretical and Applied Genetics**, Berlin, v.85, n.2-3, p.353-359, 1992.

DEKKERS, J.C.M.; HOSPITAL, F. The use of molecular genetics in the improvement of agricultural populations. **Nature Reviews Genetics**, New York, v.3, n.3, p.22-32, 2002.

DOERGE, R.W.; ZENG, Z.B.; WEIR, B.S. Statistical issues in the search for genes affecting quantitative traits in experimental populations. **Statistical Science**, Philadelphia, v.12, n.3, p.195-219, 1997.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to quantitative genetics**. 4.ed. London: Longman, 1996. 464p.

FERREIRA, A. **Mapeamento genético utilizando marcadores moleculares com distorção de segregação gamética e genotípica**. 2006. 132p. Tese (Doutorado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa.

FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3.ed. Brasília: EMBRAPA/CENARGEM, 1998. 220p.

FISHER, R.A. Average excess and average effect of a gene substitution. **Annals of Eugenics**, Cambridge, v.11, n.1, p.53-63, 1941.

GELDERMANN, H. Investigations on inheritance of quantitative characters in animals by gene markers. I. Methods. **Theoretical and Applied Genetics**, Berlin, v.46, n.7, p.319-330, 1975.

GRODZICKER, T.; WILLIAMS, J.; SHARP, P.; SAMBROOK, J. Physical mapping of temperature-sensitive mutations of adenovirus. **Cold Spring Harbor Symposium Quantitative Biology**, Woodbury, v.39, n.1, p.439-446, 1974.

HACKETT, C.A.; BROADFOOT, L.B. Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. **Heredity**, Oxford, v.90, n.1, p.33-38, 2003.

HALLAUER, A.R.; MIRANDA FILHO, J.B. **Quantitative genetics in maize breeding**. 2.ed. Ames: Iowa States University Press, 1988. 468p.

HARUSHIMA, Y.; MURATA, N.; YANO, M.; NAGAMURA, Y.; SASAKI, T.; MINOBE, Y.; NAKAGAHARA, M. Detection of segregation distortions in an *indica-japonica* rice cross using a high-resolution molecular map. **Theoretical and Applied Genetics**, Berlin, v.92, n.2, p.145-150, 1996.

HYNE, V; KEARSEY, M.J.; QTL analysis: further use of 'marker regression'. **Theoretical and Applied Genetics**, Berlin, v.91, n.3, p.471-776, 1995.

JIN, C.; FINE, J.P.; YANDELL, B.S. Unified semiparametric framework for quantitative trait loci analyses, with application to spike phenotypes. **American Statistical Association**, Alexandria, v.102, n.477, p.56-67, 2007.

KAO, C.H.; ZENG, Z.B.; TEASDALE, R.D. Multiple interval mapping for quantitative trait loci **Genetics**, Bethesda, v.152, n.3, p.1203-1216, 1999.

KEARSEY, M.J.; HYNE, V. QTL analysis: a simple 'marker regression' approach. **Theoretical and Applied Genetics**, Berlin, v.89, n.6, p.698-702, 1994.

KNAPP, S.J. Using molecular markers to map multiple quantitative trait loci: model for backcross, recombinant inbred, and double-haploid progeny. **Theoretical and Applied Genetics**, Berlin, v.81, n.3, p.333-338, 1991.

LANDE, R. The minimum number of genes contributing to quantitative variation between and within populations. **Genetics**, Bethesda, v.99, n.3-4, p.541-553, 1981.

LANDER, E.S.; BOTSTEIN, D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, Bethesda, v.121, n.1, 185-199, 1989. (erratum in 1994 *Genetics*, v.136, n.2, 705)

LIN, J.Z.; RITLAND, K. The effects of selective genotyping on estimates of proportion of recombination between linked quantitative trait loci. **Theoretical and Applied Genetics**, Berlin, v.93, n.8, p.1261-1266, 1996.

LITT, M.; LUTY, J.A. A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. **American Journal of Human Genetics**, Chicago, v.44, n.3, p.397-401, 1989.

LIU, B.H. **Statistical genomics, linkage, mapping and QTL analysis**. Boca Raton: CRC, 1998. 611p.

LORIEUX, M.; GOFFINET, B.; PERRIER, X.; GONZÁLEZ-DE-LEÓN, D.; LANAUD, C. Maximum likelihood models for mapping genetic markers showing segregation distortion. 1. Backcross populations. **Theoretical and Applied Genetics**, v.90, n.1. p.73-80, 1995.

LU, H.; ROMERO-SEVERSON, J.; BERNARDO, R. Chromosomal regions associated with segregation distortion in maize. **Theoretical and Applied Genetics**, Berlin, v.105, n.4, p.622-628, 2002.

LYNCH, M.; WALSH, B. **Genetics and analysis of quantitative traits**. Sunderland: Sinauer Associates, 1998. 980p.

MACHADO, C.F. **Repetibilidade, correlações fenotípicas e mapeamento de QTLs em populações segregantes de café arábica**. 2004. 188p. Tese (Doutorado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa.

MACKAY, T.F.C. Genetic architecture of quantitative traits. **Annual Review of Genetics**, Palo Alto, v.35, p.303-339, 2001.

MATHER, K.; JINKS, J.L. **Introdução à genética biométrica**. Tradução de DUARTE, F.A.M.; SENE, F.M.; ROTHSCHILD, H.A.; LÔBO, R.B.; MORTARI, N.; SCHLINDWEIN, A.P. Ribeirão Preto: Sociedade Brasileira de Genética, 1984. 242p.

McMILLAN, I.; ROBERTSON, A. The power of methods for the detection of major gene affecting quantitative characters. **Heredity**, Oxford, v.32, n.3, p.349-356, 1974.

MENDEL G., **Experiments in plant hybridization**, 1866 (Traduzido para o inglês por William Bateson, 1901). Disponível em <http://www.esp.org/foundations/genetics/classical/gm-65.pdf>. Acessado 27/06/06.

MURANTY, H.; GOFFINET, B.; SANTI, F. Multitrait and multipopulation QTL search using selective genotyping. **Genetical Research**, Cambridge, v.70, n.3, p.259-265, 1997.

RAMALHO, M.A.P., SANTOS, J.B. ZIMMERMANN, M.J.O. **Genética quantitativa em plantas autógamas; aplicações ao melhoramento do feijoeiro**. Goiânia: UFG, 1993. 271p.

RASMUSSEN, J.M. A contribution to the theory of quantitative character inheritance. **Hereditas**, Landskrona v.18, n.4, p.245-261, 1933.

ROCHA, R.B. **Mapeamento de QTL para características de qualidade da madeira e crescimento em híbridos (*Eucalyptus grandis* x *Eucalyptus urophylla*)**. 2004. 77p. Dissertação (Mestrado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa.

RUY, D.C.; NONES, K; BARON, E.E.; LEDUR, M.C.; MELO, C.M.R.; AMBO, M.; CAMPOS, R.L.R.; COUTINHO, L.L. Strategic marker selection to detect quantitative trait loci in chicken. **Scientia Agricola**, Piracicaba, v.62, n.2, p.111-116, 2005.

SAKIYAMA, N.S. Marcadores moleculares e as hortaliças. **Horticultura Brasileira**, Botucatu, v.11, n.2, p.204-206, 1993.

SANDLER, L.; GOLIC, K. Segregation distortion in *Drosophila*. **Trends in Genetics**, Oxford, v.1, n.1, p.181-185. 1985.

SAX, K. The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgaris*. **Genetics**, Bethesda, v.8, n.6, p.552-560, 1923.

SCHLOTTERER, C. Evolutionary dynamics of microsatellite DNA. **Chromosoma**, Wien, v.109, n.6, p.365-371, 2000.

SCHUSTER, I.; CRUZ, C.D. **Estatística genômica aplicada a populações derivadas de cruzamentos controlados**. Viçosa: UFV, 2004. 568p.

SILVA, H.D. **Aspectos biométricos da detecção de QTL's ("quantitative trait loci") em espécies cultivadas**. 2001. 166p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba.

SOLLER, M.; BRODY, T.; GENIZI, A. On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. **Theoretical and Applied Genetics**, Berlin, v.47, n.1, p.35-39, 1976.

TANKSLEY, S.D. Mapping polygenes. **Annual Review of Genetics**, Palo Alto, v.27, p.205-233, 1993.

VAN GESTEL, S.; HOUWING-DUISTERMAAT, J.J.; ADOLFSSON, R.; VAN DUIJN, C.M. CHRISTINE VAN BROECKHOVEN, C. Power of Selective Genotyping in Genetic Association Analyses of Quantitative Traits. **Behavior Genetics**, Boulder, v.30, n.2, p.141-146, 2000.

VAN OOIJEN, J.W. Accuracy of mapping quantitative trait loci in autogamous species. **Theoretical and Applied Genetics**, Berlin, v.84, n.7-8, p.803-811, 1992.

VENCOVSKY, R.; BARRIGA, P. **Genética biométrica no fitomelhoramento**. Ribeirão Preto: Sociedade Brasileira de Genética, 1992. 496p.

VIANA, J.M.S.; CRUZ, C.D.; BARROS, E.G. **Genética: Fundamentos**. 2.ed. Viçosa: UFV, 2003. v.1, 330p.

VOGL, C.; XU, S. Multipoint mapping of viability and segregation distorting loci using molecular markers. **Genetics**, Bethesda, v.155, n.3, p.1439-1447, 2000.

VOS, P.; HOGERS, R.; BLEEKER, M.; REIJANS, M.; VAN DE LEE, T.; HORNES, M.; FRITJERS, A.; POT, J.; PELEMAN, J.; KUIPER, M.; ZABEAU, M. AFLP: a new technique for DNA fingerprinting. **Nucleic Acids Research**, Oxford, v.23, n.21, p.4407-4414, 1995.

WELSH, J.; McCLELLAND, M. Fingerprinting genomes using PCR with arbitrary “primers”. **Nucleic Acids Research**, Oxford, v.18, n.24, p.7213-7218, 1990.

WILLIAMS, J.G.K; KUBELIK, A.R.; LIVAK, K.J.; RAFALSKY, J.A.; TINGEY, S.V. DNA polymorphisms amplified by arbitrary primers are useful as genetics markers. **Nucleic Acids Research**, Oxford, v.18, n.22, p.6531-6535, 1990.

XU, Y.; ZHU, L.; XIAO, J.; HUANG, N.; MCCOUCH, S.R. Chromosomal regions associated with segregation distortion of molecular markers in F₂, backcross, double haploids, and recombinant inbred populations in rice (*Oryza sativa* L.). **Molecular and General Genetics**, Heidelberg, v.253, p.535-545, 1997.

ZENG, Z.B. Precision mapping of quantitative trait loci. **Genetics**, Bethesda, v.136, n.4, p.1457-1466, 1994.

ZHANG, Y.M.; XU, S. Mapping quantitative trait loci in F₂ incorporating phenotypes of F₃ progeny. **Genetics**, Bethesda, v.166, n.4, 2004.

CAPÍTULO I

Mapeamento genético e detecção de QTL em populações F_2 simuladas

1. INTRODUÇÃO

Muitas das características econômicas desejáveis em várias espécies de plantas são controladas por grande número de genes, geralmente, influenciados pelas condições ambientais. Estas características, conhecidas como caracteres quantitativos ou poligênicos, exibem fenótipos com variação contínua, cujos valores são normalmente distribuídos (FALCONER & MACKAY, 1996). A maioria desses genes têm, teoricamente, pequeno efeito e são responsável pelo controle genético dos diversos processos metabólicos da característica. O termo “Quantitative Trait Loci” (QTL), tem sido usado para denominar regiões cromossômicas (locos), que controlam caracteres quantitativos. No estudo da herança de caracteres quantitativos são utilizadas as estimativas de parâmetros populacionais, tais como média, variância, coeficientes de regressão e correlação.

Com o surgimento dos marcadores moleculares baseados na técnica de PCR “Polimerase Chain Reaction”, como o RAPD (WILLIAMS et al., 1990), o AFLP (VOS et al., 1995) e os microssatélites (LITT & LUTY, 1989), a construção de mapas genéticos de alta cobertura genômica tornou-se possível para, praticamente, todas as espécies. A obtenção destes mapas possibilita o mapeamento de QTL, ou seja, a localização em grupos de ligação, a quantificação e a caracterização de seus efeitos, o número de locos envolvidos e a respectiva distribuição no genoma. Este conjunto de informações genéticas é um recurso poderoso no estudo da herança

desses caracteres, pois, viabilizam novas perspectivas para o aperfeiçoamento dos métodos de seleção e melhoramento genético.

Diversas metodologias para o mapeamento de QTL estão disponíveis. Inicialmente foi desenvolvida a análise de marcas simples (KEARSEY & HYNE, 1994) e em seguida, o mapeamento por intervalo simples (LANDER & BOTSTEIN, 1989), o mapeamento por intervalo composto (ZENG, 1994; JANSEN & STAM, 1994), o mapeamento por intervalo composto expandido para múltiplos ambientes ou caracteres (JIANG & ZENG, 1995) e o mapeamento de múltiplos intervalos (KAO et al., 1999), com sucessivo avanço tanto em relação à obtenção de informações quanto à precisão das mesmas. Metodologias mais recentes possibilitam, além da obtenção das informações essenciais ao mapeamento como a presença, a localização e efeitos dos QTL, a determinação do efeito da interação QTL x ambientes e de efeitos epistáticos entre QTL (SILVA, 2001; BENTO, 2006). Uma das principais aplicações da tecnologia dos marcadores moleculares consiste na utilização de informações, obtidas com o mapeamento de QTL, no processo de seleção de genótipo. Contudo, o sucesso na detecção de QTL por meio de marcadores moleculares depende da magnitude do efeito do QTL, do tamanho da população em estudo, do nível de saturação do mapa genético, da herdabilidade da característica e da utilização de métodos e modelos estatísticos adequados. Estudos envolvendo estes fatores são fundamentais para a escolha de metodologias mais adequadas à sua detecção, bem como para orientação na utilização de recursos para prover as situações mais apropriadas em análises genômicas.

2. OBJETIVOS

Este trabalho teve como objetivo geral:

Investigar, por meio de simulação, tomando como referência populações F_2 , supostamente sem restrição quanto ao tamanho, a influência dos níveis de herdabilidade, da ação gênica e da saturação dos grupos de ligação no mapeamento genético e na detecção e localização de QTL.

Os objetivos específicos foram:

- Avaliar a eficácia do mapeamento de marcadores codominantes, em população F_2 , por meio da avaliação da segregação mendeliana de locos individuais, distância entre pares de marcadores, ordenamento e estabelecimento do número de grupos de ligações esperados.
- Avaliar a eficácia da análise genômica na detecção da presença de QTL para característica fenotípica, em vários cenários.
- Avaliar a eficácia da análise genômica na mensuração da contribuição do QTL na variação do caráter quantitativo, em vários cenários.
- Avaliar a eficácia da análise genômica na estimação do efeito gênico exercido pelos QTL sobre a característica, em vários cenários.

3. MATERIAL E MÉTODOS

3.1. Simulação de dados

Para a geração dos dados de populações F_2 , com marcadores de natureza codominante, de uma espécie diplóide hipotética ($2n = 2x = 12$), utilizou-se o módulo de simulação do aplicativo computacional GQMOL (CRUZ, 2007), o qual permite gerar informações sobre genomas, genótipos de genitores, indivíduos de diferentes tipos de populações e dados de características quantitativas.

3.1.1. Simulação do genoma, níveis de saturação e tipos de marcas

Foi gerado um genoma com seis grupos de ligação, com 100 cM de tamanho cada um e, portanto, com comprimento total de 600 cM. Visando avaliar indivíduos recombinantes, provenientes de uma configuração genômica mais verossímil, para as análises de recuperação da informação biológica, foram gerados grupos de ligação com três níveis de saturação (pouco saturado, medianamente saturado e saturado), com intervalos entre marcas adjacentes de 20, 10 e 5 cM, respectivamente (Figura 1). O número de marcas nos três pares de grupos de ligação foi de 6, 11 e 21, respectivamente. Foram geradas marcas de natureza codominante, com vantagem de

identificar os indivíduos heterozigotos. Os grupos de ligação 1, 3 e 5 possuem locos, que alteram a expressão da característica quantitativa (QTL), identificados como Q_{11} , Q_{31} e Q_{51} . Os grupos de ligação 2, 4, e 6 não apresentam locos, que alteram a expressão da característica quantitativa, sendo utilizados como controle.

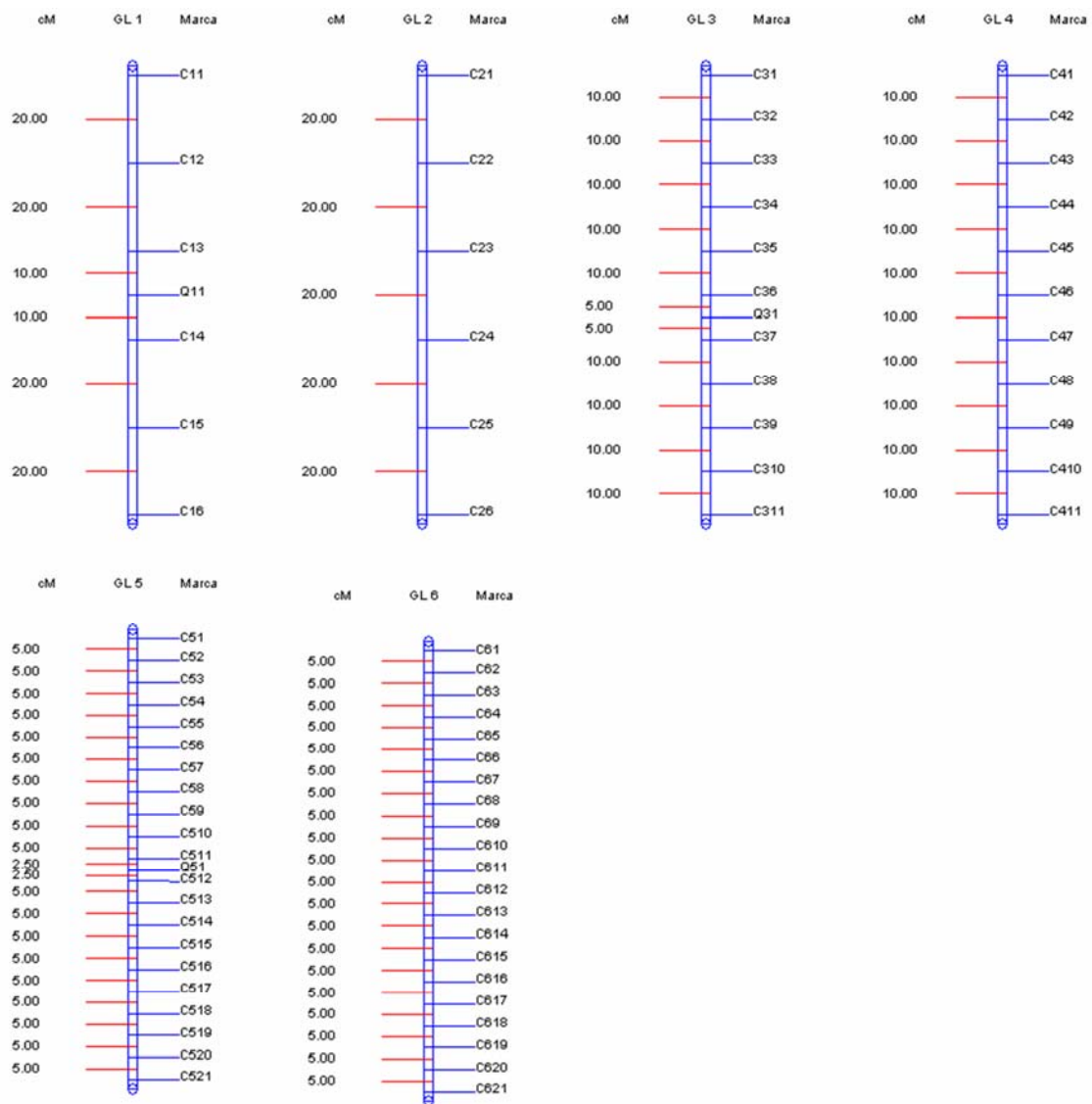


Figura 1 - Genoma com seis grupos de ligação (GL) de saturações de 20, 10 e 5 cM, contendo 6, 11 e 21 marcas, respectivamente. Total de 76 marcas do genoma simulado.

3.1.3. Procedimento de simulação dos indivíduos da população

Para cada indivíduo, foram produzidos os dados genotípicos referentes aos marcadores de acordo com a informação do genoma. A estratégia básica de simulação é baseada em permutas nos intervalos entre marcas adjacentes em cada cromossomo, de acordo com as distâncias dos marcadores, conforme apresentado a seguir:

i) A partir do genoma simulado, foram construídos os genótipos parentais homocigotos e contrastantes para os marcadores, de forma que a geração F_1 estivesse em fase de acoplamento para todos os pares de marcadores;

ii) A partir do genótipo da geração F_1 , foram gerados os gametas para a formação dos indivíduos das populações F_2 . A produção de gametas foi feita, simulando-se o pareamento dos homólogos e realizando permutas ao longo dos cromossomos, nas regiões delimitadas por dois marcadores adjacentes. A probabilidade de ocorrência de recombinação, numa região entre marcadores adjacentes, foi dada pela distância desses marcadores no genoma simulado. Após decidir, por meio da simulação de evento probabilístico, se ocorreria ou não a recombinação nesta região, passava-se para a próxima região delimitada pelo segundo e terceiro marcadores.

O procedimento prosseguiu até que todas as regiões entre marcadores adjacentes no cromossomo fossem sido alcançadas. Para a formação de cada indivíduo nas populações, foram simulados vários gametas do genitor F_1 , sendo sorteado apenas um gameta para formação de cada indivíduo, sendo, que o genótipo do indivíduo F_2 foi obtido pela junção dos dois gametas tomados, aleatoriamente, do conjunto de gametas do genitor F_1 .

3.1.4. Simulação das populações e da característica quantitativa

Foram simulados 36 cenários diferentes de populações F_2 , para a característica quantitativa. Para cada uma das situações, foram geradas 100 populações de 1.000 indivíduos. A característica foi controlada por vários locos, porém, apenas um QTL foi identificado em cada população. Estabeleceu-se que o caráter quantitativo apresentava média igual a 1.000 unidades, avaliado em condições experimentais com coeficiente de variação de 10%, controle gênico, caracterizado por grau médio de dominância igual zero. A variação fenotípica explicada, pelos QTL (Q_{11} , Q_{31} e Q_{51}), foi avaliada em quatro níveis de ação gênica (5, 10, 20 e 40%) e também três níveis de herdabilidade (80, 60 e 20%), totalizando 12 situações diferentes para cada QTL (Tabela 1). Foram, portanto, gerados 3.600 genomas analisados provenientes de um genoma simulado (três níveis de herdabilidades x quatro níveis de ação gênica x três saturações para o grupo de ligação com presença do QTL simulado x 100 repetições).

Tabela 1 - Disposição e identificação das populações F_2 em diferentes níveis de herdabilidade, de ação gênica para cada saturação do QTL (Q_{11} , Q_{31} e Q_{51})

h^2	Ação gênica	Q_{11}	Q_{31}	Q_{51}
80%	5%	POP 1	POP 13	POP 25
	10%	POP 2	POP 14	POP 26
	20%	POP 3	POP 15	POP 27
	40%	POP 4	POP 16	POP 28
60%	5%	POP 5	POP 17	POP 29
	10%	POP 6	POP 18	POP 30
	20%	POP 7	POP 19	POP 31
	40%	POP 8	POP 20	POP 32
20%	5%	POP 9	POP 21	POP 33
	10%	POP 10	POP 22	POP 34
	20%	POP 11	POP 23	POP 35
	40%	POP 12	POP 24	POP 36

h^2 = herdabilidade em porcentagem; POP = População F_2 simulada; Q_{11} = QTL do grupo de ligação um; Q_{31} = QTL do grupo de ligação três; Q_{51} = QTL do grupo de ligação cinco.

3.2. Análises genômicas - mapeamento genético

3.2.1. Análise de segregação de locos individuais - teste de qui-quadrado

Foram aplicados testes de qui-quadrado (PEARSON, 1932), para verificar a razão de segregação de cada marca em todas as populações geradas. A estatística qui-quadrado é dada por

$$\chi^2 = \sum_{i=1}^n \left[\frac{(O_i - E_i)^2}{E_i} \right]$$

em que

χ^2 é valor de qui-quadrado calculado;

O_i e E_i , são os valores observado e esperado, para a i -ésima classe fenotípica ($i= 1, 2, \dots, n$), respectivamente.

A hipótese (H_0) de segregação do loco, de 1:2:1 ($A_1A_1:A_1A_2:A_2A_2$), foi testada ao nível de 5% de probabilidade (erro tipo I). Quando o valor de probabilidade calculado era inferior ao preestabelecido, a hipótese H_0 era rejeitada, ou seja, a segregação não ocorreu de acordo com o esperado.

3.2.2. Análise de pares de marcas - estimação da porcentagem de recombinação

Após a aplicação dos testes de segregação, seguiu-se a etapa de estimação da porcentagem de recombinação entre pares de marcas. Para esta estimação, utilizou-se o método da máxima verossimilhança.

3.2.3. Análise de pares de marcas - determinação dos grupos de ligação e ordenamento das marcas

Os grupos de ligação foram obtidos pelo agrupamento, baseado na frequência máxima de recombinação de 30% e o LOD mínimo de 3. Após a formação de um grupo de ligação preliminar, verificou-se a melhor ordem dos marcadores constituintes do grupo pelo método SARF “Sum of Adjacent Recombination Fractions” (FALK, 1989; LIU, 1998). A soma das frações de recombinação adjacentes, como critério de estabelecimento do melhor ordenamento, é considerada um método robusto e foi utilizada neste trabalho, para determinar a ordem dos marcadores nos grupos de ligação.

3.2.4. Comparação de genomas

Nesse trabalho, o termo genoma simulado refere-se, genericamente, aos genomas simulados, conforme descrito nos itens 3.1.1 e 3.1.2, devendo ser referenciado como o paramétrico. O termo genoma analisado refere-se aos genomas, construídos a partir das populações simuladas, o qual é variável em razão do processo de estimação, amostragem, critério de agrupamento e outros.

Nos genomas analisados, estimaram-se, em cada situação, o número de grupos de ligação obtidos, o número de marcas por grupo, o tamanho dos grupos de ligação, a distância média entre marcadores adjacentes nos grupos de ligação, as variâncias das distâncias entre marcas adjacentes nos grupos de ligação e o estresse (KRUSKAL, 1964). Também foi verificado se houve, ou não, inversão na ordem dos marcadores, quantificando o evento pela correlação de Spearman (SPEARMAN, 1904). Todas essas comparações foram realizadas com a utilização do módulo de “Comparação de genomas” do aplicativo computacional GQMOL (CRUZ, 2007).

3.3. Análises genômicas - detecção de QTL

Os métodos utilizados neste trabalho para detecção de QTL, nas 3.600 populações F_2 e marcadores codominantes, foram o mapeamento por marca simples e de intervalo simples. Todas as análises moleculares foram realizadas, utilizando-se o programa GQMOL, 2007.1.2 (<http://www.ufv.br/dbg/gqmol/gqmol.htm>) em desenvolvimento na Universidade Federal de Viçosa - UFV.

3.3.1. Método de mapeamento de QTL por marca simples

O método da marca simples foi empregado, para avaliar a associação entre genótipos, que possuem uma forma alélica específica, utilizando-se regressão linear, em relação à significância da diferença entre grupos que apresentam determinadas marcas moleculares. Considerando duas formas alélicas A_1 e A_2 de um único loco A e valores de médias, respectivamente $\mu_{A_1A_1}$, $\mu_{A_1A_2}$ e $\mu_{A_2A_2}$ dos grupos portadores de cada uma das formas.

A hipótese nula de regressão $H_0: \beta_1=0$ foi testada contra a hipótese alternativa $H_a: \beta_1 \neq 0$ e $\beta_2 \neq 0$, aplicando-se o modelo

$$Y_j = \beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + \varepsilon_j$$

em que

Y_j = vetor de observações da característica para o j -ésimo indivíduo da população F_2 ;

X_{1j} = código do marcador para o efeito aditivo ($A_1A_1 = 1$; $A_1A_2 = 0$; $A_2A_2 = -1$);

X_{2j} = Código do marcador para o efeito dominante ($A_1A_1 = 0$; $A_1A_2 = 1$; $A_2A_2 = 0$);

β_0 = intercepto da regressão (média da característica);

β_1 = inclinação da reta para o efeito aditivo;

β_2 = inclinação da reta para o efeito de dominância; e

ε_j = erro aleatório na j -ésima observação, em que $\varepsilon_j \sim \text{NID}(0, \sigma^2)$.

Segundo SCHUSTER & CRUZ (2004),

$$E(\hat{\beta}_0) = \mu + 2r(1-r)d;$$

$$E(\hat{\beta}_1) = (1-2r)a; \text{ e}$$

$$E(\hat{\beta}_2) = (1-2r)^2 d.$$

Sendo $\mu+a$, $\mu+d$ e $\mu-d$ os valores genotípicos, associados aos genótipos QQ, Qq e qq, respectivamente.

Como neste estudo estabeleceu-se que a média genotípica da população F_2 (μ_g) fosse igual a 1.000 unidades e o grau médio de dominância nulo, então, para uma população F_2 ($p=q=0,5$) tem-se

$$E(\hat{\beta}_0) = \mu_g = 1.000, \text{ pois } \mu_g = \mu + \frac{1}{2}d \text{ e } E(\hat{\beta}_2) = 0$$

Também foi estabelecido o coeficiente de variação igual a 10%, o que equivale a:

$$CV = \frac{\sigma}{\mu} 100\% = 10\%$$

Logo, espera-se um valor de magnitude de variação residual igual a:

$$\sigma^2 = 10^4$$

Os valores paramétricos das herdabilidades possibilitam obter, indiretamente, a variação genotípica esperada na população sob influência do ambiente. Assim, sabe-se que

$$h^2 = \frac{\sigma_g^2}{\sigma^2 + \sigma_g^2}, \text{ logo, } \sigma_g^2 = \frac{h^2 \sigma^2}{100 - h^2}; \text{ de forma que:}$$

para $h^2 = 80\%$, deve-se ter $\sigma_g^2 = 4,00 \cdot 10^4$;

para $h^2 = 60\%$, deve-se ter $\sigma_g^2 = 1,50 \cdot 10^4$; e

para $h^2 = 20\%$, deve-se ter $\sigma_g^2 = 0,25 \cdot 10^4$.

Uma população F_2 apresenta valor esperado, para a variância genotípica de

$$\sigma_g^2 = \frac{1}{2} \tilde{a}^2 + \frac{1}{4} \tilde{d}^2$$

Sendo \tilde{a} e \tilde{d} a contribuição total de todos os locos. Neste estudo, foi estabelecido apenas um loco, sob avaliação, controlando uma proporção “p” da variância genotípica equivalente a 5, 10, 20 e 40%. Assim, o valor de “a” associado ao loco sob investigação é dado por

$$a = \sqrt{2p\sigma_g^2}$$

Têm-se, portanto, os seguintes valores esperados de “a”, apresentados no Tabela 2.

Tabela 2 - Valores esperados de “a” para os níveis de herdabilidade (h^2), proporção do controle gênico na variação genotípica (p) dado sua variância genética (σ_g^2)

h^2	σ_g^2	p			
		5%	10%	20%	40%
80%	$4,00.10^4$	63,24	89,44	126,49	178,89
60%	$1,50.10^4$	38,73	54,77	77,46	109,54
20%	$0,25.10^4$	15,81	22,36	31,62	44,72

3.3.2. Método de mapeamento de QTL por intervalo simples

O método de detecção de QTL por intervalo foi proposto, com a finalidade de incorporar a informação dos marcadores, que flanqueiam o suposto loco controlador da característica, para cada intervalo de um determinado grupo de ligação, dentro de um modelo de regressão. Este modelo de regressão, para populações F_2 , é dado por

$$Y_j = \mu + ax_j^* + dz_j^* + \varepsilon_j$$

em que

Y_j = é o valor mensurado para a característica Y no indivíduo j;

μ = é a média da característica na população;

a = efeito aditivo do loco que está sendo estudado sobre a característica;

d = efeito de dominância do loco que está sendo estudado sobre a característica;

x_j^* e z_j^* = variáveis condicionadoras e dependentes dos genótipos marcadores que flanqueiam o QTL, no indivíduo j ; e

ε_j = erro aleatório na j -ésima observação em que $\varepsilon_j \sim \text{NID}(0, \sigma^2)$.

O teste estatístico é realizado, considerando-se a hipótese de nulidade $H_0: a=0$ e $d=0$ contra a hipótese alternativa $H_a: a \neq 0$ e $d \neq 0$. Assim, havendo efeito significativo, conclui-se sobre a importância dos efeitos genéticos aditivos e atribuídos aos desvios da dominância. Neste estudo, entretanto, foi estabelecido grau de dominância nula ou ausência do efeito da dominância.

3.3.3. Obtenção do limite crítico

O teste estatístico, utilizado para verificação da existência de associação entre um marcador e um QTL, foi o teste de razão de verossimilhança (TRV). O TRV fornece o nível crítico, pois, é considerado assintoticamente distribuído com “ p ” graus de liberdade, segundo a distribuição de qui-quadrado (LYNCH & WALSH, 1998; SCHUSTER & CRUZ, 2004). O número de graus de liberdade para uma população F_2 foi de 2. Para a presença do QTL na análise de marca simples, o valor crítico foi de $\chi_{5\%,2}^2 = 5,99$, que corresponde ao “LOD score” de 1,3 para todo genoma. Para o mapeamento de QTL, realizado por intervalo simples, foram estimados os valores críticos em “LOD score” para cada grupo de ligação (Tabela 3). O nível de significância para os grupos de ligação foi de 5%, com 2 graus de liberdade.

Tabela 3 - Valores dos níveis críticos do qui-quadrado e do “LOD score” para os grupos de ligação, considerando 2 graus de liberdade

GL	NM	NI	γ	α	χ^2	LOD _{score}
1	6	5	5%	1,02%	9,17	1,99
3	11	10	5%	0,51%	10,55	2,29
5	21	20	5%	0,27%	11,93	2,59

GL = grupo de ligação; NM = número de marcas; NI = número de intervalo; γ = nível de significância para cada grupo de ligação; α = nível de significância para cada intervalo de um particular grupo de ligação.

4. RESULTADOS E DISCUSSÃO

Neste estudo, avaliou-se a dificuldade em estabelecer mapas genéticos e, principalmente, em detectar QTL em diferentes situações, correspondentes à saturação do grupo de ligação, controle gênico e herdabilidade da característica. Admitiu-se a situação em que, provavelmente, não haveria restrições quanto ao tamanho da população sobre as estimativas obtidas, uma vez que todas as análises foram realizadas a partir de informações de 1.000 indivíduos constituintes de uma população F_2 , com utilização de marcadores codominantes em 36 cenários para a característica quantitativa.

4.1. Mapeamento genético em populações F_2

Estudos de mapeamento genético já foram realizados com populações F_2 e marcadores codominantes, por meio de simulação, inferindo-se que um número mínimo de 200 indivíduos seria suficiente para o estabelecimento de mapas fidedignos (CRUZ, 2006). Entretanto, há carência de informação sobre a adequabilidade destes mapas, quando utilizados para detecção e localização de QTL.

4.1.1. Análise de segregação de locos individuais

Previamente à realização do mapeamento genético nas populações F_2 simuladas, foram feitos testes de qui-quadrado a fim de verificar a razão de segregação de locos individuais. Para as 3.600 populações, observou-se que, em média menos de 2% dos marcadores não segregavam de acordo com o esperado, que é de 1:2:1 ($P < 0,05$), dados não apresentados. Estes resultados serviram como um indicativo da boa qualidade dos dados gerados no processo de simulação, o que foi verificado também por SILVA (2005), CRUZ (2006) e FERREIRA et al. (2006).

4.1.2. Número de grupos de ligação e ordenamento das marcas

O número de grupos de ligação estabelecidos pelo processo de mapeamento de todas as populações F_2 , esteve de acordo com esperado, tendo em vista o genoma originalmente estabelecido por simulação, dados não apresentados. Portanto, recuperaram-se os seis grupos de ligação e não foram observadas marcas não ligadas a estes grupos. Após a formação do grupo de ligação, verificou-se pelo método SARF a melhor ordem dos marcadores constituintes de cada grupo.

Devido ao fato das populações F_2 serem efetivamente grandes (SOLLER et al., 1976, CRUZ, 2006), não foi encontrada inversão entre as marcas adjacentes, mediante comparação ao se comparar com o genoma simulado, CRUZ (2006), trabalhando com populações F_2 e marcadores codominantes, concluiu que, para intervalos entre marcas adjacentes entre 5 a 20 cM, foi desnecessário o emprego de amostras a partir de 400 indivíduos para obtenção de mapas genéticos acurados.

Problemas de mapeamento referentes a número de grupo de ligação, tamanho do grupo de ligação, distância entre pares de marcas, marcas não ligadas, invertidas ou alocadas em outros

grupos são relatados, quando o tamanho da população é inadequado para este tipo de estudo (SILVA, 2005; CRUZ, 2006; FERREIRA et al., 2006).

4.1.3. Construção do mapa de ligação em populações F₂

Os mapas de ligação, obtidos com os 1.000 indivíduos para as 3.600 populações segregantes F₂ reproduziram os 76 marcadores codominantes, distribuídos em seis grupos de ligação. Um dos mapas de ligação, tomado ao acaso aqueles vários obtidos das populações simuladas, é apresentado na Figura 3. Neste mapa de ligação, além de recuperar os seis grupos de ligação, conforme esperado, todos os marcadores encontram-se ligados e seu ordenamento está correto. Os grupos de ligação 1 e 2; 3 e 4; 5 e 6 tiveram, em média, distâncias entre os marcadores adjacentes em torno de 20, 10 e 5 cM, respectivamente.

O número de marcadores moleculares necessários à construção de um mapa de ligação varia, em função do tamanho do genoma, do número de cromossomos e da frequência de recombinação genética. Um mapa pode ser considerado saturado, quando o número de grupos de ligação, obtidos na análise dos marcadores, for igual ao número de cromossomos gaméticos do organismo e quando todos os marcadores genéticos mapeados estiverem ligados, indicando que todas as regiões do genoma estão representadas (FERREIRA & GRATTAPAGLIA, 1998; LANZA et al., 2000). Vários estudos de mapeamento genético têm sido realizados em diversas culturas vegetais, possibilitando a construção de mapas de ligação em milho (KANTETY et al., 1995, BENTO, 2006), arroz (BRONDANI, 2000), soja (SOARES, 2000; MIRANDA, 2002; CERVIGNI, 2003) e feijão (CORRÊA, 1999).

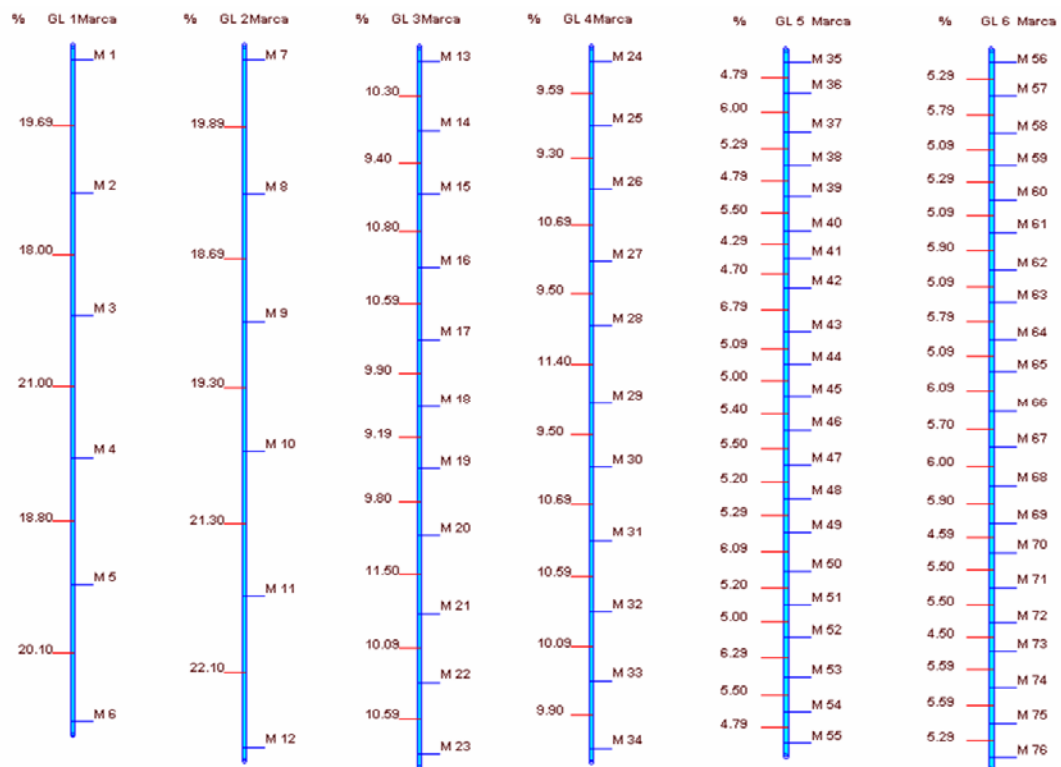


Figura 3 - Mapa genético com 76 marcadores distribuídos nos seis grupos de ligação (GL) de uma das populações F_2 . A distância entre os marcadores foi expressa em centiMorgan (cM).

A partir do mapa de ligação, torna-se possível a obtenção de informações importantes para o melhoramento genético de uma espécie, desde a associação de marcadores moleculares com características quantitativas e sua localização nos grupos de ligação até a identificação de regiões genômicas, associadas às características quantitativas (FERREIRA & GRATTAPAGLIA, 1998). Embora não seja o único método para o estudo da associação de característica quantitativa com marcadores moleculares (TANKSLEY et al., 1982; EDWARDS et al., 1987; OSBORN et al., 1987), o desenvolvimento do mapa genético oferece condições para avaliação dos marcadores, distribuídos por todo o genoma da espécie em estudo, maximizando a probabilidade

de encontrar associações significativas. Uma vez encontradas essas associações, os marcadores podem ser utilizados na seleção indireta de características de interesse agrônômico.

A construção de um mapa gera, ainda, um grande número de informações sobre a estrutura e organização do genoma da espécie estudada, tais como padrões de distorção de segregação mendeliana de segmentos cromossômicos, ou a presença de inversões, translocações e duplicações de segmentos de DNA (TANKSLEY et al., 1982; FERREIRA, 2006).

4.1.4. Comparação dos genomas

Os 3.600 mapas de ligação, obtidos neste estudo, foram comparados ao genoma simulado. Portanto, 21.600 grupos de ligação foram analisados, com seu respectivo grupo do genoma simulado.

O tamanho de cada grupo de ligação foi obtido, por meio de sua média aritmética, nas 3.600 repetições das populações F_2 segregantes. Os valores de tamanho médio do grupo de ligação (TAM) e seus respectivos desvios-padrão encontram-se na Tabela 4. O valor do TAM esperado de cada grupo de ligação foi de 100 cM, de acordo com o genoma simulado (Figura 1). As saturações de 5 e 10 cM tiveram as maiores médias para TAM, em comparação com a saturação de 20 cM. Por outro lado, as variâncias para os TAMs foram bem semelhantes, sendo que a razão entre a maior e menor variância foi de 1,58. A falta de aditividade das distâncias expressas em frequências de recombinação pode ser associada, como uma das causas da diferença para os TAMs em diferentes níveis de saturação. Para minimizar esta limitação, foram propostas várias funções de mapeamento para conversão da frequência de recombinação, expressas em unidades de mapa (cM = centiMorgan), em medidas de distâncias com propriedades mais interessantes para o ordenamento (SCHUSTER & CRUZ, 2004). As funções de

mapeamento mais conhecidas são as de MORGAN (1917), de HALDANE (1919) e de KOSAMBI (1944).

O valor esperado da distância média entre marcas adjacentes (DIM) foi de 5, 10 ou 20 cM (Figura 1) para o genoma das populações segregantes geradas. Observou-se que, na medida em que se reduziu o nível de saturação do grupo de ligação, aumentaram-se os valores dos desvios-padrão de DIM, isto é, quanto mais saturado o grupo de ligação, mais preciso foi a DIM (Tabela 4).

Os valores da variância média das distâncias entre marcas adjacentes (VAD) e seus desvios-padrão também se encontram na Tabela 4. Observou-se que, quanto menor os valores de variância, mais equidistantes estavam os marcadores dentro do grupo de ligação. Assim, a obtenção de VAD de menor magnitude e DIM próximos aos valores esperados, em cada nível de saturação do grupo de ligação, indicam boa recuperação do genoma com o mapeamento das populações segregantes em relação ao genoma simulado (SILVA, 2005).

Tabela 4 - Tamanho médio de grupos de ligação, distância média entre marcas adjacentes, variância média das distâncias entre marcas adjacentes, estresse e seus respectivos desvios-padrão entre parênteses das 3.600 populações F₂ analisadas, para cada grupo de ligação

GL	Saturação	TAM	DIM	VAD	EST
GL 1	20 cM	100,14 (1,93)	19,93 (1,07)	1,71 (1,16)	6,34 (4,59)
GL 2		100,55 (1,88)	20,02 (1,16)	1,85 (1,22)	6,56 (4,68)
GL 3	10 cM	102,67 (2,36)	10,27 (0,24)	0,56 (0,26)	7,74 (1,88)
GL 4		102,45 (2,11)	10,14 (0,27)	0,49 (0,21)	7,66 (1,85)
GL 5	5 cM	102,69 (2,18)	5,13 (0,11)	0,25 (0,09)	10,23 (1,86)
GL 6		102,98 (2,32)	5,26 (0,09)	0,22 (0,12)	10,31 (1,82)

GL = grupo de ligação; TAM = tamanho médio do grupo de ligação em centiMorgan (cM); DIM = distância média entre marcas adjacentes em centiMorgan (cM); VAD = variância média das distâncias entre marcas adjacentes; EST = estresse médio em porcentagem (%).

Os valores de estresse médio (EST) e seu desvio-padrão para cada grupo de ligação são também, apresentados Tabela 4. O estresse médio é uma média aritmética dos valores de estresse obtidos das repetições para cada grupo de ligação.

O valor do estresse expressa o grau de concordância dos valores de distância entre cada par de marcas adjacentes, no grupo de ligação obtido no mapeamento das populações simuladas, com as distâncias dos respectivos pares de marcas no genoma, utilizado para simulação das populações. As estimativas de estresse são utilizadas, para avaliar as distorções das escalas multidimensionais pela simplificação do espaço p-dimensional para bi ou tridimensional (KRUSKAL, 1964; CRUZ & CARNEIRO, 2003, SILVA, 2005).

Apesar de valores médios de estresses apresentarem diferenças entre os níveis de saturação, o valor de estresse possui significados diferentes quando se comparam genomas com níveis de saturação distintos, conforme reportado por SILVA (2005). Dentro de um mesmo tamanho de população, os valores de estresse são maiores para os genomas mais saturados. Contudo, SILVA (2005) decodificou em valores de desvio médio aproximado, verificando que os menores valores de desvio estão associados aos genomas mais saturados, em populações de RILs. Assim, deve-se ter cautela para fazer qualquer inferência sobre o estresse de diferentes saturações.

Os valores de correlação de Spearman foram iguais a uma unidade, indicando que a ordem das marcas no grupo de ligação, obtido no mapeamento das populações segregantes, não foi alterada em relação à ordem previamente conhecida no genoma, o qual foi utilizado para a geração das populações.

4.2. Detecção de QTL em populações F₂

A capacidade de detectar e localizar um QTL é função da magnitude do seu efeito sobre a característica, do tamanho da população segregante avaliada, da frequência de recombinação entre marcador e QTL, bem como da herdabilidade da característica analisada. Portanto, quanto maior o efeito, o tamanho da população e a herdabilidade, e mais próximo o marcador do QTL, mais fácil será a sua detecção (FERREIRA & GRATTAPAGLIA, 1998; LANZA et al., 2000).

Para a maioria dos caracteres, têm sido identificados QTL com efeitos maiores, mas a maioria é de pequeno efeito. Este resultado é evidente, quando se considera que a maioria dos genes segregantes em uma população, provavelmente, têm algum pequeno efeito em outros caracteres. Este pequeno efeito de um QTL pode ainda ser detectado pelos marcadores, dependendo dos seguintes fatores: a) distância entre marcador e o QTL - quanto mais próximos estiverem, menor será a recombinação e, portanto mais estável é este marcador; b) tamanho da população segregante - quanto maior a população, maior a probabilidade de detecção de QTL com menor efeito fenotípico; c) herdabilidade da característica quantitativa - quanto maior o efeito do ambiente em um caráter, isto é, menor a herdabilidade, menor será a probabilidade de um QTL de grande efeito ser detectado; d) nível de probabilidade de erro usado para declarar um efeito do QTL como significativo - quanto mais alta a precisão exigida, menores as probabilidades de identificar QTL de efeito menor. Em milho, usando uma população F₂ com 1.700 indivíduos e um nível de probabilidade de erro de 5%, EDWARDS (1987) detectou um QTL que contribuiu com apenas 0,3% da variância fenotípica. Em experimentos com populações pequenas e alto nível de probabilidade de erro, normalmente, não se detectam, QTL que expliquem menos de 3% da variância fenotípica (TANKSLEY, 1993).

A literatura mostra alguns casos, em que se observou uma grande proporção da variação quantitativa, explicada pela segregação de poucos QTL de grandes efeitos. A proporção

cumulativa, explicada pelos vários QTL detectados experimentalmente, tem variado de 30 até 70%, dependendo de uma série de fatores, tais como o cruzamento analisado (BEAVIS et al., 1991), a característica avaliada (EDWARDS et al., 1992), o delineamento experimental e a resolução do mapa em termos de número de marcadores (EDWARDS, 1987). Não é raro encontrar QTL, que explica mais de 20% da variação fenotípica em uma população, sendo que valores tão altos quanto 42% têm sido reportados para um único QTL (DOEBLEY & STEC, 1991). Vale ressaltar que a variância fenotípica inclui a variância ambiental, de forma que a variância genética, atribuída a este QTL, deve ser maior.

A metodologia de marca simples para detecção de QTL é amplamente utilizada, pois, não é necessário o conhecimento prévio do mapa de ligação. O método de marca simples foi utilizado no mapeamento de características agronômicas em cana-de-açúcar (GUIMARÃES, 1999). Os resultados indicaram que as características número de colmos, porcentagem de colmos por inflorescência, diâmetro médio de dois colmos, peso dos colmos, porcentagem de fibra e teor de sacarose no colmo (%) explicaram entre 7 e 17% da variância fenotípica destas características, e apenas 13% das associações tiveram R^2 entre 19 e 23%, não sendo identificados QTL com efeitos maiores.

Em eucalipto, utilizando a mesma metodologia, ROCHA (2004) observou a existência de 5, 8, 7, 8, 7, 11 e 7 marcas, próximas aos possíveis locos controladores da expressão das características densidade da madeira, rendimento de polpa, teor de extrativos, teor de lignina insolúvel, teor de lignina solúvel, crescimento de madeira e altura comercial em eucalipto, respectivamente.

Neste estudo, foi proposto investigar, via simulação, a influência de três níveis de herdabilidade, quatro de ação gênica e três de saturações para o grupo de ligação na detecção do QTL. Assim, foram esquematizados 36 cenários diferentes para a característica quantitativa ou fenotípica, para populações segregantes do tipo F_2 com 1.000 indivíduos e replicadas 100 vezes,

totalizando 3.600 populações. A característica quantitativa teve média de 1.000 unidades, coeficiente de variação de 10% e grau médio de dominância igual a zero. Os cenários foram divididos em quatro níveis de ação gênica (5, 10, 20 e 40%), três níveis de herdabilidade (80, 60 e 20%), totalizando 12 situações. Foram simulados três locos controladores de característica quantitativa, o Q_{11} , Q_{31} e Q_{51} para os grupos de ligação 1, 3 e 5, respectivamente. A característica fenotípica simulada foi controlada por vários locos; no entanto, apenas um QTL foi declarado presente em cada cenário. Assim, os cenários das POP 1 a 12, das POP 13 a 24 e das POP 25 a 36 tiveram a presença do loco Q_{11} , Q_{31} e Q_{51} , respectivamente.

As 3.600 populações F_2 foram avaliadas quanto à presença ou não do QTL, dentro de cada cenário, utilizando-se a metodologia de marca simples e intervalo simples. Em relação ao método de marca simples, optou-se pela análise de regressão, uma vez que esta permite a obtenção de estimativas da contribuição do marcador para a variação da característica, além de permitir a quantificação dos efeitos aditivos e devido à dominância do marcador. Quanto ao método de intervalo simples, foi realizado também por meio de análise de regressão, pois, a obtenção das estimativas foi muito mais rápida, em relação à análise de máxima verossimilhança.

4.2.1. Detecção da presença do QTL

Foram feitas análises para detecção de QTL por meio de regressão linear, pelo método de marca simples nas 3.600 populações F_2 geradas. A detecção do QTL foi baseada no teste de associação dos marcadores moleculares, focalizando as marcas flanqueadoras do QTL simulado. Sendo as marcas C_{13} e C_{14} para as POP 1 a 12, C_{36} e C_{37} para as POP 13 a 24 e C_{511} e C_{512} para as POP 25 a 36 com a característica fenotípica simulada de acordo com os cenários do Tabela 1.

Para a verificação e declaração estatística da presença do QTL, utilizou-se na análise para cada marcador, o nível crítico de 5,99 equivalente a um “LOD score” de 1,30 e o nível de significância de 5% de probabilidade com 2 graus de liberdade para uma população F_2 (SCHUSTER & CRUZ, 2004).

Foram feitas comparações entre o número de repetições, dentro de cada cenário, que possuía LOD menor que o “LOD score”. Assim, estimou-se o erro tipo I para os marcadores flanqueadores do QTL (Tabela 5).

Quando a herdabilidade da característica simulada foi de 80%, observou-se que, independente da ação gênica e da saturação do grupo de ligação, foi possível detectar o QTL em todas as populações, dentro de seu respectivo cenário, utilizando-se o método de marca simples. Para a herdabilidade de 60%, apenas, quando ação gênica foi de 5%, observou-se populações com o LOD inferior ao “LOD score” de 1,30 (Tabela 5). Nas populações com herdabilidade de 20% para o caráter, com níveis de ação gênica de 5 e 10%, independente da saturação, observou-se que o erro tipo I estimado foi maior ou igual a 4%. Neste nível de herdabilidade, à medida que aumentou o nível de saturação do grupo de ligação, foi observada tendência na diminuição da magnitude do erro tipo I.

Tabela 5 - Número de repetições nas quais foi observada a presença do QTL, com LOD menor que o “LOD score” de 1,30 pelo método de marca simples, nos cenários das POP 1 a 12, das POP 13 a 24 e das POP 25 a 36, nos marcadores flanqueadores do QTL simulado nas populações F₂ analisadas. O número total de repetições avaliadas foi 100 para cada marcador, em cada cenário

h ²	Ação gênica	Saturação					
		20 cM		10 cM		5 cM	
		C ₁₃	C ₁₄	C ₃₆	C ₃₇	C ₅₁₁	C ₅₁₂
80%	5%	0	0	0	0	0	0
	10%	0	0	0	0	0	0
	20%	0	0	0	0	0	0
	40%	0	0	0	0	0	0
60%	5%	4	3	1	0	1	0
	10%	0	0	0	0	0	0
	20%	0	0	0	0	0	0
	40%	0	0	0	0	0	0
20%	5%	42	47	28	30	20	25
	10%	12	12	6	5	4	5
	20%	0	0	0	0	0	0
	40%	0	0	0	0	0	0

h² = herdabilidade em porcentagem; C₁₃ = marcador três do grupo de ligação um; C₁₄ = marcador quatro do grupo de ligação um; C₃₆ = marcador seis do grupo de ligação três; C₃₇ = marcador sete do grupo de ligação três; C₅₁₁ = marcador 11 do grupo de ligação cinco; C₅₁₂ = marcador 12 do grupo de ligação cinco.

Foi utilizado o nível crítico mais restritivo (LOD = 3), a fim de verificar a ligação entre um marcador e a característica. Verificou-se, em todos os níveis de herdabilidade, a ocorrência de pelo menos um cenário com mais de 5% das observações com LOD menor que três (Tabela 6).

Quando o nível de herdabilidade era de 80%, apenas a situação da POP 1, com ação gênica de 5% e saturação 20 cM, teve erro tipo I superior a 5%, porém a magnitude foi menor que 10%. Nas populações com o nível 60% de herdabilidade, verificou-se em qualquer nível de saturação, para 5% de ação gênica, nos cenários das POP 5, POP 17 e POP 19, em que estas tiveram mais de 5% dos LODs menores que três. Quando o nível de herdabilidade era de 20%,

somente as situações com nível de ação gênica de 40% tiveram menos de 5% dos LODs abaixo de três. Neste nível de herdabilidade em relação à Tabela 5, que teve o nível crítico de 1,30, incluiu-se apenas a situação da POP 11, que teve mais 5% dos valores dos LODs superiores a três.

Tabela 6 - Número de repetições nas quais foi observada a presença do QTL, com LOD menor que três, pelo método de marca simples, nos cenários das POP 1 a 12, das POP 13 a 24 e das POP 25 a 36, nos marcadores flanqueadores do QTL simulado nas populações F₂ analisadas. O número total de repetições avaliadas foi 100 para cada marcador, em cada cenário

h ²	Ação gênica	Saturação					
		20 cM		10 cM		5 cM	
		C ₁₃	C ₁₄	C ₃₆	C ₃₇	C ₅₁₁	C ₅₁₂
80%	5%	5	9	4	4	0	1
	10%	1	1	0	0	0	0
	20%	0	0	0	0	0	0
	40%	0	0	0	0	0	0
60%	5%	32	30	14	12	5	6
	10%	0	1	0	1	0	0
	20%	0	0	0	0	0	0
	40%	0	0	0	0	0	0
20%	5%	85	88	83	78	73	76
	10%	57	50	38	35	26	30
	20%	14	11	0	0	3	1
	40%	0	0	0	0	0	0

h² = herdabilidade em porcentagem; C₁₃ = marcador três do grupo de ligação um; C₁₄ = marcador quatro do grupo de ligação um; C₃₆ = marcador seis do grupo de ligação três; C₃₇ = marcador sete do grupo de ligação três; C₅₁₁ = marcador 11 do grupo de ligação cinco; C₅₁₂ = marcador 12 do grupo de ligação cinco.

Na Tabela 6, além da inclusão de cinco novos cenários, cujos erros tipo I foram superiores a 5%, verificou-se também o aumento na magnitude deste nos demais cenários. No entanto, assim como na Tabela 5, à medida que aumenta o nível de saturação do grupo de ligação, há tendência na diminuição da magnitude do erro tipo I.

4.2.2. Detecção da presença QTL no grupo de ligação

Para efeito de comparação de metodologias, a detecção da presença do QTL foi realizada também por meio de regressão linear, pelo método de intervalo simples nas 3.600 populações F_2 . Este método baseia-se no teste de associação de pares de marcas moleculares adjacentes. Foram analisados, somente, os grupos de ligação 1 para as POP 1 a 12, 3 para as POP 13 a 24 e, 5 para as POP 25 a 36, com a característica quantitativa.

Para verificação da presença do QTL, utilizou-se o nível significância de 5% para cada grupo de ligação. Os níveis críticos e seus respectivos “LOD score” encontram-se na Tabela 3.

Em cada cenário, foram feitas comparações entre o número de repetições, que possuíam LOD máximo ($LOD_{máx}$) menor que o “LOD score” de 1,99; 2,29 e 2,59, para os grupos de ligação 1, 3 e 5, respectivamente. Estes números possibilitaram estimar o erro tipo I na detecção do QTL, em cada grupo de ligação (Tabela 7).

Nos níveis de herdabilidade de 80 e 60% observou-se que, independente da ação gênica e da saturação do grupo de ligação, todos os cenários tiveram menos de 5% dos $LOD_{máx}$ inferiores aos “LOD score”, considerado para cada grupo de ligação (Tabela 7). Quando o nível de herdabilidade para o caráter foi de 20%, para os níveis de ação gênica de 5 e 10%, independente da saturação, observou-se que o erro tipo I estimado foi maior ou igual a 5%. Entretanto, ao contrário do que foi verificado nas análises de marca simples, na medida em que aumentou o nível de saturação do grupo de ligação não foi observada tendência na diminuição da magnitude do erro tipo I. Este fato foi associado à adoção de “LOD score” diferentes para cada grupo de ligação (HALEY & KNOTT, 1992; ZENG, 1994; XU & ATCHLEY, 1995).

Tabela 7 - Número de repetições nas quais foi observado o efeito do QTL pela análise de intervalo simples com LOD máximo menor que “LOD score” 1,99; 2,29 e 2,59, nos grupos de ligação 1 (POP 1 a 12), 3 (POP 13 a 24) e 5 (POP 25 a 36), respectivamente. O número total de repetições avaliadas foi 100 para cada cenário

h ²	Ação gênica	Saturação		
		20 cM	10 cM	5 cM
80%	5%	0	0	0
	10%	0	0	0
	20%	0	0	0
	40%	0	0	0
60%	5%	3	1	1
	10%	0	0	0
	20%	0	0	0
	40%	0	0	0
20%	5%	39	43	41
	10%	12	11	15
	20%	0	0	0
	40%	0	0	1

h² = herdabilidade em porcentagem; Q₁₁ = QTL do grupo de ligação um; Q₃₁ = QTL do grupo de ligação três; Q₅₁ = QTL do grupo de ligação cinco.

Utilizou-se também o nível crítico mais restritivo (LOD = 3), para verificar o erro tipo I na detecção de QTL (Tabela 8). Verificou-se o mesmo padrão da Tabela 7, com exceção do cenário da POP 5, que teve para erro tipo I valor igual a 10% (Tabela 8). No entanto, as magnitudes do erro tipo I foram superiores aquelas observadas na Tabela 7.

Tabela 8 - Número de repetições nas quais foi observado o efeito do QTL pela análise de intervalo simples com LOD máximo menor que três, nos grupos de ligação 1 (POP 1 a 12), 3 (POP 13 a 24) e 5 (POP 25 a 36). O número total de repetições avaliadas foi 100 para cada cenário

h ²	Ação gênica	Saturação		
		20 cM	10 cM	5 cM
80%	5%	1	1	0
	10%	0	0	0
	20%	0	0	0
	40%	0	0	0
60%	5%	10	3	3
	10%	0	0	0
	20%	0	0	0
	40%	0	0	0
20%	5%	77	68	59
	10%	36	20	22
	20%	4	0	0
	40%	0	0	1

h² = herdabilidade em porcentagem; Q₁₁ = QTL do grupo de ligação um; Q₃₁ = QTL do grupo de ligação três; Q₅₁ = QTL do grupo de ligação cinco.

4.2.3. Detecção do QTL pelo método de marca simples e sua contribuição para a variação da característica

As análises de regressão pelo método de marca simples, para os marcadores flanqueadores dos QTL dos grupos de ligação 1, 3 e 5, possibilitaram a estimação dos valores do LOD e do coeficiente de determinação (R²). Os resultados encontram-se nas tabelas 9, 10 e 11, respectivamente.

Observou-se que a razão entre o LOD médio de uma dada herdabilidade pelo LOD médio de outra herdabilidade, para o mesmo nível de ação gênica, foi aproximadamente igual à razão entre as herdabilidades (Tabelas 9, 10 e 11). Por exemplo, considerando o LOD médio do marcador C₅₁₁ igual a 8,62 (Tabela 11) com nível de herdabilidade de 80% e ação gênica de 5% e o LOD médio do mesmo marcador igual a 6,37, porém com nível de herdabilidade de 60% e

mesma ação gênica, a razão foi de 1,35 e a razão entre as herdabilidades foi 1,33. Pode-se inferir que o valor médio do LOD é estimável, para um nível de herdabilidade intermediário de um dado nível de ação gênica e saturação do grupo de ligação.

Tabela 9 - Estimativas médias do LOD e R^2 do modelo de regressão linear por marca simples, e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica para os marcadores C_{13} e C_{14} do grupo de ligação 1. O número total de repetições avaliadas foi de 100 para cada situação (POP 1 a POP 12)

h^2	Ação gênica	\overline{LOD}		$\overline{R^2}$	
		C_{13}	C_{14}	C_{13}	C_{14}
80%	5%	6,17 (2,40)	5,92 (2,33)	2,80 (1,07)	2,69 (1,04)
	10%	11,38 (3,27)	10,99 (3,27)	5,10 (1,43)	4,92 (1,43)
	20%	23,93 (4,98)	23,70 (5,09)	10,41 (2,06)	10,32 (2,10)
	40%	47,53 (7,98)	48,18 (7,85)	19,61 (2,94)	19,85 (2,88)
60%	5%	4,22 (2,03)	4,28 (1,93)	1,92 (0,91)	1,95 (0,87)
	10%	8,66 (2,39)	8,78 (2,47)	3,90 (1,05)	3,96 (1,09)
	20%	16,80 (4,02)	16,23 (4,05)	7,43 (1,71)	7,19 (1,72)
	40%	34,50 (5,98)	34,37 (6,15)	14,66 (2,34)	14,61 (2,41)
20%	5%	1,79 (1,21)	1,67 (1,22)	0,82 (0,55)	0,76 (0,55)
	10%	3,08 (1,74)	3,01 (1,57)	1,41 (0,79)	1,38 (0,71)
	20%	5,54 (2,32)	5,68 (2,38)	2,51 (1,04)	2,58 (1,07)
	40%	10,92 (3,34)	11,11 (3,52)	4,89 (1,45)	4,98 (1,54)

h^2 = herdabilidade em porcentagem; \overline{LOD} = média do LOD; $\overline{R^2}$ = média do coeficiente de determinação; C_{13} = marcador três do grupo de ligação um; C_{14} = marcador quatro do grupo de ligação um.

Tabela 10 - Estimativas médias do LOD e R^2 do modelo de regressão linear por marca simples, e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica para os marcadores C_{36} e C_{37} do grupo de ligação 3 (POP 13 a 24)

h^2	Ação gênica	\overline{LOD}		$\overline{R^2}$	
		C_{36}	C_{37}	C_{36}	C_{37}
80%	5%	7,45 (2,86)	7,39 (2,65)	3,36 (1,27)	3,34 (1,17)
	10%	14,06 (3,92)	13,94 (3,75)	6,26 (1,69)	6,21 (1,61)
	20%	29,73 (5,33)	29,50 (5,85)	12,77 (2,13)	12,67 (2,33)
	40%	62,88 (9,36)	62,70 (8,66)	25,08 (3,20)	25,03 (2,97)
60%	5%	5,38 (2,05)	5,35 (2,15)	2,44 (0,92)	2,43 (0,96)
	10%	10,57 (3,14)	10,22 (3,05)	4,74 (1,37)	4,59 (1,33)
	20%	22,25 (4,85)	21,91 (4,41)	9,72 (2,00)	9,58 (1,82)
	40%	45,46 (6,66)	45,71 (6,84)	18,86 (2,48)	18,95 (2,52)
20%	5%	2,04 (1,25)	2,13 (1,35)	0,94 (0,57)	0,98 (0,61)
	10%	3,63 (1,61)	3,81 (1,77)	1,66 (0,73)	1,74 (0,80)
	20%	7,85 (2,65)	7,59 (2,74)	3,54 (1,17)	3,43 (1,22)
	40%	14,64 (3,70)	14,67 (3,60)	6,51 (1,59)	6,52 (1,54)

h^2 = herdabilidade em porcentagem; \overline{LOD} = média do LOD; $\overline{R^2}$ = média do coeficiente de determinação; C_{36} = marcador seis do grupo de ligação três; C_{37} = marcador sete do grupo de ligação três.

Tabela 11 - Estimativas médias do LOD e R^2 do modelo de regressão linear por marca simples, e seus desvios-padrão, entre parênteses, do nível de herdabilidade e ação gênica para os marcadores C_{511} e C_{512} do grupo de ligação 5 (POP 25 a 36)

h^2	Ação gênica	\overline{LOD}		$\overline{R^2}$	
		C_{511}	C_{512}	C_{511}	C_{512}
80%	5%	8,62 (3,08)	8,49 (2,92)	3,88 (1,36)	3,83 (1,28)
	10%	15,75 (3,82)	15,50 (3,94)	6,98 (1,63)	6,88 (1,68)
	20%	32,93 (6,38)	33,10 (6,40)	14,04 (2,50)	14,10 (2,51)
	40%	72,64 (9,09)	72,19 (9,44)	28,38 (2,97)	28,22 (3,09)
60%	5%	6,37 (2,20)	6,27 (2,25)	2,89 (0,98)	2,84 (1,01)
	10%	11,90 (3,55)	11,91 (3,44)	5,32 (1,54)	5,33 (1,50)
	20%	24,28 (5,11)	24,18 (4,78)	10,56 (2,10)	10,52 (1,96)
	40%	51,95 (8,24)	52,46 (8,25)	21,23 (2,98)	21,41 (2,98)
20%	5%	2,32 (1,34)	2,35 (1,45)	1,06 (0,61)	1,08 (0,66)
	10%	4,34 (2,06)	4,26 (2,14)	1,97 (0,92)	1,94 (0,96)
	20%	8,23 (2,78)	8,22 (2,90)	3,71 (1,23)	3,71 (1,28)
	40%	16,10 (3,76)	16,10 (3,93)	7,13 (1,60)	7,13 (1,68)

h^2 = herdabilidade em porcentagem; \overline{LOD} = média do LOD; $\overline{R^2}$ = média do coeficiente de determinação; C_{511} = marcador 11 do grupo de ligação cinco; C_{512} = marcador 12 do grupo de ligação cinco.

Independente da posição do marcador, se anterior ou posterior ao QTL simulado, em qualquer grupo de ligação, todas as médias e desvios-padrão do LOD e R^2 foram muito semelhantes em magnitude e tendência (Tabelas 9, 10 e 11). Esse fato foi devido à posição equidistante entre os marcadores e o QTL simulado.

Verificaram-se também perdas nos valores dos R^2 médios, estimados em relação ao nível de ação gênica paramétrico, em todos os cenários estudados, independentemente do grupo de ligação analisado. Para o grupo de ligação 1, as perdas médias do R^2 nos níveis de herdabilidade de 80, 60 e 20 % foram, respectivamente, 50, 62 e 86%. A razão entre o maior desvio-padrão do R^2 médio e o menor foi de 5,34.

Em relação ao grupo de ligação 3, detectou-se uma perda do R^2 em torno de 36% quando o nível de herdabilidade foi de 80%. No entanto, para o nível de herdabilidade de 60%, a perda média foi de 52% e no nível de 20 % a perda foi de aproximadamente 83%. A razão entre o maior desvio-padrão do R^2 médio e o menor foi de 5,61.

No grupo de ligação 5, quando o nível de herdabilidade foi de 80%, houve uma perda média de 36%, para o nível de herdabilidade de 60% a perda foi de 46% e no nível de 20 % a perda chegou a 81%. A razão entre o maior desvio-padrão do R^2 médio e o menor foi de 5,06.

Além disso, pode-se inferir que o R^2 médio para herdabilidade de 20%, em todos os grupos de ligação, teve uma perda superior a 80%, em relação ao nível de ação gênica paramétrico, mesmo para os níveis de ação gênica de 20 e 40%. Por outro lado, independente do grupo de ligação, nos níveis de herdabilidades de 60 e 80%, os R^2 médios tiveram suas perdas reduzidas, significativamente, à medida que o nível de saturação foi aumentado. Considerando os cenários das POP 1 a 12 com o QTL simulado presente no grupo de ligação 1 (20 cM) com referência, verificou-se que o incremento médio do R^2 foi 27 e 47% para os cenários das POP 13 a 14 grupo de ligação 3 (10 cM) e, das POP 25 a 36 grupo de ligação 5 (5 cM), respectivamente.

4.2.4. Detecção e localização do QTL pelo método de intervalo simples e sua contribuição para a variação da característica

Foram realizadas análises de regressão pelo método de intervalo simples, para os grupos de ligação 1, 3 e 5, pois, além de permitir a detecção do QTL e sua contribuição para a variação da característica (R^2), possibilitou estimar a localização do QTL ao longo do grupo de ligação. Os valores médios de LOD máximo, da posição relativa do QTL e do coeficiente de determinação e seus desvios-padrão, referentes aos grupos de ligação 1, 3 e 5 são apresentados nas tabelas 12, 13 e 14, respectivamente.

Tabela 12 - Estimativas médias do LOD máximo, da posição relativa do QTL e do coeficiente de determinação e seus respectivos desvios-padrão, entre parênteses, obtidos por regressão linear, utilizando o método de intervalo simples, em função do nível de herdabilidade e de ação gênica do grupo de ligação 1 (POP 1 a 12)

h^2	Ação gênica	$\overline{LOD}_{\text{máx}}$	$\bar{r}_{\text{máx}}$	\overline{R}^2
80%	5%	7,79 (2,67)	50,10 (6,65)	3,52 (1,18)
	10%	14,30 (3,96)	50,59 (3,42)	6,37 (1,71)
	20%	31,84 (6,70)	51,11 (2,10)	13,61 (2,67)
	40%	69,93 (12,02)	51,36 (1,99)	27,45 (3,97)
60%	5%	5,49 (2,24)	51,24 (8,90)	2,49 (1,01)
	10%	11,16 (2,81)	51,38 (4,22)	5,01 (1,23)
	20%	21,60 (4,87)	50,92 (3,02)	9,46 (2,02)
	40%	47,62 (8,17)	51,20 (1,95)	19,65 (3,01)
20%	5%	2,47 (1,30)	47,46 (19,56)	1,13 (0,59)
	10%	4,03 (2,21)	48,64 (13,46)	1,84 (0,99)
	20%	7,16 (2,73)	50,76 (7,44)	3,24 (1,22)
	40%	14,19 (4,14)	51,11 (3,30)	6,31 (1,78)

h^2 = herdabilidade em porcentagem; $\overline{LOD}_{\text{máx}}$ = média do LOD máximo; $\bar{r}_{\text{máx}}$ = média da posição relativa do QTL; \overline{R}^2 = média do coeficiente de determinação.

Tabela 13 - Estimativas médias do LOD máximo, da posição relativa do QTL e do coeficiente de determinação e seus respectivos desvios-padrão, entre parênteses, obtidos por regressão linear, utilizando o intervalo simples, em função do nível de herdabilidade e de ação gênica do grupo de ligação 3 (POP 13 a 24)

h^2	Ação gênica	$\overline{LOD}_{\text{máx}}$	$\bar{r}_{\text{máx}}$	\overline{R}^2
80%	5%	8,64 (3,03)	56,22 (3,15)	3,89 (1,33)
	10%	16,25 (4,46)	56,46 (2,33)	7,20 (1,90)
	20%	35,47 (7,62)	56,50 (1,91)	15,03 (2,98)
	40%	83,23 (15,66)	55,53 (1,40)	31,69 (5,07)
60%	5%	6,34 (2,27)	56,12 (3,55)	2,87 (1,01)
	10%	12,06 (3,50)	56,57 (2,85)	5,40 (1,52)
	20%	26,24 (5,70)	56,55 (2,01)	11,37 (2,30)
	40%	57,86 (9,31)	55,04 (1,18)	23,34 (3,23)
20%	5%	2,72 (1,41)	56,59 (9,75)	1,24 (0,64)
	10%	4,45 (1,82)	55,53 (5,04)	2,02 (0,82)
	20%	8,98 (2,96)	56,46 (3,20)	4,05 (1,31)
	40%	17,02 (4,23)	56,22 (3,15)	7,53 (1,79)

h^2 = herdabilidade em porcentagem; $\overline{LOD}_{\text{máx}}$ = média do LOD máximo; $\bar{r}_{\text{máx}}$ = média da posição relativa do QTL; \overline{R}^2 = média do coeficiente de determinação.

Tabela 14 - Estimativas médias e seus respectivos desvios-padrão, entre parênteses, do LOD máximo, da posição relativa do QTL e do coeficiente de determinação, obtidos por regressão linear, utilizando o método de intervalo simples, em função do nível de herdabilidade e de ação gênica do grupo de ligação 5 (POP 25 a 36)

h^2	Ação gênica	$\overline{LOD}_{\text{máx}}$	$\bar{r}_{\text{máx}}$	\overline{R}^2
80%	5%	9,58 (3,19)	53,84 (4,31)	4,31 (1,40)
	10%	17,43 (4,32)	53,64 (2,58)	7,70 (1,83)
	20%	38,11 (7,82)	54,20 (1,98)	16,06 (2,99)
	40%	92,58 (13,64)	54,35 (2,60)	34,61 (4,04)
60%	5%	7,03 (2,28)	52,96 (6,11)	3,18 (1,02)
	10%	13,18 (3,87)	54,50 (3,75)	5,88 (1,67)
	20%	27,32 (5,94)	54,19 (2,61)	11,80 (2,42)
	40%	62,99 (11,62)	54,27 (1,74)	25,10 (4,02)
20%	5%	2,98 (1,47)	53,02 (18,29)	1,36 (0,67)
	10%	4,93 (2,24)	53,77 (11,53)	2,24 (1,00)
	20%	9,14 (2,98)	54,49 (4,66)	4,12 (1,31)
	40%	17,71 (4,57)	54,00 (2,61)	7,82 (1,94)

h^2 = herdabilidade em porcentagem; $\overline{LOD}_{\text{máx}}$ = média do LOD máximo; $\bar{r}_{\text{máx}}$ = média da posição relativa do QTL; \overline{R}^2 = média do coeficiente de determinação.

O valor do $LOD_{m\acute{a}x}$ médio foi 23% superior, aproximadamente, em relação às médias dos LODs dos marcadores flaqueadores do QTL, obtidos pelo método de marca simples, em qualquer cenário, independente da saturação do grupo de ligação. Isto justifica o fato de haver mais cenários com erro tipo I na Tabela 6 do que na Tabela 8.

Em relação ao grupo de ligação 1 (POP 1 a POP 12), a média do $r_{m\acute{a}x}$ variou de 47,46 cM a 51,38 cM (Tabela 12), sendo que a frequência de recombinação esperada (r) foi de 50 cM. Para o grupo de ligação 3 (POP 13 a 24), a média $r_{m\acute{a}x}$ variou entre 55,04 cM e 56,59 cM (Tabela 13), sendo que o valor esperado do “ r ” foi de 55 cM. No grupo de ligação 5 (POP 25 a 36), a variação da média do $r_{m\acute{a}x}$ ficou entre 52,96 cM e 54,50 cM (Tabela 14) e “ r ” esperado foi de 52,5 cM.

Como nas análises de marca simples, detectou-se também, pelo método de intervalo simples, perdas do R^2 médio (referente ao $LOD_{m\acute{a}x}$) estimado em relação aos níveis de ação gênica paramétrico, em todos os cenários e independente do grupo de ligação analisado. Para o grupo de ligação 1, nos níveis de herdabilidade 80, 60 e 20%, foram observadas perdas de 32, 51, e 83%, respectivamente. A razão entre o maior desvio-padrão do R^2 médio e o menor foi de 6,72. No grupo de ligação 3, as perdas nos valores de R^2 médio para as herdabilidades de 80, 60 e 20% foram, respectivamente, 23, 43 e 80%. A razão entre o maior desvio-padrão do R^2 médio e o menor foi de 7,92. Em relação ao grupo de ligação 5, as perdas nos valores médios de R^2 nos níveis de herdabilidade de 80, 60 e 20% foram de 16, 39 e 79%, respectivamente. A razão entre o maior desvio-padrão do R^2 médio e o menor foi de 6,02. O valor médio do R^2 esteve mais próximo ao valor paramétrico da ação gênica, para o nível de herdabilidade alto e maior saturação do grupo de ligação (Tabela 13). À medida que diminui o nível de herdabilidade, os valores da média do R^2 tendem a ser mais próximos, independentemente da saturação do grupo de ligação. Em relação aos desvios do R^2 médio, ao contrário dos desvios da média do $r_{m\acute{a}x}$, à medida que o nível de ação gênica aumentava, também os valores dos desvios aumentavam (Tabelas 12, 13 e 14).

5.2.5. Estimação dos efeitos dos QTL sobre a característica

Em se tratando de um único loco gênico com dois alelos (Q e q), é possível encontrar, numa população, os genótipos QQ, Qq e qq com os respectivos valores genotípicos $\mu+a$, $\mu+d$ e $\mu-d$. Se a constante μ for o ponto médio (média aritmética dos valores genotípicos dos homocigotos), tem-se, em vez de três valores desconhecidos, apenas dois: “a” e “d” (SCHUSTER & CRUZ, 2004). Ao considerar uma população em equilíbrio de Hardy-Weinberg e a escala de valores genotípicos, obtêm-se os estimadores da média genotípica ($\mu_g = \mu + \frac{1}{2}d$) e da variância genotípica ($\sigma_g^2 = \frac{1}{2}\tilde{a}^2 + \frac{1}{4}\tilde{d}^2$) para população F_2 ($p = q = 0,5$).

O efeito aditivo é definido como o desvio da média da população daqueles indivíduos, que receberam um determinado alelo (por exemplo, o Q, para o caso de α_Q). Se os alelos Q e q ocorrem com frequência p e q, respectivamente, o efeito aditivo do alelo Q será α_Q que é igual a $q\alpha$; enquanto o efeito aditivo do alelo q será α_q que é igual a $-p\alpha$. Dessa forma, entende-se que α é o efeito médio de substituição gênica, proporcionado pela diferença entre os efeitos aditivos dos alelos considerados, ou seja, $\alpha = \alpha_Q - \alpha_q$ (CRUZ, 2005). A partir do modelo aditivo-dominante estimam-se também os desvios atribuídos à dominância (δ). Para as interações intra-alélicas os estimadores são: $\delta_{QQ} = -2q^2d$, $\delta_{Qq} = 2pqd$ e $\delta_{qq} = -2p^2d$. Observa-se, que as interações intra-alélicas, ou desvios atribuídos à dominância, são todas expressas, em função, apenas do valor genotípico do heterocigoto, “d”, e da frequência gênica (SCHUSTER & CRUZ, 2004).

A utilização de análise de detecção QTL por regressão, pelo método de marca simples, possibilitou obter estimativas associadas à média genotípica, ao efeito aditivo e ao efeito devido à dominância, para os marcadores próximos ao QTL, nas populações F_2 simuladas. A média

geral foi por meio das estimativas de $\hat{\beta}_0$. No entanto, os efeitos aditivos e devido à dominância estão associados às estimativas de $\hat{\beta}_1$ e $\hat{\beta}_2$, pois: $\hat{\beta}_1 = (1-2r)a$ e $\hat{\beta}_2 = (1-2r)^2 d$.

No caso particular de populações F_2 “a” é igual ao efeito de substituição gênica (α), pois, $p = q = 0,5$ $\alpha = a + d(q-p) = a$. Além disso, nas populações F_2 simuladas o grau médio de dominância foi nulo e indiretamente o valor de “d” também será próximo de zero, pois, todas as estimativas de interação intra-alélicas dependem diretamente do valor de “d”.

As estimativas dos $\hat{\beta}_0$ e $\hat{\beta}_1$ e $\hat{\beta}_2$ encontram-se nas tabelas 15, 16 e 17, para os cenários das POP 1 a 12, POP 13 a 24 e, POP 25 a 36, respectivamente. Nestas tabelas, observa-se que médias das estimativas $\hat{\beta}_0$ variaram de 998,66 a 1002,03 unidades, de 999,22 a 1000,79 unidades e 999,03 a 1001,20 respectivamente. Portanto, independente da saturação do grupo de ligação onde o QTL foi simulado, a média geral estava em torno da paramétrica. Em relação à média da estimativa do $\hat{\beta}_1$, que reflete linearmente o efeito aditivo esperado, observou-se nos cenários que as médias do efeito aditivo foram similares aos efeitos aditivos paramétricos (Tabela 2), utilizando-se a conversão $a = \overline{\hat{\beta}_1} (1-2r)^{-1}$. Os marcadores C_{13} e C_{14} , C_{311} e C_{312} e, C_{511} e C_{512} , estavam afastados do QTL simulado de 10, 5, 2,5 cM, respectivamente. Verificou-se que as médias das estimativas $\hat{\beta}_1$ variaram de 11,94 a 142,27, de 13,96 a 160,46 e de 15,12 a 171,20 para as saturações do grupo de ligação de 20, 10 e 5 cM, respectivamente. A média da estimativa $\hat{\beta}_2$ está associada, diretamente, ao valor de “d” e conseqüentemente ao efeito devido à dominância. Assim, as médias das estimativas de $\hat{\beta}_2$ dos cenários das POP 1 a 36 ficaram em torno de zero, variando de -3,88 a 2,85. O grau médio de dominância realizado pela razão entre as médias das estimativas dos efeitos devido à dominância e dos efeitos aditivos foi, praticamente zero, em qualquer cenário.

Tabelas 15 - Médias das estimativas de $\hat{\beta}_0$, $\hat{\beta}_1$ e $\hat{\beta}_2$ dos marcadores C₁₃ e C₁₄ e, seus respectivos desvios-padrão, entre parênteses, obtidos por regressão linear, utilizando o método de marca simples em função do nível de herdabilidade e de ação gênica (POP 1 a 12) das 3.600 populações F₂ analisadas com 1.000 indivíduos

M	h ²	AG	$\bar{\hat{\beta}}_0$	$\bar{\hat{\beta}}_1$	$\bar{\hat{\beta}}_2$
C ₁₃	80%	5%	999,01 (7,37)	51,69 (10,14)	1,94 (14,62)
		10%	999,79 (6,36)	71,40 (10,31)	0,38 (12,40)
		20%	1002,03 (7,01)	101,91 (10,56)	-3,88 (13,10)
		40%	1000,85 (6,01)	141,72 (9,97)	-1,64 (12,07)
	60%	5%	999,79 (4,75)	30,10 (8,25)	0,44 (9,23)
		10%	999,50 (4,88)	43,60 (5,88)	1,00 (9,54)
		20%	999,77 (4,42)	61,49 (7,04)	0,46 (8,72)
		40%	1000,25 (5,46)	86,67 (6,01)	-0,44 (10,35)
	20%	5%	1000,24 (3,84)	12,36 (5,12)	-0,54 (7,54)
		10%	1000,08 (3,02)	17,92 (5,13)	-0,12 (6,08)
		20%	1000,56 (3,38)	23,99 (5,36)	-1,06 (6,60)
		40%	1000,71 (3,50)	34,74 (4,44)	-1,37 (6,92)
C ₁₄	80%	5%	1000,09 (7,08)	50,54 (10,99)	-0,01 (13,97)
		10%	999,93 (6,91)	69,86 (10,39)	0,41 (13,61)
		20%	1001,31 (6,92)	101,29 (10,98)	-2,18 (13,36)
		40%	1000,40 (7,81)	142,27 (10,52)	-0,36(15,35)
	60%	5%	999,06 (5,15)	30,13 (7,60)	2,01 (10,17)
		10%	998,66 (4,99)	43,73 (6,48)	2,85 (10,07)
		20%	1000,72 (4,82)	60,28 (6,99)	-1,17 (9,37)
		40%	1000,23 (4,87)	86,41 (5,99)	-0,20 (9,27)
	20%	5%	1000,11 (3,51)	11,94 (5,23)	-0,19 (6,93)
		10%	999,78 (3,66)	17,35 (5,34)	0,50 (7,28)
		20%	1000,44 (3,53)	24,23 (5,33)	-0,75 (7,06)
		40%	1000,53 (3,34)	34,96 (5,21)	-0,92 (6,44)

M = marcador; C₁₃ = marcador três do grupo de ligação um; C₁₄ = marcador quatro do grupo de ligação um; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem.

Tabelas 16 - Médias das estimativas de $\hat{\beta}_0$, $\hat{\beta}_1$ e $\hat{\beta}_2$ dos marcadores C₃₆ e C₃₇ e, seus respectivos desvios-padrão, entre parênteses, obtidos por regressão linear, utilizando o método de marca simples em função do nível de herdabilidade e de ação gênica (POP 13 a 24) das 3.600 populações F₂ analisadas com 1.000 indivíduos

M	h ²	AG	$\overline{\hat{\beta}}_0$	$\overline{\hat{\beta}}_1$	$\overline{\hat{\beta}}_2$
C ₃₆	80%	5%	1000,79 (6,96)	57,22 (11,26)	-1,57 (13,81)
		10%	999,58 (6,34)	79,03 (11,05)	0,88 (12,18)
		20%	1000,21 (6,42)	114,21 (9,92)	-0,22 (12,18)
		40%	999,84 (5,98)	160,46 (8,88)	0,55 (11,37)
	60%	5%	999,60 (4,72)	34,21 (6,93)	0,81 (9,45)
		10%	1000,19 (4,69)	49,07 (7,31)	-0,32 (9,03)
		20%	1000,60 (4,64)	69,78 (6,65)	-1,03 (9,26)
		40%	999,60 (4,27)	98,63 (6,60)	0,92 (8,06)
	20%	5%	999,92 (3,38)	13,96 (5,26)	0,19 (6,76)
		10%	999,86 (3,26)	19,33 (4,96)	0,30 (6,48)
		20%	999,70 (3,36)	29,57 (5,01)	0,62 (6,69)
		40%	1000,28 (3,35)	39,79 (4,77)	-0,45 (6,47)
C ₃₇	80%	5%	999,32 (6,86)	45,00 (11,00)	1,17 (13,57)
		10%	999,34 (6,53)	78,59 (10,56)	1,26 (12,84)
		20%	999,78 (7,33)	113,38 (10,29)	0,36 (14,16)
		40%	1000,20 (6,87)	160,09 (8,52)	-0,39 (13,25)
	60%	5%	999,84 (4,97)	34,02 (6,90)	0,33 (10,02)
		10%	1000,26 (4,46)	48,18 (6,76)	-0,60 (8,94)
		20%	1000,55 (5,39)	69,18 (5,98)	-1,10 (10,47)
		40%	999,74 (4,55)	98,67 (6,08)	0,51 (8,67)
	20%	5%	999,85 (3,77)	14,04 (5,38)	0,32 (7,54)
		10%	1000,05 (3,34)	19,74 (5,34)	-0,11 (6,67)
		20%	999,22 (3,19)	28,97 (5,15)	1,52 (6,35)
		40%	999,98 (3,73)	39,76 (4,74)	0,06 (7,39)

M = marcador; C₃₆ = marcador seis do grupo de ligação três; C₃₇ = marcador sete do grupo de ligação três; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem.

Tabelas 17 - Média das estimativas de $\hat{\beta}_0$, $\hat{\beta}_1$ e $\hat{\beta}_2$ dos marcadores C₅₁₁ e C₅₁₂ e, seus respectivos desvios-padrão, entre parênteses, obtidos por regressão linear, utilizando o método de marca simples em função do nível de herdabilidade e de ação gênica (POP 25 a 36) das 3.600 populações F₂ analisadas com 1.000 indivíduos

M	h ²	AG	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
C ₅₁₁	80%	5%	1001,20 (7,06)	61,17 (10,24)	-2,33 (14,02)
		10%	1000,45 (7,17)	83,76 (9,08)	-0,68 (14,20)
		20%	1000,89 (5,87)	117,83 (9,58)	-1,63 (11,74)
		40%	1000,11 (5,50)	171,42 (8,47)	0,07 (10,70)
	60%	5%	999,03 (5,35)	37,46 (6,70)	2,01 (10,61)
		10%	1000,42 (4,56)	50,96 (7,07)	-0,68 (9,04)
		20%	1000,25 (4,67)	73,85 (6,78)	-0,40 (9,23)
		40%	999,50 (4,34)	103,62 (6,44)	1,16 (8,79)
	20%	5%	1000,18 (3,42)	15,12 (4,89)	-0,34 (6,76)
		10%	1000,35 (3,70)	21,60 (5,06)	-0,64 (7,37)
		20%	1000,37 (3,22)	29,88 (5,24)	-0,70 (6,31)
		40%	999,81 (3,62)	42,40 (4,62)	0,47 (7,13)
C ₅₁₂	80%	5%	1001,02 (7,16)	60,90 (9,88)	-2,02 (14,15)
		10%	1000,06 (7,29)	82,95 (9,61)	-0,06 (14,40)
		20%	1000,64 (6,67)	118,05 (10,12)	-1,20 (13,21)
		40%	999,63 (6,51)	170,85 (8,90)	0,77 (12,74)
	60%	5%	999,06 (5,54)	37,03 (7,00)	1,88 (10,86)
		10%	1000,14 (4,90)	51,12 (6,92)	-0,21 (9,67)
		20%	1000,45 (4,77)	73,77 (6,85)	-0,86 (9,38)
		40%	999,35 (4,92)	104,02 (6,51)	1,31 (9,76)
	20%	5%	1000,08 (3,54)	15,15 (5,19)	-0,13 (7,04)
		10%	1000,11 (3,78)	21,31 (5,32)	-0,18 (7,46)
		20%	1000,25 (3,32)	29,78 (5,49)	-0,53 (6,57)
		40%	1000,03 (3,76)	42,33 (5,00)	0,00 (7,45)

M = marcador; C₅₁₁ = marcador 11 do grupo de ligação cinco; C₅₁₂ = marcador 12 do grupo de ligação cinco; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem.

5. CONCLUSÕES

- O mapa de ligação foi, satisfatoriamente, reconstituído a partir da genotipagem simulada dos 1.000 indivíduos da população F_2 . O ordenamento e o número de grupos de ligação, obtidos pela análise genômica, foram exatamente os mesmos estabelecidos no genoma simulado.

- Cenários das populações F_2 com herdabilidade de 20% e ações gênicas de 5 e 10% não possibilitaram identificar facilmente o QTL simulado, mesmo com utilização da metodologia de intervalo simples.

- A contribuição do QTL para variação da característica dependeu muito da herdabilidade da característica, sendo que, para herdabilidade de 20%, não se observou diferença mínima do coeficiente de determinação, entre os cenários com diferentes saturações.

- Em relação ao nível de ação gênica, pode-se inferir que quanto maior seu valor, mais precisa foi a estimativa da frequência de recombinação máxima, independente da saturação do grupo de ligação e do nível de herdabilidade.

- As médias das estimativas do efeito aditivo, em todos os cenários, estavam em torno do esperado. O grau médio de dominância, observado pela razão da média do efeito aditivo e do efeito devido à dominância, foi praticamente zero e a média geral da característica aproximou-se de 1.000, conforme esperado.

6. REFERÊNCIAS BIBLIOGRÁFICAS

BEAVIS, W.D.; GRANT, D.; ALBERTSEN, M.; FINCHER, R. Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. **Theoretical and Applied Genetics**, Berlin, v.83, n.2, p.141-145, 1991.

BENTO, D.A.V. **Mapeamento de QTLs para produção de grãos e seus componentes em uma população de milho tropical**. 2006. 133p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba.

BRONDANI, C. **Desenvolvimento de marcadores microssatélites, construção de mapa genético interespecífico de *Oryza glumaepatula* x *Oryza sativa* e análise de QTLs para caracteres de importância agronômica**. 2000. 226p. Tese (Doutorado em Biologia Molecular) - Universidade de Brasília, Brasília.

CERVIGNI, G.D.L. **Análise genômica da resistência às raças 3 e 9 do nematóide de cisto e avaliação da variabilidade genética de caracteres agronômicos em soja**. 2003. 82p. Tese (Doutorado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa.

CORRÊA, R.X. **Genes de resistência a doenças do feijoeiro: identificação de marcadores moleculares, organização e identificação de análogos.** 1999. 116p. Tese (Doutorado em Genética e Melhoramento), Universidade Federal de Viçosa, Viçosa.

CRUZ, C.D. **Princípios de genética quantitativa.** Viçosa: UFV, 2005. 394p.

CRUZ, C. D. **Programa para análises de dados moleculares e quantitativos - GQMOL.** Viçosa: UFV, 2007. (Software em desenvolvimento).

CRUZ, C.D.; CARNEIRO, P.C.S. **Modelos biométricos aplicados ao melhoramento genético.** Viçosa: UFV, 2003. v.2, 585p.

CRUZ, E.M. **Efeito da saturação e do tamanho de populações F₂ e de retrocruzamento sobre a acurácia do mapeamento genético.** 2006. 119p. Tese (Doutorado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa.

DOEBLEY, J.; STEC, A. Genetic analysis of the morphological differences between maize and teosinte. **Genetics**, Bethesda, v.129, n.1, p.285-295, 1991.

EDWARDS, M.D.; STUBER, C.W.; WENDEL, J.F. Molecular-marker-facilitated investigations of quantitative trait loci in maize. I - Numbers, genomic distribution and types of gene action. **Genetics**, Bethesda , v.130, n.1, p.113-125, 1987.

EDWARDS, M.D.; HELENTJARIS, T.; WRIGHT, S.; STUBER, C.W. Molecular-marker-facilitated investigations of quantitative trait loci in maize. **Theoretical and Applied Genetics**, Berlin, v.83, n.6-7, p.765-774, 1992.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to quantitative genetics**. 4.ed. London: Longman, 1996. 464p.

FALK, C.T. A simple scheme for preliminary ordering of multiple loci: application to 45 CF families. In: ELSTON, R.C.; SPENCE, M.A.; HODGE, S.E.; MACCLUER, J.W. (Eds.). **Multipoint mapping and linkage based upon affected pedigree members**. Genetic Workshop 6, New York: Liss, 1989. p.17-22.

FERREIRA, A. **Mapeamento genético utilizando marcadores moleculares com distorção de segregação gamética e genotípica**. 2006. 132p. Tese (Doutorado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa.

FERREIRA, A.; SILVA, M.F.; SILVA, L.C.; CRUZ, C.D. Estimating the effects of population size and type on the accuracy of genetic maps, **Genetics and Molecular Biology**, Ribeirão Preto, v.29, n.1., p.187-192, 2006.

FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3.ed. Brasília: EMBRAPA/CENARGEM, 1998. 220p.

GUIMARÃES, C.T. **Mapeamento comparativo e detecção de QTLs em cana-de-açúcar utilizando marcadores moleculares.** 1999. 70p. Tese (Doutorado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa.

HALDANE, J.B.S. The combination of linkage values and the calculation of distance between the loci of linked factors. **Journal of Genetics**, Bangalore, v.8, p.299-309, 1919.

HALEY, C.; KNOTT, S. A simple regression method for mapping quantitative trait loci of linked factors. **Journal of Genetics**, Bangalore, v.8, n.2, p.299-309, 1992.

JANSEN, R.C.; STAM, P. High resolution of quantitative traits into multiple loci via interval mapping. **Genetics**, Bethesda, v.136, n.4, p.1447-1455, 1994.

JIANG, C.; ZENG, Z. Multiple trait analysis of genetic mapping for quantitative trait loci. **Genetics**, Bethesda, v.140, n.3, p.1111-1127, 1995.

KANTETY, R.V.; ZENG, X.; BENNETZEN, J.L.; ZEHR, B.E. Assessment of genetic diversity in dent and popcorn (*Zea mays* L.) inbred lines using inter-simple sequence repeat (ISSR) amplification. **Molecular Breeding**, Dordrecht, v.1, n.4, p.365-373, 1995.

KAO, C.H.; ZENG, Z.B.; TEASDALE, R.D. Multiple interval mapping for quantitative trait loci **Genetics**, Bethesda, v.152, n.3, p.1203-1216, 1999.

KOSAMBI, D.D. The estimation of map distance from recombination values. **Annals of Eugenics**, Cambridge, v.12, n.1, p.172-175, 1944.

KRUSKAL, J.B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. **Psychometrika**, New York, v.29, n.1, p.1-27, 1964.

KEARSEY, M.J.; HYNE, V. QTL analysis: a simple 'marker regression' approach. **Theoretical and Applied Genetics**, Berlin, v.89, n.6, p.698-702, 1994.

LANDER, E.S., BOTSTEIN, D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, Bethesda, v.121, n.1, 185-199, 1989. (erratum in 1994 *Genetics*, v.136, n.2, 705)

LANZA, M.A.; GUIMARÃES, C.T.; SCHUSTER, I. Aplicações de marcadores moleculares no melhoramento genético. **Informe Agropecuário**, Belo Horizonte, v.21, n.204, p.97-108, 2000.

LITT, M.; LUTY, J.A. A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. **American Journal of Human Genetics**, Chicago, v.44, n.3, p.397-401, 1989.

LIU, B.H. **Statistical genomics, linkage, mapping and QTL analysis**. Boca Raton: CRC, 1998. 611p.

LYNCH, M.; WALSH, B. **Genetics and analysis of quantitative traits**. Sunderland: Sinauer Associates, 1998. 980p.

MIRANDA, F.D. **Uso de marcadores RAPD para mapeamento de QTLs que determinam teor de proteína em soja**. 2002. 112p. Dissertação (Mestrado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa.

MORGAN, T.H. The theory of the gene. **American Naturalist**, Chicago, v.51, n.609, p.513-544, 1917.

OSBORN, T.C.; ALEXANDER, D.C.; FOBES, J.F. Identification of restriction fragment length polymorphisms linked to genes controlling soluble solids content in tomato fruit. **Theoretical and Applied Genetics**, Berlin, v.73, n.3, p.350-356, 1987.

PEARSON, K. Experimental discussion of the (χ^2 , p) test for goodness of fit. **Biometrika**, Oxford; v.24, n.3-4, p.351-381, 1932.

ROCHA, R.B. **Mapeamento de QTL para características de qualidade da madeira e crescimento em híbridos (*Eucalyptus grandis* x *Eucalyptus urophylla*)**. 2004. 77p. Dissertação (Mestrado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa.

SCHUSTER, I.; CRUZ, C.D. **Estatística genômica aplicada a populações derivadas de cruzamentos controlados**. Viçosa: UFV, 2004. 568p.

SILVA, H.D. **Aspectos biométricos da detecção de QTL's ("quantitative trait loci") em espécies cultivadas**. 2001. 166p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba.

SILVA, L.C. **Simulação do tamanho da população e da saturação do genoma para mapeamento genético de RILs**. 2005. 120p. Dissertação (Mestrado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa.

SOARES, T.C.B. **Mapeamento de locos que controlam o conteúdo de proteína em soja**. 2000. 58p. Dissertação (Mestrado em Agroquímica) - Universidade Federal de Viçosa, Viçosa.

SOLLER, M.; BRODY, T.; GENIZI, A. On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. **Theoretical and Applied Genetics**, Berlin, v.47, n.1, p.35-39, 1976.

SPEARMAN, C. General intelligence, objectively determined and measured. **American Journal of Psychology**, Chicago, v.15, n.1, p.201-293, 1904.

TANKSLEY, S.D. Mapping polygenes. **Annual Review of Genetics**, Palo Alto, v.27, p.205-233, 1993.

TANKSLEY, S.D.; MEDINA-FILHO, H.; RICK, C.M. Use of naturally occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. **Heredity**, Oxford, v.49, n.1, p.11-25, 1982.

VOS, P.; HOGERS, R.; BLEEKER, M.; REIJANS, M.; VAN DE LEE, T.; HORNES, M.; FRITJERS, A.; POT, J.; PELEMAN, J.; KUIPER, M.; ZABEAU, M. AFLP: a new technique for DNA fingerprinting. **Nucleic Acids Research**, Oxford, n.21, v.23, p.4407-4414, 1995.

WILLIAMS, J.G.K; KUBELIK, A.R.; LIVAK, K.J.; RAFALSKY, J.A.; TINGEY, S.V. DNA polymorphisms amplified by arbitrary primers are useful as genetics markers. **Nucleic Acids Research**, Oxford, v.18, n.22, p.6531-6535, 1990.

XU, S. A comment on the simple regression method for interval mapping. **Genetics**, Bethesda, v.141, n.4, p.1657-1659, 1995.

XU, S.; ATCHLEY, R.W. A random model approach to interval mapping of quantitative trait loci. **Genetics**, Bethesda, v.141, n.3, p. 1189-1197, 1995.

ZENG, Z.B. Precision mapping of quantitative trait loci. **Genetics**, Bethesda, v.136, n.4, p.1457-1466, 1994.

CAPÍTULO II

Genotipagem seletiva em populações F_2 simuladas para fins de mapeamento genético e detecção de QTL

1. INTRODUÇÃO

A associação entre uma característica quantitativa com marcadores moleculares de segregação mendeliana é a idéia principal, no mapeamento genético de locos controladores de tal caráter. Inicialmente proposta por SAX (1923), esta idéia tem sido usada por muitos estudiosos para detecção e localização dos locos controladores da característica quantitativa (QTL) em populações segregantes de espécies animais e vegetais.

Populações controladas são usadas rotineiramente para o mapeamento de QTL, porém um grande número de indivíduos é necessário para atingir poder de detecção do QTL (SOLLER et al. 1976). Os problemas são que os efeitos dos QTL são usualmente baixos, dificultando ainda mais a sua mensuração. A probabilidade de detectar um QTL depende também do desequilíbrio de ligação entre o marcador molecular e o QTL. Além disso, a acurácia do mapeamento de QTL depende de um grande número de fatores, como a herdabilidade da característica, o número de genes envolvidos, as interações dos genes, a distribuição dos genes no genoma, o tipo de população segregante estudada, o tamanho desta população e o número de marcadores moleculares empregados, assim como sua distribuição sobre o genoma (VAN OOIJEN, 1992).

A impossibilidade de obter todas as combinações de fatores envolvidos na acurácia do mapeamento de QTL, devido ao desconhecimento dos verdadeiros parâmetros, pode ser

solucionada por meio de simulação. Na simulação de dados via computador, os parâmetros são conhecidos, podendo ser usados na comparação da eficiência dos diversos cenários.

No mapeamento de QTL, existem estratégias como a genotipagem seletiva, que permitem reduzir, significativamente, o número de indivíduos genotipados para um dado poder de detecção do QTL, a partir de um grande número de indivíduos fenotipados (LEBOWITZ et al., 1987; LANDER & BOTSTEIN, 1989; DARVASI & SOLLER; 1992).

Esta técnica tem sido empregada por diversos pesquisadores, em várias espécies vegetais, para aumentar o poder de detecção de QTL. Por exemplo, a genotipagem seletiva foi utilizada por: GROOVER et al. (1994) em Pínus; NADI et al. (1997) em arroz; WINGBERMUEHLE et al. (2004) e VALES et al. (2005), em cevada. Entretanto, estudos de simulação com genotipagem seletiva em combinações de fatores, que afetam a detecção de QTL, bem como as comparações com outras estratégias de seleção são necessários para melhor compreensão do efeito da seleção e sua relação em diversos cenários. Conseqüentemente, novos modelos e programas de computador são imprescindíveis para o desenvolvimento de algoritmos, que incorporem o método clássico de mapeamento de QTL com genotipagem seletiva.

2. OBJETIVOS

Os objetivos gerais foram:

Investigar a influência da intensidade de seleção em amostras, obtidas pela genotipagem seletiva, nas análises genômicas de mapeamento genético, detecção e localização do QTL. Comparar os resultados obtidos, pela genotipagem seletiva e outras duas estratégias experimentais de amostragem (sistemática e mista), com a amostragem aleatória de indivíduos, para intensidade de seleção de 20%.

Os objetivos específicos foram:

- Avaliar a eficácia do mapeamento genético, em populações amostradas pela genotipagem seletiva, por meio da avaliação da segregação mendeliana de locos individuais, distância entre pares de marcadores, ordenamento e estabelecimento do número de grupos de ligação esperados.
- Avaliar a eficácia da análise genômica na detecção de QTL e sua contribuição na variação do caráter quantitativo, em populações amostradas.
- Avaliar a eficácia da análise genômica na estimação do efeito gênico, exercido pelo QTL sobre a característica, em populações amostradas.

3. MATERIAL E MÉTODOS

3.1. Populações amostradas

Uma espécie hipotética diplóide ($2n = 2x = 12$) foi utilizada como modelo para simulação de genomas com três níveis de saturações nos grupos de ligação. Foram utilizados os seguintes níveis de saturação: pouco saturado (grupos de ligação 1 e 2), mediamente saturado (grupos de ligação 3 e 4) e saturado (grupos de ligação 5 e 6), com distâncias entre marcadores adjacentes de 20 cM, 10 cM e 5 cM, respectivamente. O comprimento dos grupos de ligação foi de 100 cM e, portanto, um comprimento total do genoma de 600 cM. O genoma construído foi utilizado para geração de diferentes populações simuladas do tipo F_2 , com 1.000 indivíduos. Foram simulados 36 cenários diferentes, de populações F_2 com 1.000 indivíduos e replicadas 100 vezes, para uma característica quantitativa, totalizando 3.600 populações. Admitiu-se um caráter fenotípico com média de 1.000 unidades, coeficiente de variação de 10% e grau médio de dominância igual zero. Os cenários foram divididos em quatro níveis de ação gênica (5, 10, 20 e 40%) e três níveis de herdabilidade (80, 60 e 20%), totalizando 12 situações. Foram simulados três QTL (Q_{11} , Q_{31} e Q_{51}) para os grupos de ligação 1, 3 e 5, respectivamente. A característica simulada foi controlada por 10 locos, porém apenas um QTL foi declarado presente em cada

cenário. Assim, nos cenários das POP 1 a 12, POP 13 a 24 e POP 25 a 36 estiveram presentes os locos Q_{11} , Q_{31} e Q_{51} , respectivamente.

Das 3.600 populações F_2 geradas, cinco populações foram amostradas, aleatoriamente, em cada um dos 36 cenários, totalizando 180 populações de 1.000 indivíduos.

3.2. Genotipagem seletiva

As 180 populações F_2 de 1.000 indivíduos foram submetidas à genotipagem seletiva. A intensidade de seleção foi de 50, 30, 20 e 10%, amostrando 500, 300, 200 e 100 indivíduos, respectivamente, presentes nos extremos superior e inferior da distribuição normal dos fenótipos de cada população.

3.3. Comparação das estratégias experimentais de amostragem

Para verificar a eficácia da genotipagem seletiva (GS), outras duas estratégias de amostragem fenotípica, ou seja, sistemática (AS) e mista (AM) foram realizadas e comparadas com amostragem aleatória (AA). Na primeira estratégia, a distribuição de frequência do fenótipo foi dividida em dez classes equidistantes, sendo que 20 indivíduos foram amostrados dentro de cada classe. Na segunda estratégia, metade dos indivíduos foi obtida pela estratégia de genotipagem seletiva, enquanto a outra metade foi amostrada de acordo com a primeira estratégia. Portanto, além das 180 populações de 200 indivíduos, obtidas por genotipagem seletiva, outras 540 populações foram obtidas por meio de outras estratégias de amostragem.

3.4. A Análises genômicas - mapeamento genético em populações amostradas

Nas populações selecionadas ou amostradas das 180 populações F₂ de referência, realizaram-se análises genômicas de mapeamento genético. As análises foram constituídas de teste de segregação das marcas, pelo teste de qui-quadrado, seguindo-se a reconstrução do mapa de ligação, sendo comparadas, posteriormente, com o genoma simulado. Para a comparação, utilizaram-se os seguintes critérios: 1) correlação de Spearman; 2) tamanho de grupos de ligação; 3) distância média entre marcadores adjacentes; 4) variância das distâncias entre marcas adjacentes; e 5) estresse.

3.5. Análises genômicas - detecção e mapeamento de QTL em populações amostradas

As análises de detecção de QTL foram realizadas por meio do método de marca simples em todas as populações, que dispensa o conhecimento *a priori* do mapa de ligação. Utilizou-se, também, o método de intervalo simples para detecção e localização do QTL, quando foi possível a reconstrução do mapa de ligação com o mesmo número de grupos de ligação, em relação ao genoma simulado. Os métodos utilizados neste estudo, para detecção de QTL, foram descritos nos itens 3.3., 3.3.1, 3.3.2. e 3.3.3 da Seção de material e métodos, do Capítulo I.

4. RESULTADOS E DISCUSSÃO

Com a intenção de avaliar o efeito da intensidade de seleção por genotipagem seletiva na reconstrução de mapas de ligação, detecção e localização do QTL, foram amostradas, aleatoriamente, cinco populações de cada um dos 36 cenários nas 3.600 populações F_2 , totalizando 180 populações de 1.000 indivíduos. Após essa amostragem, as populações foram submetidas à genotipagem seletiva com intensidade de seleção de 50, 30, 20 e 10%. Formaram-se, então, populações amostradas de 500, 300, 200 e 100 indivíduos, totalizando 720 populações. Para as análises genômicas de mapeamento genético e detecção de QTL, realizadas nas populações amostradas, foram utilizados os modelos e métodos para populações F_2 codominante, bem como utilizado nas populações originais de 1.000 indivíduos.

4.1. Mapeamento genético em populações estabelecidas por genotipagem seletiva

4.1.1. Análise de segregação de locos individuais

Em todas as populações F_2 amostradas por genotipagem seletiva, foram aplicados testes de qui-quadrado para verificar a razão de segregação de locos individuais, com nível de 5% de probabilidade. Verificou-se maior número de distorções nos grupos de ligação, que possuíam a

presença do QTL simulado. O número médio de distorção, em geral, foi maior à medida que se aumentava o nível de ação gênica, de herdabilidade e saturação do grupo de ligação, no qual o QTL estava presente (Tabela 1). Considerando todos os cenários (180 populações) dentro de cada tamanho da população, as médias das distorções foram 3,77 (4,96%), 3,18 (4,19%), 3,51 (4,61%) e 2,93 (3,85%) para os tamanhos de 500, 300, 200 e 100 indivíduos, respectivamente. A distorção média máxima foi de 9,40 (12,37%), encontrada nas populações com 200 indivíduos, enquanto a mínima foi de 0,40 (0,53%), para aquelas com 300 indivíduos. Os cenários com o nível de herdabilidade de 80% e o nível de ação gênica de 40% apresentaram as maiores médias de distorção. Conforme esperado, os cenários com grupos de ligação mais saturados apresentaram mais distorções, realçadas principalmente no nível de ação gênica mais elevada. Em relação às intensidades de amostragem, não se observou nenhuma tendência nas médias das distorções.

Deve-se ressaltar que, para as populações estabelecidas por genotipagem seletiva, não há expectativa quanto a apresentarem estrutura idêntica a uma população F_2 , obtida por amostragem aleatória, pois, à medida que a variação explicada pelo QTL aumenta, em relação à variação fenotípica, a amostragem baseada no caráter quantitativo leva a distorção dos marcadores próximos ao QTL. A distorção ocorre, devido ao efeito conhecido como efeito “carona”. Assim, as análises genômica possibilitaram comparar as populações F_2 com as populações amostradas.

Tabela 1 - Médias de distorções e seus respectivos desvios-padrão, entre parênteses, de acordo com o tamanho da população amostrada por genotipagem seletiva (N), nível de herdabilidade (h^2), nível de ação gênica, nível de saturação para o grupo de ligação em que QTL simulado foi declarado presente

N	h^2	Ação gênica	Saturação		
			20 cM	10 cM	5 cM
500	80%	5%	1,60 (1,34)	6,80 (5,02)	2,60 (2,30)
		10%	2,40 (3,36)	3,20 (4,09)	3,80 (2,68)
		20%	5,80 (5,93)	3,00 (2,12)	5,20 (3,42)
		40%	5,60 (4,04)	4,40 (3,21)	8,80 (3,83)
	60%	5%	3,00 (2,55)	1,80 (1,48)	3,20 (2,95)
		10%	2,60 (4,22)	4,40 (6,43)	2,00 (2,35)
		20%	3,80 (3,90)	5,40 (2,97)	1,80 (1,48)
		40%	4,00 (2,83)	5,40 (3,05)	4,60 (2,88)
	20%	5%	2,80 (2,59)	3,20 (2,95)	2,00 (1,22)
		10%	1,20 (1,30)	2,80 (2,28)	3,20 (3,56)
		20%	6,20 (6,69)	3,20 (3,27)	2,60 (2,70)
		40%	4,40 (5,32)	3,60 (1,14)	5,40 (6,11)
300	80%	5%	1,20 (1,64)	4,20 (3,42)	3,20 (4,55)
		10%	3,80 (4,76)	3,00 (3,39)	1,60 (0,89)
		20%	4,80 (4,02)	3,20 (2,77)	6,40 (1,95)
		40%	4,80 (2,68)	6,80 (2,05)	8,00 (2,12)
	60%	5%	2,60 (2,61)	1,40 (1,14)	2,40 (1,52)
		10%	2,20 (3,83)	4,20 (5,67)	1,40 (2,07)
		20%	3,40 (4,22)	4,20 (2,77)	2,80 (2,17)
		40%	2,40 (1,52)	5,20 (1,64)	7,00 (5,29)
	20%	5%	0,40 (0,55)	3,40 (2,61)	1,60 (0,55)
		10%	1,20 (0,84)	2,00 (1,87)	2,80 (2,17)
		20%	2,60 (2,30)	2,60 (1,34)	0,60 (1,34)
		40%	3,60 (3,58)	1,20 (1,30)	2,40 (2,07)
200	80%	5%	2,20 (3,90)	3,80 (3,35)	1,60 (1,14)
		10%	4,00 (4,64)	4,60 (2,51)	6,00 (4,95)
		20%	4,40 (2,88)	2,40 (3,05)	6,60 (4,04)
		40%	6,20 (4,09)	6,60 (2,88)	9,40 (3,36)
	60%	5%	2,80 (2,59)	1,80 (1,10)	2,60 (2,30)
		10%	3,40 (3,97)	2,60 (2,19)	2,40 (3,29)
		20%	2,80 (2,95)	3,00 (1,87)	1,60 (1,34)
		40%	2,40 (1,52)	4,80 (4,97)	6,00 (5,66)
	20%	5%	0,80 (0,84)	0,60 (0,89)	2,40 (1,67)
		10%	2,60 (1,82)	4,00 (4,53)	2,40 (2,30)
		20%	2,60 (1,67)	4,60 (3,36)	2,40 (2,30)
		40%	3,20 (3,27)	1,60 (1,67)	5,00 (3,46)
100	80%	5%	2,40 (2,61)	2,20 (2,28)	2,80 (2,39)
		10%	3,20 (3,49)	2,60 (1,52)	2,80 (1,92)
		20%	2,60 (2,70)	2,60 (2,30)	2,80 (2,17)
		40%	4,40 (1,14)	8,60 (4,45)	3,40 (2,51)
	60%	5%	1,60 (1,14)	2,60 (2,07)	2,00 (1,22)
		10%	1,80 (1,30)	3,00 (3,08)	1,20 (1,30)
		20%	1,80 (1,48)	5,20 (2,39)	3,00 (2,92)
		40%	4,40 (2,19)	5,00 (2,74)	5,60 (5,46)
	20%	5%	1,80 (1,64)	2,20 (1,79)	1,60 (1,14)
		10%	2,40 (2,30)	3,80 (2,05)	1,40 (0,89)
		20%	1,40 (0,89)	3,20 (3,35)	3,20 (2,39)
		40%	1,20 (0,84)	1,80 (1,79)	3,80 (3,11)

4.1.2. Número de grupos de ligação e ordenamento das marcas

Para as populações obtidas por genotipagem seletiva com 500, 300 e 200 indivíduos, o número de grupos de ligação de todas as populações F_2 analisadas esteve de acordo com o esperado, mesmo na presença de distorções. Portanto, os seis grupos de ligação foram recuperados e não se observaram marcas não ligadas. Após a formação do grupo de ligação, verificou-se pelo método SARF a melhor ordem dos “n” marcadores constituintes de cada grupo. Não foi detectada inversão de marcas, mesmo nas populações com 200 indivíduos, pois, a correlação de Spearman foi igual a um.

No entanto, para as populações de 100 indivíduos amostradas por genotipagem seletiva, o mapeamento genético foi comprometido em, praticamente, todos os cenários estudados. Verificou-se inversão de marcas, marcas não ligadas e formação de mais de seis grupos de ligação, em muitas populações. Então, mesmo com a menor média de distorção entre as intensidades de amostragem, não foi possível a reconstrução do mapa de ligação igual ao genoma simulado em muitas populações. Existem algumas hipóteses, que corroboraram com este fato. Segundo uma dessas hipóteses, o tamanho da população sob amostragem da característica era insuficiente para recuperação do mapa de ligação, conforme esperado. Segundo a outra, a distorção causada pela intensidade de amostragem de 10% pode ter sido mais severa. Assim, essas populações com 100 indivíduos não foram comparadas no mapeamento genético.

4.1.3. Construção do mapa de ligação

Duas populações com 200 indivíduos cada uma e contrastantes, em relação ao número de marcadores com distorção, foram escolhidas e seus mapas de ligação foram reconstruídos. A primeira população, com 80% de herdabilidade e 40% de ação gênica, apresentou 13 marcas

distorcidas, sendo que 11 marcas estavam presentes no grupo de ligação 5 no qual estava presente o QTL simulado Q_{51} (Figura 1a). A segunda população com 20% de herdabilidade e 5% de ação gênica não apresentou marcas distorcidas, sendo que o QTL simulado foi o Q_{11} (Figura 1b). Embora apresentando um número de distorção elevado a primeira população não comprometeu a formação do mapa de ligação para o tamanho de 200 indivíduos. O viés que poderia existir na estimação da frequência de recombinação em populações obtidas por genotipagem seletiva, segundo DARVASI & SOLLER (1992), pode ser por meio da técnica de máxima verossimilhança para estimação da frequência de recombinação.

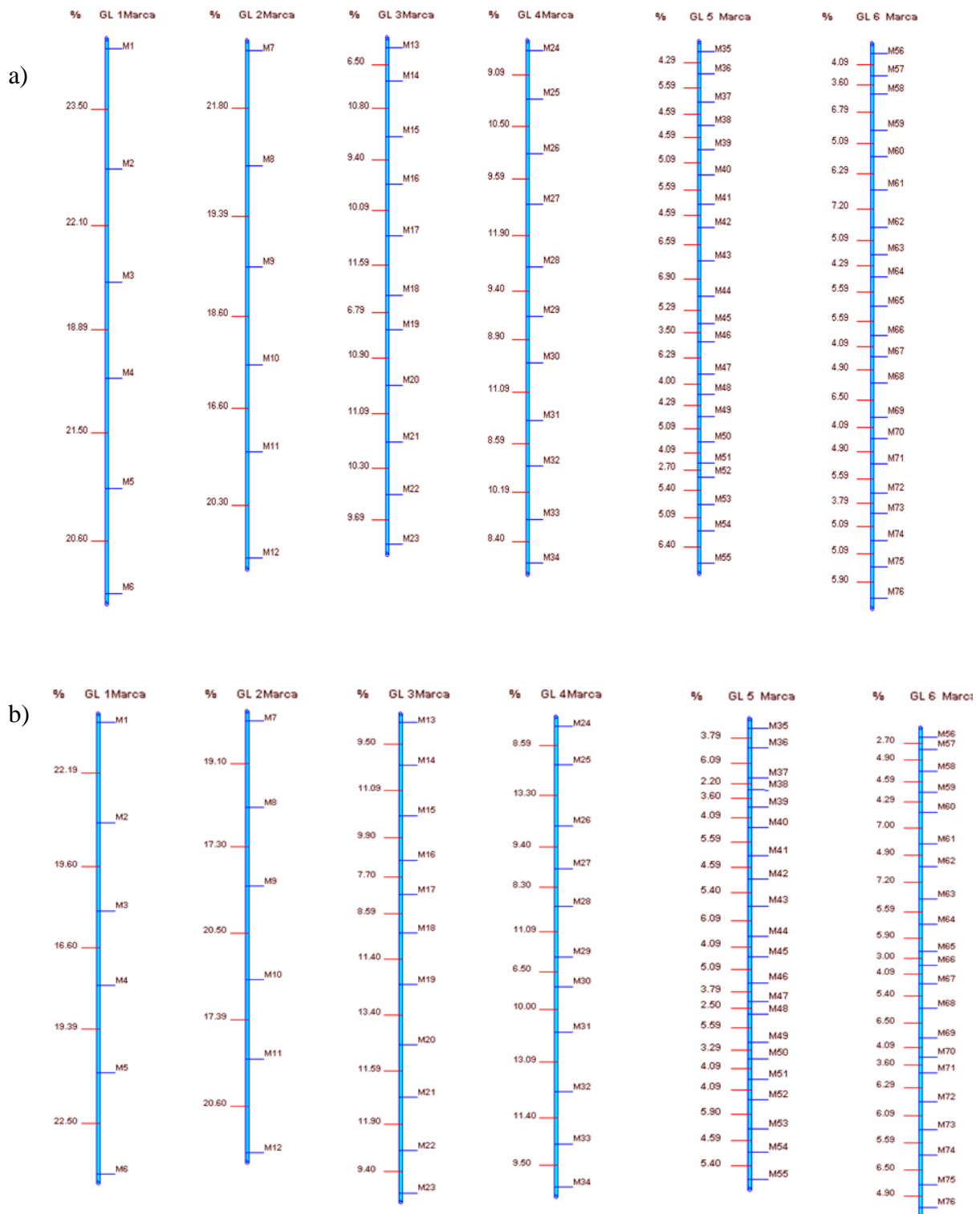


Figura 1 - Mapas genéticos com 76 marcadores distribuídos nos seis grupos de ligação (GL) de duas populações, amostradas por genotipagem seletiva com 200 indivíduos. a) Mapa de ligação da população com 80% de herdabilidade e 40% de ação, que apresentou 13 marcas distorcidas. b) Mapa de ligação da população com 20% de herdabilidade e 5% de ação gênica, não apresentando marcas distorcidas.

4.1.4. Comparação dos genomas

O tamanho médio do grupo de ligação (TAM), da distância média entre marcas adjacentes (DIM), da variância média das distâncias entre marcas adjacentes (VAD) e do estresse médio (EST) foram calculados nos grupos de ligação 1, 3 e 5, obtidos das populações amostradas por meio de genotipagem seletiva. Os resultados foram comparados com aqueles obtidos das 3.600 populações de 1.000 indivíduos (Tabela 2). Para comparação, foram escolhidos os grupos de ligação 1, 3 e 5 devido à presença do QTL nos cenários das POP 1 a POP 12, POP 13 a POP 24 e, POP 25 a POP 36, respectivamente. Em seguida, nas 180 populações amostradas dos 36 cenários para cada grupo de ligação, as análises foram separadas em duas partes. A parte “a” foi constituída pelas 60 populações, nas quais o QTL estava presente, enquanto a parte “b” foi constituída por 120 populações, nas quais o QTL estava ausente no respectivo grupo de ligação.

Não houve grande diferença entre as partes “a” e “b” para os valores de TAM, DIM, VAD e EST nos grupos de ligação, dentro de cada tamanho populacional. Observou-se que os TAMs e DIMs entre populações com diferentes tamanhos foram bem semelhantes, quanto a seu respectivo grupo de ligação. Entretanto, à medida que diminui o tamanho da população por genotipagem seletiva, verificou-se aumento nos desvios-padrão de TAM e DIM, em relação às populações de referência de 1.000 indivíduos. Para as variáveis VAD e EST, os valores e seus desvios-padrão aumentaram à medida que intensificava amostragem por genotipagem seletiva. Portanto, as precisões das estimativas de TAM, DIM, VAD e EST diminuíram com a redução no tamanho da população.

Tabela 2 - Tamanho médio de grupos de ligação, distância média entre marcas adjacentes, variância média das distâncias entre marcas adjacentes, estresse e seus respectivos desvios-padrão entre parênteses das 3.600 populações F₂ analisadas com 1.000 indivíduos e das 180 populações amostradas por genotipagem seletiva com 500, 300 e 200 indivíduos, para os grupos de ligação 1, 3 e 5,

N	GL	TAM	DIM	VAD	EST
1.000	GL 1	100,14 (1,93)	19,93 (1,07)	1,71 (1,16)	4,94 (1,59)
	GL 3	102,67 (2,36)	10,27 (0,24)	0,56 (0,26)	6,98 (1,58)
	GL 5	102,69 (2,18)	5,13 (0,11)	0,25 (0,09)	9,91 (1,60)
500	GL 1 ^a	99,51 (2,70)	19,90 (0,54)	2,45 (1,82)	7,13 (2,36)
	GL 1 ^b	100,00 (2,41)	20,00 (0,48)	2,44 (1,60)	7,07 (2,17)
	GL 3 ^a	103,04 (2,95)	10,30 (0,30)	1,18 (0,46)	10,93 (2,19)
	GL 3 ^b	102,77 (3,35)	10,28 (0,33)	1,16 (0,47)	10,02 (2,11)
	GL 5 ^a	101,59 (2,59)	5,08 (0,14)	0,44 (0,17)	14,20 (2,36)
	GL 5 ^b	101,68 (2,77)	5,12 (0,16)	0,45 (0,15)	13,93 (2,15)
300	GL 1 ^a	99,21 (3,35)	19,84 (0,67)	4,28 (3,18)	9,34 (3,19)
	GL 1 ^b	100,18 (3,41)	20,04 (0,69)	3,68 (2,30)	8,84 (2,77)
	GL 3 ^a	103,41 (3,97)	10,34 (0,40)	1,97 (0,87)	13,00 (2,97)
	GL 3 ^b	102,27 (4,92)	10,23 (0,49)	1,83 (0,77)	12,60 (2,97)
	GL 5 ^a	101,73 (3,57)	5,09 (0,17)	0,76 (0,22)	18,23 (2,51)
	GL 5 ^b	101,58 (3,75)	5,08 (0,20)	0,75 (0,26)	18,06 (2,93)
200	GL 1 ^a	99,15 (5,05)	19,90 (0,88)	6,50 (6,16)	11,80 (4,38)
	GL 1 ^b	99,96 (4,78)	19,99 (0,95)	5,14 (3,36)	10,73 (3,31)
	GL 3 ^a	102,81 (5,05)	10,28 (0,51)	2,79 (1,08)	15,84 (3,29)
	GL 3 ^b	101,90 (5,78)	10,19 (0,57)	2,78 (1,15)	15,60 (3,30)
	GL 5 ^a	101,87 (3,89)	5,09 (0,22)	1,22 (0,39)	22,64 (3,98)
	GL 5 ^b	101,70 (5,37)	5,08 (0,26)	1,18 (0,44)	22,22 (3,58)

N= número de indivíduos; GL = grupo de ligação; ^a grupo de ligação na presença do QTL; ^b grupo de ligação na ausência dos QTL; TAM = tamanho médio do grupo de ligação em centiMorgan (cM); DIM = distância média entre marcas adjacentes em centiMorgan (cM); VAD = variância média das distâncias entre marcas adjacentes; EST = estresse médio em porcentagem (%).

4.2. Detecção de QTL em populações amostradas por genotipagem seletiva

A principal dificuldade no mapeamento de QTL é o fato intrínseco, em que vários fatores genéticos e ambientais afetam a expressão final do fenótipo. Os delineamentos experimentais envolvem, assim, a medição e genotipagem de várias centenas de indivíduos para permitir a

resolução desejada. Uma população F_2 de aproximadamente 2.000 indivíduos seria necessária, para detectar QTL com efeito igual a 2 a 3% em média (SOLLER et al. 1976).

Portanto, novas estratégias experimentais de mapeamento de QTL têm sido propostas e utilizadas. Tais estratégias permitem reduzir o número de indivíduos genotipados, a fim de obter uma mesma potência estatística de detecção. Para isto, aumenta-se o número de indivíduos avaliados em relação à característica quantitativa. Entre as novas estratégias experimentais, destaca-se a genotipagem seletiva, proposta por LANDER & BOTSTEIN (1989), cujo princípio foi descrito, inicialmente, por STUBER et al. (1980). Baseia-se no fato de que uma diferença significativa na frequência alélica nos QTL, nos extremos alto e baixo uma distribuição fenotípica, pode ser utilizada para um teste de ligação entre o marcador e QTL. Fenótipos com uma unidade de desvio-padrão acima ou abaixo da média constituem 33% da população e contribuem com 81% da informação genética (LANDER & BOTSTEIN, 1989). Nesta estratégia, portanto, um número maior de indivíduos é gerado e têm o fenótipo avaliado para o caráter quantitativo. Entretanto, em geral, somente 10 a 20% em cada extremo da distribuição possuem os genótipos, identificados em cada loco marcador. Métodos convencionais de regressão superestimam os efeitos do QTL (FERREIRA & GRATTAPAGLIA, 1998).

As populações, amostradas por genotipagem seletiva, foram submetidas às análises de detecção de QTL por regressão linear pelo método de marca simples e intervalo simples. As estimativas do LOD e do coeficiente de determinação (R^2) obtidas a partir do método de marca simples e as estimativas de LOD máximo ($LOD_{m\acute{a}x}$), posição relativa do QTL ($r_{m\acute{a}x}$) e o R^2 , obtidas do método de intervalo simples, foram comparadas com as respectivas estimativas das populações F_2 com 1.000 indivíduos, consideradas como referência. Assim, foi possível verificar o efeito da intensidade de amostragem por genotipagem seletiva nos 36 cenários para a característica quantitativa, tanto no método de marca simples quanto no método de intervalo simples.

4.2.1. Detecção do QTL pelo método de marca simples e sua contribuição para a variação da característica

Análises de regressão foram realizadas pelo método de marca simples, para os marcadores flanqueadores dos QTL simulado dos grupos de ligação 1, 3 e 5, para as populações amostradas por genotipagem seletiva, sendo os dados comparados às populações originais. Estimaram-se os valores do LOD (Tabelas 3, 4 e 5) e de R^2 (Tabelas 6, 7 e 8), respectivamente.

Como nas populações originais, nas populações submetidas à genotipagem seletiva verificou-se que, independente da posição do marcador, se anterior ou posterior ao QTL simulado, em qualquer grupo de ligação, todas as médias e desvios-padrão do LOD e R^2 foram muito similares quanto à magnitude e tendência.

Observou-se redução do LOD médio à medida que se diminuía o número de indivíduos selecionados por genotipagem seletiva. Os desvios-padrão do LOD médio mantiveram-se praticamente, constantes para cada situação, independente da intensidade de amostragem. Por outro lado, os valores dos R^2 médios e os desvios-padrão aumentaram, significativamente, à medida que diminuía o tamanho da população. Portanto, mesmo com a redução do LOD médio, a genotipagem seletiva aumentou a associação entre a característica quantitativa, verificado em razão do aumento do R^2 médio.

Tabela 3 - Estimativas médias do LOD para os marcadores C₁₃ e C₁₄ e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F₂ analisadas com 1.000 indivíduos e das 180 populações amostradas (cinco de cada cenário) por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h ²	AG	LOD			
			1.000	500	300	200
C ₁₃	80%	5%	6,17 (2,40)	4,77 (2,27)	4,44 (1,29)	3,41 (1,36)
		10%	11,38 (3,27)	11,10 (2,64)	9,49 (3,35)	8,68 (3,57)
		20%	23,93 (4,98)	23,53 (3,10)	21,08 (4,62)	17,62 (6,48)
		40%	47,53 (7,98)	45,45 (6,18)	38,10 (5,96)	29,97 (7,53)
	60%	5%	4,22 (2,03)	2,25 (1,68)	1,91 (1,71)	1,85 (1,82)
		10%	8,66 (2,39)	6,19 (0,86)	4,15 (0,85)	3,63 (1,54)
		20%	16,80 (4,02)	17,52 (2,91)	15,59 (2,91)	12,85 (3,83)
		40%	34,50 (5,98)	31,50 (3,48)	27,54 (3,75)	23,55 (5,07)
	20%	5%	1,79 (1,21)	0,74 (0,64)	0,86 (0,92)	0,86 (0,75)
		10%	3,08 (1,74)	2,33 (1,35)	2,50 (0,66)	2,24 (1,42)
		20%	5,54 (2,32)	5,33 (2,01)	4,10 (1,66)	3,97 (1,56)
		40%	10,92 (3,34)	11,12 (4,23)	9,36 (2,80)	7,69 (3,05)
C ₁₄	80%	5%	5,92 (2,33)	3,86 (1,80)	3,66 (1,27)	3,42 (1,05)
		10%	10,99 (3,27)	10,49 (3,86)	9,35 (2,74)	8,06 (2,60)
		20%	23,70 (5,09)	23,41 (2,77)	22,10 (5,92)	18,99 (7,83)
		40%	48,18 (7,85)	49,04 (8,29)	41,03 (7,09)	33,53 (4,14)
	60%	5%	4,28 (1,93)	2,78 (1,45)	2,72 (1,51)	2,50 (1,89)
		10%	8,78 (2,47)	5,99 (1,86)	4,31 (1,74)	3,91 (2,10)
		20%	16,23 (4,05)	15,22 (5,95)	13,71 (4,70)	10,82 (3,36)
		40%	34,37 (6,15)	30,81 (3,32)	26,47 (3,33)	23,06 (5,00)
	20%	5%	1,67 (1,22)	1,07 (0,99)	1,32 (1,39)	1,15 (1,11)
		10%	3,01 (1,57)	2,28 (1,18)	2,14 (0,77)	2,08 (1,62)
		20%	5,68 (2,38)	6,13 (1,70)	5,10 (2,10)	4,44 (0,97)
		40%	11,11 (3,52)	12,00 (3,15)	9,77 (2,16)	8,24 (1,07)

M = marcador; C₁₃ = marcador três do grupo de ligação um; C₁₄ = marcador quatro do grupo de ligação um; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem e \overline{LOD} = média do LOD.

Tabela 4 - Estimativas médias do LOD para os marcadores C₃₆ e C₃₇ e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F₂ analisadas com 1.000 indivíduos e das 180 populações amostradas (cinco de cada cenário) por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h ²	AG	LOD			
			1.000	500	300	200
C ₃₆	80%	5%	7,45 (2,86)	6,54 (2,97)	5,52 (2,76)	4,12 (1,73)
		10%	14,06 (3,92)	12,39 (1,61)	11,11 (3,55)	10,07 (3,81)
		20%	29,73 (5,33)	31,65 (9,08)	27,07 (7,60)	21,72 (6,26)
		40%	62,88 (9,36)	56,18 (7,15)	48,21 (7,87)	37,93 (7,41)
	60%	5%	5,38 (2,05)	4,26 (1,76)	3,96 (3,13)	3,24 (3,24)
		10%	10,57 (3,14)	11,42 (2,70)	10,67 (3,22)	9,99 (2,54)
		20%	22,25 (4,85)	22,89 (3,33)	20,09 (3,35)	17,89 (3,64)
		40%	45,46 (6,66)	37,58 (4,76)	32,46 (6,41)	26,22 (4,98)
	20%	5%	2,04 (1,25)	0,85 (0,45)	1,13 (0,60)	1,29 (0,52)
		10%	3,63 (1,61)	4,25 (2,24)	4,67 (2,66)	4,45 (2,19)
		20%	7,85 (2,65)	7,86 (2,86)	8,08 (2,75)	6,73 (1,85)
		40%	14,64 (3,70)	11,84 (3,38)	9,70 (2,18)	8,56 (1,66)
C ₃₇	80%	5%	7,39 (2,65)	7,14 (3,02)	6,07 (2,49)	4,69 (2,39)
		10%	13,94 (3,75)	11,19 (1,88)	10,16 (3,86)	9,52 (4,01)
		20%	29,50 (5,85)	30,40 (8,22)	25,49 (6,37)	20,13 (6,19)
		40%	62,70 (8,66)	58,73 (9,41)	50,56 (8,15)	41,25 (7,40)
	60%	5%	5,35 (2,15)	5,47 (3,01)	4,68 (3,82)	4,01 (3,82)
		10%	10,22 (3,05)	10,31 (4,32)	9,75 (4,68)	8,74 (3,28)
		20%	21,91 (4,41)	22,70 (4,48)	20,38 (5,30)	19,18 (6,59)
		40%	45,71 (6,84)	38,40 (6,42)	29,52 (14,38)	27,66 (9,17)
	20%	5%	2,13 (1,35)	1,07 (0,79)	1,35 (1,04)	1,44 (0,97)
		10%	3,81 (1,77)	4,08 (2,45)	4,35 (2,78)	4,66 (2,51)
		20%	7,59 (2,74)	7,90 (3,82)	8,33 (3,20)	6,55 (2,09)
		40%	14,67 (3,60)	14,35 (2,84)	11,46 (1,75)	10,17 (1,42)

M = marcador; C₃₆ = marcador seis do grupo de ligação três; C₃₇ = marcador sete do grupo de ligação três; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem e $\overline{\text{LOD}}$ = média do LOD.

Tabela 5 - Estimativas médias do LOD para os marcadores C₅₁₁ e C₅₁₂ e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F₂ analisadas com 1.000 indivíduos e das 180 populações amostradas (cinco de cada cenário) por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h ²	AG	LOD			
			1.000	500	300	200
C ₅₁₁	80%	5%	8,62 (3,08)	8,78 (3,20)	7,90 (2,10)	8,05 (1,63)
		10%	15,75 (3,82)	17,15 (5,01)	13,71 (5,17)	12,07 (4,94)
		20%	32,93 (6,38)	31,21 (8,96)	24,37 (6,24)	20,60 (8,52)
		40%	72,64 (9,09)	70,95 (12,19)	57,45 (8,37)	43,79 (2,62)
	60%	5%	6,37 (2,20)	7,53 (0,79)	6,61 (1,22)	5,41 (1,24)
		10%	11,90 (3,55)	9,86 (5,88)	10,28 (7,51)	9,64 (6,11)
		20%	24,28 (5,11)	24,22 (7,03)	20,08 (6,78)	16,04 (5,04)
		40%	51,95 (8,24)	49,63 (10,97)	43,52 (9,80)	36,31 (9,93)
	20%	5%	2,32 (1,34)	2,89 (1,33)	2,73 (1,59)	2,35 (1,76)
		10%	4,34 (2,06)	4,76 (2,86)	3,50 (2,03)	3,41 (1,72)
		20%	8,23 (2,78)	5,55 (2,67)	5,93 (1,80)	4,64 (1,79)
		40%	16,10 (3,76)	14,51 (4,05)	14,13 (3,67)	11,18 (2,70)
C ₅₁₂	80%	5%	8,49 (2,92)	9,24 (2,44)	8,46 (0,81)	8,68 (2,86)
		10%	15,50 (3,94)	16,38 (4,50)	13,25 (4,87)	10,70 (3,00)
		20%	33,10 (6,40)	30,91 (8,21)	24,78 (5,59)	17,70 (5,83)
		40%	72,19 (9,44)	64,09 (7,29)	53,85 (10,23)	41,21 (4,35)
	60%	5%	6,27 (2,25)	7,53 (1,22)	6,47 (1,89)	5,32 (1,12)
		10%	11,91 (3,44)	10,42 (5,41)	9,75 (8,20)	10,32 (5,89)
		20%	24,18 (4,78)	23,55 (4,47)	19,37 (4,66)	15,53 (3,05)
		40%	52,46 (8,25)	48,76 (7,36)	43,43 (9,46)	35,39 (9,69)
	20%	5%	2,35 (1,45)	3,27 (1,81)	2,74 (1,71)	2,60 (1,54)
		10%	4,26 (2,14)	4,56 (2,60)	3,46 (2,28)	3,33 (1,75)
		20%	8,22 (2,90)	5,45 (2,67)	5,92 (1,19)	4,55 (1,20)
		40%	16,10 (3,93)	13,46 (4,03)	12,95 (3,82)	10,64 (3,16)

M = marcador; C₅₁₁ = marcador 11 do grupo de ligação cinco; C₅₁₂ = marcador 12 do grupo de ligação cinco; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem e $\overline{\text{LOD}}$ = média do LOD.

Tabela 6 - Estimativas médias de R^2 para os marcadores C_{13} e C_{14} e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F_2 analisadas com 1.000 indivíduos e das 180 populações amostradas (cinco de cada cenário) por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h^2	AG	\bar{R}^2			
			1.000	500	300	200
C_{13}	80%	5%	2,80 (1,07)	4,28 (1,99)	6,58 (1,86)	7,52 (2,87)
		10%	5,10 (1,43)	9,70 (2,23)	13,46 (4,37)	17,90 (6,77)
		20%	10,41 (2,06)	19,46 (2,28)	27,50 (5,13)	32,74 (10,42)
		40%	19,61 (2,94)	34,12 (3,68)	44,10 (4,84)	49,26 (8,25)
	60%	5%	1,92 (0,91)	2,04 (1,51)	2,86 (2,54)	4,11 (3,96)
		10%	3,90 (1,05)	5,54 (0,75)	6,16 (1,23)	7,98 (3,27)
		20%	7,43 (1,71)	14,88 (2,29)	21,22 (3,43)	25,38 (6,44)
		40%	14,66 (2,34)	25,15 (2,38)	34,39 (3,78)	41,54 (6,88)
	20%	5%	0,82 (0,55)	0,68 (0,58)	1,30 (1,38)	1,96 (1,70)
		10%	1,41 (0,79)	2,12 (1,22)	3,76 (0,98)	4,98 (3,06)
		20%	2,51 (1,04)	4,78 (1,77)	6,08 (2,40)	8,68 (3,33)
		40%	4,89 (1,45)	9,68 (3,49)	13,32 (3,69)	16,06 (5,76)
C_{14}	80%	5%	2,69 (1,04)	3,48 (1,59)	5,45 (1,85)	7,55 (2,22)
		10%	4,92 (1,43)	9,17 (3,27)	13,31 (3,64)	16,81 (4,87)
		20%	10,32 (2,10)	19,38 (2,06)	28,55 (6,45)	34,62 (11,61)
		40%	19,85 (2,88)	36,20 (4,68)	46,48 (5,58)	53,53 (4,28)
	60%	5%	1,95 (0,87)	2,52 (1,31)	4,06 (2,22)	5,53 (4,07)
		10%	3,96 (1,09)	5,36 (1,64)	6,38 (2,53)	8,53 (4,43)
		20%	7,19 (1,72)	14,98 (4,80)	18,82 (5,99)	21,85 (6,21)
		40%	14,61 (2,41)	24,68 (2,27)	33,32 (3,43)	40,89 (6,66)
	20%	5%	0,76 (0,55)	0,98 (0,90)	1,98 (2,09)	2,60 (2,48)
		10%	1,38 (0,71)	2,07 (1,06)	3,86 (1,15)	4,63 (3,52)
		20%	2,58 (1,07)	5,48 (1,48)	7,50 (2,98)	9,70 (2,02)
		40%	4,98 (1,54)	10,43 (2,58)	13,89 (2,85)	17,24 (2,01)

M = marcador; C_{13} = marcador três do grupo de ligação um; C_{14} = marcador quatro do grupo de ligação um; h^2 = herdabilidade em porcentagem; AG = ação gênica em porcentagem e \bar{R}^2 = média do coeficiente de determinação.

Tabela 7 - Estimativas médias de R^2 para os marcadores C_{36} e C_{37} e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F_2 analisadas com 1.000 indivíduos e das 180 populações amostradas (cinco de cada cenário) por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h^2	AG	\bar{R}^2			
			1.000	500	300	200
C_{36}	80%	5%	3,36 (1,27)	5,82 (2,58)	8,06 (3,91)	9,00 (3,63)
		10%	6,26 (1,69)	10,78 (1,32)	15,58 (4,59)	20,46 (6,76)
		20%	12,77 (2,13)	25,08 (6,13)	33,64 (7,78)	38,86 (8,49)
		40%	25,08 (3,20)	40,29 (3,98)	52,01 (5,81)	57,76 (7,18)
	60%	5%	2,44 (0,92)	3,84 (1,54)	5,81 (4,49)	6,98 (6,83)
		10%	4,74 (1,37)	9,96 (2,22)	15,02 (4,23)	20,44 (4,79)
		20%	9,72 (2,00)	18,98 (2,53)	26,46 (3,83)	33,58 (5,64)
		40%	18,86 (2,48)	29,20 (3,15)	39,00 (6,36)	45,04 (6,08)
	20%	5%	0,94 (0,57)	0,78 (0,41)	1,72 (0,91)	2,92 (1,16)
		10%	1,66 (0,73)	3,82 (2,01)	6,86 (3,87)	9,64 (4,71)
		20%	3,54 (1,17)	6,96 (2,46)	11,60 (3,75)	14,30 (3,69)
		40%	6,51 (1,59)	10,30 (2,80)	13,80 (2,92)	17,84 (3,17)
C_{37}	80%	5%	3,34 (1,17)	6,34 (2,62)	8,84 (3,50)	10,13 (4,92)
		10%	6,21 (1,61)	9,78 (1,55)	14,33 (4,94)	19,43 (7,09)
		20%	12,67 (2,33)	24,24 (5,65)	32,11 (6,51)	36,58 (8,71)
		40%	25,03 (2,97)	41,62 (5,04)	53,69 (5,94)	60,84 (6,65)
	60%	5%	2,43 (0,96)	4,89 (2,58)	6,82 (5,30)	8,55 (7,72)
		10%	4,59 (1,33)	9,01 (3,53)	13,74 (6,10)	18,06 (6,15)
		20%	9,58 (1,82)	18,81 (3,36)	26,69 (5,94)	35,18 (9,56)
		40%	18,95 (2,52)	29,70 (4,04)	35,30 (15,15)	46,32 (10,18)
	20%	5%	0,98 (0,61)	0,98 (0,72)	2,04 (1,56)	3,24 (2,15)
		10%	1,74 (0,80)	3,67 (2,18)	6,39 (4,02)	10,06 (5,27)
		20%	3,43 (1,22)	6,98 (3,28)	11,92 (4,35)	13,92 (4,11)
		40%	6,52 (1,54)	12,35 (2,28)	16,10 (2,24)	20,82 (2,57)

M = marcador; C_{36} = marcador seis do grupo de ligação três; C_{37} = marcador sete do grupo de ligação três; h^2 = herdabilidade em porcentagem; AG = ação gênica em porcentagem e \bar{R}^2 = média do coeficiente de determinação.

Tabela 8 - Estimativas médias de \bar{R}^2 para os marcadores C_{511} e C_{512} e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F_2 analisadas com 1.000 indivíduos e das 180 populações amostradas (cinco de cada cenário) por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h^2	AG	\bar{R}^2			
			1.000	500	300	200
C_{511}	80%	5%	3,88 (1,36)	7,74 (2,72)	11,38 (2,87)	16,88 (3,17)
		10%	6,98 (1,63)	14,54 (3,94)	18,77 (6,45)	23,88 (8,18)
		20%	14,04 (2,50)	24,78 (6,16)	30,96 (6,50)	36,84 (11,63)
		40%	28,38 (2,97)	47,72 (5,63)	58,34 (5,06)	63,46 (2,20)
	60%	5%	2,89 (0,98)	6,70 (0,68)	9,64 (1,69)	11,68 (2,52)
		10%	5,32 (1,54)	8,58 (4,82)	14,16 (9,17)	19,30 (10,54)
		20%	10,56 (2,10)	19,86 (5,30)	26,21 (7,55)	30,52 (7,85)
		40%	21,23 (2,98)	36,42 (6,68)	48,26 (7,84)	55,80 (9,19)
	20%	5%	1,06 (0,61)	2,62 (1,20)	4,08 (2,34)	5,20 (3,80)
		10%	1,97 (0,92)	4,26 (2,51)	5,20 (2,96)	7,50 (3,67)
		20%	3,71 (1,23)	4,96 (2,33)	8,68 (2,49)	10,08 (3,60)
		40%	7,13 (1,60)	12,46 (3,26)	19,40 (4,54)	22,58 (4,68)
C_{512}	80%	5%	3,83 (1,28)	8,14 (2,05)	12,16 (1,09)	17,98 (5,31)
		10%	6,88 (1,68)	13,94 (3,56)	18,21 (6,11)	21,65 (5,39)
		20%	14,10 (2,51)	24,60 (5,70)	31,43 (5,86)	32,89 (9,07)
		40%	28,22 (3,09)	44,44 (3,78)	55,84 (6,59)	61,13 (4,04)
	60%	5%	2,84 (1,01)	6,70 (1,04)	9,42 (2,64)	11,50 (2,31)
		10%	5,33 (1,50)	9,06 (4,40)	13,39 (10,21)	20,59 (9,90)
		20%	10,52 (1,96)	19,43 (3,35)	25,53 (5,35)	29,89 (4,83)
		40%	21,41 (2,98)	36,04 (4,39)	48,23 (7,28)	54,88 (9,26)
	20%	5%	1,08 (0,66)	2,95 (1,62)	4,10 (2,54)	5,75 (3,36)
		10%	1,94 (0,96)	4,09 (2,31)	5,13 (3,31)	7,32 (3,75)
		20%	3,71 (1,28)	4,88 (2,32)	8,67 (1,66)	9,91 (2,45)
		40%	7,13 (1,68)	11,61 (3,28)	17,92 (4,86)	21,58 (5,57)

M = marcador; C_{511} = marcador 11 do grupo de ligação cinco; C_{512} = marcador 12 do grupo de ligação cinco; h^2 = herdabilidade em porcentagem; AG = ação gênica em porcentagem e \bar{R}^2 = média do coeficiente de determinação.

Nas populações originais, observou-se que o R^2 médio para herdabilidade de 20%, em todos os grupos de ligação, teve uma perda superior a 80% em relação ao nível de ação gênica paramétrico, mesmo para os níveis de ação gênica de 20 e 40%. Entretanto, quando foram selecionados 200 indivíduos, independente do grupo de ligação, as perdas dos R^2 médios foram reduzidas para 50%, em média. Assim, foi realçada a associação entre o marcador e o caráter quantitativo, mesmo com o nível de herdabilidade de 20%, por genotipagem seletiva.

Em relação às médias das estimativas do LOD e R^2 das populações com 1.000 indivíduos, o decréscimo médio do LOD foi de 0,93, 0,81, 0,68 para os tamanhos de 500, 300 e 200, respectivamente. Por outro lado, houve acréscimo médio do R^2 para as intensidades de amostragem de 50, 30 e 20% de 1,76, 2,36 e 2,80, respectivamente. Assim, a resolução de detecção de QTL, com LOD maior que 3, em populações que sofreram genotipagem seletiva, sendo que a média de seus LOD foi similar às populações com 1.000 indivíduos.

4.2.2. Detecção e localização do QTL pelo método de intervalo simples e sua contribuição para a variação da característica

A detecção de QTL pelo método de intervalo foi realizada por meio de regressão linear nas populações amostradas por genotipagem seletiva com 500, 300 e 200 indivíduos e comparadas com as 3.600 populações F_2 com 1.000 indivíduos. Este processo foi baseado no teste de associação de pares de marcas moleculares adjacentes, considerando-se apenas os grupos de ligação 1 para os cenários das POP 1 a 12; 3 para as POP 13 a 24; e 5 para as POP 25 a 36 com sua respectiva característica quantitativa.

Os valores médios para LOD máximo, para a posição relativa do QTL e do coeficiente de determinação e seus desvios-padrão, referentes aos grupos de ligação 1, 3 e 5, encontram-se nas tabelas 9, 10 e 11.

Em todos os cenários das populações amostradas por genotipagem seletiva, verificou-se que os valores de $LOD_{máx}$ e seus desvios-padrão foram semelhantes, independente da intensidade de amostragem. No entanto, em estudo com estratégia de amostragem aleatória de indivíduos, realizado por LI et al. (2006), verificou-se uma redução linear entre o $LOD_{máx}$ da população original com a amostrada em populações F_2 . Assim, se a amostragem fosse aleatória, haveria em média uma redução de 50, 70 e 80% nos $LOD_{máx}$, para os tamanhos de 500, 300 e 200, respectivamente (LI et al., 2006). Entretanto, para as populações amostradas por genotipagem seletiva de tamanho de 500, 300 e 200 indivíduos, observou-se uma redução média dos $LOD_{máx}$ de 1, 3 e 12%, respectivamente, em relação aos $LOD_{máx}$ das populações de 1.000 indivíduos.

Tabela 9 - Estimativas médias do LOD máximo e seus respectivos desvios-padrão, entre parênteses, obtidos por regressão linear pelo método de intervalo simples em função do nível de herdabilidade e de ação gênica, nos grupos de ligação 1 (POP 1 a 12), 3 (POP 13 a 24) e 5 (POP 25 a 36) das 3.600 populações F₂ analisadas com 1.000 indivíduos e das 180 populações amostradas (cinco de cada cenário) por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

GL	h ²	AG	$\overline{\text{LOD}}_{\text{máx}}$			
			1.000	500	300	200
GL 1	80%	5%	7,79 (2,67)	5,62 (2,31)	5,31 (1,02)	4,80 (0,83)
		10%	14,30 (3,96)	14,13 (3,98)	13,04 (4,57)	11,86 (4,64)
		20%	31,84 (6,70)	33,14 (4,41)	33,54 (10,10)	31,51 (15,85)
		40%	69,93 (12,02)	82,95 (13,61)	77,65 (17,44)	68,06 (18,82)
	60%	5%	5,49 (2,24)	3,40 (1,79)	3,28 (1,95)	2,97 (2,25)
		10%	11,16 (2,81)	8,09 (1,26)	5,59 (1,48)	5,09 (2,23)
		20%	21,60 (4,87)	22,20 (5,13)	20,90 (4,72)	17,51 (4,98)
		40%	47,62 (8,17)	46,52 (4,24)	44,23 (6,50)	41,33 (11,68)
	20%	5%	2,47 (1,30)	2,70 (2,26)	1,92 (1,09)	1,62 (0,98)
		10%	4,03 (2,21)	2,98 (1,40)	3,03 (0,37)	2,82 (1,79)
		20%	7,16 (2,73)	7,55 (1,97)	6,19 (2,10)	5,68 (1,26)
		40%	14,19 (4,14)	15,30 (5,22)	12,97 (3,69)	11,30 (3,09)
GL 3	80%	5%	8,64 (3,03)	8,16 (3,54)	7,06 (3,10)	5,43 (2,70)
		10%	16,25 (4,46)	13,89 (1,75)	13,00 (4,56)	12,50 (5,36)
		20%	35,47 (7,62)	41,45 (12,20)	37,97 (11,43)	32,02 (12,26)
		40%	83,23 (15,66)	91,42 (18,55)	82,61 (22,53)	81,06 (20,87)
	60%	5%	6,34 (2,27)	6,18 (3,04)	5,66 (4,01)	5,06 (4,07)
		10%	12,06 (3,50)	13,03 (4,56)	12,60 (5,10)	11,88 (3,79)
		20%	26,24 (5,70)	28,53 (5,20)	27,26 (6,94)	28,03 (9,87)
		40%	57,86 (9,31)	51,79 (9,10)	51,95 (11,27)	45,82 (20,37)
	20%	5%	2,72 (1,41)	1,20 (0,83)	1,67 (1,01)	1,80 (1,01)
		10%	4,45 (1,82)	5,04 (2,53)	5,78 (2,69)	5,80 (2,91)
		20%	8,98 (2,96)	9,31 (3,87)	9,83 (3,75)	8,02 (2,47)
		40%	17,02 (4,23)	16,43 (3,81)	13,55 (2,44)	12,43 (2,01)
GL 5	80%	5%	9,58 (3,19)	10,85 (2,46)	10,11 (0,53)	10,50 (2,92)
		10%	17,43 (4,32)	19,85 (6,48)	16,01 (6,58)	14,69 (6,95)
		20%	38,11 (7,82)	38,89 (12,32)	32,65 (9,15)	30,55 (14,26)
		40%	92,58 (13,64)	95,17 (9,30)	96,69 (22,85)	78,63 (6,59)
	60%	5%	7,03 (2,28)	8,48 (0,98)	7,34 (1,65)	6,20 (1,27)
		10%	13,18 (3,87)	11,59 (6,18)	12,95 (9,30)	12,71 (8,11)
		20%	27,32 (5,94)	29,02 (6,84)	25,28 (8,82)	21,39 (7,17)
		40%	62,99 (11,62)	67,62 (13,82)	67,62 (18,55)	62,79 (26,39)
	20%	5%	2,98 (1,47)	3,74 (1,57)	3,33 (1,41)	3,21 (1,30)
		10%	4,93 (2,24)	5,44 (3,29)	4,06 (2,33)	4,28 (1,77)
		20%	9,14 (2,98)	6,22 (3,07)	6,78 (1,97)	5,37 (1,78)
		40%	17,71 (4,57)	15,70 (4,65)	15,99 (4,64)	13,04 (3,62)

GL = grupo de ligação; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem e $\overline{\text{LOD}}_{\text{máx}}$ = média do LOD máximo.

Tabela 10 - Estimativas médias da posição relativa do QTL e seus respectivos desvios-padrão, entre parênteses, obtidos por regressão linear pelo método de intervalo simples em função do nível de herdabilidade e de ação gênica, nos grupos de ligação 1 (POP 1 a 12), 3 (POP 13 a 24) e 5 (POP 25 a 36) das 3.600 populações F₂ analisadas com 1.000 indivíduos e das 180 populações amostradas (cinco de cada cenário) por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

GL	h ²	AG	$\bar{r}_{\text{máx}}$			
			1.000	500	300	200
GL 1	80%	5%	50,10 (6,65)	39,10 (8,81)	40,55 (6,47)	42,47 (9,25)
		10%	50,59 (3,42)	42,19 (6,91)	45,84 (7,47)	45,64 (9,45)
		20%	51,11 (2,10)	48,19 (2,17)	48,92 (1,00)	49,75 (1,52)
		40%	51,36 (1,99)	57,49 (2,41)	55,77 (1,95)	54,34 (2,88)
	60%	5%	51,24 (8,90)	44,90 (7,46)	47,88 (5,45)	49,00 (3,96)
		10%	51,38 (4,22)	42,68 (7,60)	44,10 (3,96)	45,85 (8,77)
		20%	50,92 (3,02)	42,53 (6,73)	45,15 (1,00)	44,25 (1,95)
		40%	51,20 (1,95)	50,42 (1,10)	50,21 (2,30)	49,39 (0,71)
	20%	5%	47,46 (19,56)	56,78 (19,74)	59,15 (20,07)	57,77 (14,27)
		10%	48,64 (13,46)	40,33 (9,15)	41,00 (10,53)	45,14 (8,08)
		20%	50,76 (7,44)	47,42 (6,43)	46,87 (5,79)	45,78 (5,63)
		40%	51,11 (3,30)	45,38 (2,70)	46,33 (7,87)	47,88 (9,20)
GL 3	80%	5%	56,22 (3,15)	49,03 (3,27)	49,34 (4,09)	46,74 (4,74)
		10%	56,46 (2,33)	47,48 (1,92)	48,83 (2,74)	49,75 (4,97)
		20%	56,50 (1,91)	54,08 (4,93)	52,83 (5,54)	52,00 (3,67)
		40%	55,53 (1,40)	62,57 (1,64)	60,27 (1,48)	60,01 (2,17)
	60%	5%	56,12 (3,55)	48,36 (16,74)	46,61 (16,12)	34,01 (30,43)
		10%	56,57 (2,85)	47,84 (2,05)	47,77 (3,83)	46,98 (4,64)
		20%	56,55 (2,01)	52,42 (1,48)	52,38 (2,70)	54,17 (2,79)
		40%	55,04 (1,18)	55,30 (2,00)	51,82 (9,04)	53,47 (2,45)
	20%	5%	56,59 (9,75)	61,20 (17,11)	46,95 (11,45)	47,13 (12,32)
		10%	55,53 (5,04)	40,51 (17,97)	44,96 (9,65)	47,47 (7,01)
		20%	56,46 (3,20)	45,39 (9,32)	48,80 (2,07)	48,17 (3,03)
		40%	56,22 (3,15)	53,74 (1,92)	52,76 (1,95)	51,91 (3,56)
GL 5	80%	5%	53,84 (4,31)	46,47 (3,29)	43,35 (7,38)	43,42 (7,58)
		10%	53,64 (2,58)	47,14 (1,82)	47,17 (2,41)	47,45 (5,15)
		20%	54,20 (1,98)	51,48 (2,00)	50,11 (1,82)	49,43 (4,09)
		40%	54,35 (2,60)	60,19 (3,11)	60,78 (3,36)	57,61 (5,77)
	60%	5%	52,96 (6,11)	47,25 (2,92)	44,56 (5,79)	44,87 (6,23)
		10%	54,50 (3,75)	46,93 (1,87)	47,66 (3,13)	46,78 (3,13)
		20%	54,19 (2,61)	48,34 (2,59)	47,71 (1,92)	48,23 (4,04)
		40%	54,27 (1,74)	55,94 (2,19)	56,77 (2,61)	55,63 (3,27)
	20%	5%	53,02 (18,29)	49,46 (9,84)	41,06 (8,96)	44,04 (10,35)
		10%	53,77 (11,53)	50,07 (6,78)	47,51 (6,06)	47,55 (5,45)
		20%	54,49 (4,66)	45,87 (4,32)	46,13 (2,74)	45,39 (8,73)
		40%	54,00 (2,61)	46,95 (3,27)	46,16 (2,59)	46,67 (2,07)

GL = grupo de ligação; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem; $\bar{r}_{\text{máx}}$ = média da posição relativa do QTL.

Tabela 11 - Estimativas médias do coeficiente de determinação do modelo de regressão linear por marca simples e seus desvios-padrão, entre parênteses, obtidos por regressão linear pelo método de intervalo simples em função do nível de herdabilidade e de ação gênica, nos grupos de ligação 1 (POP 1 a 12), 3 (POP 13 a 24) e 5 (POP 25 a 36) das 3.600 populações F₂ analisadas com 1.000 indivíduos e das 180 populações amostradas (cinco de cada cenário) por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

GL	h ²	AG	\bar{R}^2			
			1.000	500	300	200
GL 1	80%	5%	3,52 (1,18)	5,03 (2,00)	7,83 (1,45)	10,47 (1,71)
		10%	6,37 (1,71)	12,16 (3,30)	18,00 (5,55)	23,58 (7,90)
		20%	13,61 (2,67)	26,28 (2,97)	39,70 (9,23)	49,04 (18,16)
		40%	27,45 (3,97)	53,17 (5,57)	68,87 (7,22)	77,71 (8,23)
	60%	5%	2,49 (1,01)	3,08 (1,60)	4,88 (2,85)	6,53 (4,80)
		10%	5,01 (1,23)	7,18 (1,08)	8,21 (2,10)	10,98 (4,61)
		20%	9,46 (2,02)	18,44 (3,85)	27,31 (5,27)	32,85 (7,61)
		40%	19,65 (3,01)	34,84 (2,54)	49,12 (5,14)	60,33 (10,13)
	20%	5%	1,13 (0,59)	2,44 (2,01)	2,89 (1,62)	3,65 (2,17)
		10%	1,84 (0,99)	2,71 (1,25)	4,54 (0,54)	6,22 (3,81)
		20%	3,24 (1,22)	6,71 (1,70)	9,04 (2,93)	12,24 (2,56)
		40%	6,31 (1,78)	13,08 (4,14)	17,96 (4,60)	22,78 (5,32)
GL 3	80%	5%	3,89 (1,33)	7,21 (3,05)	10,20 (4,31)	11,62 (5,49)
		10%	7,20 (1,90)	12,01 (1,41)	17,95 (5,65)	24,59 (8,81)
		20%	15,03 (2,98)	31,43 (7,42)	43,52 (9,73)	50,76 (12,59)
		40%	31,69 (5,07)	56,45 (7,45)	70,57 (9,67)	83,13 (7,78)
	60%	5%	2,87 (1,01)	5,51 (2,60)	8,19 (5,54)	10,69 (8,07)
		10%	5,40 (1,52)	11,26 (3,64)	17,40 (6,41)	23,72 (6,81)
		20%	11,37 (2,30)	23,06 (3,74)	33,92 (7,01)	46,51 (11,87)
		40%	23,34 (3,23)	37,79 (5,05)	54,44 (7,94)	62,58 (13,34)
	20%	5%	1,24 (0,64)	1,10 (0,76)	2,53 (1,50)	4,04 (2,20)
		10%	2,02 (0,82)	4,52 (2,23)	8,43 (3,81)	12,34 (6,08)
		20%	4,05 (1,31)	8,18 (3,29)	13,90 (5,01)	16,77 (4,75)
		40%	7,53 (1,79)	14,01 (3,01)	18,76 (3,05)	24,85 (3,49)
GL 5	80%	5%	4,31 (1,40)	9,50 (2,04)	14,39 (0,69)	21,36 (5,18)
		10%	7,70 (1,83)	16,60 (4,96)	21,49 (7,90)	28,02 (10,59)
		20%	16,06 (2,99)	29,77 (7,93)	38,97 (8,50)	48,52 (15,29)
		40%	34,61 (4,04)	58,29 (3,54)	76,36 (6,87)	83,52 (2,47)
	60%	5%	3,18 (1,02)	7,52 (0,83)	10,64 (2,28)	13,28 (2,55)
		10%	5,88 (1,67)	10,02 (4,96)	17,42 (10,71)	24,41 (12,61)
		20%	11,80 (2,42)	23,35 (4,88)	31,71 (8,91)	38,29 (9,41)
		40%	25,10 (4,02)	46,03 (7,08)	63,53 (9,49)	73,53 (11,56)
	20%	5%	1,36 (0,67)	3,38 (1,39)	4,97 (2,06)	7,10 (2,77)
		10%	2,24 (1,00)	4,86 (2,89)	6,00 (3,37)	9,33 (3,76)
		20%	4,12 (1,31)	5,55 (2,67)	9,87 (2,69)	11,58 (3,54)
		40%	7,82 (1,94)	13,41 (3,70)	21,62 (5,60)	25,76 (5,98)

GL = grupo de ligação; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem e \bar{R}^2 = média do coeficiente de determinação.

Em relação ao grupo de ligação 1 (POP 1 a POP 12), o valor esperado da posição do QTL simulado foi de 50 cM, sendo que o marcador anterior foi posicionado a 40 cM e o posterior na posição de 60 cM. Para o grupo de ligação 3 (POP 13 a 24), o valor da posição do QTL foi de 55 cM, sendo que a posição dos marcadores flanqueadores estava localizada a 50 e 60 cM. No grupo de ligação 5 (POP 25 a 36), a posição do QTL foi de 52,5 cM e os limites da posição dos marcadores foram de 50 cM e 55 cM.

A precisão da localização do QTL simulado em qualquer cenário selecionado por genotipagem seletiva depende, assim como nas populações de referências, do nível de ação gênica. Além disso, foi verificado que a redução no tamanho da população também influenciou a precisão e acurácia da localização do QTL, assim como a amostragem aleatória de indivíduos (LI et al., 2006).

Em relação aos valores do R^2 médio (Tabela 11), verificou-se que à medida que a genotipagem seletiva foi intensificada, ocorreram aumentos nestes, conforme observado no método de marcas simples para os marcadores flanqueadores do QTL simulado. Em relação às populações de referência, o acréscimo médio do R^2 foi de 1,84, 2,58 e 3,16 para os tamanhos das populações amostradas por genotipagem seletiva de 500, 300 e 200, respectivamente.

Baseado na relação quantitativa, observada entre o R^2 e o tamanho da população amostrada por genotipagem seletiva, propôs-se uma correção empírica para o R^2 . Assim, ao invés de calcular o R^2 para o tamanho da população amostrada (Tabela 11), o R^2 foi obtido para o tamanho da população original de 1.000 indivíduos (Tabela 12). Portanto, verificou-se que a média do R^2 corrigido foi similar àquela do R^2 médio, obtido nas populações de 1.000 indivíduos, em qualquer cenário. O viés da média do R^2 corrigido, em relação às estimativas das populações de referência de 1.000 indivíduos, foi de 2,8, 3,6 e 7,6% para os tamanho de 500, 300 e 200 indivíduos, respectivamente.

Tabela 12 - Estimativas médias do coeficiente de determinação do modelo de regressão linear por marca simples e seus desvios-padrão, entre parênteses, obtidos por regressão linear pelo método de intervalo simples, em função do nível de herdabilidade e de ação gênica, nos grupos de ligação 1 (POP 1 a 12), 3 (POP 13 a 24) e 5 (POP 25 a 36) das 3.600 populações F₂ analisadas com 1.000 indivíduos e das 180 populações amostradas (cinco de cada cenário) por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

GL	h ²	AG	\bar{R}^2			
			1.000	500	300	200
GL 1	80%	5%	3,52 (1,18)	2,19 (0,37)	2,42 (0,46)	2,55 (1,03)
		10%	6,37 (1,71)	5,30 (2,01)	5,82 (1,96)	6,29 (1,74)
		20%	13,61 (2,67)	13,34 (6,31)	14,25 (3,98)	14,15 (1,74)
		40%	27,45 (3,97)	26,71 (6,18)	29,91 (5,40)	31,67 (4,18)
	60%	5%	2,49 (1,01)	1,36 (1,02)	1,50 (0,88)	1,55 (0,81)
		10%	5,01 (1,23)	2,32 (1,01)	2,54 (0,67)	3,66 (0,56)
		20%	9,46 (2,02)	7,73 (2,11)	9,17 (1,98)	9,71 (2,13)
		40%	19,65 (3,01)	17,25 (4,42)	18,42 (2,45)	19,29 (1,57)
	20%	5%	1,13 (0,59)	0,75 (0,45)	0,88 (0,50)	1,23 (1,02)
		10%	1,84 (0,99)	1,29 (0,81)	1,39 (0,17)	1,36 (0,64)
		20%	3,24 (1,22)	2,58 (0,57)	2,81 (0,94)	3,42 (0,88)
		40%	6,31 (1,78)	5,07 (1,34)	5,79 (1,60)	6,79 (2,23)
GL 3	80%	5%	3,89 (1,33)	3,68 (1,58)	3,20 (1,39)	2,46 (1,21)
		10%	7,20 (1,90)	6,20 (0,75)	5,80 (1,97)	5,58 (2,31)
		20%	15,03 (2,98)	17,29 (4,56)	15,97 (4,40)	13,62 (4,78)
		40%	31,69 (5,07)	34,20 (5,64)	31,38 (7,05)	30,93 (6,58)
	60%	5%	2,87 (1,01)	2,80 (1,35)	2,56 (1,79)	2,29 (1,82)
		10%	5,40 (1,52)	5,82 (1,96)	5,62 (2,21)	5,32 (1,66)
		20%	11,37 (2,30)	12,30 (2,12)	11,77 (2,82)	12,05 (3,99)
		40%	23,34 (3,23)	21,18 (3,25)	21,21 (4,10)	18,77 (7,23)
	20%	5%	1,24 (0,64)	0,55 (0,38)	0,77 (0,46)	0,82 (0,46)
		10%	2,02 (0,82)	2,29 (1,14)	2,62 (1,21)	2,63 (1,31)
		20%	4,05 (1,31)	4,19 (1,71)	4,42 (1,66)	3,62 (1,10)
		40%	7,53 (1,79)	7,28 (1,62)	6,05 (1,06)	5,57 (0,87)
GL 5	80%	5%	4,31 (1,40)	4,87 (1,08)	4,55 (0,23)	4,72 (1,28)
		10%	7,70 (1,83)	8,71 (2,72)	7,08 (2,82)	6,51 (2,95)
		20%	16,06 (2,99)	16,31 (4,74)	13,91 (3,62)	12,99 (5,60)
		40%	34,61 (4,04)	35,47 (2,75)	35,69 (6,42)	30,38 (2,11)
	60%	5%	3,18 (1,02)	3,83 (0,43)	3,32 (0,74)	2,82 (0,57)
		10%	5,88 (1,67)	5,17 (2,66)	5,73 (3,93)	5,64 (3,45)
		20%	11,80 (2,42)	12,49 (2,77)	10,94 (3,58)	9,35 (2,95)
		40%	25,10 (4,02)	26,66 (4,75)	26,57 (6,12)	24,71 (8,59)
	20%	5%	1,36 (0,67)	1,71 (0,71)	1,52 (0,64)	1,47 (0,59)
		10%	2,24 (1,00)	2,47 (1,48)	1,85 (1,06)	1,95 (0,80)
		20%	4,12 (1,31)	2,82 (1,37)	3,08 (0,88)	2,44 (0,80)
		40%	7,82 (1,94)	6,96 (1,99)	7,09 (1,99)	5,82 (1,56)

GL = grupo de ligação; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem e \bar{R}^2 = média do coeficiente de determinação.

4.2.3. Estimação do efeito do QTL sobre a característica em populações amostradas

Para os marcadores flanqueadores do QTL simulado, o efeito do QTL sobre a característica fenotípica foi avaliado por meio de regressão linear, pelo método de marca simples nas 180 populações, amostradas por genotipagem seletiva com 500, 300 e 200 indivíduos e comparadas com as 3.600 populações F_2 com 1.000 indivíduos.

Utilizando o modelo de regressão, foi possível estimar a média geral, o efeito aditivo e o efeito devido à dominância. Para os cenários das POP 1 a 12, POP 13 a 24 e, POP 25 a 36, as médias das estimativas do modelo de regressão encontram-se nas tabelas 13, 14 e 15, nas tabelas 16, 17 e 18 e, nas tabelas 19, 20 e 21, respectivamente.

Nas tabelas 13, 16 e 19, encontram-se as estimativas da média geral para a característica quantitativa. Verificou-se que, independente da intensidade de amostragem as médias das estimativas estavam em torno da média paramétrica simulada, que era de 1.000 unidades. Os marcadores $C_{13} C_{14}$, $C_{311} C_{312}$ e $C_{511} C_{512}$ estavam afastados em 10, 5, 2,5 cM do QTL simulado, respectivamente. Para a média do efeito aditivo “ $a = \bar{\beta}_1(1-2r)^{-1}$ ”, observou-se um acréscimo de 75, 135 e 180% nas intensidades de amostragem de 50, 30 e 20%, respectivamente (Tabelas 14, 17 e 20), em comparação com as médias das estimativas do efeito aditivo das populações de referência. Nas populações com 1.000 indivíduos, observou-se que a média do efeito devido à dominância aproximou-se de zero, conforme esperado. À medida que a intensidade de amostragem aumentou, os valores médios das estimativas de $\hat{\beta}_2$ que refletem os efeitos devido à dominância foram superestimados (Tabelas 15, 18 e 21). Entretanto, o grau médio de dominância permaneceu em torno de zero e, conseqüentemente o viés médio encontrado na estimativa dos efeitos devido à dominância, em populações amostradas, foi proporcionalmente igual ao viés médio da estimativa do efeito aditivo.

Tabela 13 - Médias das estimativas de $\hat{\beta}_0$ dos marcadores C₁₃ e C₁₄ sobre a característica e seus respectivos desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F₂ analisadas com 1.000 indivíduos e das 180 populações amostradas por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h ²	AG	1.000	500	300	200
C ₁₃	80%	5%	999,01 (7,37)	1008,26 (8,97)	1008,95 (15,33)	1022,42 (22,20)
		10%	999,79 (6,36)	1020,76 (9,10)	1036,98 (12,20)	1041,48 (23,24)
		20%	1002,03 (7,01)	1001,89 (10,62)	1002,67 (18,34)	1013,28 (16,44)
		40%	1000,85 (6,01)	1002,43 (8,60)	1004,04 (13,80)	1003,56 (16,47)
	60%	5%	999,79 (4,75)	995,52 (9,45)	995,34 (18,06)	995,78 (17,36)
		10%	999,50 (4,88)	996,29 (16,70)	994,68 (24,08)	987,22 (21,84)
		20%	999,77 (4,42)	1006,88 (14,30)	1009,43 (18,98)	1020,74 (24,98)
		40%	1000,25 (5,46)	1000,07 (10,58)	999,25 (14,39)	1007,11 (27,91)
	20%	5%	1000,24 (3,84)	997,09 (3,74)	995,70 (4,65)	997,75 (7,91)
		10%	1000,08 (3,02)	1000,61 (4,53)	1000,19 (7,93)	996,26 (15,57)
		20%	1000,56 (3,38)	999,52 (5,03)	998,60 (11,06)	992,88 (11,66)
		40%	1000,71 (3,50)	997,28 (7,82)	998,64 (9,00)	1003,89 (14,38)
C ₁₄	80%	5%	1000,09 (7,08)	1011,60 (3,43)	1014,24 (9,80)	1025,69 (12,66)
		10%	999,93 (6,91)	1011,52 (7,88)	1019,30 (22,14)	1017,30 (15,41)
		20%	1001,31 (6,92)	1007,80 (13,72)	1003,73 (26,87)	1019,21 (30,11)
		40%	1000,40 (7,81)	1003,75 (10,47)	1002,51 (9,57)	1003,22 (13,18)
	60%	5%	999,06 (5,15)	999,84 (9,65)	1002,01 (24,77)	1007,11 (20,54)
		10%	998,66 (4,99)	993,47 (14,07)	994,14 (27,17)	995,19 (26,07)
		20%	1000,72 (4,82)	1005,14 (9,91)	1007,14 (17,01)	1017,73 (15,81)
		40%	1000,23 (4,87)	1009,06 (11,51)	1011,43 (12,03)	1008,54 (21,59)
	20%	5%	1000,11 (3,51)	996,16 (6,61)	997,31 (8,88)	997,80 (8,81)
		10%	999,78 (3,66)	997,26 (9,29)	996,27 (7,25)	992,73 (13,66)
		20%	1000,44 (3,53)	998,53 (7,15)	996,66 (15,81)	993,34 (23,11)
		40%	1000,53 (3,34)	1000,83 (11,45)	1003,45 (11,79)	1007,67 (21,96)

M = marcador; C₁₃ = marcador três do grupo de ligação um; C₁₄ = marcador quatro do grupo de ligação um; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem.

Tabela 14 - Médias das estimativas de $\hat{\beta}_1$ dos C_{13} e C_{14} sobre a característica e seus respectivos desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F_2 analisadas com 1000 indivíduos e das 180 populações amostradas por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h^2	AG	1.000	500	300	200
C_{13}	80%	5%	51,69 (10,14)	87,06 (17,38)	129,47 (20,02)	163,08 (20,87)
		10%	71,40 (10,31)	134,58 (16,72)	179,87 (31,79)	219,19 (44,24)
		20%	101,91 (10,56)	186,24 (20,18)	256,40 (36,65)	301,86 (61,29)
		40%	141,72 (9,97)	238,29 (11,34)	307,45 (9,92)	347,67 (16,26)
	60%	5%	30,10 (8,25)	38,90 (17,71)	52,92 (27,37)	67,31 (39,01)
		10%	43,60 (5,88)	73,00 (6,76)	92,45 (10,89)	115,52 (25,92)
		20%	61,49 (7,04)	111,33 (8,59)	156,32 (15,29)	187,98 (23,73)
		40%	86,67 (6,01)	149,10 (4,78)	202,22 (11,15)	241,43 (16,95)
	20%	5%	12,36 (5,12)	14,87 (8,89)	25,71 (14,03)	33,00 (17,88)
		10%	17,92 (5,13)	30,55 (9,05)	48,87 (5,32)	60,61 (19,07)
		20%	23,99 (5,36)	43,22 (8,77)	54,60 (11,95)	75,38 (16,41)
		40%	34,74 (4,44)	62,62 (8,19)	88,70 (10,16)	107,93 (19,17)
C_{14}	80%	5%	50,54 (10,99)	77,04 (17,78)	114,05 (21,82)	149,36 (27,16)
		10%	69,86 (10,39)	128,44 (22,53)	181,85 (29,33)	221,20 (29,42)
		20%	101,29 (10,98)	186,55 (20,21)	263,30 (40,34)	308,25 (59,04)
		40%	142,27 (10,52)	242,45 (16,70)	307,45 (15,78)	354,66 (13,70)
	60%	5%	30,13 (7,60)	45,83 (15,93)	68,90 (24,89)	82,86 (38,63)
		10%	43,73 (6,48)	72,48 (14,10)	89,75 (24,45)	114,78 (34,65)
		20%	60,28 (6,99)	101,57 (15,78)	144,47 (19,78)	173,17 (23,61)
		40%	86,41 (5,99)	148,99 (10,80)	196,48 (11,62)	232,23 (18,88)
	20%	5%	11,94 (5,23)	16,03 (16,39)	31,00 (20,75)	39,12 (26,00)
		10%	17,35 (5,34)	28,92 (9,82)	43,90 (9,69)	55,26 (24,27)
		20%	24,23 (5,33)	47,82 (5,92)	63,41 (12,64)	80,72 (6,92)
		40%	34,96 (5,21)	66,35 (7,00)	91,57 (13,50)	113,80 (13,32)

M = marcador; C_{13} = marcador três do grupo de ligação um; C_{14} = marcador quatro do grupo de ligação um; h^2 = herdabilidade em porcentagem; AG = ação gênica em porcentagem.

Tabela 15 - Médias das estimativas de $\hat{\beta}_2$ dos marcadores C₁₃ e C₁₄ sobre a característica e seus respectivos desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F₂ analisadas com 1000 indivíduos e das 180 populações amostradas por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h ²	AG	1.000	500	300	200
C ₁₃	80%	5%	1,94 (14,62)	-6,92 (16,97)	0,42 (24,61)	-15,23 (39,40)
		10%	0,38 (12,40)	-25,99 (9,05)	-59,51 (30,30)	-85,53 (32,59)
		20%	-3,88 (13,10)	-1,77 (23,37)	-18,18 (39,69)	-40,16 (44,27)
		40%	-1,64 (12,07)	-9,62 (23,96)	-10,52 (46,42)	0,67 (68,55)
	60%	5%	0,44 (9,23)	9,17 (8,74)	7,44 (15,04)	9,35 (19,62)
		10%	1,00 (9,54)	-3,30 (28,94)	-8,89 (30,69)	2,11 (26,69)
		20%	0,46 (8,72)	-9,45 (21,58)	-6,54 (25,21)	-19,97 (37,57)
		40%	-0,44 (10,35)	3,49 (19,28)	9,64 (20,28)	8,25 (32,65)
	20%	5%	-0,54 (7,54)	4,84 (7,75)	8,59 (6,03)	6,26 (22,62)
		10%	-0,12 (6,08)	-6,92 (4,92)	-8,99 (9,29)	-4,33 (16,64)
		20%	-1,06 (6,60)	-7,79 (10,20)	-12,71 (21,37)	-8,19 (13,69)
		40%	-1,37 (6,92)	1,96 (13,14)	1,91 (16,19)	-2,03 (9,71)
C ₁₄	80%	5%	-0,01 (13,97)	-12,46 (12,79)	-17,11 (20,66)	-32,17 (33,97)
		10%	0,41 (13,61)	-7,60 (12,26)	-20,17 (36,24)	-22,71 (26,47)
		20%	-2,18 (13,36)	-10,01 (20,81)	-16,57 (46,97)	-42,62 (59,11)
		40%	-0,36 (15,35)	-9,74 (32,00)	5,20 (39,65)	14,87 (69,54)
	60%	5%	2,01 (10,17)	2,30 (14,00)	-4,40 (28,51)	-7,15 (35,06)
		10%	2,85 (10,07)	2,80 (25,66)	-5,34 (47,08)	-7,04 (44,17)
		20%	-1,17 (9,37)	-4,29 (13,89)	-3,02 (19,38)	-16,55 (9,99)
		40%	-0,20 (9,27)	-6,93 (16,44)	-0,52 (17,79)	2,73 (29,63)
	20%	5%	-0,19 (6,93)	6,13 (10,51)	5,32 (15,75)	5,80 (21,71)
		10%	0,50 (7,28)	1,51 (14,12)	0,06 (13,39)	1,73 (25,19)
		20%	-0,75 (7,06)	-5,46 (12,79)	-7,80 (23,44)	-8,58 (29,95)
		40%	-0,92 (6,44)	-3,20 (17,18)	-7,36 (16,37)	-11,01 (21,58)

M = marcador; C₁₃ = marcador três do grupo de ligação um; C₁₄ = marcador quatro do grupo de ligação um; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem.

Tabela 16 - Médias das estimativas de $\hat{\beta}_0$ dos marcadores C_{36} e C_{37} sobre a característica e seus respectivos desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F_2 analisadas com 1000 indivíduos e das 180 populações amostradas por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h^2	AG	1.000	500	300	200
C_{36}	80%	5%	1000,79 (6,96)	1010,05 (17,96)	1019,11 (26,61)	1009,63 (55,51)
		10%	999,58 (6,34)	992,68 (11,81)	1005,48 (27,98)	1003,20 (40,55)
		20%	1000,21 (6,42)	1016,66 (5,31)	1009,45 (18,43)	1011,11 (23,92)
		40%	999,84 (5,98)	1004,25 (19,69)	1004,94 (23,69)	997,61 (20,81)
	60%	5%	999,60 (4,72)	993,63 (7,10)	990,98 (12,71)	989,52 (15,09)
		10%	1000,19 (4,69)	1000,35 (6,55)	1002,79 (14,04)	987,29 (25,72)
		20%	1000,60 (4,64)	992,01 (11,59)	989,95 (16,73)	989,70 (30,62)
		40%	999,60 (4,27)	1007,35 (10,99)	1007,39 (17,00)	1008,96 (30,85)
	20%	5%	999,92 (3,38)	1001,46 (1,57)	1003,89 (6,18)	1007,90 (12,32)
		10%	999,86 (3,26)	997,16 (8,47)	1003,34 (13,00)	1004,12 (14,18)
		20%	999,70 (3,36)	998,18 (7,91)	999,77 (17,47)	1008,09 (27,37)
		40%	1000,28 (3,35)	991,09 (2,17)	989,39 (7,22)	990,49 (5,53)
C_{37}	80%	5%	999,32 (6,86)	1018,88 (20,54)	1035,20 (22,60)	1028,32 (59,93)
		10%	999,34 (6,53)	997,17 (9,09)	1012,72 (23,61)	1003,64 (47,85)
		20%	999,78 (7,33)	1014,39 (11,56)	1006,78 (28,84)	1007,25 (37,24)
		40%	1000,20 (6,87)	1005,80 (11,25)	1006,88 (14,63)	998,63 (10,46)
	60%	5%	999,84 (4,97)	994,27 (7,27)	992,95 (6,13)	996,98 (15,20)
		10%	1000,26 (4,46)	1001,13 (4,67)	1002,26 (8,59)	981,86 (19,76)
		20%	1000,55 (5,39)	998,05 (11,43)	996,02 (18,09)	993,24 (35,92)
		40%	999,74 (4,55)	1005,65 (8,29)	1002,24 (16,85)	998,46 (28,10)
	20%	5%	999,85 (3,77)	1001,36 (3,63)	1007,65 (5,73)	1008,89 (10,29)
		10%	1000,05 (3,34)	998,11 (8,43)	1000,49 (9,95)	994,48 (11,92)
		20%	999,22 (3,19)	1000,17 (6,37)	1003,46 (12,08)	1008,18 (23,60)
		40%	999,98 (3,73)	991,22 (5,33)	991,07 (7,27)	991,66 (7,37)

M = marcador; C_{36} = marcador seis do grupo de ligação três; C_{37} = marcador sete do grupo de ligação três; h^2 = herdabilidade em porcentagem; AG = ação gênica em porcentagem.

Tabela 17 - Médias das estimativas de $\hat{\beta}_1$ dos marcadores C₃₆ e C₃₇ sobre a característica e seus respectivos desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F₂ analisadas com 1000 indivíduos e das 180 populações amostradas por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h ²	AG	1.000	500	300	200
C ₃₆	80%	5%	57,22 (11,26)	100,26 (25,99)	133,51 (37,82)	147,97 (48,07)
		10%	79,03 (11,05)	146,13 (8,21)	204,13 (26,73)	255,65 (37,07)
		20%	114,21 (9,92)	208,23 (17,10)	276,78 (30,79)	325,94 (26,30)
		40%	160,46 (8,88)	267,94 (7,34)	336,35 (11,30)	382,30 (9,29)
	60%	5%	34,21 (6,93)	56,32 (7,44)	78,24 (33,59)	86,14 (53,07)
		10%	49,07 (7,31)	91,82 (10,25)	128,63 (18,47)	166,06 (20,77)
		20%	69,78 (6,65)	129,04 (8,64)	179,27 (13,11)	222,09 (16,61)
		40%	98,63 (6,60)	159,84 (16,07)	214,00 (22,01)	252,58 (17,59)
	20%	5%	13,96 (5,26)	16,44 (6,27)	29,67 (14,02)	46,32 (12,31)
		10%	19,33 (4,96)	40,03 (15,49)	62,75 (26,62)	82,93 (29,69)
		20%	29,57 (5,01)	56,91 (11,26)	86,03 (13,67)	105,22 (11,89)
		40%	39,79 (4,77)	69,12 (11,61)	91,81 (14,37)	114,21 (14,22)
C ₃₇	80%	5%	45,00 (11,00)	102,84 (24,34)	138,62 (31,51)	159,70 (44,32)
		10%	78,59 (10,56)	137,47 (5,77)	194,35 (27,99)	247,12 (40,70)
		20%	113,38 (10,29)	203,57 (11,24)	271,55 (19,52)	316,51 (27,63)
		40%	160,09 (8,52)	273,06 (14,66)	342,36 (13,40)	394,44 (10,88)
	60%	5%	34,02 (6,90)	65,40 (10,40)	87,48 (25,78)	104,65 (35,26)
		10%	48,18 (6,76)	85,63 (13,29)	119,83 (20,33)	152,13 (22,50)
		20%	69,18 (5,98)	127,02 (12,00)	178,07 (19,77)	225,38 (28,64)
		40%	98,67 (6,08)	159,88 (15,89)	216,05 (24,64)	254,68 (27,14)
	20%	5%	14,04 (5,38)	19,30 (8,15)	30,88 (15,41)	47,93 (19,27)
		10%	19,74 (5,34)	38,93 (14,59)	60,10 (24,43)	83,54 (26,59)
		20%	28,97 (5,15)	56,56 (14,74)	87,92 (18,00)	102,82 (16,99)
		40%	39,76 (4,74)	77,34 (9,02)	102,26 (12,35)	127,50 (10,80)

M = marcador; C₃₆ = marcador seis do grupo de ligação três; C₃₇ = marcador sete do grupo de ligação três; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem.

Tabela 18 - Médias das estimativas de $\hat{\beta}_2$ dos marcadores C₃₆ e C₃₇ sobre a característica e seus respectivos desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F₂ analisadas com 1.000 indivíduos e das 180 populações amostradas por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h ²	AG	1.000	500	300	200
C ₃₆	80%	5%	-1,57 (13,81)	-8,97 (23,23)	-13,58 (42,45)	10,47 (96,86)
		10%	0,88 (12,18)	8,06 (12,90)	-10,50 (34,46)	4,46 (53,54)
		20%	-0,22 (12,18)	-6,21 (10,62)	16,04 (22,56)	25,06 (44,64)
		40%	0,55 (11,37)	8,41 (33,15)	2,97 (41,00)	-2,44 (39,28)
	60%	5%	0,81 (9,45)	9,57 (15,27)	10,00 (24,29)	12,45 (35,45)
		10%	-0,32 (9,03)	-2,02 (18,58)	-8,29 (34,09)	10,36 (71,53)
		20%	-1,03 (9,26)	14,91 (20,40)	17,64 (22,71)	17,81 (32,77)
		40%	0,92 (8,06)	-5,05 (13,21)	-3,85 (14,59)	6,80 (23,24)
	20%	5%	0,19 (6,76)	-6,28 (10,40)	-9,00 (17,71)	-11,20 (17,64)
		10%	0,30 (6,48)	5,11 (12,38)	-2,28 (15,38)	1,12 (15,72)
		20%	0,62 (6,69)	6,32 (7,52)	3,59 (18,92)	-5,92 (21,14)
		40%	-0,45 (6,47)	9,78 (5,64)	11,97 (12,89)	20,01 (18,74)
C ₃₇	80%	5%	1,17 (13,57)	-25,23 (30,00)	-45,37 (42,34)	-27,14 (105,23)
		10%	1,26 (12,84)	2,21 (11,82)	-16,96 (22,96)	9,08 (59,27)
		20%	0,36 (14,16)	-4,00 (19,44)	18,20 (19,83)	19,45 (32,05)
		40%	-0,39 (13,25)	9,04 (23,61)	3,58 (28,99)	-3,65 (46,77)
	60%	5%	0,33 (10,02)	8,70 (11,85)	7,59 (10,79)	5,41 (31,28)
		10%	-0,60 (8,94)	-4,30 (12,10)	-12,09 (28,11)	15,39 (58,84)
		20%	-1,10 (10,47)	2,30 (23,21)	6,55 (19,39)	6,11 (35,99)
		40%	0,51 (8,67)	-4,74 (11,69)	-2,68 (13,52)	10,53 (26,37)
	20%	5%	0,32 (7,54)	-5,33 (9,77)	-15,34 (17,12)	-11,55 (17,68)
		10%	-0,11 (6,67)	3,91 (13,42)	4,10 (8,90)	19,97 (12,20)
		20%	1,52 (6,35)	2,99 (7,61)	-4,31 (8,06)	-10,32 (14,57)
		40%	0,06 (7,39)	9,31 (11,20)	7,28 (11,67)	15,76 (31,17)

M = marcador; C₃₆ = marcador seis do grupo de ligação três; C₃₇ = marcador sete do grupo de ligação três; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem.

Tabela 19 - Médias das estimativas de $\hat{\beta}_0$ dos marcadores C₅₁₁ e C₅₁₂ sobre a característica e seus respectivos desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F₂ analisadas com 1.000 indivíduos e das 180 populações amostradas por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h ²	AG	1000	500	300	200
C ₅₁₁	80%	5%	1001,20 (7,06)	996,82 (13,05)	1011,97 (21,41)	1013,29 (27,71)
		10%	1000,45 (7,17)	1003,74 (15,78)	1016,39 (21,92)	1028,13 (21,60)
		20%	1000,89 (5,87)	990,52 (22,96)	992,01 (30,66)	992,24 (39,15)
		40%	1000,11 (5,50)	1003,46 (9,53)	997,33 (13,31)	994,42 (18,74)
	60%	5%	999,03 (5,35)	998,08 (15,53)	998,71 (23,88)	992,53 (16,14)
		10%	1000,42 (4,56)	1001,10 (4,77)	998,79 (14,27)	989,30 (19,15)
		20%	1000,25 (4,67)	1004,31 (5,29)	1001,43 (12,76)	1007,88 (11,36)
		40%	999,50 (4,34)	997,88 (7,43)	999,28 (13,39)	998,06 (28,76)
	20%	5%	1000,18 (3,42)	994,10 (3,67)	990,04 (12,84)	992,90 (16,35)
		10%	1000,35 (3,70)	1000,23 (8,11)	1006,72 (7,06)	1015,82 (19,06)
		20%	1000,37 (3,22)	1001,07 (6,92)	1005,89 (14,30)	1006,20 (23,84)
		40%	999,81 (3,62)	999,40 (6,79)	997,39 (11,06)	994,03 (10,10)
C ₅₁₂	80%	5%	1001,02 (7,16)	999,30 (17,15)	1016,34 (22,64)	1022,30 (35,19)
		10%	1000,06 (7,29)	1000,63 (11,93)	1010,87 (17,21)	1019,36 (26,08)
		20%	1000,64 (6,67)	990,76 (24,86)	988,89 (28,44)	987,07 (42,12)
		40%	999,63 (6,51)	1005,27 (13,06)	998,37 (11,59)	994,12 (19,16)
	60%	5%	999,06 (5,54)	998,95 (12,74)	1002,29 (19,55)	997,77 (14,03)
		10%	1000,14 (4,90)	1006,40 (6,90)	1005,40 (16,27)	995,63 (23,84)
		20%	1000,45 (4,77)	1005,85 (4,71)	1003,21 (16,58)	1011,96 (15,94)
		40%	999,35 (4,92)	998,46 (6,26)	1003,06 (13,42)	1001,32 (26,53)
	20%	5%	1000,08 (3,54)	992,98 (7,46)	991,51 (13,11)	993,21 (17,15)
		10%	1000,11 (3,78)	1001,64 (7,74)	1008,30 (11,15)	1020,71 (21,82)
		20%	1000,25 (3,32)	1001,34 (7,31)	1006,96 (14,16)	1007,84 (16,64)
		40%	1000,03 (3,76)	1003,01 (8,41)	1002,69 (7,55)	1000,28 (7,72)

M = marcador; C₅₁₁ = marcador 11 do grupo de ligação cinco; C₅₁₂ = marcador 12 do grupo de ligação cinco; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem.

Tabela 20 - Médias das estimativas de $\hat{\beta}_1$ dos marcadores C_{511} e C_{512} sobre a característica e seus respectivos desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F_2 analisadas com 1.000 indivíduos e das 180 populações amostradas por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h^2	AG	1.000	500	300	200
C_{511}	80%	5%	61,17 (10,24)	118,20 (22,05)	171,42 (29,15)	224,02 (34,46)
		10%	83,76 (9,08)	154,92 (13,37)	203,79 (22,52)	249,93 (18,25)
		20%	117,83 (9,58)	205,37 (19,53)	262,35 (11,30)	312,98 (28,48)
		40%	171,42 (8,47)	269,27 (15,08)	338,73 (14,03)	380,23 (18,52)
	60%	5%	37,46 (6,70)	75,68 (5,56)	102,62 (10,68)	123,97 (12,25)
		10%	50,96 (7,07)	85,55 (18,23)	130,69 (28,35)	166,92 (27,77)
		20%	73,85 (6,78)	133,71 (19,89)	181,21 (24,55)	216,57 (24,79)
		40%	103,62 (6,44)	176,77 (14,55)	228,31 (12,84)	265,18 (14,58)
	20%	5%	15,12 (4,89)	33,42 (10,02)	45,31 (20,50)	55,49 (26,23)
		10%	21,60 (5,06)	39,40 (12,46)	53,24 (17,93)	70,35 (18,72)
		20%	29,88 (5,24)	47,85 (11,13)	76,30 (10,27)	88,85 (18,13)
		40%	42,40 (4,62)	72,21 (8,60)	107,23 (13,35)	124,77 (17,94)
C_{512}	80%	5%	60,90 (9,88)	122,22 (18,01)	177,11 (16,65)	232,44 (41,27)
		10%	82,95 (9,61)	151,55 (9,97)	199,97 (20,48)	247,33 (31,21)
		20%	118,05 (10,12)	203,71 (21,03)	263,38 (14,42)	317,74 (28,18)
		40%	170,85 (8,90)	264,10 (14,48)	328,38 (14,76)	369,62 (22,68)
	60%	5%	37,03 (7,00)	76,08 (5,64)	100,60 (13,00)	123,74 (16,04)
		10%	51,12 (6,92)	87,90 (13,80)	134,00 (23,38)	172,54 (22,79)
		20%	73,77 (6,85)	134,11 (13,46)	180,05 (19,92)	214,89 (16,82)
		40%	104,02 (6,51)	175,92 (7,30)	226,35 (12,84)	262,21 (19,78)
	20%	5%	15,15 (5,19)	35,34 (11,84)	45,68 (23,06)	59,92 (25,08)
		10%	21,31 (5,32)	38,15 (13,72)	51,66 (20,42)	68,41 (22,68)
		20%	29,78 (5,49)	47,59 (10,86)	76,45 (6,53)	89,59 (12,67)
		40%	42,33 (5,00)	68,73 (7,99)	101,20 (12,76)	119,92 (18,35)

M = marcador; C_{511} = marcador 11 do grupo de ligação cinco; C_{512} = marcador 12 do grupo de ligação cinco; h^2 = herdabilidade em porcentagem; AG = ação gênica em porcentagem.

Tabela 21 - Médias das estimativas de $\hat{\beta}_2$ dos marcadores C₅₁₁ e C₅₁₂ sobre a característica e seus respectivos desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica das 3.600 populações F₂ analisadas com 1.000 indivíduos e das 180 populações amostradas por genotipagem seletiva com 500, 300 e 200 indivíduos, respectivamente

M	h ²	AG	1.000	500	300	200
C ₅₁₁	80%	5%	1,94 (14,62)	5,94 (9,25)	-7,51 (21,72)	-9,89 (31,77)
		10%	0,38 (12,40)	-2,64 (27,59)	-13,68 (34,42)	-19,93 (60,80)
		20%	-3,88 (13,10)	6,39 (29,51)	4,17 (47,73)	12,34 (38,00)
		40%	-1,64 (12,07)	-10,66 (16,22)	-28,59 (20,32)	-25,92 (20,50)
	60%	5%	0,44 (9,23)	9,91 (23,37)	13,48 (40,46)	26,81 (32,01)
		10%	1,00 (9,54)	-3,26 (9,66)	-4,83 (22,71)	7,14 (22,87)
		20%	0,46 (8,72)	1,42 (10,70)	9,14 (11,82)	4,37 (28,08)
		40%	-0,44 (10,35)	-1,98 (17,28)	-17,28 (25,15)	-13,92 (47,82)
	20%	5%	-0,54 (7,54)	7,46 (6,28)	17,37 (13,93)	14,15 (20,21)
		10%	-0,12 (6,08)	-0,15 (14,23)	-5,65 (13,00)	-9,93 (26,70)
		20%	-1,06 (6,60)	-3,60 (10,97)	-6,22 (16,72)	6,11 (21,31)
		40%	-1,37 (6,92)	-2,90 (13,25)	0,97 (25,28)	6,69 (25,77)
C ₅₁₂	80%	5%	-0,01 (13,97)	3,89 (28,25)	-10,48 (40,59)	-14,85 (81,64)
		10%	0,41 (13,61)	2,23 (21,96)	-7,12 (29,79)	-5,66 (44,41)
		20%	-2,18 (13,36)	9,60 (31,37)	13,92 (46,34)	30,19 (36,66)
		40%	-0,36 (15,35)	-15,30 (25,83)	-34,56 (41,95)	-31,86 (45,52)
	60%	5%	2,01 (10,17)	8,41 (15,92)	8,77 (29,17)	21,55 (20,82)
		10%	2,85 (10,07)	-14,14 (13,36)	-16,54 (25,69)	-3,00 (31,23)
		20%	-1,17 (9,37)	-0,59 (11,59)	4,42 (25,97)	-2,09 (42,48)
		40%	-0,20 (9,27)	-1,73 (9,84)	-20,50 (17,73)	-15,47 (34,98)
	20%	5%	-0,19 (6,93)	9,73 (12,61)	15,20 (17,29)	14,46 (25,02)
		10%	0,50 (7,28)	-2,81 (13,08)	-8,80 (17,04)	-18,72 (27,70)
		20%	-0,75 (7,06)	-4,42 (13,82)	-7,89 (21,08)	2,05 (16,37)
		40%	-0,92 (6,44)	-10,81 (14,75)	-10,22 (23,60)	-7,96 (29,15)

M = marcador; C₅₁₁ = marcador 11 do grupo de ligação cinco; C₅₁₂ = marcador 12 do grupo de ligação cinco; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem.

4.3. Outras estratégias de amostragem

Para a intensidade de amostragem de 20%, além da genotipagem seletiva, foram comparadas três estratégias de amostragem fenotípica. Na primeira estratégia, a distribuição normal dos fenótipos foi dividida em dez classes equidistantes, sendo amostrados 20 indivíduos dentro de cada classe. Na segunda estratégia, metade dos indivíduos foi obtida pela estratégia de genotipagem seletiva e a outra metade foi amostrada de acordo com a segunda estratégia. Na terceira estratégia, foram amostrados 200 indivíduos, aleatoriamente.

4.3.1. Mapeamento genético em populações obtidas a partir de diferentes estratégias de amostragem

A construção de um mapa genético torna possível a aquisição de informações importantes para o melhoramento genético de uma espécie. Tais informações variam desde a associação de marcadores, com caracteres qualitativos e localização dos mesmos nos grupos de ligação até a identificação de regiões genômicas associadas a caracteres quantitativos (FERREIRA & GRATTAPAGLIA, 1998). No entanto, informações de mapeamento genético em diferentes estratégias experimentais tornam-se dispendiosas, na prática. Assim, estudos de simulação foram viáveis para avaliação do efeito de quatro estratégias de amostragem no mapeamento genético, sendo que estas possibilitaram comparações entre as populações amostradas.

A estratégia de genotipagem seletiva (GS) e outras duas estratégias de amostragem (AS e AM), executadas em populações F_2 com 1.000 indivíduos nos 36 cenários, foram comparadas à amostragem aleatória de indivíduos (AA) em relação ao mapeamento genético e detecção de QTL. Escolheu-se uma intensidade de amostragem de 20%, pois, muitos estudos sugerem que o

número mínimo de 200 indivíduos deve ser considerado para o desenvolvimento de mapas de ligação mais fidedignos, para intervalos entre marcadores moleculares de 5 a 20 cM ao longo do genoma (MACKAY, 2001; SCHUSTER & CRUZ, 2004; SILVA, 2005, CRUZ, 2006).

4.3.1.1. Análise de segregação de locos individuais

Além das 180 populações amostradas por genotipagem seletiva, em outras 540 populações amostradas, por meio de outras três estratégias de amostragem, foram aplicados testes de qui-quadrado para verificar a razão de segregação de locos individuais, com nível de 5% de probabilidade. Considerando todas as populações de todos cenários dentro de cada estratégia de amostragem, as médias das distorções foram de 1,97 (2,60%), 2,54 (3,34%), 2,12 (2,79%) e 3,51 (4,61%) para as AS, AM, AA e GS, respectivamente. Portanto, a genotipagem seletiva e a segunda estratégia de amostragem apresentaram maior número de distorção. Os valores médios das distorções em todos os cenários, para cada estratégia de amostragem, encontram-se na Tabela 22.

4.3.1.2. Número de grupos de ligação e ordenamento das marcas

Em todas as estratégias de amostragem, o número de grupos de ligação de todas as 180 populações amostradas esteve conforme o esperado. Foram recuperados os seis grupos de ligação, com todas as marcas simuladas. Após da formação do grupo de ligação, a melhor ordem dos “n” marcadores constituintes de cada grupo foi realizada por meio do, método SARF. Não foi detectada inversão de marcas, nem mesmo nas populações com 200 indivíduos.

Tabela 22 - Valores médios de distorções e seus respectivos desvios-padrão, entre parênteses, de acordo com o estratégia de amostragem, nível de herdabilidade (h^2), nível de ação gênica, nível de saturação para o grupo de ligação, em que QTL simulado estava presente na população

ES	h^2	Ação gênica	Saturação		
			20 cM	10 cM	5 cM
AS	80%	5%	0,40 (0,55)	1,60 (1,14)	1,80 (1,10)
		10%	3,00 (2,24)	2,00 (1,58)	2,00 (2,35)
		20%	3,00 (4,00)	2,00 (3,46)	2,40 (1,52)
		40%	1,20 (1,79)	0,40 (0,55)	1,80 (1,48)
	60%	5%	2,60 (2,30)	2,00 (1,87)	3,80 (3,90)
		10%	2,20 (2,77)	3,20 (4,44)	1,00 (0,71)
		20%	3,40 (3,58)	2,60 (1,67)	1,60 (2,51)
		40%	2,00 (3,46)	0,40 (0,55)	1,80 (2,05)
	20%	5%	2,20 (1,64)	0,80 (0,45)	0,40 (0,89)
		10%	0,80 (0,84)	1,40 (1,52)	2,60 (2,07)
		20%	2,00 (2,92)	2,60 (3,13)	1,40 (1,67)
		40%	4,60 (5,27)	2,20 (2,28)	1,80 (1,10)
AM	80%	5%	1,40 (1,34)	1,80 (0,84)	1,80 (3,49)
		10%	1,80 (1,48)	2,00 (1,22)	1,40 (1,34)
		20%	2,20 (1,79)	2,60 (1,14)	1,80 (2,05)
		40%	2,00 (1,58)	5,20 (2,49)	3,20 (2,17)
	60%	5%	3,40 (4,93)	1,20 (1,64)	2,60 (2,30)
		10%	2,20 (1,92)	2,60 (1,67)	4,20 (5,07)
		20%	3,20 (3,49)	3,60 (3,29)	1,40 (1,67)
		40%	5,00 (4,06)	4,00 (2,55)	5,00 (4,18)
	20%	5%	1,80 (2,49)	2,40 (3,21)	1,00 (1,22)
		10%	1,00 (1,73)	4,00 (3,46)	1,40 (2,19)
		20%	1,60 (2,51)	1,80 (1,92)	4,40 (2,79)
		40%	1,60 (2,07)	2,20 (1,92)	2,60 (1,82)
AA	80%	5%	1,80 (1,30)	1,40 (1,34)	1,40 (1,14)
		10%	2,00 (1,58)	2,40 (2,19)	1,80 (1,64)
		20%	0,60 (0,89)	1,20 (1,10)	1,80 (1,48)
		40%	0,80 (0,45)	2,20 (2,77)	2,20 (4,38)
	60%	5%	1,80 (0,84)	1,00 (1,41)	0,60 (0,89)
		10%	2,60 (3,78)	1,60 (1,14)	3,40 (2,30)
		20%	3,40 (4,88)	1,00 (1,00)	3,20 (4,97)
		40%	2,40 (2,61)	1,00 (1,22)	3,00 (2,12)
	20%	5%	2,80 (2,39)	3,00 (3,32)	2,20 (1,10)
		10%	4,20 (3,49)	3,00 (2,35)	1,40 (1,34)
		20%	2,40 (1,52)	3,80 (2,39)	3,40 (2,61)
		40%	2,00 (1,22)	1,80 (1,92)	1,60 (1,67)
GS	80%	5%	2,20 (3,90)	3,80 (3,35)	1,60 (1,14)
		10%	4,00 (4,64)	4,60 (2,51)	6,00 (4,95)
		20%	4,40 (2,88)	2,40 (3,05)	6,60 (4,04)
		40%	6,20 (4,09)	6,60 (2,88)	9,40 (3,36)
	60%	5%	2,80 (2,59)	1,80 (1,10)	2,60 (2,30)
		10%	3,40 (3,97)	2,60 (2,19)	2,40 (3,29)
		20%	2,80 (2,95)	3,00 (1,87)	1,60 (1,34)
		40%	2,40 (1,52)	4,80 (4,97)	6,00 (5,66)
	20%	5%	0,80 (0,84)	0,60 (0,89)	2,40 (1,67)
		10%	2,60 (1,82)	4,00 (4,53)	2,40 (2,30)
		20%	2,60 (1,67)	4,60 (3,36)	2,40 (2,30)
		40%	3,20 (3,27)	1,60 (1,67)	5,00 (3,46)

4.3.1.3. Construção do mapa de ligação

Os mapas de ligação obtidos dos 200 indivíduos, independente da estratégia experimental, reproduziram as 76 marcas, distribuídas em seis grupos de ligação. Portanto, todas as marcas encontravam-se ligadas e seu ordenamento foi correto. Os grupos de ligação de 1 a 6, em média foram bem similares ao genoma simulado, apesar da ocorrência de maiores distorções dos marcadores nos grupos de ligação, que tinha a presença do QTL simulado.

4.3.1.4. Comparação dos genomas

Além dos 180 mapas de ligação, obtidos por genotipagem seletiva, outros 540 mapas obtidos das populações, amostradas por meio de outras estratégias de amostragem, foram comparados com o genoma simulado. Portanto, 3.240 grupos de ligação foram comparados com seu respectivo grupo do genoma simulado.

O tamanho médio do grupo de ligação (TAM), a distância média entre marcas adjacentes (DIM), a variância média das distâncias entre marcas adjacentes (VAD) e o estresse médio (EST) foram comparados para os grupos de ligação 1, 3 e 5, obtidos das populações amostradas por meio das estratégias AS, AM, AA e GS (Tabela 23). Os grupos de ligação 1, 3 e 5 foram escolhidos para comparação devido, à presença do QTL simulado nos cenários das POP 1 a POP 12, POP 13 a POP 24 e POP 25 a POP 36, respectivamente. As 180 populações amostradas, para cada grupo de ligação, foram separadas em duas partes. A parte “a” constituía-se de 60 populações e tinha o QTL simulado, enquanto a parte “b” constituía-se de 120 populações, mas não tinha o QTL no respectivo grupo de ligação. Os valores de TAM, DIM, VAD e EST nos grupos de ligação, dentro e entre cada estratégia de amostragem, foram bem similares. Assim, pode-se observar que, independente da estratégia de amostragem para as populações de 200

indivíduos, foi possível reconstruir os mapas de ligação, mesmo sob amostragem da característica quantitativa (AS, AM e GS).

Tabela 23 - Tamanho médio de grupos de ligação, distância média entre marcas adjacentes, variância média das distâncias entre marcas adjacentes, estresse e seus respectivos desvios-padrão entre parênteses das populações amostradas por genotipagem seletiva e outras três estratégias de amostragem com 200 indivíduos, para os grupos de ligação 1, 3 e 5

ES	GL	TAM	DIM	VAD	EST
GS	GL 1 ^a	99,15 (5,05)	19,90 (0,88)	6,50 (6,16)	11,80 (4,38)
	GL 1 ^b	99,96 (4,78)	19,99 (0,95)	5,14 (3,36)	10,73 (3,31)
	GL 3 ^a	102,81 (5,05)	10,28 (0,51)	2,79 (1,08)	15,84 (3,29)
	GL 3 ^b	101,90 (5,78)	10,19 (0,57)	2,78 (1,15)	15,60 (3,30)
	GL 5 ^a	101,87 (3,89)	5,09 (0,20)	1,22 (0,39)	22,64 (3,98)
	GL 5 ^b	101,70 (5,37)	5,08 (0,26)	1,18 (0,44)	22,22 (3,58)
AS	GL 1 ^a	100,13 (5,23)	20,03 (1,05)	6,17 (4,39)	10,64 (3,88)
	GL 1 ^b	99,56 (4,74)	19,91 (0,95)	5,18 (3,78)	10,66 (3,52)
	GL 3 ^a	103,24 (5,10)	10,32 (0,51)	2,82 (1,49)	16,46 (4,34)
	GL 3 ^b	101,98 (5,50)	10,20 (0,55)	2,65 (1,48)	15,96 (4,21)
	GL 5 ^a	101,34 (4,63)	5,07 (0,22)	1,30 (0,47)	22,37 (4,00)
	GL 5 ^b	101,55 (5,27)	5,08 (0,26)	1,22 (0,42)	22,87 (3,84)
AM	GL 1 ^a	100,45 (4,90)	20,09 (0,98)	6,38 (5,11)	11,32 (4,18)
	GL 1 ^b	99,91 (4,60)	19,98 (0,92)	5,55 (3,97)	10,89 (3,67)
	GL 3 ^a	102,25 (5,16)	10,23 (0,52)	2,36 (1,21)	15,09 (4,01)
	GL 3 ^b	102,05 (5,08)	10,20 (0,51)	2,64 (1,29)	15,93 (3,76)
	GL 5 ^a	102,73 (4,44)	5,14 (0,22)	1,19 (0,40)	22,63 (3,60)
	GL 5 ^b	101,12 (5,02)	5,06 (0,24)	1,18 (0,41)	22,47 (3,73)
AA	GL 1 ^a	99,48 (4,38)	19,79 (1,35)	5,16 (3,57)	10,72 (3,64)
	GL 1 ^b	99,98 (4,45)	19,68 (1,43)	5,21 (3,43)	10,51 (3,71)
	GL 3 ^a	102,02 (4,79)	10,20 (0,48)	2,53 (1,18)	15,67 (3,49)
	GL 3 ^b	102,22 (4,88)	10,26 (0,49)	2,53 (1,18)	15,58 (3,53)
	GL 5 ^a	102,33 (5,54)	5,07 (0,28)	1,18 (0,44)	22,59 (3,74)
	GL 5 ^b	102,63 (5,49)	5,13 (0,32)	1,22 (0,42)	22,52 (3,67)

GS = genotipagem seletiva; AS = amostragem sistemática; AM = amostragem mista; AA = amostragem aleatória; GL = grupo de ligação; ^a grupo de ligação na presença do QTL; ^b grupo de ligação na ausência dos QTL; TAM = tamanho médio do grupo de ligação em centiMorgan (cM); DIM = distância média entre marcas adjacentes em centiMorgan (cM); VAD = variância média das distâncias entre marcas adjacentes; EST = estresse médio em porcentagem (%).

4.3.2. Detecção de QTL em populações amostradas por meio de diferentes estratégias de amostragem

4.3.2.1. Detecção do QTL pelo método de marca simples e sua contribuição para a variação da característica

Análises de regressão pelo método de marca simples foram realizadas, para os marcadores anteriores e posteriores dos QTL simulados dos grupos de ligação 1, 3 e 5, para as populações amostradas por meio das estratégias AS, AM e GS, sendo os dados comparados com as populações amostradas aleatoriamente (AA). Estimaram-se os valores do LOD (Tabelas 24, 25 e 26) e de R^2 (Tabelas 27, 28 e 29).

Nas populações amostradas por meio de diferentes estratégias, observou-se que independente da posição do marcador, se anterior ou posterior ao QTL simulado, em qualquer grupo de ligação, todas as médias e desvios-padrão do LOD e R^2 apresentaram magnitude semelhante.

Em todos os cenários, independente da saturação do grupo de ligação, a relação média do LOD entre as estratégias de amostragem, considerando-se a AA como padrão, foi de 1; 1,65; e 2,02 para AS, AM e GS, respectivamente. Então, o valor LOD da AS, em média, não diferenciou da AA. No entanto, as AM e GS tiveram aumentos significativos nos LOD médios, podendo chegar ao dobro em relação à estratégia de amostragem aleatória (AA). Além disso, em relação a AA, a média do R^2 foi a 1; 1,55; e 1,80 para as AS, AM e GS, respectivamente. Assim, as melhores estratégias de amostragem para detecção de QTL, pelo método de marca simples, foram as AM e GS.

Tabela 24 - Estimativas médias do LOD para os marcadores C₁₃ e C₁₄ e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica. Cinco populações foram avaliadas, para cada estratégia de amostragem com seu respectivo cenário (POP 1 a 12)

M	h ²	AG	$\overline{\text{LOD}}$			
			AS	AM	AA	GS
C ₁₃	80%	5%	1,55 (1,24)	3,90 (2,16)	1,47 (0,99)	3,41 (1,36)
		10%	2,88 (0,83)	6,38 (1,64)	3,38 (0,71)	8,68 (3,57)
		20%	5,88 (1,10)	12,94 (4,30)	6,74 (2,71)	17,62 (6,48)
		40%	11 (2,74)	21,73 (6,36)	8,56 (2,18)	29,97 (7,53)
	60%	5%	0,93 (0,70)	2,43 (1,78)	0,41 (0,34)	1,85 (1,82)
		10%	1,11 (1,15)	3,25 (1,64)	1,40 (1,22)	3,63 (1,54)
		20%	3,99 (0,77)	11,15 (1,67)	4,85 (1,93)	12,85 (3,83)
		40%	7,77 (3,82)	15,43 (2,63)	6,06 (3,26)	23,55 (5,07)
	20%	5%	0,49 (0,35)	0,77 (0,71)	0,61 (0,44)	0,86 (0,75)
		10%	0,85 (0,64)	2,37 (1,14)	0,68 (0,38)	2,24 (1,42)
		20%	1,46 (1,01)	3,68 (0,94)	0,74 (0,49)	3,97 (1,56)
		40%	3,48 (1,21)	6,81 (2,64)	2,73 (1,77)	7,69 (3,05)
C ₁₄	80%	5%	0,94 (0,83)	3,11 (1,87)	1,30 (0,97)	3,42 (1,05)
		10%	2,77 (1,93)	6,70 (1,45)	3,21 (0,52)	8,06 (2,60)
		20%	5,79 (1,61)	13,81 (4,80)	6,49 (2,90)	18,99 (7,83)
		40%	11,95 (3,12)	19,47 (8,05)	10,90 (1,12)	33,53 (4,14)
	60%	5%	1,24 (0,76)	2,09 (1,47)	0,72 (0,59)	2,50 (1,89)
		10%	1,81 (0,24)	3,20 (0,91)	1,41 (1,36)	3,91 (2,10)
		20%	3,77 (1,66)	8,71 (2,44)	3,43 (1,50)	10,82 (3,36)
		40%	6,65 (2,02)	17,43 (3,24)	5,94 (1,67)	23,06 (5,00)
	20%	5%	0,68 (0,52)	0,95 (0,61)	0,52 (0,25)	1,15 (1,11)
		10%	0,36 (0,19)	2,62 (1,52)	1,01 (0,46)	2,08 (1,62)
		20%	1,15 (0,83)	4,20 (1,36)	1,57 (1,64)	4,44 (0,97)
		40%	3,75 (1,92)	7,83 (1,84)	2,36 (0,73)	8,24 (1,07)

M = marcador; C₁₃ = marcador três do grupo de ligação um; C₁₄ = marcador quatro do grupo de ligação um; h² = herdabilidade em porcentagem; $\overline{\text{LOD}}$ = média do LOD.

Tabela 25 - Estimativas médias do LOD para os marcadores C₃₆ e C₃₇ e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica. Cinco populações foram avaliadas, para cada estratégia de amostragem com seu respectivo cenário (POP 13 a POP 24)

M	h ²	AG	$\overline{\text{LOD}}$				
			AS	AM	AA	GS	
C ₃₆	80%	5%	1,59 (0,58)	4,76 (0,89)	1,92 (1,14)	4,12 (1,73)	
		10%	2,99 (1,42)	8,33 (5,58)	3,87 (1,38)	10,07 (3,81)	
		20%	6,25 (1,60)	16,82 (2,87)	7,01 (3,77)	21,72 (6,26)	
		40%	9,60 (1,39)	27,60 (6,16)	11,30 (3,74)	37,93 (7,41)	
	60%	5%	1,05 (0,96)	2,75 (1,86)	1,34 (1,08)	3,24 (3,24)	
		10%	2,58 (0,86)	7,44 (1,59)	2,17 (0,85)	9,99 (2,54)	
		20%	4,98 (1,55)	14,16 (2,16)	5,25 (1,18)	17,89 (3,64)	
		40%	9,61 (0,84)	21,23 (4,01)	9,87 (1,74)	26,22 (4,98)	
	20%	5%	0,46 (0,09)	0,78 (0,45)	0,43 (0,46)	1,29 (0,52)	
		10%	1,27 (1,05)	3,43 (1,95)	1,85 (1,29)	4,45 (2,19)	
		20%	2,02 (0,66)	4,88 (1,24)	2,00 (1,08)	6,73 (1,85)	
		40%	2,68 (1,12)	7,18 (2,59)	3,74 (2,29)	8,56 (1,66)	
	C ₃₇	80%	5%	2,06 (1,21)	4,95 (1,23)	2,30 (1,50)	4,69 (2,39)
			10%	2,69 (1,46)	7,01 (5,43)	3,69 (1,51)	9,52 (4,01)
			20%	6,38 (2,19)	15,74 (4,01)	6,82 (3,72)	20,13 (6,19)
			40%	10,37 (2,42)	30,93 (7,75)	11,30 (2,18)	41,25 (7,40)
60%		5%	1,60 (1,63)	2,95 (2,10)	2,26 (1,16)	4,01 (3,82)	
		10%	2,52 (1,24)	6,31 (1,87)	1,89 (0,92)	8,74 (3,28)	
		20%	4,45 (1,55)	15,80 (4,30)	5,85 (0,71)	19,18 (6,59)	
		40%	9,12 (1,77)	21,08 (4,45)	9,92 (2,28)	27,66 (9,17)	
20%		5%	0,45 (0,26)	0,86 (0,84)	0,55 (0,48)	1,44 (0,97)	
		10%	1,53 (1,39)	3,40 (2,53)	1,63 (1,09)	4,66 (2,51)	
		20%	1,89 (0,85)	5,32 (2,27)	2,18 (1,42)	6,55 (2,09)	
		40%	3,76 (1,24)	7,65 (1,74)	3,68 (1,54)	10,17 (1,42)	

M = marcador; C₃₆ = marcador seis do grupo de ligação três; C₃₇ = marcador sete do grupo de ligação três; h² = herdabilidade em porcentagem; h² = herdabilidade em porcentagem; $\overline{\text{LOD}}$ = média do LOD.

Tabela 26 - Estimativas médias do LOD para os marcadores C₅₁₁ e C₅₁₂ e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica. Cinco populações foram avaliadas, para cada estratégia de amostragem com seu respectivo cenário (POP 25 a POP 36)

h ²	AG	LOD				
		AS	AM	AA	GS	
C ₅₁₁	80%	5%	1,80 (1,54)	6,03 (3,27)	1,53 (0,25)	8,05 (1,63)
		10%	4,75 (1,50)	8,35 (2,50)	3,64 (0,35)	12,07 (4,94)
		20%	7,73 (2,45)	17,94 (4,57)	5,91 (2,84)	20,60 (8,52)
		40%	16,43 (3,62)	30,20 (3,65)	15,59 (2,45)	43,79 (2,62)
	60%	5%	1,60 (1,01)	5,38 (2,32)	1,52 (0,39)	5,41 (1,24)
		10%	2,10 (0,98)	6,71 (3,48)	2,56 (1,23)	9,64 (6,11)
		20%	5,11 (2,78)	12,21 (5,27)	5,92 (2,54)	16,04 (5,04)
		40%	11,53 (3,29)	26,14 (6,31)	10,31 (2,29)	36,31 (9,93)
	20%	5%	0,80 (0,36)	1,43 (0,95)	1,01 (0,90)	2,35 (1,76)
		10%	0,90 (0,60)	2,35 (1,39)	0,79 (0,67)	3,41 (1,72)
		20%	2,16 (1,35)	3,70 (1,69)	2,54 (0,99)	4,64 (1,79)
		40%	4,17 (1,19)	8,70 (4,02)	4,66 (1,74)	11,18 (2,70)
C ₅₁₂	80%	5%	2,14 (1,09)	6,51 (3,62)	1,49 (0,69)	8,68 (2,86)
		10%	4,65 (0,79)	8,07 (3,80)	3,66 (1,22)	10,70 (3,00)
		20%	8,10 (2,09)	17,20 (3,56)	6,34 (2,14)	17,70 (5,83)
		40%	15,40 (2,65)	28,55 (4,45)	14,56 (4,29)	41,21 (4,35)
	60%	5%	1,46 (0,81)	5,17 (1,88)	1,65 (0,51)	5,32 (1,12)
		10%	2,50 (0,62)	7,42 (4,04)	3,50 (2,02)	10,32 (5,89)
		20%	5,11 (2,89)	11,69 (4,31)	5,71 (3,06)	15,53 (3,05)
		40%	11,22 (1,96)	25,21 (5,50)	10,08 (1,57)	35,39 (9,69)
	20%	5%	3,24 (6,58)	1,59 (1,38)	1,07 (0,97)	2,60 (1,54)
		10%	1,07 (0,73)	2,46 (1,43)	0,83 (0,76)	3,33 (1,75)
		20%	2,11 (1,30)	3,74 (1,24)	3,07 (0,90)	4,55 (1,20)
		40%	3,81 (0,90)	8,24 (3,89)	4,04 (1,03)	10,64 (3,16)

M = marcador; C₅₁₁ = marcador 11 do grupo de ligação cinco; C₅₁₂ = marcador 12 do grupo de ligação cinco; h² = herdabilidade em porcentagem; h² = herdabilidade em porcentagem; $\overline{\text{LOD}}$ = média do LOD.

Tabela 27 - Estimativas médias de R^2 do modelo de regressão linear por marca simples e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica para os marcadores C_{13} e C_{14} do grupo de ligação 1. Cinco populações foram avaliadas, para cada estratégia de amostragem com seu respectivo cenário (POP 1 a POP 12)

	h^2	AG	\bar{R}^2			
			AS	AM	AA	GS
C_{13}	80%	5%	3,48 (2,73)	8,50 (4,49)	3,30 (2,20)	7,52 (2,87)
		10%	6,40 (1,79)	13,62 (3,28)	7,48 (1,52)	17,90 (6,77)
		20%	12,63 (2,20)	25,46 (7,55)	14,24 (5,16)	32,74 (10,42)
		40%	22,25 (4,89)	38,84 (8,83)	17,80 (4,10)	49,26 (8,25)
	60%	5%	2,10 (1,58)	5,38 (3,85)	0,95 (0,77)	4,11 (3,96)
		10%	2,50 (2,55)	7,16 (3,51)	3,14 (2,71)	7,98 (3,27)
		20%	8,76 (1,62)	22,59 (3,01)	10,49 (3,97)	25,38 (6,44)
		40%	16,13 (7,23)	29,80 (4,31)	12,83 (6,55)	41,54 (6,88)
	20%	5%	1,12 (0,80)	1,74 (1,60)	1,39 (1)	1,96 (1,70)
		10%	1,93 (1,45)	5,28 (2,51)	1,56 (0,85)	4,98 (3,06)
		20%	3,29 (2,22)	8,12 (1,99)	1,69 (1,10)	8,68 (3,33)
		40%	7,68 (2,57)	14,39 (5,19)	6,02 (3,75)	16,06 (5,76)
C_{14}	80%	5%	2,13 (1,87)	6,83 (3,93)	2,93 (2,16)	7,55 (2,22)
		10%	6,11 (4,11)	14,26 (2,86)	7,12 (1,12)	16,81 (4,87)
		20%	12,44 (3,26)	26,89 (8,02)	13,74 (5,55)	34,62 (11,61)
		40%	23,90 (5,19)	35,23 (13,53)	22,14 (1,98)	53,53 (4,28)
	60%	5%	2,81 (1,70)	4,65 (3,19)	1,63 (1,34)	5,53 (4,07)
		10%	4,07 (0,53)	7,07 (1,96)	3,16 (2,98)	8,53 (4,43)
		20%	8,27 (3,57)	18,05 (4,58)	7,53 (3,19)	21,85 (6,21)
		40%	14,08 (4,01)	32,90 (4,98)	12,70 (3,29)	40,89 (6,66)
	20%	5%	1,54 (1,19)	2,16 (1,38)	1,19 (0,57)	2,60 (2,48)
		10%	0,81 (0,42)	5,81 (3,30)	2,31 (1,03)	4,63 (3,52)
		20%	2,59 (1,85)	9,19 (2,87)	3,51 (3,58)	9,70 (2,02)
		40%	8,22 (4,00)	16,43 (3,50)	5,28 (1,59)	17,24 (2,01)

M = marcador; C_{13} = marcador três do grupo de ligação um; C_{14} = marcador quatro do grupo de ligação um; h^2 = herdabilidade em porcentagem; AG = ação gênica em porcentagem e \bar{R}^2 = média do coeficiente de determinação.

Tabela 28 - Estimativas médias de R^2 do modelo de regressão linear por marca simples e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica para os marcadores C_{36} e C_{37} do grupo de ligação 3. Cinco populações foram avaliadas, para cada estratégia de amostragem com seu respectivo cenário (POP 13 a POP 24)

h^2	AG	\bar{R}^2				
		AS	AM	AA	GS	
C_{36}	80%	5%	3,59 (1,29)	10,37 (1,86)	4,29 (2,48)	9,00 (3,63)
		10%	6,61 (3,02)	16,93 (9,91)	8,48 (2,87)	20,46 (6,76)
		20%	13,37 (3,19)	32 (4,49)	14,65 (7,21)	38,86 (8,49)
		40%	19,80 (2,55)	46,59 (8,19)	22,69 (6,28)	57,76 (7,18)
	60%	5%	2,38 (2,16)	6,07 (4,04)	3,02 (2,38)	6,98 (6,83)
		10%	5,75 (1,86)	15,71 (3,09)	4,86 (1,87)	20,44 (4,79)
		20%	10,78 (3,22)	27,75 (3,58)	11,37 (2,42)	33,58 (5,64)
		40%	19,83 (1,56)	38,45 (5,89)	20,28 (3,23)	45,04 (6,08)
	20%	5%	1,05 (0,20)	1,79 (1,03)	0,97 (1,04)	2,92 (1,16)
		10%	2,86 (2,33)	7,52 (4,15)	4,13 (2,82)	9,64 (4,71)
		20%	4,53 (1,43)	10,60 (2,54)	4,49 (2,38)	14,30 (3,69)
		40%	5,97 (2,43)	15,12 (5,07)	8,15 (4,80)	17,84 (3,17)
C_{37}	80%	5%	4,61 (2,66)	10,75 (2,54)	5,11 (3,23)	10,13 (4,92)
		10%	5,96 (3,14)	14,39 (10,06)	8,11 (3,22)	19,43 (7,09)
		20%	13,57 (4,32)	30,17 (6,52)	14,29 (7,40)	36,58 (8,71)
		40%	21,15 (4,40)	50,29 (9,45)	22,81 (3,88)	60,84 (6,65)
	60%	5%	3,57 (3,54)	6,48 (4,56)	5,04 (2,51)	8,55 (7,72)
		10%	5,61 (2,71)	13,46 (3,70)	4,23 (2,05)	18,06 (6,15)
		20%	9,68 (3,29)	30,26 (6,60)	12,60 (1,42)	35,18 (9,56)
		40%	18,89 (3,33)	38,21 (6,36)	20,33 (4,21)	46,32 (10,18)
	20%	5%	1,04 (0,60)	1,95 (1,88)	1,26 (1,09)	3,24 (2,15)
		10%	3,43 (3,07)	7,41 (5,33)	3,65 (2,41)	10,06 (5,27)
		20%	4,25 (1,87)	11,45 (4,53)	4,86 (3,13)	13,92 (4,11)
		40%	8,25 (2,61)	16,09 (3,33)	8,06 (3,26)	20,82 (2,57)

M = marcador; C_{36} = marcador seis do grupo de ligação três; C_{37} = marcador sete do grupo de ligação três; h^2 = herdabilidade em porcentagem; h^2 = herdabilidade em porcentagem; \bar{R}^2 = média do coeficiente de determinação.

Tabela 29 - Estimativas médias de R^2 do modelo de regressão linear por marca simples e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica para os marcadores C_{511} e C_{512} do grupo de ligação 5. Cinco populações foram avaliadas, para cada estratégia de amostragem com seu respectivo cenário (POP 25 a POP 36)

M	h^2	AG	\bar{R}^2			
			AS	AM	AA	GS
C_{511}	80%	5%	4,01 (3,38)	12,77 (6,52)	3,46 (0,56)	16,88 (3,17)
		10%	10,32 (3,10)	17,39 (4,80)	8,04 (0,74)	23,88 (8,18)
		20%	16,21 (4,76)	33,55 (6,73)	12,58 (5,81)	36,84 (11,63)
		40%	31,30 (5,83)	49,97 (4,13)	30,07 (3,94)	63,46 (2,20)
	60%	5%	3,60 (2,22)	11,56 (4,56)	3,43 (0,87)	11,68 (2,52)
		10%	4,69 (2,15)	14,10 (6,73)	5,70 (2,62)	19,30 (10,54)
		20%	10,95 (5,61)	24,07 (9,15)	12,62 (5,13)	30,52 (7,85)
		40%	23,14 (5,70)	44,76 (7,92)	21,03 (4,26)	55,80 (9,19)
	20%	5%	1,81 (0,82)	3,22 (2,12)	2,28 (2,03)	5,20 (3,80)
		10%	2,04 (1,35)	5,24 (2,96)	1,78 (1,50)	7,50 (3,67)
		20%	4,82 (2,95)	8,11 (3,52)	5,67 (2,16)	10,08 (3,60)
		40%	9,13 (2,54)	17,87 (7,20)	10,13 (3,55)	22,58 (4,68)
C_{512}	80%	5%	4,79 (2,39)	13,69 (7,03)	3,36 (1,54)	17,98 (5,31)
		10%	10,13 (1,64)	16,73 (7,23)	8,05 (2,58)	21,65 (5,39)
		20%	16,93 (4,04)	32,51 (5,40)	13,49 (4,28)	32,89 (9,07)
		40%	29,73 (4,31)	47,96 (5,21)	28,23 (6,90)	61,13 (4,04)
	60%	5%	3,30 (1,80)	11,15 (3,78)	3,72 (1,12)	11,50 (2,31)
		10%	5,58 (1,35)	15,44 (7,51)	7,67 (4,19)	20,59 (9,90)
		20%	10,95 (5,80)	23,28 (7,67)	12,16 (6,11)	29,89 (4,83)
		40%	22,69 (3,44)	43,66 (6,95)	20,67 (2,88)	54,88 (9,26)
	20%	5%	6,49 (12,63)	3,55 (3,03)	2,42 (2,18)	5,75 (3,36)
		10%	2,42 (1,64)	5,47 (3,04)	1,87 (1,70)	7,32 (3,75)
		20%	4,72 (2,82)	8,22 (2,60)	6,80 (1,92)	9,91 (2,45)
		40%	8,37 (1,92)	17,02 (7,06)	8,85 (2,14)	21,58 (5,57)

M = marcador; C_{511} = marcador 11 do grupo de ligação cinco; C_{512} = marcador 12 do grupo de ligação cinco; h^2 = herdabilidade em porcentagem; h^2 = herdabilidade em porcentagem; \bar{R}^2 = média do coeficiente de determinação.

4.3.2.2. Detecção e localização QTL pelo método de intervalo simples e sua contribuição para a variação da característica

As análises de detecção de QTL foram realizadas por meio de regressão linear, pelo método de intervalo nas 180 populações amostradas, por meio das estratégias AS, AM, AA e GS, todos com 200 indivíduos. Foram avaliados apenas os grupos de ligação 1, para os cenários das POP 1 a 12, 3 para as POP 13 a 24 e 5 para as POP 25 a 36 com suas respectivas características quantitativas.

Para as estimativas médias dos LOD dos cenários das POP 1 a 36 (Tabela 30), em cada estratégia de amostragem, foram avaliados os cenários que apresentaram “LOD score” menor que três. Verificou-se quatro cenários menores que três para GS e AM, mas 9 e 16 cenários para a AS e AA, respectivamente. Os valores do LOD médios variaram de 1,03 a 21,29; de 1,25 a 50,87; de 1,24 a 19,7; de 1,62 a 81,06, para as estratégias AS, AM, AA e GS, respectivamente. Às médias dos LOD, em todas as 180 populações de cada estratégia de amostragem, foram comparadas com a AA, sendo observado um acréscimo de 2, 153 e 266% para as AS, AM e GS, respectivamente.

Em relação ao grupo de ligação 1 (POP 1 a POP 12), as médias do $r_{\text{máx}}$ variaram entre 38,7 a 63,3 cM; 40,5 a 49,7 cM; 36,9 a 66,2 cM; e 42,5 a 57,8 cM para as AS, AM, AA e GS, respectivamente, sendo que o a frequência de recombinação (r) esperada foi de 50 cM. Para o grupo de ligação 3 (POP 13 a 24), a média $r_{\text{máx}}$ variou de 34,5 a 59,7 cM; 35,5 a 55,8 cM; 32,7 a 53,8 cM; e 34,0 a 60,0 cM, para as AS, AM, AA e GS, respectivamente, enquanto o valor esperado do “ r ” foi de 55 cM. No grupo de ligação 5 (POP 25 a 36), a variação média do $r_{\text{máx}}$ ficou entre 34,5 a 59,7 cM; 35,5 a 55,8 cM; 32,7 a 53,8 cM; e 34,0 a 60,0 cM para as AS, AM, AA e GS, respectivamente, enquanto “ r ” esperado foi de 52,5 cM (Tabela 31).

Como o valor do R^2 está associado ao LOD (LIU, 1998; SCHUSTER & CRUZ, 2004), verificou-se o mesmo padrão de variação do LOD médio, entre os cenários e entre as estratégias experimentais de amostragem com o R^2 médio (Tabela 32). O acréscimo médio no R^2 , proporcionado pela amostragem, foi de 2, 108 e 165% para a AS, AM e GS, respectivamente, em relação a AA, considerando todas as 180 populações amostradas em cada estratégia.

Tabela 30 - Estimativas médias do LOD máximo e seus respectivos desvios-padrão, entre parênteses, obtidos por regressão linear pelo método de intervalo simples, em função do nível de herdabilidade e de ação gênica para o grupo de ligação 1 (POP 1 a 12)

GL	h ²	AG	$\overline{\text{LOD}}_{\text{máx}}$			
			AS	AM	AA	GS
GL 1	80%	5%	1,87 (1,15)	4,56 (2,39)	2,21 (1,05)	4,80 (0,83)
		10%	3,81 (1,63)	9,25 (1,26)	4,19 (0,69)	11,86 (4,64)
		20%	8,03 (1,80)	21,65 (8,51)	8,59 (4,46)	31,51 (15,85)
		40%	17,64 (4,90)	41,03 (11,94)	14,80 (2,22)	68,06 (18,82)
	60%	5%	1,78 (0,84)	3 (1,65)	1,24 (0,30)	2,97 (2,25)
		10%	2,37 (0,68)	4,28 (1,46)	2,07 (1,54)	5,09 (2,23)
		20%	5,22 (1,02)	14,28 (2,41)	5,61 (2,43)	17,51 (4,98)
		40%	10,17 (4,57)	26,68 (3,54)	8,48 (2,95)	41,33 (11,68)
	20%	5%	1,03 (0,48)	1,46 (0,51)	1,29 (0,97)	1,62 (0,98)
		10%	1,64 (0,81)	3,33 (1,55)	1,67 (0,46)	2,82 (1,79)
		20%	3,30 (1,99)	5,44 (1,22)	1,93 (1,10)	5,68 (1,26)
		40%	5,04 (2,07)	10,29 (3,13)	3,47 (1,78)	11,30 (3,09)
GL 3	80%	5%	2,47 (1,11)	7 (1,51)	2,71 (1,49)	5,43 (2,70)
		10%	3,63 (1,25)	9,92 (7,53)	4,52 (1,46)	12,50 (5,36)
		20%	7,93 (2,53)	22,56 (5,65)	8,73 (4,87)	32,02 (12,26)
		40%	12,95 (2,70)	50,87 (15,93)	14,96 (4,38)	81,06 (20,87)
	60%	5%	2,11 (1,66)	3,99 (1,63)	2,61 (1,18)	5,06 (4,07)
		10%	3,42 (1,58)	8,46 (2,16)	2,36 (0,94)	11,88 (3,79)
		20%	5,62 (1,75)	21,09 (6,55)	6,71 (0,87)	28,03 (9,87)
		40%	11,74 (1,37)	31,99 (8,04)	12,57 (2,73)	45,82 (20,37)
	20%	5%	3,03 (2,30)	1,25 (0,57)	2,45 (2,52)	1,80 (1,01)
		10%	2,16 (1,25)	4,35 (2,50)	2,68 (2,02)	5,80 (2,91)
		20%	2,81 (0,72)	6,16 (2,37)	2,48 (1,10)	8,02 (2,47)
		40%	4,40 (1,24)	9,28 (2,61)	4,40 (2,36)	12,43 (2,01)
GL 5	80%	5%	3,09 (0,87)	7,97 (4,27)	2,07 (0,56)	10,50 (2,92)
		10%	5,74 (1,21)	9,92 (3,97)	4,47 (0,80)	14,69 (6,95)
		20%	9,43 (2,79)	23,40 (6,79)	7,89 (2,36)	30,55 (14,26)
		40%	21,29 (4,56)	44,97 (7,63)	19,74 (5,01)	78,63 (6,59)
	60%	5%	2,17 (0,78)	6,08 (2,64)	2,13 (0,25)	6,20 (1,27)
		10%	3,03 (0,59)	8,81 (4,80)	4,30 (1,79)	12,71 (8,11)
		20%	6,06 (2,85)	15,10 (6,80)	7,43 (3,33)	21,39 (7,17)
		40%	14,09 (3,75)	38,46 (11,23)	12,59 (1,92)	62,79 (26,39)
	20%	5%	1,93 (1,34)	2,16 (1,38)	1,99 (1,01)	3,21 (1,30)
		10%	1,67 (0,80)	2,96 (1,51)	1,54 (0,70)	4,28 (1,77)
		20%	3,20 (1,75)	4,46 (1,45)	3,78 (1,35)	5,37 (1,78)
		40%	4,63 (1,37)	9,89 (5,14)	5,45 (2,08)	13,04 (3,62)

GL = grau de liberdade; h² = herdabilidade em porcentagem; AG = ação gênica; $\overline{\text{LOD}}_{\text{máx}}$ = média do LOD máximo.

Tabela 31 - Estimativa média da frequência de recombinação associada ao LOD máximo do grupo de ligação e seus respectivos desvios-padrão, entre parênteses, obtidos por regressão linear pelo método de intervalo simples, em função do nível de herdabilidade e de ação gênica do grupo de ligação 3 (POP 13 a 24)

GL	h ²	AG	$\bar{r}_{\text{máx}}$			
			AS	AM	AA	GS
GL 1	80%	5%	45,81 (34,01)	43,76 (15,27)	36,87 (31,57)	42,47 (9,25)
		10%	45,30 (2,88)	42,04 (11,78)	44,53 (4,83)	45,64 (9,45)
		20%	41,67 (11,01)	47,78 (3,61)	42,10 (2,68)	49,75 (1,52)
		40%	46,11 (4,21)	49,67 (3,44)	47,80 (6,27)	54,34 (2,88)
	60%	5%	39,96 (37,21)	40,50 (8,83)	46,21 (23,42)	49,00 (3,96)
		10%	52,40 (19,79)	45,55 (8,70)	55,84 (27,13)	45,85 (8,77)
		20%	46,37 (8,50)	40,55 (9,39)	38,27 (7,22)	44,25 (1,95)
		40%	46,86 (13,94)	49,45 (3,24)	41,75 (8,26)	49,39 (0,71)
	20%	5%	43,84 (32,04)	48,58 (13,11)	54,38 (34,81)	57,77 (14,27)
		10%	63,27 (29,17)	45,56 (11,05)	60,45 (18,55)	45,14 (8,08)
		20%	38,72 (18,02)	48,74 (6,77)	66,15 (15,48)	45,78 (5,63)
		40%	45,84 (10,95)	48,21 (6,53)	47,08 (8,98)	47,88 (9,20)
GL 3	80%	5%	51,24 (22,97)	54 (17,24)	44,62 (7,52)	46,74 (4,74)
		10%	52,44 (22,79)	46,99 (3,36)	48,75 (10,31)	49,75 (4,97)
		20%	48,32 (3,51)	50,76 (1,67)	47,79 (6,02)	52,00 (3,67)
		40%	49,99 (4,28)	55,81 (2,59)	47,16 (3,13)	60,01 (2,17)
	60%	5%	47,68 (14,92)	35,50 (31,37)	53,77 (8,09)	34,01 (30,43)
		10%	45,90 (10,88)	44,91 (4,92)	48,06 (11,34)	46,98 (4,64)
		20%	47,77 (4,27)	53,02 (1,14)	49,78 (2,97)	54,17 (2,79)
		40%	45,62 (5,73)	53 (3,70)	47,76 (3,56)	53,47 (2,45)
	20%	5%	59,67 (33,06)	42,54 (22,50)	32,74 (34,37)	47,13 (12,32)
		10%	34,53 (26,33)	49,56 (12,16)	46,95 (26,38)	47,47 (7,01)
		20%	40,30 (31,20)	47,86 (2,59)	42,75 (14,31)	48,17 (3,03)
		40%	54,07 (4,95)	52,38 (6,28)	48,90 (2,28)	51,91 (3,56)
GL 5	80%	5%	51,18 (18,90)	42,33 (8,29)	45,51 (22,79)	43,42 (7,58)
		10%	44,62 (8,08)	47,49 (4,85)	42,74 (5,55)	47,45 (5,15)
		20%	48,41 (5,07)	47,23 (1,22)	48,65 (6,38)	49,43 (4,09)
		40%	46,05 (2)	51,83 (4,66)	48,46 (4,60)	57,61 (5,77)
	60%	5%	39,36 (30,42)	44,85 (6,99)	39,86 (30,26)	44,87 (6,23)
		10%	52,01 (7,09)	49,13 (5,54)	48,88 (3,27)	46,78 (3,13)
		20%	42,34 (13,79)	49,68 (2,88)	48,07 (10,06)	48,23 (4,04)
		40%	45,68 (4)	50,41 (3,03)	46,26 (4,12)	55,63 (3,27)
	20%	5%	41,66 (18,96)	51,03 (15,75)	33,83 (25,01)	44,04 (10,35)
		10%	46,28 (23,53)	49,16 (5,55)	39,09 (28,27)	47,55 (5,45)
		20%	47,70 (11,97)	47,74 (10,83)	48,80 (8,44)	45,39 (8,73)
		40%	43,27 (4,64)	45,65 (2,49)	40,74 (7,05)	46,67 (2,07)

GL = grau de liberdade; h² = herdabilidade em porcentagem; AG = ação gênica; $\bar{r}_{\text{máx}}$ = média da posição relativa do QTL.

Tabela 32 - Estimativas médias de R^2 do modelo de regressão linear por marca simples e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica para os marcadores do grupo de ligação 5. O número total de repetições foi cinco, para cada situação (POP 25 a POP 36)

GL	h^2	AG	\bar{R}^2			
			AS	AM	AA	GS
GL 1	80%	5%	4,18 (2,51)	9,88 (4,84)	4,94 (2,28)	10,47 (1,71)
		10%	8,36 (3,40)	19,17 (2,34)	9,20 (1,45)	23,58 (7,90)
		20%	16,84 (3,43)	38,34 (12,40)	17,63 (7,96)	49,04 (18,16)
		40%	33,07 (7,22)	60,03 (10,16)	28,83 (3,59)	77,71 (8,23)
	60%	5%	4,01 (1,89)	6,64 (3,52)	2,82 (0,68)	6,53 (4,80)
		10%	5,31 (1,49)	9,36 (3,08)	4,60 (3,33)	10,98 (4,61)
		20%	11,31 (2,10)	27,96 (4,07)	12,02 (4,89)	32,85 (7,61)
		40%	20,55 (8,03)	45,80 (4,26)	17,61 (5,61)	60,33 (10,13)
	20%	5%	2,35 (1,09)	3,30 (1,14)	2,92 (2,17)	3,65 (2,17)
		10%	3,69 (1,80)	7,35 (3,31)	3,77 (1,02)	6,22 (3,81)
		20%	7,26 (4,17)	11,77 (2,47)	4,32 (2,41)	12,24 (2,56)
		40%	10,90 (4,23)	20,95 (5,66)	7,63 (3,72)	22,78 (5,32)
GL 3	80%	5%	5,50 (2,42)	14,86 (2,92)	6,02 (3,18)	11,62 (5,49)
		10%	7,99 (2,62)	19,53 (12,59)	9,85 (3,01)	24,59 (8,81)
		20%	16,58 (4,90)	40,15 (7,79)	17,82 (9,14)	50,76 (12,59)
		40%	25,70 (4,63)	67,18 (13,62)	28,88 (6,84)	83,13 (7,78)
	60%	5%	4,69 (3,61)	8,73 (3,43)	5,82 (2,54)	10,69 (8,07)
		10%	7,53 (3,36)	17,63 (4,10)	5,27 (2,07)	23,72 (6,81)
		20%	12,10 (3,59)	37,96 (8,76)	14,31 (1,72)	46,51 (11,87)
		40%	23,67 (2,41)	51,48 (9,29)	25,03 (4,78)	62,58 (13,34)
	20%	5%	6,65 (4,86)	2,83 (1,28)	5,37 (5,32)	4,04 (2,20)
		10%	4,83 (2,72)	9,42 (5,12)	5,90 (4,26)	12,34 (6,08)
		20%	6,26 (1,55)	13,14 (4,65)	5,53 (2,38)	16,77 (4,75)
		40%	9,62 (2,58)	19,15 (4,83)	9,53 (4,85)	24,85 (3,49)
GL 5	80%	5%	6,87 (1,86)	16,46 (8,07)	4,66 (1,23)	21,36 (5,18)
		10%	12,36 (2,44)	20,18 (7,20)	9,77 (1,66)	28,02 (10,59)
		20%	19,40 (5,23)	41,14 (8,53)	16,53 (4,61)	48,52 (15,29)
		40%	38,51 (6,56)	64,10 (6,16)	36,22 (7,17)	83,52 (2,47)
	60%	5%	4,86 (1,69)	12,94 (5,09)	4,79 (0,55)	13,28 (2,55)
		10%	6,74 (1,27)	17,99 (8,64)	9,38 (3,65)	24,41 (12,61)
		20%	12,89 (5,56)	28,71 (11,09)	15,54 (6,55)	38,29 (9,41)
		40%	27,51 (6,05)	57,71 (10,50)	25,12 (3,31)	73,53 (11,56)
	20%	5%	4,31 (2,93)	4,81 (2,97)	4,45 (2,22)	7,10 (2,77)
		10%	3,76 (1,76)	6,55 (3,22)	3,48 (1,57)	9,33 (3,76)
		20%	7,04 (3,67)	9,72 (2,97)	8,30 (2,85)	11,58 (3,54)
		40%	10,09 (2,87)	19,96 (8,84)	11,72 (4,20)	25,76 (5,98)

GL = grau de liberdade; h^2 = herdabilidade em porcentagem; AG = ação gênica; \bar{R}^2 = média do coeficiente de determinação.

4.3.2.3. Estimação do efeito do QTL sobre a característica

Avaliou-se o efeito do QTL sobre a característica fenotípica, por meio de análise de regressão pelo método de marca simples para as populações amostradas utilizando as estratégias AS, AM, AA e GS, para intensidade de amostragem de 20%. As estimativas de $\hat{\beta}_0$ e $\hat{\beta}_1$ e $\hat{\beta}_2$ para os marcadores flanqueadores do QTL, simulado nos cenários das POP 1 a 12, POP 13 a 24 e, POP 25 a 36, encontram-se nas tabelas 33, 34 e 35, nas tabelas 36, 37 e 38 e, nas tabelas 39, 40 e 41, respectivamente.

Nas tabelas 33, 36 e 39, as estimativas da média geral para a característica quantitativa, independente da estratégia de amostragem, encontram-se em torno da média paramétrica simulada, que foi de 1.000 unidades. Os valores das médias das estimativas de $\hat{\beta}_1$ foram maiores na GS e AM, em relação a AS e AA.

Tabela 33 - Médias das estimativas de $\hat{\beta}_0$ dos marcadores C_{13} e C_{14} sobre a característica e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica. Cinco populações foram avaliadas para cada estratégia de amostragem com seu respectivo cenário (POP 1 a 12)

M	h^2	AG	AS	AM	AA	GS
C_{13}	80%	5%	1003,11 (22,60)	1011,95 (19,90)	995,53 (19,54)	1022,42 (22,20)
		10%	1006,78 (21,08)	1012,02 (29,03)	992,55 (15,52)	1041,48 (23,24)
		20%	1007,15 (16,38)	990,96 (22,86)	991,08 (19,80)	1013,28 (16,44)
		40%	997,96 (6,80)	1004,80 (24,74)	992,24 (22,01)	1003,56 (16,47)
	60%	5%	999,99 (4,54)	993,76 (23,71)	1010,06 (11,56)	995,78 (17,36)
		10%	1002,54 (11,20)	984,47 (26,41)	999,45 (10,74)	987,22 (21,84)
		20%	1003,24 (5,85)	1011,80 (12,74)	1005,69 (20,07)	1020,74 (24,98)
		40%	997,02 (17,21)	1001 (18,64)	998,03 (14,64)	1007,11 (27,91)
	20%	5%	995,56 (5,62)	998,95 (14,43)	998,58 (13,43)	997,75 (7,91)
		10%	1006,64 (7,97)	999,60 (13,60)	1000,94 (6,61)	996,26 (15,57)
		20%	1003,18 (12,40)	997,19 (7,08)	1001,67 (10,35)	992,88 (11,66)
		40%	1000,89 (6,37)	1002,60 (13,15)	1000,91 (10,32)	1003,89 (14,38)
C_{14}	80%	5%	995,37 (7,15)	1022,40 (40,39)	1001,65 (12,50)	1025,69 (12,66)
		10%	1003,36 (10,12)	1000,80 (26,64)	997 (17,84)	1017,30 (15,41)
		20%	1012,32 (9,42)	998,79 (14,10)	990,89 (26,11)	1019,21 (30,11)
		40%	1000,52 (18,52)	1006,12 (31,31)	994,40 (21,93)	1003,22 (13,18)
	60%	5%	1000,84 (16,08)	1005,99 (12,84)	1009,70 (17,93)	1007,11 (20,54)
		10%	991,68 (16,16)	993,81 (27,11)	999,93 (6,29)	995,19 (26,07)
		20%	1006,59 (12,07)	1004,11 (12,65)	1004,32 (10,19)	1017,73 (15,81)
		40%	999,12 (11,49)	1004,49 (18,79)	1000,09 (10,79)	1008,54 (21,59)
	20%	5%	992,68 (7,20)	1001,39 (10,47)	997,32 (13,08)	997,80 (8,81)
		10%	999,39 (3,40)	994,45 (20,34)	999,97 (12,11)	992,73 (13,66)
		20%	1004,31 (8,68)	1002,29 (16,82)	1002,07 (10,20)	993,34 (23,11)
		40%	1002,73 (8,07)	1012,73 (14,11)	1004,67 (8,96)	1007,67 (21,96)

M = marcador; C_{13} = marcador três do grupo de ligação um; C_{14} = marcador quatro do grupo de ligação um; h^2 = herdabilidade em porcentagem; AG = ação gênica em porcentagem; AS = amostragem sistemática; AM = amostragem mista; AA = amostragem aleatória; GS = genotipagem seletiva.

Tabela 34 - Médias das estimativas de $\hat{\beta}_1$ dos marcadores C_{13} e C_{14} sobre a característica e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica. Cinco populações foram avaliadas para cada estratégia de amostragem com seu respectivo cenário (POP 1 a 12)

M	h^2	AG	AS	AM	AA	GS
C_{13}	80%	5%	51,04 (23,52)	137,20 (39,21)	52,52 (21,89)	163,08 (20,87)
		10%	78,13 (13,20)	179,93 (19,19)	84,23 (13,50)	219,19 (44,24)
		20%	109,90 (8,01)	239,94 (33,96)	112,67 (26,08)	301,86 (61,29)
		40%	147,68 (17,45)	277,97 (23,70)	122,80 (13,91)	347,67 (16,26)
	60%	5%	25,52 (21,93)	69,51 (29,49)	17,34 (12)	67,31 (39,01)
		10%	32,53 (17,63)	95,79 (27,38)	34,98 (20,18)	115,52 (25,92)
		20%	64,17 (5,23)	154,35 (14,87)	68,82 (8,62)	187,98 (23,73)
		40%	89,50 (15,37)	180,21 (7,02)	77,65 (21,89)	241,43 (16,95)
	20%	5%	11,94 (5,04)	20,72 (12,88)	11,20 (12,05)	33 (17,88)
		10%	14,85 (10,66)	55,87 (17,53)	18,64 (6,99)	60,61 (19,07)
		20%	22,51 (7,25)	67,03 (7,23)	11,52 (10,97)	75,38 (16,41)
		40%	42,69 (5,07)	89,35 (14,45)	33,46 (9,99)	107,93 (19,17)
C_{14}	80%	5%	37,04 (24,27)	117,93 (35,65)	50,35 (23,15)	149,36 (27,16)
		10%	75,49 (24,89)	178,65 (30,60)	81,02 (12,24)	221,20 (29,42)
		20%	112,04 (15,53)	242,41 (28,74)	111,14 (27,14)	308,25 (59,04)
		40%	150,72 (14,54)	286,52 (17,24)	142,80 (8,96)	354,66 (13,70)
	60%	5%	32,81 (17,20)	63,29 (25,10)	24,44 (16,44)	82,86 (38,63)
		10%	45,71 (6,10)	94,50 (10,63)	39,36 (16,32)	114,78 (34,65)
		20%	59,82 (17,15)	132,71 (13,67)	56,49 (11,51)	173,17 (23,61)
		40%	84,81 (11,37)	190,18 (16,29)	81 (17,79)	232,23 (18,88)
	20%	5%	12,56 (8,52)	28,10 (20,92)	10,32 (7,85)	39,12 (26)
		10%	10,78 (5,50)	57,56 (23,07)	22,53 (8,58)	55,26 (24,27)
		20%	16,92 (14,04)	68,01 (10,49)	19,28 (14,84)	80,72 (6,92)
		40%	42,87 (7,96)	97,45 (10,68)	33,51 (3,38)	113,80 (13,32)

M = marcador; C_{13} = marcador três do grupo de ligação um; C_{14} = marcador quatro do grupo de ligação um; h^2 = herdabilidade em porcentagem; AG = ação gênica em porcentagem; AS = amostragem sistemática; AM = amostragem mista; AA = amostragem aleatória; GS = genotipagem seletiva.

Tabela 35 - Médias das estimativas de $\hat{\beta}_2$ dos marcadores C₁₃ e C₁₄ sobre a característica e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica. Cinco populações foram avaliadas para cada estratégia de amostragem com seu respectivo cenário (POP 1 a 12)

M	h ²	AG	AS	AM	AA	GS
C ₁₃	80%	5%	-4,38 (45,25)	-15,83 (34,50)	21,56 (23,77)	-15,23 (39,40)
		10%	-10,41 (41,44)	-22,43 (44,93)	0,05 (39,23)	-85,53 (32,59)
		20%	-16,77 (32,87)	7,05 (54,35)	-3,04 (14,49)	-40,16 (44,27)
		40%	-1,24 (22,43)	-8,42 (66,17)	5,66 (27,29)	0,67 (68,55)
	60%	5%	3,49 (10,05)	13,41 (36,95)	-0,06 (14,87)	9,35 (19,62)
		10%	-4,04 (20,65)	14,83 (33,66)	-6,36 (18,22)	2,11 (26,69)
		20%	-8,90 (14,43)	-19,64 (24,09)	-4,01 (36,94)	-19,97 (37,57)
		40%	11,40 (31,68)	17,84 (26,54)	-0,90 (13,12)	8,25 (32,65)
	20%	5%	11,09 (13)	4,38 (32,20)	4,84 (17,79)	6,26 (22,62)
		10%	-13,43 (15,76)	-8,70 (16,20)	0,93 (10,82)	-4,33 (16,64)
		20%	-7,59 (22,93)	-6,40 (14,07)	-2,55 (22,39)	-8,19 (13,69)
		40%	-0,34 (11,47)	-0,11 (18,58)	6,48 (19,94)	-2,03 (9,71)
C ₁₄	80%	5%	12,88 (15,11)	-36,61 (59,15)	11,78 (22,11)	-32,17 (33,97)
		10%	-3,91 (16,66)	-4,26 (41,60)	-0,86 (37,74)	-22,71 (26,47)
		20%	-18,27 (18,23)	8,28 (37,56)	3,88 (27,28)	-42,62 (59,11)
		40%	2,49 (33,04)	-5,74 (68,89)	0,01 (32,66)	14,87 (69,54)
	60%	5%	-0,90 (28,11)	-11,69 (31,40)	0,25 (23,38)	-7,15 (35,06)
		10%	16,08 (30,42)	4,29 (45,06)	-3,80 (11,10)	-7,04 (44,17)
		20%	-9,33 (22,67)	-6,11 (26,66)	-0,08 (24,57)	-16,55 (9,99)
		40%	2,69 (21,07)	-1,55 (21,27)	-7,62 (15,46)	2,73 (29,63)
	20%	5%	15,83 (15,54)	-0,03 (24,11)	7,84 (17,90)	5,80 (21,71)
		10%	1,53 (6,70)	1,39 (27,87)	2,51 (15,32)	1,73 (25,19)
		20%	-7,42 (17,58)	-15,72 (29,19)	-3,29 (14,96)	-8,58 (29,95)
		40%	-5,18 (16,80)	-13,31 (18,66)	-0,86 (14,98)	-11,01 (21,58)

M = marcador; C₁₃ = marcador três do grupo de ligação um; C₁₄ = marcador quatro do grupo de ligação um; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem; AS = amostragem sistemática; AM = amostragem mista; AA = amostragem aleatória; GS = genotipagem seletiva.

Tabela 36 - Médias das estimativas $\hat{\beta}_0$ dos marcadores C_{36} e C_{37} sobre a característica e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica. Cinco populações foram avaliadas para cada estratégia de amostragem com seu respectivo cenário (POP 13 a 24)

M	h^2	AG	AS	AM	AA	GS	
C_{36}	80%	5%	1006,13 (11,35)	1024,65 (25,01)	1002,04 (25,62)	1009,63 (55,51)	
		10%	1002,19 (10,80)	1003,37 (16)	1001,14 (17,73)	1003,20 (40,55)	
		20%	1001,10 (15,33)	1013,38 (11,66)	995,29 (14,21)	1011,11 (23,92)	
		40%	992,96 (13,24)	1006,92 (17,47)	999,91 (26,23)	997,61 (20,81)	
	60%	5%	998,28 (3,31)	999,67 (9,56)	1004,42 (18,80)	989,52 (15,09)	
		10%	1001,55 (8,76)	997,22 (15,57)	999,13 (9,50)	987,29 (25,72)	
		20%	993,44 (8,74)	996,70 (22,69)	1003,25 (15,64)	989,70 (30,62)	
		40%	1002,45 (10,40)	997,05 (14,39)	1002,49 (14,28)	1008,96 (30,85)	
	20%	5%	1001,50 (6,77)	999,38 (15,82)	1008,21 (7,08)	1007,90 (12,32)	
		10%	1002,01 (6,78)	1002,94 (18,49)	1007,14 (7,44)	1004,12 (14,18)	
		20%	1002,17 (5,24)	998,78 (8,29)	996,26 (6,80)	1008,09 (27,37)	
		40%	1002,29 (8,08)	993,87 (9,52)	1000,94 (10,54)	990,49 (5,53)	
	C_{37}	80%	5%	1007,29 (9,95)	1026,96 (23,64)	1011,21 (23,96)	1028,32 (59,93)
			10%	1003,65 (10,48)	1022,92 (18,96)	1001,80 (15,11)	1003,64 (47,85)
			20%	998,17 (8,97)	1021,47 (15,69)	998,50 (7,84)	1007,25 (37,24)
			40%	1001,52 (14,30)	998,89 (15,03)	994,22 (33,22)	998,63 (10,46)
60%		5%	999,03 (8,81)	1004,69 (17,73)	994,18 (22,85)	996,98 (15,20)	
		10%	1002,03 (7,79)	995,40 (5,28)	1001,35 (6,69)	981,86 (19,76)	
		20%	997,80 (11,35)	1000,92 (27,83)	1002,12 (14,42)	993,24 (35,92)	
		40%	998,72 (8)	997,60 (12,12)	999,71 (13,14)	998,46 (28,10)	
20%		5%	997,16 (8,29)	999,21 (11,35)	1009,34 (8,24)	1008,89 (10,29)	
		10%	998,90 (5,34)	998,60 (16,01)	1003,62 (5,37)	994,48 (11,92)	
		20%	999,55 (4,75)	1006,83 (12,68)	999,75 (9,79)	1008,18 (23,60)	
		40%	1000,69 (3,82)	994,35 (11,26)	998,83 (12,22)	991,66 (7,37)	

M = marcador; C_{36} = marcador seis do grupo de ligação três; C_{37} = marcador sete do grupo de ligação três; h^2 = herdabilidade em porcentagem; AG = ação gênica em porcentagem; AS = amostragem sistemática; AM = amostragem mista; AA = amostragem aleatória; GS = genotipagem seletiva.

Tabela 37 - Médias das estimativas de $\hat{\beta}_1$ dos marcadores C₃₆ e C₃₇ sobre a característica e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica. Cinco populações foram avaliadas para cada estratégia de amostragem com seu respectivo cenário (POP 13 a 24)

GL	h ²	AG	AS	AM	AA	GS
C ₃₆	80%	5%	57,67 (12,86)	147,88 (26,60)	58,02 (17,42)	147,97 (48,07)
		10%	83,69 (16,56)	203,30 (48,91)	94,65 (14,71)	255,65 (37,07)
		20%	112,37 (21,53)	264,59 (14,87)	114,50 (28,72)	325,94 (26,30)
		40%	146,43 (7,57)	307,88 (17,40)	147,92 (12,67)	382,30 (9,29)
	60%	5%	29,85 (18,22)	74,79 (37,52)	33,93 (16,54)	86,14 (53,07)
		10%	52,19 (9,23)	128,54 (19,63)	46,15 (9,02)	166,06 (20,77)
		20%	70,90 (13,62)	173,90 (7,19)	76,37 (8,10)	222,09 (16,61)
		40%	101,76 (10,07)	208,16 (25,98)	105,37 (10,47)	252,58 (17,59)
	20%	5%	13,05 (5,83)	23,87 (12,02)	6,81 (12,99)	46,32 (12,31)
		10%	23,15 (13,80)	60,39 (26,73)	29,52 (10,42)	82,93 (29,69)
		20%	32,13 (2,88)	79,94 (10,21)	30,22 (9,52)	105,22 (11,89)
		40%	38,50 (8,88)	92,31 (17,41)	44,04 (14,93)	114,21 (14,22)
C ₃₇	80%	5%	65,10 (20,73)	153,16 (30,48)	60,11 (17,90)	159,70 (44,32)
		10%	77,62 (20,21)	180,61 (61,15)	88,87 (15,38)	247,12 (40,70)
		20%	117,76 (13,41)	253,93 (30,86)	112,28 (33,40)	316,51 (27,63)
		40%	149,77 (15,83)	325,22 (21,31)	150,23 (14,78)	394,44 (10,88)
	60%	5%	37,90 (18,50)	85,20 (31,54)	49,02 (12,27)	104,65 (35,26)
		10%	51,04 (14,83)	116,22 (20,27)	41,56 (12,27)	152,13 (22,50)
		20%	69,25 (15,36)	179,89 (11,14)	83,78 (10,14)	225,38 (28,64)
		40%	98,89 (9,18)	205,71 (23,68)	103,96 (13,82)	254,68 (27,14)
	20%	5%	10,26 (4,78)	25,02 (19,77)	10,56 (12,32)	47,93 (19,27)
		10%	26,84 (15,13)	60,55 (26,18)	26,17 (13,44)	83,54 (26,59)
		20%	31,92 (7,80)	82,07 (16,96)	28,79 (10,38)	102,82 (16,99)
		40%	45,88 (5,32)	98,41 (12,15)	43,87 (12,67)	127,50 (10,80)

M = marcador; C₃₆ = marcador seis do grupo de ligação três; C₃₇ = marcador sete do grupo de ligação três; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem; AS = amostragem sistemática; AM = amostragem mista; AA = amostragem aleatória; GS = genotipagem seletiva.

Tabela 38 - Médias das estimativas de $\hat{\beta}_2$ dos marcadores C₃₆ e C₃₇ sobre a característica e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica. Cinco populações foram avaliadas para cada estratégia de amostragem com seu respectivo cenário (POP 13 a 24)

M	h ²	AG	AS	AM	AA	GS
C ₃₆	80%	5%	-7,35 (24,11)	-30,50 (37,75)	-17,61 (39,18)	10,47 (96,86)
		10%	-0,45 (21,42)	8,05 (30,63)	-5,79 (17,56)	4,46 (53,54)
		20%	-8,14 (35,65)	0,01 (42,96)	15,21 (23,57)	25,06 (44,64)
		40%	21,35 (21,69)	2,68 (35,75)	6,92 (25,87)	-2,44 (39,28)
	60%	5%	5,01 (7,49)	-7,45 (16,06)	-5,46 (31,92)	12,45 (35,45)
		10%	-2,64 (15,78)	-2,89 (42,72)	3,25 (8,19)	10,36 (71,53)
		20%	11,39 (15,70)	6,53 (28,80)	1,56 (19,15)	17,81 (32,77)
		40%	-3,61 (16,09)	4,57 (25,19)	0,10 (18,11)	6,80 (23,24)
	20%	5%	-3,75 (13,92)	1,29 (32,82)	-10,38 (10,28)	-11,20 (17,64)
		10%	-4,82 (13,83)	0,36 (32,82)	-8,65 (19,15)	1,12 (15,72)
		20%	-5,49 (10,76)	11,46 (8,09)	7,27 (11,13)	-5,92 (21,14)
		40%	-2,36 (13,32)	9,87 (21,77)	-1,01 (14,97)	20,01 (18,74)
C ₃₇	80%	5%	-11,36 (22,61)	-36,65 (54)	-37,55 (35,21)	-27,14 (105,23)
		10%	-5,24 (21,74)	-20,32 (22,76)	-7,16 (28,77)	9,08 (59,27)
		20%	-0,50 (9,21)	-25,72 (42,05)	8,66 (19,97)	19,45 (32,05)
		40%	3,69 (30,69)	14,55 (37,13)	13,97 (34,03)	-3,65 (46,77)
	60%	5%	2,66 (17,97)	-11,09 (15,31)	11,03 (41,08)	5,41 (31,28)
		10%	-2,27 (13,42)	-1,98 (24,76)	-1,05 (8,02)	15,39 (58,84)
		20%	4,26 (20,72)	0,75 (32,93)	3 (13,32)	6,11 (35,99)
		40%	5,33 (14,10)	-5,11 (25,86)	2,56 (19,05)	10,53 (26,37)
	20%	5%	6,78 (17,69)	3,06 (28)	-11,94 (11,87)	-11,55 (17,68)
		10%	2,28 (10,13)	8,53 (24,76)	-1,05 (10,78)	19,97 (12,20)
		20%	1,55 (10,33)	-3,87 (23,70)	-1,68 (18,35)	-10,32 (14,57)
		40%	1,44 (9,55)	7,94 (18,84)	3,13 (19,68)	15,76 (31,17)

M = marcador; C₃₆ = marcador seis do grupo de ligação três; C₃₇ = marcador sete do grupo de ligação três; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem; AS = amostragem sistemática; AM = amostragem mista; AA = amostragem aleatória; GS = genotipagem seletiva.

Tabela 39 - Médias das estimativas de $\hat{\beta}_0$ dos marcadores C₅₁₁ e C₅₁₂ sobre a característica e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica. Cinco populações foram avaliadas para cada estratégia de amostragem com seu respectivo cenário (POP 25 a 36)

M	h ²	AG	AS	AM	AA	GS
C ₅₁₁	80%	5%	997,49 (7,79)	1013,87 (12,16)	995,35 (19,86)	1013,29 (27,71)
		10%	997,25 (15,03)	998,01 (16,75)	1003,88 (12,63)	1028,13 (21,60)
		20%	995,43 (10,74)	985,04 (29,78)	1008,72 (15,53)	992,24 (39,15)
		40%	1001,47 (17,72)	984,61 (6,58)	994,14 (15,86)	994,42 (18,74)
	60%	5%	991,55 (9,40)	984,32 (19,27)	986,15 (6,67)	992,53 (16,14)
		10%	1003,70 (6,46)	997,75 (9,42)	999,92 (14,29)	989,30 (19,15)
		20%	997,11 (7,49)	1006,71 (8,59)	1004,09 (17,99)	1007,88 (11,36)
		40%	998,75 (7,69)	983,76 (11,35)	993,70 (16,27)	998,06 (28,76)
	20%	5%	998,56 (5,63)	998,88 (14,01)	994,05 (11,83)	992,90 (16,35)
		10%	998,61 (5,76)	1009,50 (11,09)	1000,24 (12,91)	1015,82 (19,06)
		20%	997,52 (7,14)	1000,15 (21,97)	997,99 (9,57)	1006,20 (23,84)
		40%	998,67 (3,87)	997,17 (14,40)	997,45 (6,86)	994,03 (10,10)
C ₅₁₂	80%	5%	997,85 (21,76)	1011,11 (20,79)	995,29 (20,22)	1022,30 (35,19)
		10%	997,06 (14,78)	995 (15,87)	999,23 (18,15)	1019,36 (26,08)
		20%	1007,42 (15,40)	981,62 (25,64)	1009,41 (21,41)	987,07 (42,12)
		40%	1003,95 (17,65)	986,11 (12,57)	1000,24 (21,07)	994,12 (19,16)
	60%	5%	990,53 (5,70)	992,55 (17,93)	984,02 (8,64)	997,77 (14,03)
		10%	1005,17 (8,49)	1003,33 (7,41)	1006,22 (10,05)	995,63 (23,84)
		20%	1000,06 (5,88)	1011,81 (21,46)	1004,11 (17,12)	1011,96 (15,94)
		40%	996,28 (6,39)	987,20 (6,36)	993,06 (15,16)	1001,32 (26,53)
	20%	5%	997,53 (7,49)	997,14 (15,31)	992,68 (12,90)	993,21 (17,15)
		10%	999,25 (8,01)	1012,56 (14,89)	1001,89 (12,86)	1020,71 (21,82)
		20%	1002,81 (9,76)	1005,95 (17,60)	1001,68 (12,87)	1007,84 (16,64)
		40%	1003,50 (3,89)	1002,38 (10,29)	1002,53 (6,03)	1000,28 (7,72)

M = marcador; C₅₁₁ = marcador 11 do grupo de ligação cinco; C₅₁₂ = marcador 12 do grupo de ligação cinco; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem; AS = amostragem sistemática; AM = amostragem mista; AA = amostragem aleatória; GS = genotipagem seletiva.

Tabela 40 - Médias das estimativas de $\hat{\beta}_1$ dos marcadores C_{511} e C_{512} sobre a característica e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica. Cinco populações foram avaliadas para cada estratégia de amostragem com seu respectivo cenário (POP 25 a 36)

M	h^2	AG	AS	AM	AA	GS
C_{511}	80%	5%	59,09 (27,02)	170,04 (44,65)	54,11 (9,77)	224,02 (34,46)
		10%	98,43 (15,81)	189,95 (14,56)	84,32 (7,78)	249,93 (18,25)
		20%	125,46 (14,71)	277,72 (17,30)	112,14 (25,24)	312,98 (28,48)
		40%	182,21 (17,73)	319,49 (10,86)	174,48 (9,54)	380,23 (18,52)
	60%	5%	38,10 (6,24)	104,16 (20,88)	30,22 (19,68)	123,97 (12,25)
		10%	47,83 (11,46)	123,17 (22,58)	53,39 (7,52)	166,92 (27,77)
		20%	73,70 (22,54)	168,84 (32,47)	80,07 (20,89)	216,57 (24,79)
		40%	109,97 (14,12)	217,78 (22,60)	98,22 (10,39)	265,18 (14,58)
	20%	5%	19,51 (4,80)	38,73 (19,25)	15,05 (16,80)	55,49 (26,23)
		10%	18,31 (7,19)	53,79 (9,93)	16,90 (7,19)	70,35 (18,72)
		20%	32,88 (13,62)	70,51 (17,49)	36,22 (6,78)	88,85 (18,13)
		40%	47,61 (9,83)	97,59 (22,21)	48,43 (3,80)	124,77 (17,94)
C_{512}	80%	5%	64,39 (18,27)	174,90 (43,09)	51,46 (18,92)	232,44 (41,27)
		10%	99,25 (9,15)	179,75 (32,63)	83,29 (16,98)	247,33 (31,21)
		20%	129,15 (15,22)	270,45 (19,25)	112,30 (28,42)	317,74 (28,18)
		40%	174,37 (16,18)	310,22 (11,56)	166 (17,15)	369,62 (22,68)
	60%	5%	36,12 (7,71)	103,79 (17,25)	31,98 (14,72)	123,74 (16,04)
		10%	53,30 (7,42)	126,81 (16,93)	61,46 (11,39)	172,54 (22,79)
		20%	77,11 (20,70)	166,47 (28,92)	78,09 (23,69)	214,89 (16,82)
		40%	110,60 (9,27)	213,56 (23,08)	98,32 (7,13)	262,21 (19,78)
	20%	5%	19,50 (4,95)	40,29 (22,85)	13 (19,09)	59,92 (25,08)
		10%	19,11 (8,95)	53,65 (11,51)	18,79 (7,49)	68,41 (22,68)
		20%	31,92 (10,64)	71,13 (12,88)	40,94 (4,49)	89,59 (12,67)
		40%	44,28 (8,61)	94,77 (22,64)	44,33 (5,34)	119,92 (18,35)

M = marcador; C_{511} = marcador 11 do grupo de ligação cinco; C_{512} = marcador 12 do grupo de ligação cinco; h^2 = herdabilidade em porcentagem; AG = ação gênica em porcentagem; AS = amostragem sistemática; AM = amostragem mista; AA = amostragem aleatória; GS = genotipagem seletiva.

Tabela 41 - Média das estimativas $\hat{\beta}_2$ dos marcadores C₅₁₁ e C₅₁₂ sobre a característica e seus desvios-padrão, entre parênteses, em função do nível de herdabilidade e de ação gênica. Cinco populações foram avaliadas para cada estratégia de amostragem com seu respectivo cenário (POP 25 a 36)

M	h ²	AG	AS	AM	AA	GS	
C ₅₁₁	80%	5%	4,29 (12,81)	-5,55 (24,05)	10,75 (27,63)	-9,89 (31,77)	
		10%	7,54 (26,29)	14,34 (58,82)	-6,59 (40,02)	-19,93 (60,80)	
		20%	6,35 (21,61)	18,80 (52,11)	-2,87 (26,22)	12,34 (38)	
		40%	-8,76 (30,34)	14,48 (34,43)	12,72 (20,50)	-25,92 (20,50)	
	60%	5%	16,32 (18,58)	37,75 (39,50)	23,10 (20,06)	26,81 (32,01)	
		10%	-4,24 (10,30)	-6,29 (11,06)	-5,97 (14,23)	7,14 (22,87)	
		20%	3,54 (9,97)	-1,71 (8,58)	-2,10 (24,95)	4,37 (28,08)	
		40%	2,40 (16,29)	17,98 (16,22)	7,09 (8,64)	-13,92 (47,82)	
	20%	5%	3,77 (10,96)	4,35 (13,27)	2,10 (17,23)	14,15 (20,21)	
		10%	3,28 (12,42)	-0,64 (11,57)	1,34 (10,55)	-9,93 (26,70)	
		20%	6,48 (16,15)	7,80 (24,23)	-9,48 (20,72)	6,11 (21,31)	
		40%	0,67 (6,75)	3,07 (27,38)	7,05 (11,70)	6,69 (25,77)	
	C ₅₁₂	80%	5%	2,56 (43,47)	2,91 (42,97)	13,37 (21,26)	-14,85 (81,64)
			10%	8,99 (23,27)	20,85 (61,11)	2,63 (47,19)	-5,66 (44,41)
			20%	-11,97 (33,61)	32,54 (53,63)	-0,88 (44,37)	30,19 (36,66)
			40%	-10,63 (25,73)	13,25 (31,25)	2,07 (27,20)	-31,86 (45,52)
60%		5%	17,41 (15,08)	22,81 (30,35)	28,03 (22,63)	21,55 (20,82)	
		10%	-7,63 (15,98)	-19,24 (20,07)	-17,63 (21,15)	-3 (31,23)	
		20%	1,39 (9,53)	-8,47 (27)	-3,83 (25,21)	-2,09 (42,48)	
		40%	9,11 (12,64)	16,78 (13,54)	7,17 (3,01)	-15,47 (34,98)	
20%		5%	5,27 (14,92)	8,61 (18,03)	5,59 (18,12)	14,46 (25,02)	
		10%	1,36 (16,13)	-7,83 (20,07)	-1,72 (7,40)	-18,72 (27,70)	
		20%	-3,55 (21,47)	-2,54 (16,01)	-14,90 (22,15)	2,05 (16,37)	
		40%	-8,81 (7,17)	-6,37 (24,65)	-2,23 (13,97)	-7,96 (29,15)	

M = marcador; C₅₁₁ = marcador 11 do grupo de ligação cinco; C₅₁₂ = marcador 12 do grupo de ligação cinco; h² = herdabilidade em porcentagem; AG = ação gênica em porcentagem; AS = amostragem sistemática; AM = amostragem mista; AA = amostragem aleatória; GS = genotipagem seletiva.

5. CONCLUSÕES

- O mapa de ligação, mesmo com distorção de segregação dos locos individuais, foi satisfatoriamente reconstituído nas populações selecionadas por genotipagem seletiva, com suas respectivas intensidades de amostragem de 50, 30 e 20%. O ordenamento e o número de grupos de ligação, obtidos por análise genômica, estiveram de acordo conforme esperado. No entanto, para intensidade de seleção de 10%, o que equivale a uma amostragem de 100 indivíduos, em todos os cenários, pelo menos uma população apresentou marcas não-ligadas e, ou inversão de marcar, maior número de grupos de ligação, em relação ao genoma simulado.

- Os cenários que apresentaram limitação na identificação de QTL nas populações amostradas por genotipagem seletiva, em relação ao “LOD score”, foi praticamente a mesma para uma população F_2 de 1.000 indivíduos, embora a ligeira diminuição no valor da média “LOD score”, a medida em aumenta-se a intensidade de amostragem.

- A contribuição do QTL para variação da característica, obtida por meio de análises de regressão pelo método de intervalo simples, foi em média superestimada, porém, aplicando-se a correção empírica para o tamanho da população original. Portanto, o valor calculado para coeficiente de determinação foi muito próximo ao que encontraria numa população F_2 de 1.000 indivíduos.

- As médias da posição relativa do QTL estavam entre as distâncias dos marcadores flanqueadores do QTL simulado. Entretanto, à medida a intensidade de amostragem foi aumentada, também os desvios foram aumentados.

- As médias das estimativas do efeito aditivo “ $a = \hat{\beta}_1(1-2r)^{-1}$ ”, em todos os cenários, foram superestimadas. O grau médio de dominância observado pela razão da média do efeito aditivo e do efeito devido à dominância, foi praticamente nulo. Assim, o viés médio das interações intra-alélicas foi proporcional ao viés médio do efeito aditivo superestimado. A média geral da característica esteve em torno de 1.000.

Mesmo com as maiores médias de distorções observadas nos locos individuais, os resultados obtidos das análises genômicas, em estratégias experimentais de amostragem mista e genotipagem seletiva, foram satisfatórios. Em média, foram mantidos mapas de ligação próximos ao esperado, com alto poder de detecção de QTL, em comparação com resultados obtidos em amostragens aleatória e sistemática.

6. REFERÊNCIAS BIBLIOGRÁFICAS

CRUZ, E.M. **Efeito da saturação e do tamanho de populações F₂ e de retrocruzamento sobre a acurácia do mapeamento genético.** 2006. 119p. Tese (Doutorado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa.

DARVASI, A.; SOLLER, M. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. **Theoretical and Applied Genetics**, Berlin, v.85, n.2-3, p.353-359, 1992.

FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética.** 3.ed. Brasília: EMBRAPA/CENARGEM, 1998. 220p.

GROOVER, A.; DEVEY, M.; FIDDLER, T.; LEE, J.; MEGRAW, R.; MITCHEL-OLDS, T.; SHERMAN, B.; VUJCIC, S.; WILLIAMS, C.; NEALE D. Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. **Genetics**, Bethesda, v.138, n.4, p.1293-1300, 1994.

LANDER, E.S., BOTSTEIN, D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, Bethesda, v.121, n.1, 185-199, 1989. (erratum in 1994 **Genetics**, v.136, n.2, 705)

LEBOWITZ, R.J.; SOLLER, M.; BECKMANN, J.S. Trait-based analyses for the detection of linkage between marker loci and quantitative trait loci in crosses between inbred lines, **theoretical and Applied Genetics**, Berlin, v.73, n.4, p.556-562, 1987.

LI, XINMIN; QUIGG, R.J.; ZHOU, J.; XU, S.; MASINDE, G.; MOHAN, S.; BAYLINK, D.J. A critical evaluation of the effect of population size and phenotypic measurement on QTL detection and localization using a large F₂ murine mapping population. **Genetics and Molecular Biology**, Ribeirão Preto, v.29, n.1., p.166-173, 2006.

LIU, B.H. **Statistical genomics, linkage, mapping and QTL analysis**. Boca Raton: CRC, 1998. 611p.

MACKAY, T.F.C. Genetic architecture of quantitative traits. **Annual Review of Genetics**, Palo Alto, v.35, p.303-339, 2001.

NANDI, S.; SUBUDHI, P.K.; SENADHIRA, D.; MANIGBAS; N.L.; SEN-MANDI, S.; HUANG, N. Mapping QTLs for submergence tolerance in rice by AFLP analysis and selective genotyping. **Molecular and General Genetics**, Heidelberg, v.255, n.1, p.1-8, 1997.

SAX, K. The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgaris*. **Genetics**, Bethesda, v.8, n.6, p.552-560, 1923.

SCHUSTER, I.; CRUZ, C.D. **Estatística genômica aplicada a populações derivadas de cruzamentos controlados**. Viçosa: UFV, 2004. 568p.

SILVA, L.C. **Simulação do tamanho da população e da saturação do genoma para mapeamento genético de RILs**. 2005. 120p. Dissertação (Mestrado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa.

SOLLER, M.; BRODY, T.; GENIZI, A. On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. **Theoretical and Applied Genetics**, Berlin, v.47, n.1, p.35-39, 1976.

STUBER, C.W.; MOLL, R.H.; GOODMAN, M.M.; SCHAFFER, H.E.; WEIR, B.E. Allozyme frequency changes associated with selection for increased grain yield in maize (*Zea mays* L.). **Genetics**, Bethesda, v.95, n.1, p.225–236, 1980.

VALES, M.I.; SCHÖN, C.C.; CAPETTINI, F.; CHEN, X.M.; COREY, A.E.; MATHER, D.E.; MUNDT, C.C.; RICHARDSON, K.L.; SANDOVAL-ISLAS, J.S.; UTZ, H.F.; HAYES, P.M. Effect of population size on the estimation of QTL: a test using resistance to barley stripe rust. **Theoretical and Applied Genetics**, Berlin, v.111, n.7, p.1260-1270, 2005.

VAN OOIJEN, J.W. Accuracy of mapping quantitative trait loci in autogamous species. **Theoretical and Applied Genetics**, Berlin, v.84, n.7-8, p.803-811, 1992.

WINGBERMUEHLE, W.J.; GUSTUS, C.; SMITH, K.P. Exploiting selective genotyping to study genetic diversity of resistance to Fusarium head blight in barley. **Theoretical and Applied Genetics**, Berlin, v.109, n.6, p.1160-1168, 2004.