

**BORIS BRIÑEZ RODRIGUEZ**

**“DESENVOLVIMENTO DA PLATAFORMA DART E  
MAPEAMENTO DE LOCOS ASSOCIADOS COM TOLERÂNCIA  
À SECA EM FEIJÃO (*Phaseolus vulgaris* L.)”**

**CAMPINAS**

**2013**





**UNIVERSIDADE ESTADUAL DE CAMPINAS**

**INSTITUTO DE BIOLOGIA**

**BORIS BRIÑEZ RODRIGUEZ**

**“DESENVOLVIMENTO DA PLATAFORMA DART E  
MAPEAMENTO DE LOCOS ASSOCIADOS COM TOLERÂNCIA À  
SECA EM FEIJÃO (PHASEOLUS VULGARIS L)”**

Este exemplar corresponde à redação final  
da tese defendida pelo(a) candidato (a)  
BORIS BRIÑEZ RODRIGUEZ  
Luciana Benchimol Reis  
e aprovada pela Comissão Julgadora.

Tese apresentada ao Instituto de Biologia da  
UNICAMP para obtenção do Título de Doutor  
em Genética e Biologia Molecular, na área de  
Genética Vegetal e Melhoramento .

Orientadora: Dra. Luciana Lasry Benchimol Reis (Orientadora)  
Coorientador: Dr. Matthew Ward Blair

CAMPINAS,  
2013

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Biologia  
Mara Janaina de Oliveira - CRB 8/6972

B771d Briñez Rodriguez, Boris, 1975-  
Desenvolvimento da plataforma DART e mapeamento de locos associados com tolerância à seca em feijão (*Phaseolus vulgaris* L.) / Boris Briñez Rodriguez. – Campinas, SP : [s.n.], 2013.

Orientador: Luciana Lasry Benchimol Reis.  
Coorientador: Matthew Ward Blair.  
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Desidratação. 2. Mapeamento de QTL. 3. Microssatélites (Genética). 4. Polimorfismo de nucleotídeo único. 5. Linhagem endogâmica recombinante. I. Benchimol-Reis, Luciana Lasry, 1970-. II. Blair, Matthew Ward. III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título.

Informações para Biblioteca Digital

**Título em inglês:** Development of DART platform and quantitative trait loci identification associated to drought tolerance in common bean (*Phaseolus vulgaris* L.)

**Palavras-chave em inglês:**

Dehydration

QTL mapping

Microssatellites (Genetics)

Polymorphism single nucleotide

Recombinant inbred lines

**Área de concentração:** Genética Vegetal e Melhoramento

**Titulação:** Doutor em Genética e Biologia Molecular

**Banca examinadora:**

Luciana Lasry Benchimol Reis [Orientador]

Rosana Pereira Vianello

Siu Mui Tsai

José Bladin Pinheiro

Julieta Andrea Silva de Almeida

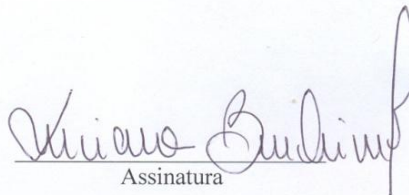
**Data de defesa:** 06-06-2013

**Programa de Pós-Graduação:** Genética e Biologia Molecular

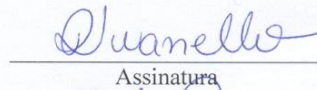
Campinas, 6 de junho de 2013

**BANCA EXAMINADORA**

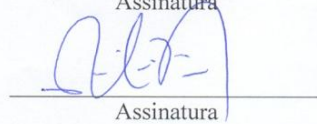
Dra. Luciana Lasry Benchimol Reis (Orientadora)  
(orientadora)

  
Assinatura

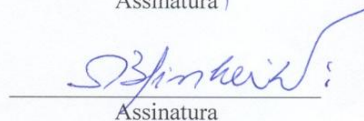
Dra. Rosana Pereira Vianello

  
Assinatura

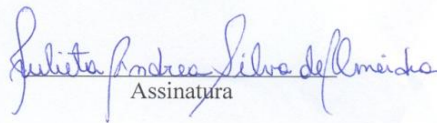
Dra. Siu Mui Tsai

  
Assinatura

Dr. José Baldin Pinheiro

  
Assinatura

Dra. Julieta Andrea Silva de Almeida

  
Assinatura

Dra. Tammy Aparecida Manabe Kiihl

\_\_\_\_\_  
Assinatura

Dr. Jorge Mauricio Costa Mondego

\_\_\_\_\_  
Assinatura

Dr. Carlos Augusto Colombo

\_\_\_\_\_  
Assinatura

## SUMÁRIO

RESUMO .....	viii
ABSTRACT .....	x
LISTA DE SIGLAS .....	xi
1. INTRODUÇÃO GERAL.....	1
2. REVISÃO DE LITERATURA.....	4
2.1 O feijoeiro comum.....	4
2.2 Respostas do feijoeiro ao estresse hídrico .....	5
2.3 Marcadores Moleculares.....	9
2.3.1 Marcadores Microssatélites (SSRs).....	10
2.3.2 Marcadores SNPs ( <i>Single Nucleotide Polimorphism</i> ).....	11
2.3.3 Marcadores moleculares: DArTs ( <i>Diversity Arrays Tecchnology</i> ) .....	13
2.4 Populações de Mapeamento .....	16
2.5 Construção de mapas genéticos .....	17
2.6 Identificação de QTL.....	22
3. OBJETIVO .....	26
3.1 Objetivo Geral .....	26
3.2 Objetivos específicos.....	26
4. MATERIAL E MÉTODOS .....	27
4.1 Material Vegetal .....	27
4.2 Avaliações quanto às Características Relacionadas ao Estresse Hídrico .....	27
4.2.1 Fenotipagem em plantas .....	27
4.2.2 Avaliação do sistema radicular quanto a comprimento, volume e diâmetro.....	29
4.3 Análise de dados .....	30
4.4 Extração e Quantificação de DNA .....	31
4.5 Desenvolvimento da plataforma DArTs para feijão comum e genotipagem na população AS com marcadores DArTs.....	31
4.5.1 Preparação dos microarranjos .....	35

4.5.2 Análise da diversidade genética. ....	37
4.5.3 Análises de coordenadas principais.....	37
4.5.4 Genotipagem com marcadores DArTs .....	38
4.6 Genotipagem com marcadores SNPs. ....	38
4.7 Caracterização e genotipagem com marcadores microssatélites .....	39
4.8 Construção do mapa genético AS com marcadores SNPs e SSRs .....	40
4.9 Mapeamento dos locos associados à tolerância ao estresse hídrico no mapa SSR-SNP	41
4.10 Construção do mapa genético AS com marcadores SNPs – SSRs – DArTs.....	41
5 RESULTADOS E DISCUSSÃO.....	42
5.1 Desenvolvimento da plataforma DArT para feijão comum .....	42
5.1.1 Seleção do método de redução da complexidade .....	42
5.1.2 Desenvolvimento do <i>Array</i> .....	43
5.1.3 Diversidade Genética de Feijão Comum.....	44
5.1.4 Análise de coordenadas principais .....	44
5.2 Construção do Mapa AS.....	49
5.2.1 Mapa com marcadores SSRs – SNPs .....	49
5.2.2 Mapa com marcadores SSRs – SNPs – DArTs.....	51
5.3 Fenotipagem em vasos .....	54
5.3.1 Identificação de QTLs em vasos no mapa SSR-SNPs. ....	57
5.3.2 Identificação de QTLs em vasos no mapa SSR-SNPs-DArTs.....	62
5.4 Fenotipagem em rizotron.....	64
5.4.1 Identificação de QTLs em rizotron no mapa SSR-SNPs.....	67
5.4.2 Identificação de QTLs em rizotron a partir do mapa SSR-SNPs-DArT. ....	73
CONSIDERAÇÕES FINAIS .....	76
CONCLUSÕES .....	79
REFERÊNCIAS .....	80
ANEXOS .....	94

## RESUMO

O feijão comum (*Phaseolus vulgaris* L.) é uma cultura importante economicamente tanto para o consumo nacional como para a exportação. A seca é um dos principais estresses abióticos em todo o mundo e afeta cerca de 60% da área de cultivo de feijão. O avanço nas tecnologias de marcadores moleculares oferecem poderosos métodos para examinar as relações entre as características, gerando um grande volume de informações potencialmente úteis para assessorar os programas de melhoramento. O presente projeto teve como objetivo o desenvolvimento da Plataforma DArT para feijão comum junto à empresa DArT Pty Ltd, e o mapeamento destes marcadores juntamente com microssatélites e SNPs na população AND 277 x SEA 5 proveniente do CIAT (Colômbia), a fim de localizar os QTLs associados à tolerância à seca. O genitor SEA 5 é uma linhagem avançada do BAT 477, é tolerante à seca e de origem Mesoamericano e o genitor AND 277 é um genótipo resistente à mancha angular e antracnose e de origem Andina. Um total de 4.468 marcadores DArTs, 288 marcadores SNPs e 180 marcadores microssatélites polimórficos foram identificados na população e utilizados na genotipagem para construir um mapa genético saturado. A fenotipagem das 105 linhagens endogâmicas recombinantes (RILs) na geração F<sub>8</sub> mais os dois genitores foi realizada avaliando 18 características associadas a tolerância a seca utilizando um delineamento inteiramente casualizado com quatro repetições, aplicando um estresse terminal na fase vegetativa V3/V4. Dois mapas foram construídos, um integrando 80 SSR e 251 SNPs e outro com cinco SSR, 91 SNPs e 4.468 DArTs. A identificação dos QTLs foi realizada através da análise de mapeamento por intervalo composto (CIM) para o mapa SSR – SNPs e mapeamento de precisão (SML) para o mapa SSR-SNPs-DArT. Um total de 12 QTLs foram identificados para o tratamento não irrigado e 29 QTLs para o tratamento irrigado pela análise CIM. Para as análises SML, 23 QTLs foram identificados para o tratamento não irrigado e 11 QTLs para o irrigado. QTLs de maior efeito foram encontrados para clorofila, biomassa fresca do caule e da folha, Massa seco da folia, temperatura da folha, número de vagens, número de sementes, massa de sementes, dias para florescimento, massa seca das vagens e produtividade nos dois tratamentos. Todos os QTLs detectados sob condições de seca apresentaram o alelo do genitor SEA 5. Este estudo é importante para o melhoramento genético não só para entender melhor a herança genética de uma característica tão complexa como a tolerância à



seca, bem como para encontrar ferramentas moleculares a serem utilizados para a seleção assistida por marcadores.

**Palavras-chave:** SSRs, SNPs, RILs, estresse hídrico, QTL.

## ABSTRACT

Common bean (*Phaseolus vulgaris* L.) is the most important food legume for consumption and for exportation. Drought is one of the main abiotic stresses in the world and affects about 60% of bean growing area across the world. The advance in technologies of molecular markers provide a powerful method to examine the relationships between traits, generating large amount of potentially useful information to assist the breeding programs. The objective of this project was the development of DArT platform for common beans with DArT Pty Ltd and the mapping of these markers with microsatellites and SNPs in the population AND 277 x SEA 5 from CIAT (Colombia), in order to locate the QTLs associated with drought tolerance. The SEA 5 parent is a drought tolerant advanced line (Mesoamerican) and the AND 277 is resistant to the angular leaf spot and antracnose (Andean). A total of 4,468 DArT markers, 288 SNP and 180 SSR polymorphic markers were identified in the population and used in genotyping to constructed a saturated genetic map. Phenotyping of 105 recombinant inbred lines (RILs) in  $F_8$  generation plus the genitors were performed evaluating 18 traits associated with drought tolerance using a completely randomized design with four replicates, applying terminal stress at vegetative phase V3/V4. Two maps were constructed, one integrating 80 SSR and 251 SNPs and another with five SSR, 91 SNPs and 4,468 DArTs. The identification of QTL analysis was performed by composite interval mapping (CIM) for the SSR – SNPs map and the precision mapping (SML) to map DArT-SSR-SNPs. A total of 12 QTLs were identified for the non-irrigated treatment and 29 QTLs for the irrigated treatment by CIM analysis. For SML analysis, 23 QTLs were identified for the non-irrigated and 11 QTLs for irrigated treatment. QTLs of major effect was found for chlorophyll, fresh biomass of stem and leaf dry weight, leaf temperature, number of pods, number of seeds, seed weight, days to flowering, dry weight of pods and yield in both treatments. All QTLs detected under dry conditions showed the allele of parent SEA 5. This study is important for genetic improvement not only to better understand the genetic inheritance of a trait as complex as drought tolerance, as well as to find molecular tools to be used for marker assisted selection.

**Keywords:** SSRs, SNPs, RILs, water stress, QTL.

## LISTA DE SIGLAS

AFI/AFNI - Área foliar (Irigado/Não irrigado)  
AFLP - *Amplified Fragment Length Polymorphism*  
APRI/APRNI - Área projetada de raízes (Irigado/Não irrigado)  
AS – “AND277 x SEA5”  
ASO - alelo-específico  
ASRI/ASRNI - Área superficial de raízes (Irigado/Não irrigado)  
BIC - *Baysian Information Criterion*  
BJ – “BAT93 x Jalo EEP558”  
CFM – “Carioca x Flor de Mayo”  
CI/CNI - Clorofila (Irigado/Não irrigado)  
cM - centiMorgans  
CRI/CRNI - Comprimento de raízes (Irigado/Não irrigado)  
DArTseq - *Diversity Arrays Technology*  
DRI/DRNI - Diâmetro de raízes (Irigado/Não irrigado)  
FI/FNI - Dias até florescimento (Irigado/Não irrigado)  
LODscore - *Logarithm of Odds*  
LRT – *Linkelihood Ratio test*  
LSO - loco-específico  
MFFI/MFFNI – Massa fresca folha (Irigado/Não irrigado)  
MFCI/MFCNI - Massa fresca caule (Irigado/Não irrigado)  
MSCI/MSANI - Massa seca caule (Irigado/Não irrigado)  
MSFI/MCFNI - Massa seca folha (Irigado/Não irrigado)  
NSI/NSNI - Número de sementes (Irigado/Não irrigado)  
NVI/NVNI - Número de vagens (Irigado/Não irrigado)  
OPA - *Oligo Pool Assay*  
PCR – *Polimerase Chain Reaction*  
PSI/PSNI - Massa de 100 sementes (Irigado/Não irrigado)  
PSSI/PSSNI – Produtividade (Irigado/Não irrigado)  
PSVI/PSVNI - Massa seca de vagens (Irigado/Não irrigado)

QTL - *Quantitative Trait Loci*  
RAPD - *Random Amplified Polymorphic DNA*  
RFLP - *Restriction Fragment Length Polymorphism*  
RIL - *Recombinant Inbred Line*  
SAM - Seleção Assistida por Marcadores  
SAR – *Sum of Adjacent Recombination Coefficients*  
SMA - Análise por Marcas Individuais  
SNP - *Single Nucleotide Polimorphism*  
SPAD - *Soil-Plant Analysis Development Section*  
SSD - *Single Seed Descent*  
SSR – *Simple Sequence Repeat*  
STS - *sequence-tagged sites*  
SVI/SVNI - Número de sementes por vagem (Irrigado/Não irrigado)  
TII/TINI - Temperatura da folha (Irrigado/Não irrigado)  
UC – “IAC-UNA x CAL-143”  
VRI/VRNI - Volume de raízes (Irrigado/Não irrigado)

## 1. INTRODUÇÃO GERAL

O feijoeiro comum é uma leguminosa anual domesticada a mais de sete milhões de anos, com dois principais centros de origem: Andino e Mesoamericano (Gepts, 1984). Pesquisas em todo o mundo vêm sendo realizadas visando desenvolver novas cultivares mais tolerantes a períodos de estresse, mais produtivas e com resistência a doenças.

Pode ser considerado uma planta sensível ao estresse hídrico, principalmente em virtude da baixa capacidade de recuperação após a deficiência hídrica e sistema radicular pouco desenvolvido (Guimarães, 1992). Estima-se que a seca afeta 60% da área de produção de feijão no mundo (Wortmann et al. 1998; Beebe et al. 2010). Existem áreas importantes de produção como México, América Central, sul da África e o nordeste do Brasil que recebem em média menor número de chuvas (Beebe et al. 2011). A resposta da planta quando submetida a déficit hídrico é diferente dependendo do padrão de seca o qual é especialmente variável nos trópicos, onde os padrões de chuva podem variar amplamente em poucos quilômetros (Beebe et al., 2012). Quatro padrões de seca têm sido definidos: início tardio das chuvas; cessação precoce das chuvas ou seca terminal; seca intermitente, ou pouca chuva durante toda a temporada (Fischer et al., 2003). Genótipos de feijão comum apresentam diferentes reações quando expostos a estresse hídrico. Didonet & Silva (2004) relataram que a cultura do feijoeiro quando submetida ao estresse hídrico, apresenta redução da área foliar e aumento da resistência estomática. A fase reprodutiva é a mais sensível à deficiência hídrica com vulnerabilidade do início da floração até o início da formação de vagens (Fageria et al., 1991), sendo que o período de 5 a 10 dias antes da antese é o mais crítico, podendo levar a uma redução de produtividade superior a 50% no rendimento (Norman et al., 1995).

O desenvolvimento de cultivares com qualidade de grão, com resistência a estresse biótico e abiótico é o principal objetivo dos programas de melhoramento de feijão aqui e no mundo. Enquanto a resposta a estresse biótico tem sido mais bem estudada, a resistência a estresse abiótico é fisiologicamente e geneticamente de natureza mais complexa e geralmente tem herança poligênica, pois pode ser condicionado por múltiplos mecanismos interagindo ao mesmo tempo (Miklas et al., 2006). A resposta da planta a um estresse pode ser condicionada pela presença ou ausência de outros estresses. Sendo assim, a baixa

fertilidade do solo pode reduzir o crescimento da raiz e assim limitar o potencial da planta em expressar a tolerância à seca.

Seca, alta salinidade e frio diminuem o status de água nas células das plantas causando danos, podendo levar numa condição severa a morte da planta. Muitas são as respostas moleculares e mecanismos de transdução de sinal disparados em decorrência de um estresse hídrico.

A principal abordagem destas pesquisas é a utilização de dados genômicos para garantir a eficiência e agilidade dos programas de melhoramento do feijoeiro (Kelly et al., 2003). Uma técnica que tem sido pouco utilizada no melhoramento até agora é a seleção assistida por marcadores moleculares (Benchimol et al., 2005), onde a seleção dos melhores genótipos é feita através da obtenção de marcadores moleculares ligados a características de interesse (*QTL - Quantitative trait loci*) obtidas pela construção de mapas genéticos saturados.

O conhecimento genético bem como da fisiologia vegetal permite associar os marcadores genéticos as respostas específicas ao estresse. Avanços nas tecnologias de marcadores moleculares oferecem métodos poderosos como alternativas para examinar as relações entre estas características. Usando estas técnicas fica claro que mesmo para as características altamente complexas, um pequeno número de QTLs podem explicar grande parte da variabilidade genética (Edwards *et al.*, 1992). Informações de vários mapas genéticos podem ser integrados para facilitar a comparação entre QTLs detectados e importantes genes conhecidos nos mapas genéticos convencionais (Yano *et al.*, 1997). A aplicação de marcadores moleculares para a identificação e exploração de potenciais alelos favoráveis existentes nos genitores das populações, proporciona uma oportunidade em curto prazo para alcançar os avanços necessários no desempenho das culturas ( Tanksley & McCouch, 1997; Miflin, 2000). A plataforma de análise molecular DArT, recentemente desenvolvida para feijão comum, como parte deste projeto (Briñez *et al.*, 2012), compreende muitas vantagens em relação a outros marcadores moleculares existentes e utilizados até agora no mapeamento de características complexas. Outros marcadores moleculares que tem sido amplamente utilizados para estudar polimorfismos entre sequências de DNA, são os microssatélites ou SSRs (*Simple Sequence Repeats*, Tautz & Renz, 1984) e os SNPs (Single Nucleotide Polymorphism, Gupta *et al.*, 2001). O

polimorfismo de um único nucleotídeo ou SNP, é um dos tipos mais comuns de variação genética existente nos genomas, e pelo fato de serem conservados evolutivamente, tem sido utilizado em estudos de identificação de características de interesse (QTL – *Quantitative trait loci*).

Este trabalho teve como objetivo o desenvolvimento da plataforma DArT para feijão comum e sua utilização na genotipagem junto com marcadores de tipo SSR e SNPs para a construção de um mapa genético saturado e a identificação de QTLs de efeito maior envolvidos no controle da tolerância ao estresse hídrico.

## 2. REVISÃO DE LITERATURA

### 2.1 O feijoeiro comum

As leguminosas (Fabaceae) incluem um grupo de espécies de maior importância agrícola, depois dos cereais. Com cerca de 20.000 espécies, as leguminosas são a terceira maior família de plantas superiores (Gepts, 2005). O gênero *Phaseolus* é originário da América e contém 55 espécies, das quais cinco são amplamente cultivadas: *P. vulgaris* L., *P. lunatus* L., *P. coccineus* L., *P. acutifolius* A., Gray var. *latifolius* Freeman e *P. polyanthus* Greenman (Debouck, 1993). Dessas cinco espécies, o feijão comum (*Phaseolus vulgaris* L.) é a mais importante devido ao seu elevado teor de proteínas e é considerada a principal cultura alimentar na África e na América Latina (Yu & Bliss, 1978).

O feijão comum é uma planta diplóide ( $2n = 2x = 22$ ) e autógama, com taxa de fecundação cruzada estimada entre 3 a 5% (Burle et al., 2010). Com base em informações morfológicas, fisiológicas e bioquímicas, Singh et al. (1991) estabeleceram a hipótese de três centros de origem para o feijão (México e na América Central- Mesoamericano, Andino e Colombiano). Maciel et al (2003) sugeriram alguma mistura entre os dois principais *pools* gênicos (Mesoamericano e Andino), pois identificaram a faseolina tipo ‘T’ no grupo mesoamericano, sabidamente encontrada somente em feijões andinos.

Por meio de estudos com marcadores moleculares, Burle et al. (2010) observaram que, apesar do Brasil não ser um centro primário de diversidade do feijoeiro, há uma ampla gama de variabilidade genética da espécie no país que é justificada pelo histórico da domesticação da cultura em épocas de pré e pós colonização européia. Dada a grande variabilidade observada, os autores sugerem que o Brasil deve ser considerado um centro secundário de diversidade da espécie. Evidências obtidas nos últimos anos sugerem que o feijão comum foi domesticado na Mesoamérica e nos Andes, mas seu centro secundário de diversidade genética se estendeu para o Brasil, China e Europa (Santalla et al., 2010). Recentemente, Bitocchi et al. (2011) sugeriram que os *pools* gênicos Mesoamericano e Andino se originaram de diferentes eventos migratórios de populações mesoamericanas com características do México Central.

O feijão comum é um alimento rico sob o aspecto nutricional pelo fato de possuir cerca de 20 a 30% de proteínas (Yin et. al., 2010), 65% de carboidratos (Ribeiro et. al.,



2009), vitaminas do complexo B e minerais (Ca, Fe, Cu, Zn, P, K e Mg) (Armelin et. al., 2007), sendo bom constituinte de alimentos que contém proteína de origem animal. O consumo diário de feijão está entre 50 a 100 g por dia/pessoa, contribuindo com 28% de proteínas e 12% de calorias ingeridas (Sgarbieri, 1980)

Segundo o Instituto Brasileiro de Geografia e Estatística - IBGE (2010), o Brasil é o segundo produtor mundial de feijão do gênero *Phaseolus* e o primeiro na espécie *Phaseolus vulgaris* L.

São realizadas três colheitas anuais, num sistema contínuo de cultivo, de acordo com o zoneamento ecológico das regiões e as épocas de semeadura (safrinha das águas – setembro a novembro, safra da seca ou “safrinha” – janeiro a março, safra de inverno – maio a julho). A produção de feijão no Brasil na temporada 2011/2012 em função das adversidades climáticas ocorridas principalmente na região Nordeste, atingiu 1,24 milhão de toneladas, representando uma redução de 26,2% comparada com a safra passada (CONAB, 2012).

A produtividade média da safra está estimada em 896 kg/ha, apresentando uma redução de 4,2%. A produção consolidada das três safras está estimada em 2,93 milhões de toneladas, representando um decréscimo de 21,4%, face ao observado na temporada anterior.

Os estádios fenológicos da cultura do feijoeiro foram propostos por Gepts & Fernandez (1982) e apresentada por Dourado Neto e Fancelli (2000) conforme a tabela 1.

## **2.2 Respostas do feijoeiro ao estresse hídrico**

As previsões ambientais apontam para o aumento do aquecimento global nas próximas décadas e consequentemente um aumento dos períodos de seca certamente acompanharão esse fenômeno. Assim, o desenvolvimento de cultivares mais tolerantes a períodos de estresse hídrico, bem como novas tecnologias que auxiliem as plantas a tolerar períodos prolongados de estiagem serão essenciais na manutenção da produção agrícola brasileira e mundial em níveis que possam alimentar a população em constante crescimento (Nepomuceno et al., 2001).

Tabela 1. Escala fenológica do feijoeiro comum (*Phaseolus vulgaris* L.). (Dourado Neto & Fancelli, 2000)

<b>Estádio</b>	<b>Descrição</b>
<b>Fase Vegetativa</b>	
V <sub>0</sub>	Germinação/Emergência
V <sub>1</sub>	Cotilédones ao nível do solo
V <sub>2</sub>	Folhas primárias expandidas
V <sub>3</sub>	Primeira folha trifoliada
V <sub>4</sub>	Terceira folha trifoliada
<b>Fase Reprodutiva</b>	
R <sub>5</sub>	Botões florais
R <sub>6</sub>	Abertura da primeira flor
R <sub>7</sub>	Aparecimento das primeiras vagens
R <sub>8</sub>	Primeiras vagens cheias
R <sub>9</sub>	Modificação da cor das vagens (ponto de maturidade fisiológica)

O estresse pode ser definido como um fator externo, que exerce uma influência desvantajosa sobre a planta e está relacionado ao conceito de tolerância a seca, que é a aptidão da planta para enfrentar um ambiente desfavorável (Taiz & Zeiger, 2009). Os autores relatam três mecanismos para tolerância a seca que são: retardo na dessecação ou capacidade de manter a hidratação do tecido; tolerância a dessecação que é a capacidade de desempenhar as funções fisiológicas enquanto desidratada, e escape da seca, que é a capacidade da planta para completar seu ciclo durante a estação úmida, antes do início da seca.

As plantas utilizam três linhas principais de defesa contra a seca, sendo a primeira, a diminuição da área foliar, pois assim ocorre uma menor transpiração, conservando de forma efetiva um suprimento de água limitado no solo por um período mais longo. A segunda linha de defesa é o alongamento das raízes, as quais abrangem camadas mais profundas do solo para resistirem as condições de deficiência hídrica, implicando a alocação de fotoassimilados para os ápices das raízes em crescimento e quando as condições de seca ocorrem no período reprodutivo, há uma competição entre os frutos e o sistema radicular

por fotoassimilados. Quando a restrição hídrica ocorre após a planta obter plena área foliar, a terceira linha de defesa, é o fechamento estomático, que protege a planta contra a desidratação reduzindo a evapotranspiração (Taiz & Zeiger, 2009).

O feijoeiro é considerado uma planta sensível ao estresse hídrico, principalmente em virtude da baixa capacidade de recuperação após a deficiência hídrica e sistema radicular pouco desenvolvido (Guimarães, 1992). De acordo com Moreira et al. (1988), o consumo de água pelo feijoeiro depende do estágio de desenvolvimento, das condições do solo, época de cultivo e das condições climáticas durante todo o ciclo. A necessidade de água do feijoeiro com ciclo de 60 a 120 dias, varia entre 300 a 500 mm para obtenção de alta produtividade (Doorenbos & Kassam, 1979).

Didonet & Silva (2004) relataram que a cultura do feijoeiro, quando submetida ao estresse hídrico, apresenta redução da área foliar e aumento da resistência estomática. A fase reprodutiva é a mais sensível à deficiência hídrica, com vulnerabilidade do início da floração até o início da formação de vagens (Fageria et al., 1991), sendo que o período de 5 a 10 dias antes da antese é o mais crítico, podendo levar a uma redução de produtividade superior a 50% (Norman et al., 1995).

Em avaliações de cultivares para resistência à seca, verificou-se que a produção de grãos diminui à medida que o número de dias de estresse aumenta (Silveira et al., 1981). A redução na produção do feijoeiro sob estresse hídrico decorre da baixa porcentagem de vingamento das flores, quando ocorre floração, e do abortamento de óvulos produzindo vagens chochas (Ramirez-Vallejo & Kelly, 1998). O número de dias em que o feijoeiro fica sujeito à redução de água no solo influencia a redução de produtividade, proporcionando decréscimos de até 20% no rendimento (Magalhães & Millar, 1978). Assim sendo, existe correlação positiva alta entre disponibilidade de água no solo e alguns componentes de rendimento (Aguiar et al., 2008).

A resposta da planta a um estresse pode ser condicionada pela presença ou ausência de outros estresses. Sendo assim, a baixa fertilidade do solo pode reduzir o crescimento da raiz e assim limitar o potencial da planta em expressar a tolerância à seca.

As estratégias de resistência à seca podem variar com o clima e as condições de solo e ainda, algumas plantas possuem adaptações que lhes permitem explorar ambientes mais áridos. Alguns mecanismos de resistência à seca estão divididos entre retardo da

dessecação e tolerância à dessecação, que são referidas como tolerância à seca sob potenciais hídricos alto e baixo, respectivamente, e escape da seca, que engloba as plantas que completam seu ciclo durante a estação úmida, antes do início da seca. A inibição da expansão foliar e o alongamento de raízes são os mecanismos mais sensíveis aos déficits hídricos, além dos quais incluem o aumento da resistência ao fluxo da água, aumento de depósito de cera na folha, fechamento estomático, pela indução do ácido abscísico, que age reduzindo a perda de água pela planta e alteração da dissipação de energia das folhas. O sorgo e a soja, por exemplo, possuem padrões de crescimento indeterminado que permitem que essas espécies sejam beneficiadas com chuvas de ocorrência tardia (Taiz & Zeiger, 2009).

A degradação de proteínas por proteases é fundamental para a resposta da planta a estresses abióticos como a seca. Existe um grande número de genes que codificam proteases em plantas. Apenas alguns dos genes envolvidos na resposta à seca foram caracterizados, e a sua regulação ainda é mal compreendida. Foram identificadas duas novas subtilases das folhas da cultivar Zorin de *Phaseolus vulgaris* L., a PvSLP1 e a PvSLP2, resultados que sugerem um papel específico na resposta à seca e à senescência de feijão comum (Budic et al., 2013).

Hoje em dia tem se investido muito na transformação da proteína DREB “Dehydration Responsive Element Binding Protein” – Proteína de ligação ao elemento responsivo à desidratação; também conhecido como CBF, “C-repeat-binding fator” – fator de ligação a repetições C) que interage especificamente com o elemento DRE, e induz a expressão de genes envolvidos na resposta ao estresse por desidratação (Kasuga et al., 2004). A função de alguns produtos desses genes parece estar relacionada à mecanismos de manutenção das estruturas e funções celulares básicas durante déficit hídrico, baixas temperaturas e alta salinidade (Shinozaki & Yamaguchi-Shinozaki, 1996). Trabalhos científicos (Pellegrineschi, et al., 2002; Kasuga et al., 2004; Qin et al., 2007; Chen et al., 2007) têm demonstrado o potencial da manipulação genética dos genes DREB no desenvolvimento de plantas geneticamente modificadas (PGM) tolerantes à seca, à alta salinidade e ao frio. Entretanto, a expressão de um gene não garante que o produto desse gene específico melhore a habilidade da planta de sobreviver ao estresse (Nepomuceno et al., 2001). Fatores de transcrição desempenham papéis vitais na tradução de sinal no

estresse e na modulação da expressão genica durante o crescimento e desenvolvimento das plantas. Um destes fatores é *OsAP21*, que pertence a família do arroz *AP2/ERF* e subfamília *FSC/DREB*. Plantas que expressam o gene *OsAP21* apresentam crescimento mais forte do que plantas de tipo selvagem sob estresse salino ou de seca (Jin et al., 2013). Outro gene, o *PeDREB2a*, classificado na subfamília do grupo A-5 de *DREB*, quando expressado, as plantas apresentam maior comprimento de raiz, e tolerância à seca e sal (Zhou et al., 2012). Vários outros genes tem sido induzidos e expressos sob estresse em plantas transgênicas como o fator *DREB/CBF* em trigo e cevada que aumenta a tolerância seca e frio; o *TSRF1* que é um fator de transcrição de tomate que melhora a tolerância a seca em arroz e o fator responsivo ERF que aumenta a tolerância a seca e sal em plantas de tabaco (Trujillo et al., 2008; Priyanka et al., 2010; Quan et al., 2010; Morran et al 2011).

### **2.3 Marcadores Moleculares**

As diferenças na sequência de DNA ao longo do genoma de uma determinada espécie são traduzidas pelos marcadores moleculares. Estas variações são chamadas de polimorfismos, sendo originadas por rearranjos cromossômicos, como inserções, duplicações, deleções, mutações, entre outros. O polimorfismo dos marcadores moleculares pode ser herdado de forma Mendeliana e ser acompanhado através das gerações (Varshney et al., 2010).

As diferenças nas sequências de DNA podem ser devidas a um único nucleotídeo (SNPs), a inserção ou deleções (INDELs), ou rearranjos (translocações ou inversões). Apesar destes métodos de genotipagem terem contribuído grandemente para a compreensão da organização genômica e variação genética, eles têm a aquisição de dados limitadas num dado intervalo de tempo. Os métodos de detecção de polimorfismo podem envolver o uso de uma endonuclease de restrição, hibridização de DNA, ou ainda amplificação de DNA (Nayak et al., 2010).

A escolha do tipo de marcador molecular é de grande importância na produção de um mapa genético. O marcador influencia no método a ser adotado de genotipagem da população de mapeamento, e principalmente no tipo de informação gerada no mapa. Marcadores codominantes, que permitem a identificação do heterozigoto, favorecem estudos de interações gênicas, revelando interações aditivas e de dominância. Por outro

lado, marcadores dominantes, que permitem a identificação apenas de homozigotos, revelam somente interações gênicas de aditividade (Liu et al., 1998).

Os marcadores moleculares de DNA são abundantemente utilizados para estimar os níveis de diversidade e heterozigosidade entre diferentes materiais, estimar o tipo de distribuição espacial e temporal de populações em relação ao fluxo gênico (Schuster et al., 2007), para mapeamento de QTLs (Liu et al., 2008), para introgressão gênica (Aerts et al., 2013), entre outras aplicações. Além disso, podem ser utilizados para estimar a proporção relativa de alogamia e autofecundação de uma espécie, assim como determinar seus ancestrais no processo de evolução (Bernardo, 2008).

A influência do número de marcadores genéticos está relacionada até quanto à resolução pré-estabelecida pelo número de genótipos identificados será, de fato, satisfeita. Tanto o grau de cobertura do genoma como a densidade do mapa a serem construídos são funções diretas do número de marcadores (Varshney, 2010).

### **2.3.1 Marcadores Microssatélites (SSRs)**

Marcadores moleculares são ferramentas fundamentais em estudos de variação genética. Avanços nas tecnologias de genética molecular tem gerado numerosos métodos de marcadores ao nível de DNA e alguns dependentes da PCR (Varshney et al., 2010).

Alguns destes marcadores amplamente utilizados incluem SSRs ou microssatélites (Tautz, 1989), AFLPs (Vos et al., 1995), SNPs (Kwok, 2003) e DArTs (Jaccoud et al., 2001).

Os marcadores microssatélites ou SSRs (*Simple Sequence Repeats*), por varias razões práticas tem sido o principal marcador molecular utilizado em aplicações genéticas (Gupta & Varshney, 2000). Estes marcadores são combinações de 1 a 6 bases repetidas em *tandem*. A técnica é baseada na amplificação das seqüências repetidas, utilizando-se *primers* ancorados às regiões flangeadoras das repetições, que são altamente conservadas e únicas. O comprimento do microssatélite varia normalmente entre 15 e 50 pb (Petes et al., 1997). As repetições podem ser simples perfeitas, simples imperfeitas, compostas perfeitas ou compostas imperfeitas (Reddy et al., 2001). Os microssatélites baseiam-se no uso de pares de *primers* na reação de PCR para detectar variações em locos destas seqüências repetitivas. A técnica de SSRs revela polimorfismo em um loco devido a diferenças no

número de vezes em que um motivo se repete naquele loco. O conteúdo genético informativo de um loco SSR é bastante alto, por se tratar de seqüências de alta taxa evolutiva. A sua natureza multialélica e codominante permite o estabelecimento da genotipagem de indivíduos dentro de populações ou entre genótipos relacionados, por isso, a técnica de microssatélites tem sido usada para identificação individual, análise de diversidade, estudos de evolução e de estrutura de populações de espécies relacionadas (Blair et al., 2009). Além destas aplicações, os marcadores microssatélites são amplamente utilizados no mapeamento genético e na análise de QTLs. Muitos microssatélites já foram desenvolvidos para feijão comum (Yu et al., 1999; 2000; Gaitán-Solís, et al., 2002; Métais et al., 2002; Blair et al., 2003; Guerra-Sanz, 2004; Buso et al., 2006, Grisi et al., 2007; Hanai et al., 2007; Benchimol et al., 2007, Campos et al., 2007; Cardoso et al., 2008; Blair et al., 2008). Entretanto, embora o número de marcadores microssatélites tenha aumentado, o polimorfismo revelado em populações de genitores oriundos do mesmo *pool* gênico é restrito, o que dificulta o mapeamento de locos de interesse nas mesmas (Teles, 2007).

Por isso, é muito importante o uso de populações segregantes oriundas do cruzamento de genitores advindos de *pools* gênicos contrastantes como Andino x Mesoamericano. Neste sentido, na população do mapa UC (Campos et.al., 2011) oriunda do cruzamento da variedade IAC-UNA (Mesoamericana) com a linhagem CAL-143 (Andina), foi encontrado 28,5% de polimorfismo entre os genitores.

### **2.3.2 Marcadores SNPs (*Single Nucleotide Polymorphism*)**

Os SNPs, definidos como variações em uma única posição nucleotídica em uma única fita de DNA, compreendem uma nova geração de marcadores (Ganal et al., 2009). Estas variantes ocorrem entre diferentes indivíduos dentro de uma mesma espécie. Em geral, os SNPs são a forma mais comum de polimorfismo presente entre os alelos no DNA. Marcadores SNPs são úteis em uma variedade de aplicações, incluindo a construção de mapas de alta resolução, traços genéticos de mapeamento, diagnóstico genéticos, a análise da estrutura genética de populações e análise filogenética (Rafalski, 2002). Seu aspecto binário e sua estabilidade de geração em geração, tornam os SNPs passíveis de automatização gerando grande volume de resultados, e os tornam atraentes para estudos de

mapeamento de QTLs e de seleção assistida por marcadores em programas de melhoramento de plantas.

Os avanços em robótica e tecnologia tem permitido a utilização dos SNPs em plantas. Colaborações internacionais tem gerado grande quantidade de SNPs e plataformas de genotipagem em arroz ([www.ricesnp.org](http://www.ricesnp.org)) (McNally et al., 2009), milho ([www.maizegenetics.net](http://www.maizegenetics.net)) (Yan et al., 2009; Gore et al., 2009), trigo (<http://wheat.pw.usda.gov/SNP/new/index.shtml>) (Akhunov et al., 2009), cevada (Close et al., 2009) e soja (Hyten et al., 2008). Duas abordagens principais tem sido utilizadas para a descoberta de SNPs, a primeira baseada na disponibilidade de sequencias EST e STR e a segunda no resequenciamento de genes candidatos ou genoma aleatório.

Gaitán-Solís et al. (2008) identificaram SNPs em *P. vulgaris* comparando sequências de regiões expressas e de regiões não-expressas obtidas no GenBank e no DNA genômico do feijão. A análise das sequências foi conduzida utilizando 10 genótipos cultivados e genótipos selvagens, sendo que estes genótipos são de origem Mesoamericana e Andina. Para os 10 genótipos avaliados, um total de 20.964 pb foram analisadas em cada genótipo e comparadas. Esta avaliação resultou na descoberta de 372 novos SNPs em 41 STSs (*sequence-tagged sites*) diferentes.

Cortés et al. (2011) desenvolveram 94 SNPs e testaram estes marcadores em 70 genótipos de feijão comum que têm sido utilizados como genitores de populações de mapeamento e que foram previamente avaliados com marcadores SSRs. Os autores avaliaram 84 regiões gênicas e 10 locos não gênicos usando a tecnologia KASPar (Hiremath et al., 2012). SNPs apresentaram níveis elevados de diversidade genética e um excesso de polimorfismo de frequência média. Este conjunto de marcadores foi útil para distinguir genótipos andino e mesoamericano, mas menos útil para distinção dentro de cada *pool* gênico. Os autores discutem que o marcador genético ideal para fazer o mapeamento, diversidade e estudos de associação em feijão deve considerar uma mistura de marcadores SNPs e marcadores microssatélites.

No estudo de Souza et al. (2012), fragmentos de DNA que abrigam SNPs foram identificados em amplicons de seis genótipos de *P. vulgaris* contrastantes dos *pools* gênicos Andino (Jalo EEP 558, G 19833, e E 277) e Mesoamericano (BAT 93, DOR 364 e Rudá). O total de 677 SNPs foram identificados, incluindo 555 mudanças de uma única base (295



transições e 260 transversões) e 122 pequenas inserções/eliminações de nucleotídeos (indels). Estes SNPs são úteis para análise da diversidade e estudos de microsintenia entre as espécies de leguminosas.

Hyten et al. (2010) descobriram em feijão comum, utilizando uma representação reduzida de biblioteca, o total de 3487 SNPs que continha 2795 sequências genômicas flanqueadoras suficientes para o desenvolvimento de ensaio SNP. Além disso, foi elaborado um ensaio GoldenGate que continha 1.050 SNPs dos 3487 previstos. Um total de 827 dos 1050 SNPs produziram um ensaio GoldenGate de trabalho (79%). Concluíram que através da combinação de duas técnicas de sequenciamento de nova geração, foi desenvolvido um método que permite a descoberta high-throughput de SNPs em qualquer organismo diplóide, sem a necessidade da sequência de genoma completo ou da criação de bibliotecas de cDNA normalizadas.

### **2.3.3 Marcadores moleculares: DArTs (*Diversity Arrays Tecchnology*)**

A tecnologia do DArT (“Diversity Arrays Tecchnology”, Jaccoud *et al.*, 2001) é baseada em hibridização em uma plataforma que utiliza a tecnologia de microarranjos para análise de polimorfismo. É um método muito eficiente em termos da relação custo/benefício. Um único ensaio de DArT simultaneamente genotipa centenas a milhares de SNPs e INDELs ao longo do genoma de uma espécie. O procedimento do DArT envolve basicamente cinco passos: (1) construção de uma biblioteca genômica (representação genômica); (2) impressão da biblioteca genômica nos microarranjos; (3) marcação das representações genômicas; (4) hibridização da representação genômica marcada sobre o microarranjo seguida de lavagens e (5) registro (por *scanner*) e análise dos dados. O DArT compreende muitas vantagens em relação a outros marcadores moleculares existentes:

(1) DArT é capaz de realizar análises em paralelo ao invés de séries por dado de marcador: muitas tecnologias de marcadores são dependentes de eletroforese em gel, resultando em baixo desempenho. Em poliacrilamida, entre 50 a 150 fragmentos podem ser amostrados por vez, enquanto num arranjo altas densidades (+ de 10.000) dados de marcadores podem ser obtidas; (2) DArT não requer informação prévia de sequência: muitos métodos (SSRs, SNPs) requerem informação de sequência *a priori*. DArT é independente de investimento em sequenciamento e é particularmente interessante para espécies com poucos recursos ou

sem outras tecnologias disponíveis; (3) Marcadores DArT são genotipados com alta eficácia: um *software* específico (DArTsoft) analisa uma grande quantidade de dados gerados para cada experimento de DArT. O *software* analisa as imagens de microarranjos e subsequentemente identifica e genotipa os marcadores (Wenzl *et al.*, 2004). Os parâmetros deste programa para qualidade são estabelecidos objetivamente pelo usuário que pode selecionar um conjunto de marcadores com alta qualidade e reprodutibilidade; (4) DArT é uma plataforma aberta: DArT Pty/Ltd. tem estabelecido uma rede de usuários de DArT ([www.diversityarrays.com/dartnetwork.html](http://www.diversityarrays.com/dartnetwork.html)), para desenvolver e melhorar ainda mais esta tecnologia; (5) A plataforma DArT permite flexibilidade de aplicações: as bibliotecas de DArT são preparadas com genomas individuais ou misturados (por exemplo: metagenomas) de indivíduos que melhor se ajustem a uma desejada aplicação. Para estudos de mapeamento, geralmente se utilizam os genitores da população segregante, já para estudos de diversidade o DNA pode ser derivado de variedades cultivadas e/ou materiais selvagens. A plataforma de microarranjo é também flexível. Em experimentos iniciais, marcadores podem ser identificados na biblioteca genômica (arranjos de descoberta). Estes marcadores podem ser rearranjados em novos slides (arranjos de genotipagem) e servem para alto fluxo de detecção de centenas a milhares de marcadores em amplas populações.

A tecnologia DArT tem demonstrado ser eficiente em estudos de diversidade genética em coleções de germoplasma além de ser marcadores adequados em programas de melhoramento molecular (Kilian *et al.*, 2005). A tecnologia foi desenvolvida primeiramente para arroz (Xie, 2006), mandioca (Xia *et al.*, 2005), cevada (Wenzl *et al.*, 2004), trigo (Akbari *et al.*, 2006, Semagn *et al.*, 2006), ervilha de pombo (Yang *et al.*, 2006), eucalipto (Lezar *et al.*, 2004), sambambaia e musgo (James *et al.*, 2006), sorgo, banana, coco (<http://www.generationcp.org>), e *Arabidopsis* (Wittenberg *et al.*, 2005). A plataforma tem se expandido para muitas outras espécies incluindo o feijão comum (Tabela 2).

Tabela 2. Leguminosas para as quais a plataforma DArT foi desenvolvida

<b>Cultura</b>	<b>Genótipos utilizados para desenvolver as Bibliotecas</b>	<b>Arranjos</b>	<b>Painel de diversidade</b>	<b>Marcas Polimórficas</b>	<b>Marcas no mapa genético</b>
Feijão comum	150	>22,000	150	2,500	NA*
Grao de bico	250	21,500	300	5,400	>700
Amendoim	164	15,360	300	>5,000	~200
Guandu	300	29,000	400	>5,000	>400

NA\*: Não disponível

Fonte: DArT Pty Ltd., Australia ([www.diversityarrays.com](http://www.diversityarrays.com)).

Uma das vantagens dos marcadores DArTs é que estes podem ser sequenciados imediatamente, o que pode ser uma vantagem para integração de mapas genéticos e físicos através da ancoragem de BACs, por exemplo. Esta integração já foi testada com sucesso para *Triticale* (Alheit et al., 2011). Em cana de açúcar, uma cultura de difícil genotipagem com outros marcadores moleculares devido a diferentes ploidias (alo, auto, aneuplóides), com número de cromossômico elevado (80 a 120), marcadores DArTs se mostraram muito adequados e com herança Mendeliana. A plataforma DArT também foi adequada para análise em BSA (*Bulk Segregant Analysis*) de características quantitativas em cevada e em trigo (Wenzl et al., 2007). Neste último estudo, DArTs foram usados em trigo para fazer BSA (*Bulk Segregant Analysis*) quantitativo, se mostrando adequado e identificando locos associados com tolerância a alumínio com precisão menor a 0,8cM, independente da ploidia das plantas.

De acordo com Wenzl et al. (2006), um mapa consenso para cevada pôde ser construído associando marcadores SSRs, RFLPs, STS e DArTs, compreendendo 2.935 locos, dos quais 2.085 foram de DArTs e 850 foram de outros marcadores. A cobertura do mapa foi de 1.161 cM. Mais de 98% do mapa foi coberto por um único ensaio de DArT. Na média,  $14 \pm 9$  locos de DArT foram identificados entre 5cM em ambos os lados de SSR, RFLP e STS precisamente identificados ladeando regiões ligadas a caracteres de interesse agronômicos.

Recentemente, outros mapas tem sido construídos utilizando os marcadores DArTs como o descrito por Raman et al. (2012) onde foi construído um mapa de ligação para nabo (*Brassica napus* L.) inserindo 437 DArT, 135 SSR, e 6 marcadores gênicos com um comprimento total de 2.288 cM. Petroli et al. (2012) estabeleceram um mapa para eucalipto

com 2.274 marcadores DArTs ancorados a 210 microssatélites. Lu et al. (2013) desenvolveram a tecnologia DArT para tabaco e construíram um mapa de ligação com 238 DArTs e 613 SSR. Marone et al. (2012) estudaram as características genéticas e funcionais dos marcadores DArTs em trigo analisando as sequências derivadas da tecnologia DArT. As análises das posições genéticas dos marcadores correspondentes a sequências quase idênticas indicaram que a redundância de sequências é um dos fatores que explica a aglomeração desses marcadores em regiões específicas do genoma. Além disso, os autores concluíram que a atribuição da função de genes putativos, torna estes marcadores em instrumento ideal para estudos de sintenia e ou para a identificação de genes candidatos.

A tecnologia DArT foi desenvolvida como uma alternativa baseada na hibridação, frente a maioria das tecnologias que atualmente utilizam os géis para a identificação de marcadores polimórficos. Os marcadores DArTs tem sido utilizados em estudos de diversidade em bancos de germoplasma (Yang et al., 2006), mapeamento genético (Hippolyte et al., 2010) e na identificação de genes (Grewal et al., 2008), e podem ser úteis na seleção assistida por marcadores (SAM), especialmente no monitoramento na recuperação do genoma do parental recorrente (Varshney et al., 2010).

## **2.4 Populações de Mapeamento**

A seleção da população para mapeamento envolve a escolha de genitores e a determinação do tipo de cruzamento (F2, retrocruzamento, etc), sendo considerada uma etapa crítica para o sucesso da construção do mapa (Staub et al., 1996).

Um aspecto que deve ser considerado para o desenvolvimento de um mapa genético é a distância genética que separa as linhagens que darão origem à população segregante: quanto mais próximas, mais difícil a obtenção de polimorfismos. Assim, a escolha de linhagens sem ancestral comum ou mesmo cruzamentos interespecíficos para espécies de autofecundação, têm sido utilizados em maior ou menor escala para permitir ou aperfeiçoar o processo de análise de marcadores (Bonierbale et al., 1988; Tanksley et al., 1988).

De modo geral, níveis moderados de polimorfismos podem ser detectados entre *pools* gênicos mesoamericanos e andinos do feijão comum (Chase et al., 1991; Vallejos et al., 1992; Nodari et al., 1993). No trabalho de Vasconcelos et al. (1996), a diversidade genética de 28 cultivares de feijão comum foi analisada utilizando marcadores moleculares

RAPD. Com base nas distâncias genéticas, os cultivares foram separados em dois grandes grupos: o grupo A abrangendo cultivares do Centro de Domesticação Andino, e o grupo B, cultivares do Centro Mesoamericano.

A população segregante utilizada em plantas autógamas para o mapeamento é tradicionalmente derivada do cruzamento entre duas linhagens puras, o qual origina a geração F1 essencialmente heterozigota em todos os locos. A F1 é autofecundada ou retrocruzada de forma a produzir uma geração F2 ou RC1 segregante tanto para os marcadores moleculares como para as características fenotípicas de interesse (Ferreira & Grattapaglia, 1995).

Uma população de linhas endogâmicas recombinantes (*Recombinant Inbred Lines* – RILs) pode ser utilizada alternativamente, derivada de uma F2 por meio de sucessivas autofecundações ou uma população de linhagens di-haplóides obtidas a partir da cultura de anteras ou de micrósporos de planta F1 (Ferreira, 1993).

A população de RILs é obtida a partir de uma população F2, por meio de sucessivas autofecundações pelo método do descendente de uma única semente ou SSD (*Single Seed Descent*). As sucessivas autofecundações permitem que ao longo do processo ocorra a fixação gênica, ou seja, a eliminação de locos em heterozigose. Ao final do processo, todos os genes terão uma segregação de 1:1 (AA, aa). Sendo assim, é possível avaliar interação genótipo x ambiente realizando repetições dos experimentos. Permite, ainda, o estudo de diferentes características em uma mesma população. No entanto, como essa população permite identificar apenas dois genótipos, impede a identificação de relações de dominância entre os genes (Lynch & Walsh, 1998). Uma desvantagem da utilização de RILs no mapeamento é o tempo requerido para a obtenção da população segregante (6 a 8 gerações); entretanto, o desenvolvimento de RILs em plantas autógamas, como o feijoeiro comum, é extremamente fácil, uma vez que estas plantas reproduzem-se normalmente por autofecundação (Faleiro, 2000).

## **2.5 Construção de mapas genéticos**

O entendimento do funcionamento de genomas, com o profundo conhecimento dos genes que os compõe e as mais diversas interações entre eles, são o grande desafio deste século (Flint-Garcia, 2003). Os estudos direcionados ao entendimento de como genes e

alelos controlam características fenotípicas complexas (QTLs) têm sido de grande importância nesta busca. Uma ferramenta bastante usada, nesse contexto, é a obtenção de mapas genéticos.

O desenvolvimento de mapas genéticos é considerado uma das aplicações de maior impacto da tecnologia de marcadores moleculares na análise genética de espécies e, potencialmente, no melhoramento de plantas. Os mapas genéticos possibilitam a cobertura e análise completa de genomas, a decomposição de características genéticas complexas nos seus componentes mendelianos, a localização de regiões genômicas que controlam caracteres de importância, a quantificação do efeito dessas regiões na característica estudada e a canalização de toda essa informação para uso em programas de melhoramento (Faleiro et al., 2003).

O ponto central da análise de ligação entre as marcas no genoma para a construção e saturação de mapas genéticos é a existência de desequilíbrio de ligação entre estas marcas. O desequilíbrio de ligação é decorrente do processo de recombinação homóloga que ocorre entre cromátides irmãs durante a meiose, em geral de forma randômica ao longo dos cromossomos (Liu, 1998; Carneiro & Vieira, 2002). O desequilíbrio de ligação é determinado pela fração de recombinação ( $r$ ) entre as marcas, que é a razão entre as quantidades de gametas recombinantes e o total de gametas. Assim, se ' $r$ ' for menor que 0,5, quer dizer que as marcas estão em desequilíbrio. Desta forma, podemos inferir que as marcas analisadas estão ligadas. É desta relação entre ' $r$ ' e ligação física que surgem os mapas genéticos (Liu, 1998).

A construção de mapas de ligação tem como base a análise de segregação de centenas de marcadores e, por isso, é computadorizada. Programas como Mapmaker (Lander et al., 1987), Linkage 1 (Suiter et al., 1983), Gmendel (Liu e Knapp, 1992), JoinMap (Stam, 1993), Multimap (Matise et al., 1994), Outmap (Butcher et al., 2002), AntMap (Iwata & Ninomiya, 2006) e OneMap (Margarido et al., 2007) foram desenvolvidos para auxiliar na análise genética dos dados visando à construção de mapas genéticos e estão disponíveis na internet (<http://linkage.rockefeller.edu>).

Tais programas indicam os marcadores ligados, formando os respectivos cromossomo; qual a ordem correta entre essas marcas; e a distância em cM (centiMorgans) entre elas. Para isso, devem-se seguir os seguintes passos: (1) verificar o padrão de

segregação Mendeliana esperada para o tipo de população de mapeamento escolhido, chamado de teste de segregação na proporção esperada; (2) realizar o teste de ligação, verificando a frequência de recombinação entre cada par de marcas, o que possibilita identificar os desequilíbrios de ligação e consequentemente formar os cromossomo, (3) empreender a ordenação das marcas a partir das distâncias entre elas e (4) fazer a conversão das distâncias em centiMorgans (Carneiro & Vieira, 2002).

O teste de segregação mais comumente utilizado é o teste estatístico Qui-quadrado ( $\chi^2$ ), dado pela fórmula  $\chi^2 = \frac{(fo-fe)^2}{fe}$  em que: fo = frequência observada em cada uma das classes; fe = frequência esperada em cada uma das classes (proporção Mendeliana do modelo genético adotado). A hipótese ( $H_0$ ) não admite a ocorrência de distorção, ou seja, os dados genotípicos devem seguir a distribuição Mendeliana esperada para a população de mapeamento adotado. Uma forma de verificar a significância ( $\alpha$ ) estatística do  $\chi^2$  é convertê-lo em seu respectivo *p-valor*, que pode ser definido como a probabilidade de se observar uma amostra fora do padrão, assumindo que a hipótese nula é verdadeira. Quanto menor o valor do *p-valor*, mais forte é a evidência de rejeição de  $H_0$  (Liu, 1998).

O fato de realizar o  $\chi^2$  repetidas vezes, uma para cada marca, gera um problema intrínseco deste tipo de análise: erros ditos do tipo I (Benjamini & Hochberg, 1995), ou seja, falsos positivos (Liu, 1998).

Deste modo, usa-se uma correção que elimina os erros do tipo I, baseada em valores de significância dos testes de segregação derivados de cada marca. A mais usada é a correção de Bonferroni (Lynch & Walsh, 1998), que assume uma significância individual ( $\alpha$ ), a partir da razão da significância total pelo número de testes:  $\alpha/n$ , em que  $\alpha$  = nível de significância global e  $n$  = número de testes independentes, que correspondem ao número de marcadores utilizados.

As causas da distorção no desvio de segregação de locos ao longo do genoma são conhecidas, trabalhos recentes mostram a importância desses locos na identificação de QTL a partir de mapas genéticos. Xu (2008) estabelece uma teoria para o surgimento de locos com desvio de segregação, e afirma que o poder de detecção de QTL pode ser sutilmente reduzido quando se faz as análises ignorando os locos com desvio, mas também ressalta que se o mapa for bem saturado, e tais locos apresentarem distribuição randômica no genoma, as perdas são insignificantes. Finalmente, o autor defende que estes locos com

desvio de segregação sejam utilizados na obtenção de mapa genéticos e identificação de QTL.

A mensuração da frequência de gametas recombinantes entre dois locos constitui o procedimento a partir do qual se realiza o mapeamento (Staub et al., 1996). Diversos tratamentos estatísticos podem ser usados para estimar a frequência de recombinantes, entre eles, o método dos momentos, o método dos quadrados mínimos e o da máxima verossimilhança. De modo geral, pelo método dos momentos, obtêm-se os parâmetros igualando-se os valores observados às suas expectativas, em termos de esperanças matemáticas. Pelo método dos quadrados mínimos, obtêm-se as estimativas para os parâmetros de modo que a soma de quadrados dos desvios entre os valores observados e os estimados seja mínima, enquanto pelo método de máxima verossimilhança obtêm-se as estimativas de modo que os valores observados representem o caso mais provável de ocorrência (Weir, 1996; Liu, 1998).

Estatisticamente, ligação genética é definida como a associação de genes ligados em um mesmo cromossomo, ou seja, que segregam juntos durante a meiose. Assim, um grupo de ligação (GL) é formado por marcadores genéticos que possuam menos de 50% de recombinação ( $r$ ) com pelo menos um dos marcadores do grupo (Liu, 1998).

Uma vez conhecidas as frações de recombinação entre cada par de marcadores e já discriminados os cromossomos, deve-se determinar a melhor ordem para os marcadores dentro de cada grupo. Weir (1996) sugere a ordenação de três locos, A, B, e C, feita pela soma dos coeficientes de recombinação adjacentes (*sar* – *sum of adjacent recombination coefficients*). Esse processo propõe que a ordem correta entre os locos seja aquela que possibilite o mínimo valor de *sar*. O valor de *sar* sob a ordem incorreta será sempre maior que a correta. Esse método pode ser empregado para qualquer número de locos.

Estabelecidas as frações de recombinação entre cada marcador, estes valores devem ser convertidos em centiMorgans, unidade usada para distância entre locos em mapas genéticos, em homenagem a Thomas Hunt Morgan que estabeleceu a relação em frequência de recombinação e distância entre genes (Griffiths et al., 2006).

A fim de estabelecer critérios para inferir se dois locos estão, ou não, ligados é necessário definir a frequência máxima de recombinação e o LOD score mínimo. LOD score (*Logarithm of Odds* - Lynch & Walsh, 1998) é um teste baseado na razão de



verossimilhança, que testa a hipótese de dois locos estarem ligados, e utiliza o logaritmo na base 10 para facilitar o entendimento. Assim sendo, um LOD score com valor de três indica que a ocorrência de ligação é mil vezes mais provável do que a segregação independente (Carneiro & Vieira, 2002).

O teste que realiza esta análise é chamado de teste de Dois-Pontos. Para este teste a ligação é detectada testando a independência entre cada par de marcas na população segregante. A melhor forma de se estabelecer a ordem dos locos é por meio da “Busca Exaustiva”, que como o nome sugere, verifica todas as ordens possíveis e define como melhor aquela de menor comprimento. No entanto, este método apresenta uma limitação computacional, pois há um grande número de possíveis ordens quando um GL possui um grande número de locos. Assim, os métodos computacionais e/ou algorítmicos, para ordenação exata das marcas, são limitados a um pequeno número de locos por GL. Alternativa muito usual para GL grandes é o teste de Três-Pontos, baseada na ideia central do teste de Busca Exaustiva, porém ao invés de realizar os cálculos com todos os locos, dois a dois, a ordem é formada em pequenos grupos de três em três. Ou seja, o teste é feito a cada três marcas em sequencia, com sobreposição, montando-se assim a ordem final (Liu, 1998).

Embora a frequência de recombinantes indique a medida de distância entre dois locos ao longo de determinado cromossomo, o ideal é que as distâncias usadas na construção de mapas genéticos sejam totalmente aditivas, de modo que, ao atribuir novos locos (marcadores) ao mapa, não seja necessário ajustar as distâncias já estabelecidas. Entretanto, as frequências de recombinação não são aditivas; são, portanto, inadequadas como medida direta de distâncias entre dois locos a serem mapeados (Staub et al., 1996; Lynch & Walsh, 1998).

Assim, as funções de mapeamento transformam as frequências de recombinação em distâncias aditivas. A função de Haldane admite que a ocorrência de permuta se dê de modo independente (ausência de interferência), seguindo uma distribuição de Poisson (Haldane, 1919). A função de Kosambi (Kosambi, 1944) admite a ocorrência de permutas próximas como eventos não independentes. Esta função admite interferência completa entre regiões arbitrariamente próximas, que decresce com o aumento da distância, sendo nula para locos independentes (Schuster & Cruz, 2008; Hanai, 2008). Desta forma, torna-se mais

apropriado o uso desta função de mapeamento quando se trabalha com marcas distribuídas ao longo do genoma, como é o caso de marcadores moleculares.

## **2.6 Identificação de QTL**

Os caracteres quantitativos são aqueles que não permitem a classificação dos indivíduos em classes distintas, ou seja, não há uma diferenciação marcante entre os tipos segregantes. Esses caracteres são determinados por muitos genes (poligênicos), o efeito de cada alelo sobre o caráter é pequeno e, normalmente, o efeito do ambiente sobre o mesmo é grande, o que faz com que a herdabilidade destes seja pequena (Allard, 1999).

Um caráter poligênico é um sistema de locos no qual a segregação contribui para a variação genética associada à variação de uma variável contínua (Thoday, 1977). QTL (loco de um caráter quantitativo) é a denominação sugerida para qualquer loco que tem contribuído para a variância do caráter (Gelderman, 1975).

Com o desenvolvimento dos marcadores moleculares foi proporcionada a construção de mapas genéticos de ligação, os quais permitiram mapear e estimar a magnitude dos efeitos dos QTLs (Paterson et al., 1988; Paterson et al., 1991).

Uma vez detectados os marcadores ligados a uma determinada característica de interesse, é possível selecionar os indivíduos com base no fenótipo do marcador, sem que haja a necessidade de avaliar o fenótipo da característica. Essa abordagem é conhecida como seleção assistida por marcadores (SAM). É importante, principalmente, nos casos em que a característica de interesse é de avaliação difícil e cara, nos estudos de espécies perenes de ciclo longo, e para características fenotípicas de herdabilidade baixa, o que pode significar redução no tempo e nos custos de programas de melhoramento (Xu et al., 2012).

No entanto, os dados obtidos tanto do mapa genético quanto das análises fenotípicas devem ser avaliados a fim de fornecer as posições, a magnitude dos efeitos de cada região genômica e indicar as interações gênicas entre elas. Dentre os programas mais utilizados para este fim estão QTLCartographer vs.2.5 – WinQTLCart (Wang et al., 2005) e QGENE 3.05 (Nelson, 1997).

Inicialmente, a identificação dos QTLs era feita por uma relação direta entre o genótipo de marcadores dos indivíduos da população de mapeamento com suas características fenotípicas. Este método é denominado “Análise por Marcas Individuais”

(SMA), podendo ser baseado no método de máxima verossimilhança, ou ainda em análises de regressão linear. Como limitação, há o fato de quanto maior for a distância entre o QTL e o marcador, menor será a probabilidade de detectá-lo em vista da possibilidade de recombinação entre o gene marcador e o QTL. Além disso, a magnitude do QTL, pelo mesmo motivo, é normalmente subestimada. Outro inconveniente é que não é possível determinar a posição do QTL no genoma (Lynch & Walsh, 1998; Schuster & Cruz, 2008).

Como alternativa para aumentar o poder de detecção das associações, permitindo estimar o efeito e a posição dos QTLs, o método de mapeamento com base em intervalo (*interval mapping*) foi proposto por Lander & Botstein (1989). Esse método analisa a existência de QTL a cada intervalo entre dois marcadores, ao invés de a cada único marcador. Uma desvantagem atribuída ao mapeamento fundamentado em intervalo é a de que os outros QTLs fora do intervalo são ignorados, o que pode resultar na diminuição da precisão das estimativas e do poder dos testes.

O mapeamento por intervalo composto foi proposto para solucionar essas questões. Trata-se de um método em que se consideram os QTLs fora do intervalo em questão por um modelo de regressão linear múltipla. Além de proporcionar maior resolução no mapeamento, permite o uso de dados obtidos em vários ambientes, para avaliar os efeitos da interação QTLs vs. ambientes (Jansen, 1993; Zeng, 1993, 1994).

O método mais recente é o mapeamento por intervalo múltiplo (Kao et al., 1999) mas ainda não muito utilizado na maioria dos trabalhos, porque incorpora parâmetros de epistasia ao modelo e por isso apresenta algumas vantagens potenciais como maior eficiência e precisão na identificação de QTLs; seus efeitos são estimados sem viés e, ao contrário de outros modelos, pode aumentar a eficiência da seleção assistida, considerando os efeitos epistáticos.

Outra importante aplicação dos mapas é o mapeamento comparativo (*comparative mapping* ou *synteny mapping*) que se constitui na comparação das estruturas genômicas de diferentes espécies, do ponto de vista de homologia de genes, conservação de distâncias e da ordem de ligação nos cromossomos. Esta comparação contribui para o entendimento sobre a evolução dos genomas (Tanksley et al., 1988b).

### **2.6.1 Mapas de Ligação em Feijoeiro comum e mapeamento de locos associados à tolerância ao déficit hídrico**

Os trabalhos de mapeamento genético do feijoeiro-comum tiveram início com o primeiro mapa de ligação publicado por Lamprecht (1961). Outros mapas de ligação têm sido desenvolvidos para esta cultura (Basset, 1991; Vallejos et al., 1992; Nodari et al. 1993; Adam-Blondom, 1994; Freyre et al., 1998; Ariyaratne et al., 1999; Blair et al., 2003; Grisi et al., 2007; Blair et al., 2008; Blair et al., 2010; Campos et al., 2011; Asfaw & Blair, 2012). Os mapas de Vallejos et al. (1992), Gepts et al. (1993), Nodari et al. (1993) e Adam-Blondom et al. (1994) foram baseados primariamente em marcadores RFLPs. Freyre et al. (1998) descreveram o primeiro mapa consenso para feijão comum (*Core Map*), no qual 550 marcadores moleculares do tipo RFLP e RAPD foram mapeados na população BAT93 x Jalo EEP558 (BJ). Porém, nesta população existe um reduzido número de linhagens endogâmicas recombinantes (RILs), o que acaba limitando a chance de encontrar recombinações e, portanto, a robustez do mapa genético é diminuída. Algumas regiões não apresentaram saturação, como por exemplo, os grupos GL2, GL3, GL8 e GL10, havendo discrepâncias com mapas anteriores. Assim, tem-se buscado a integração dos novos marcadores moleculares a este mapa consenso (Yu et al., 2000; Blair et al., 2003; Blair et al., 2006). O mapa de Grisi et al. (2007) adicionou novos marcadores SSRs ao *Core Map* de feijão, totalizando 199 marcas SSR em 13 cromossomo, com 1.358 cM e distância mínima de 7,23 cM entre duas marcas. Os autores também obtiveram o primeiro mapa de feijão exclusivamente de microssatélites, com 106 marcas distribuídas em 606,8 cM, totalizando 12 cromossomo.

Blair et al. (2008) mapearam 45 novos locos de microssatélites na população DOR364 x G19833 e houve uma distribuição preferencial dos microssatélites ATA nos grupos GL02, GL09 e GL11.

Hanai (2008) também desenvolveu dois mapas para feijão comum, para a população CFM (população gerada a partir do cruzamento entre os genitores Carioca e Flor de Mayo) e BJ (população derivada do cruzamento entre 'Bat 93' e 'Jalo EEP558'), os quais foram alinhados. Onze marcadores SSR-EST e outros 9 SSR genômicos foram mapeados tanto numa população quanto na outra. O mapeamento comparativo dos mapas das duas

populações permitiu identificar, no mapa CFM, oito dos 11 cromossomo do mapa consenso BJ e determinar a orientação de cinco deles.

Cordoba et al. (2010) realizaram a integração do mapa físico e genético de feijão comum, cuja importância está na capacidade de localizar fisicamente *loci* que são conhecidos por serem polimórficos.

Blair et al. (2010) mapearam com marcadores AFLP, RAPD e SSR a população DOR364 e BAT477 e foram encontrados 49 QTLs para rendimento e componentes de rendimento sob condições de estresse hídrico e de irrigação.

Campos et al. (2011) mapearam a população UC oriunda do cruzamento da variedade IAC-UNA com a linhagem CAL-143 com 198 marcadores microssatélites. O comprimento total do mapa foi de 1.865,9 cM, com um comprimento médio de ligação entre os 11 grupos mapeados de 170,5 cM. Este mapa foi gerado pelos esforços do programa de melhoramento do feijoeiro do IAC e é um dos mais robustos em literatura pelo tamanho amostral da população de mapeamento (384 RILs), bem como pelo número de SSRs posicionados. A partir do mapa UC, foi possível o mapeamento com alta precisão de locos de resistência à mancha angular (Oblessuc, 2012) e à antracnose (Baroni, 2010).

### **3. OBJETIVO**

#### **3.1 Objetivo Geral**

Desenvolver a plataforma DArTs para feijão comum e construir um mapa de ligação a partir da população AS (AND 277 x SEA 5) com perfil contrastante para tolerância a seca, de tal forma a localizar os QTLs de efeito maior envolvidos no controle desta característica.

#### **3.2 Objetivos específicos**

- Desenvolver a plataforma DArT para feijão comum usando um painel de diversidade de 96 genótipos contrastantes;
- Genotipar 105 linhagens na geração  $F_8$  da população AND 277 x SEA 5 utilizando marcadores DArTs, SSRs e SNPs;
- Avaliar fenotipicamente as linhagens  $F_8$  da população AS quanto ao desenvolvimento das raízes, componentes fisiológicos, assim como também a avaliação da produtividade em experimentos contrastantes em condição irrigada e não irrigada;
- Identificar o número e a localização cromossômica bem como o tipo e magnitude dos efeitos dos QTLs envolvidos na resposta ao estresse hídrico;
- Identificar possíveis ligações e efeitos pleiotrópicos entre os locos mapeados com os perfis fenotípicos de tolerância a seca.

## 4. MATERIAL E MÉTODOS

### 4.1 Material Vegetal

A população segregante composta por 105 linhagens endogâmicas recombinantes (RILs) de feijoeiro comum foi originada no CIAT (Centro Internacional de Agricultura Tropical, Cali, Colômbia) a partir do cruzamento entre a cultivar AND 277 e a cultivar SEA 5. A população foi avançada até a geração F<sub>8</sub> utilizando o método do descendente de uma única semente ou SSD (*Single Seed Descent*). A cultivar SEA 5 é uma linhagem avançada do CIAT selecionada para tolerância à seca, por isto a denominação SE A, de série de “Seca Avançada”, e é considerada uma linhagem superior do BAT 477 em tolerância a seca (Asfaw & Blair, 2011). A tolerância a seca do SEA 5 pode ser atribuída principalmente pela capacidade que este genótipo possui de concentrar os assimilados na produção de grãos (Rao, 2001). A cultivar foi derivada a partir do cruzamento inicial (BAT 477 x SanCristobal) x (Guanajuato 31 x Rio Tibagi), tem hábito de crescimento indeterminado tipo III, tamanho pequeno e sementes de cor creme (Singh et al, 2001; Szilagyi et al, 2011). A cultivar AND 277 baseia-se no cruzamento de [Cargabello x (Pompador Checa x Línea 17) x (Línea 17 x Red Kloud)], e apresenta o alelo *Co-1*<sup>4</sup> e o gene *Phg-1* que conferem resistência à antracnose (*Colletotrichum lindemuthianum*) e mancha angular (*Pseudocercospora griseola*), respectivamente (Gonçalves-Vidigal et al., 2011), mas é suscetível à seca. SEA 5 é de origem mesoamericano e AND 277 de origem Andina.

### 4.2 Avaliações quanto às Características Relacionadas ao Estresse Hídrico

#### 4.2.1 Fenotipagem em plantas cultivadas em vasos

O experimento foi realizado nos meses de Janeiro a Abril de 2012 no Instituto Agrônomo (IAC, Campinas, São Paulo, Brazil), com coordenadas geográficas 22°52'40'' latitude sul, 47°04'72'' longitude oeste e 685 m de altitude. Duas casas de vegetação foram construídas para o experimento, a primeira foi revestida com tela de sombra e plástico de polietileno para o tratamento não irrigado e a outra só com tela de sombra para o tratamento irrigado ou controle. Em cada casa de vegetação tinha 428 vasos com capacidade para 14 litros adaptados a um sistema de irrigação individual permitindo o controle da água nas

casas de vegetação. Os vasos foram pre-enchidos com uma mistura de solo, esterco e areia na proporção 3:1:1. O solo foi analisado num laboratório comercial e fertilizado com 1,5 g de N, 6,0 g de  $P_2O_5$  e 3,5 g de  $K_2O$  por vaso. A irrigação foi fornecida por um sistema automatizado de gotejamento até o momento da imposição do déficit hídrico, bem como no período posterior a sua aplicação, duas vezes por dia num volume total de 400 ml de água com vazão de 200ml. O potencial matricial do solo ( $\psi_m$ ) foi determinado utilizando sensores instalados a 0,20 m de profundidade, com as leituras sendo realizadas pelo medidor Watermark, onde valores determinados próximos a -40 KPa de acordo com o manual de instruções são considerados satisfatórios para plantas mantidas em condições irrigadas. Valores próximos a -199 KPa, indicam escassez total de água na profundidade analisada. Assim sendo, as avaliações no tratamento com imposição do déficit hídrico na fase vegetativa foram em função da determinação do  $\psi_m$ . Um total de 30 sensores foram instalados aleatoriamente nas duas casas de vegetação com o objetivo de medir o potencial matricial da água no solo nos dois tratamentos.

As sementes foram lavadas durante um minuto em solução de 5% (v/v) de hipoclorito de sódio, lavadas com água destilada e germinadas em incubadora (BOD) por 72 horas a 25 °C. Três sementes de cada genótipo foram plantadas por vaso em um delineamento experimental inteiramente casualizado, com quatro repetições. Após emergência, duas plantas foram preservadas por vaso, uma para análise destrutiva e outra para análise de produtividade. A cada dois dias, o potencial matricial do solo foi medido, além da temperatura das folhas que continham os sensores, a temperatura ambiental e umidade relativa do ar de cada ambiente. A casa de vegetação (controle) foi regada a 80% da capacidade de campo durante todo o experimento com uma temperatura ambiental média de 34 °C e 52% de umidade relativa do ar aferidos por um termo-higrômetro. O estresse aplicado na casa de vegetação não irrigada (estresse hídrico) foi estresse terminal na fase vegetativa (V3/V4) (Tabela 1) e apresentou temperatura média de 36,4 °C e umidade relativa de 42,4%. Dias para o florescimento foram registrados todos os dias e a média das plantas floridas foi calculada por genótipo.

Após 20 dias de estresse hídrico, foi coletada uma planta de cada genótipo e replicação para as análises fenotípicas e as demais foram mantidas para medir a produtividade no final do ciclo da cultura. Temperatura da folha foi medida usando um



termômetro infravermelho (Telatemp modelo AG-42D, Telatemp, CA, EUA). Depois disso, as plantas foram cortadas na superfície do solo e separadas em folhas e caules. A estimativa do conteúdo de clorofila presente na folha foi medido utilizando o método não destrutivo SPAD-502 Plus (“Soil-Plant Analysis Development”, Konica Minolta Camera Co., Ltd., Japão), onde folhas completamente expandidas localizadas na parte mediana das plantas foram analisadas e cada leitura constou da média de cinco aferições expressas em unidades SPAD. A área foliar foi determinada usando um integrador de área foliar (LICOR modelo LI-3100C). A massa fresca da folha e do caule foram pesadas utilizando uma balança analítica (*BEL engineering*, Milano, Italia). Posteriormente, as partes das plantas foram separadas e colocadas em sacos de papel e secas em estufa com circulação de ar forçada a 60 °C de temperatura durante 48 h. Em seguida, a massa seca do caule e da folha foram determinadas. Após as avaliações morfológicas e fisiológicas, as plantas foram re-hidratados até à maturação fisiológica para a avaliação dos componentes de produtividade de grãos. Depois de três meses do ciclo da cultura, as vagens foram colhidas e o número de vagens, sementes, número de sementes por vagem foram contadas e a produtividade, Massa de grãos e vagens estimados. O índice de intensidade de seca (DII) foi calculado como  $1 - X_{ds}/X_{ns}$ , onde  $X_{ds}$  e  $X_{ns}$  são a produção de sementes média de todos os genótipos sob estresse hídrico (ds) e sem estresse (ns) (CIAT, 2012).

#### **4.2.2 Avaliação do sistema radicular quanto a comprimento, volume e diâmetro**

O experimento em casa de vegetação foi realizado entre março a setembro de 2012 no Instituto Agrônomo (IAC), Campinas, São Paulo, Brasil. O experimento foi conduzido em delineamento inteiramente casualizado com três repetições, sob condições contrastantes de abastecimento de água. Os tratamentos foram os mesmos relatados e descritos nos experimentos realizados com vasos e a irrigação foi fornecida por um sistema automatizado de gotejamento com duas irrigações diárias com duração de um minuto e vazão de  $0,225 \text{ L h}^{-1}$  no tratamento irrigado e no tratamento com déficit hídrico, até que fossem submetidos ao estresse de restrição hídrica. A diferença é que todos os tratamentos (com e sem irrigação) foram realizados numa mesma casa de vegetação onde os rizotrons foram instalados. Duzentos e quatorze rizotrons, cilindros de plástico de 60 cm de comprimento e 24 cm de diâmetro, foram preenchidos com  $0,013565 \text{ m}^3$  de mistura de solo

e areia na proporção 2:1 w/w. As sementes pre-germinadas foram previamente esterilizadas com uma solução de 5% de hipoclorito de cálcio durante 5 minutos, secas a temperatura ambiente e semeadas. Um total de 20 medidores Watermark foram instalados aleatoriamente, a uma profundidade de 20 cm para medir o potencial matricial do solo, no tratamento irrigado e não irrigado. A cada dois dias, foi medido o potencial matricial do solo e a temperatura foliar das plantas que continham os sensores “watermark”, a temperatura ambiental e a umidade relativa da casa de vegetação. O tratamento irrigado foi mantido a 80 % da capacidade de campo durante todo o experimento e a casa de vegetação sob estresse ficou sob condição de estresse terminal na fase vegetativa V3/V4, sem receber aplicação de água desde o 25º dia após o plantio. As plantas foram colhidas ao 52º dia e as características fisiológicas foram avaliadas.

A temperatura da folha, o nível de clorofila, a área foliar e a massa fresca e seca do caule e das folhas foram avaliadas utilizando a metodologia mencionada na fenotipagem dos vasos. As raízes foram recolhidas individualmente. Primeiramente, foram lavadas com água limpa, em seguida, passaram por uma solução com detergente 5%, outra com detergente 1% e, finalmente, enxaguadas em água limpa. Depois de lavadas, todas as raízes foram mantidas separadamente em potes contendo uma solução de álcool 20% até serem analisadas. As raízes foram digitalizadas num scanner Epson Expression 10000 XL e as análises das imagens foram realizadas utilizando-se o sistema WinRHIZO (“Regent Instruments Inc”.., Quebec, Canadá). O comprimento total de raízes (m) considerando-se a raiz primária e raízes secundárias, a área da superfície radicular (mm), volume de raízes (cm<sup>3</sup>) e diâmetro médio da raiz (mm) foram estimados.

#### 4.3 Análise de dados

As análises de variância foram realizadas para todas as características medidas na condição de seca e irrigada, utilizando o procedimento de modelos lineares generalizados (“General Linear Models Procedure”, GLM), utilizando o software SAS v.8.2 (“SAS Institute”, Cary, NC, EUA). A herdabilidade no sentido amplo foi estimada utilizando a fórmula 
$$h_g^2 = \frac{\delta^2_g}{\delta^2_g + \frac{\delta^2_e}{r}}$$
. Com o objetivo de confirmar as características contrastantes entre os

genótipos, análises separadas foram submetidas para os genitores e para as linhagens endogâmicas recombinantes. As análises de variância foram significativas quando  $p < 0,05$ .

#### **4.4 Extração e Quantificação de DNA**

Amostras de folhas de cada uma das 105 linhagens recombinantes e dos genitores SEA5 e AND277 foram coletadas. Depois de congeladas em nitrogênio líquido, as amostras foram liofilizadas e moídas. Em seguida, foi feita a extração do DNA a partir de 300 mg de folha liofilizada de acordo com a metodologia do CTAB descrita por Hoisington et al. (1994).

A quantificação do DNA final das amostras foi feita em gel de agarose 1% por comparação de intensidade das bandas com o padrão de fago  $\lambda$  (Invitrogen) e também pelo equipamento de quantificação NanoDrop 2000 (Thermo Scientific). Posteriormente, as amostras de DNA foram diluídas em solução tampão TE (Tris pH 8,0, 10 mM; EDTA pH 8,0, 1 mM), para uma concentração final de 100 ng/ $\mu$ l, e estocadas à 4°C.

#### **4.5 Desenvolvimento da plataforma DArTs para feijão comum e genotipagem na população AS com marcadores DArTs**

O desenvolvimento da plataforma incluiu a composição do painel de diversidade de 93 amostras (Anexo 1) sendo 32 amostras do Instituto Agrônomo (IAC), sete da Universidade Federal de Lavras (UFLA), seis do Instituto Agrônomo de Paraná (IAPAR), cinco da Universidade Federal de Viçosa (UFV), uma amostra da Escola Superior de Agricultura “Luiz de Queiroz”/USP e 42 DNAs do CIAT (Centro Internacional de Agricultura Tropical, Colômbia).

Outro conjunto com DNA de 186 amostras de feijão comum da Universidade de Cork (Cork, Irlanda) foi enviado para inclusão no desenvolvimento da plataforma incluindo um genótipo selvagem de *P. tuerckheimii*; 20 genótipos selvagens de *P. dumosus*, 21 genótipos de etnovariedades e 9 amostras de cultivares; 25 genótipos selvagens de *P. coccineus* L., 12 etnovariedades e 19 de cultivares; 4 genótipos de feijão comum identificadas como genótipo nn1, nn1, nn3 e nn4; 3 genótipos selvagens de *P. acutifolius*; um genótipo selvagem de *P. anisotrichos*; 3 genótipos selvagens de *P. costaricensis*; 2

genótipos selvagens de *P. albescens*; 42 genótipos selvagens de *P. vulgaris*, 6 cultivares; 9 amostras registradas como híbridos de *P. vulgaris* x *P. dumosus*; 2 genótipos selvagens de *P. lunatus* L. (Anexo 2 e 3). Essas amostras não foram incluídas no painel de diversidade, mas sim na construção das bibliotecas genômicas e arranjos para análises posteriores.

A qualidade dos DNAs foi verificada misturando 5µl 1X de tampão de corrida mais 1X “RE buffer” e 1µl de DNA de cada amostra. As placas de controle de qualidade com as reações foram incubadas a 37°C durante 2 horas, após incubação, as amostras foram analisadas em gel de agarose 0,8% e fotografadas.

O primeiro passo na redução da complexidade do genoma envolveu à digestão da mistura das amostras de DNAs com a enzima de corte raro *Pst*I combinada com uma outra enzima de corte freqüente (*Taq*I, *Pst*I, *Bst*NI, *Alu*I, *Bam*II, *Mse*I, *Hae*III e *Msp*I), seguida da ligação dos adaptadores da *Pst*I com a enzima T4 ligase e amplificação dos fragmentos ligados aos adaptadores. As amostras com a combinação de enzimas *Pst*I/*Bst*NI e *Pst*I/*Taq*I tiveram uma incubação a 37°C por 2 horas, depois 60°C por 2 horas e por último 80°C 20 minutos, e as amostras com a combinação de enzimas *Pst*I/*Alu*I, *Pst*I/*Bam*II, *Pst*I/*Mse*I, *Pst*I/*Hae*III e *Pst*I/*Msp*I tiveram uma incubação a 37°C por 2 horas.

Os adaptadores utilizados na ligação foram:

*Pst*I adaptador 1: 5' CACGATGGATCCAGTGCA 3'  
*Pst*I adaptador 2: 5' CTGGATCCATCGTGCA 3'

Foi feito um primeiro teste com as amostras: CAL143, IAC Alvorada e Bat-477 com as oito enzimas para verificar qual combinação de enzimas apresentava melhor digestão nos DNAs. O teste da digestão com enzimas de restrição e ligação dos adaptadores mostrou que as melhores combinações foram *Pst*I/*Taq*I e *Pst*I/*Bst*NI.

Após a avaliação dos DNAs quanto a digestibilidade com as enzimas de restrição e número de marcadores polimórficos gerados em cada combinação de enzimas, foram obtidas as representações genômicas para todas as amostras. A reação da polimerase em cadeia (PCR) foi utilizada para criar as representações genômicas contendo os fragmentos de corte primário, na maioria dos casos fragmentos *Pst*I possuem duas características

importantes: a) Não tem sitio de restrição para o corte freqüente usado na redigestão e b) podem ser amplificados eficientemente.

Um microlitro de reação de digestão/ligação de *PstI/TaqI* e *PstI/BstNI* foi adicionado a 49  $\mu$ l de mix de PCR que foi preparado da seguinte maneira: 5  $\mu$ l 10 x PCR buffer (100 mM Tris-Cl pH 8.3, 500 mM KCl, 15 mM MgCl<sub>2</sub>, 0.1 % gelatina; Sigma), 1  $\mu$ l 10 mM dNTPs (Sigma), 2  $\mu$ l 10  $\mu$ M PstI+0 primer (5' GATGGATCCAGTGCAG 3'), 2  $\mu$ l REDTaq polimerase (1 U/ $\mu$ l; Sigma) e 39  $\mu$ l água.

A Amplificação de PCR foi feita utilizando o programa com o seguinte perfil: 94°C por 1:00', 94°C por 0:20'', anelamento a 58°C por 0:40'', e extensão a 72°C por 1:00', por 29 vezes mais, seguido de uma extensão final a 72°C por 7:00' e mantido em incubação a 10°C.

Os fragmentos alvo (*targets*) amplificados foram visualizados em gel de agarose 1,2% com o objetivo de verificar a homogeneidade da digestão (*smear* dos fragmentos) e a distribuição dos tamanhos. As representações genômicas com distribuição de bandas diferentes podem ser derivadas de seqüências genômicas repetitivas e/ou DNA de organelas e, por isto não foram utilizadas.

Para o controle de qualidade dos “targets”, foram analisados 5  $\mu$ l de produto de PCR em gel de agarose 1,2% e corados com brometo de etídio 0,5  $\mu$ g/ml.

Uma vez estabelecidas as combinações de enzimas de restrição visando a redução da complexidade, estas foram utilizadas para gerar as bibliotecas genômicas.

O DNA de cada genótipo (Anexo 1) foi misturado em quantidades equimolares gerando um painel de diversidade genética para feijoeiro comum. As representações genômicas alvo geradas a partir da amplificação de PCR foram clonadas utilizando o sistema “TOPO” “TA Cloning” (Invitrogen). Foi usado o vetor “pCR 2.1-TOPO” para a ligação e “TOP10 *Electrocomp*” (Invitrogen) como células electrocompetentes para a transformação.

A ligação foi realizada misturando 2,0  $\mu$ l do produto de PCR de cada genótipo a ser clonado e depois foram ligados 4,0  $\mu$ l dos produtos de PCR misturados no vetor “pCR2.1-

TOPO”, 1,0 µl de solução de NaCl diluída 1:4 e 1,0 µl do vetor “pCR2.1-TOPO” num volume total de reação de 6,0 µl.

A reação foi incubada 15 min. a temperatura ambiente. Após este período foram adicionados 2,0 µl da reação de ligação TOPO ao tubo de células competentes “*TOP10 Electrocomp*” *E.coli*, misturadas gentilmente. Para a transformação, foi utilizado o eletroporador (*CELL-PORATOR E. coli Electroporation System*) (BioRad).

Foram transferidos 25 µl da solução da reação anterior de bactéria/ligação a uma electrocuba e colocada dentro da câmara de electroparação. Foi aplicada uma carga de 2,48 kV e depois o volume foi transferido a um tubo com 500 µl de meio SOC (2% Triptona, 0,5% extrato de levadura, 10mM NaCl, 2,5 mM KCl, 10mM MgCl<sub>2</sub>, 10mM MgSO<sub>4</sub>, 20mM de glicose), e incubado a 37°C por 45 minutos.

Depois foram colocados 30 µl da solução transformada em placas de petri contendo agar LB com ampicilina (100 ug/ml) e X-gal (40 ug/ml), incubadas a 37°C por 16 horas.

Os fragmentos de DNA individualizados pela clonagem com TOPO em transformantes bacterianos (colônias brancas) foram transferidos para um meio próprio de congelamento o (*freezing medium*) distribuído em placas de 384 poços. O “freezing medium” contém: 1 × LB contendo 4,4 % de glicerol, 8,21 g/l K<sub>2</sub>HPO<sub>4</sub>, 1,80 g/l KH<sub>2</sub>PO<sub>4</sub>, 0,50 g/l Na<sub>3</sub>-citrato, 0,10 g/l MgSO<sub>4</sub> x 7 H<sub>2</sub>O, 0,90 g/l (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 100 mg/L Ampicilina e 100mg/L kanamicina.

As placas foram incubadas a 37°C por 22 horas. Os insertos individuais foram amplificados diretamente por PCR em um meio IAM (*Insert Amplification Mix*). Este meio contém: 1x tampão “*Possum Taq*” (50 mM Tris, 6 mM HCl, 16 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1,5 mM MgCl<sub>2</sub>), 200 µM dNTPs, 0,2 µM M13 “forward primer” (5’ GTAAAACGACGGCCAG 3’), 0,2 µM M13 “reverse primer” (5’ CAGGAAACAGCTATGAC 3’), 0,1 – 2,0 µl “*Mango Taq*” polimerase (Bioline).

O programa utilizado para amplificar os insertos foi 95°C por 4:00’, 57°C por 0:35”, 72°C por 1:00’, 94°C por 0:35”, 52°C por 0:35”, 72°C por 1:00’, por 34 vezes, e uma extensão final de 72°C por 7:00’ e incubação a 10°C.

Para o controle de qualidade, foram analisados 2  $\mu$ l dos produtos de PCR das colônias de 39 poços escolhidos aleatoriamente em agarose 1,2%. As placas foram consideradas amplificações satisfatórias quando menos do 10% dos insertos não amplificaram ou tiveram múltiplas bandas

As placas de PCR das colônias que foram bem sucedidas no controle de qualidade foram processadas da seguinte forma:

Os produtos de PCR foram secados incubando as placas a 37°C durante 20 horas, depois foi adicionado 30  $\mu$ l de etanol 77%, incubadas a temperatura ambiente por 1,5 horas e centrifugadas a 3220g por 40 minutos a 30°C. Descartado o etanol, as placas foram incubadas a temperatura ambiente durante a noite inteira. Posteriormente, foram distribuídos 25  $\mu$ l de *DArT Spotter 2* por poço (50 % DMSO, 1,5 M sorbitol, 0,1 M Triethanolamine. HCl, 0,5 % (w/v) dextran.

#### **4.5.1 Preparação dos microarranjos**

##### **4.5.1.1 Impressão**

Foram impressas 40 placas de 384 poços em 96 “slides” seguindo o protocolo “Guide to printing an array” da Diversity Arrays Technology ([www.diversityarrays.com](http://www.diversityarrays.com)). Os “slides” processados asseguram que o DNA impresso sob o “array” sejam fixado a superfície de polilisina do “slide”.

##### **4.5.1.2 Gerando os “Targets” para hibridação**

O primeiro passo foi o controle de qualidade das amostras de DNA, em seguida a preparação dos “Targets” cortando com as mesmas enzimas de restrição que foram utilizadas para fazer as bibliotecas genômicas. Uma vez feita a ligação dos adaptadores e amplificação por PCR, os “Targets” foram precipitados com isopropanol. Em resumo, foram adicionados 45  $\mu$ l de isopropanol a cada reação de PCR, centrifugados a 3220g por 40 minutos a 30°C, descartados o sobrenadante e adicionado 100  $\mu$ l de etanol 77%. As amostras foram centrifugadas novamente a 3220g por 40 minutos a 30°C, descartado o sobrenadante, os “pellets” foram colocados para secar a temperatura ambiente “overnight”.

#### 4.5.1.3 Marcação dos “Targets”

Os “Targets” foram marcados pela incorporação de cy3-dUTP ou cy5-dUTP numa reação de extensão de “primers” com decâmeros aleatórios e *exo-klenow*. Ao pellet, foram adicionados 1x NEB Buffer 2, 50uM de decâmeros aleatórios, dNTPs marcados (2mM dATP, 2mM dCTP, 2mM dGTP, 0.2mM dTTP). Os “Targets” foram desnaturados a 95°C por 3 minutos e esfriados a 25°C.

As misturas dos Cy-dye consistiram em 1x NEB Buffer 2, 25uM cy3-dUTP and cy5-dUTP, 2,5 unidades de fragmento *Klenow exo-* de *E.coli* polimerase I (NEB), 5ul dos cy-dye foram adicionados aos “Targets” uma vez alcançados os 25°C. As placas foram cobertas com papel alumínio e incubadas a 37°C por 3 horas.

#### 4.5.1.4 Hibridação dos “Targets” para os microarranjos.

Foram adicionados 5 ul da mistura de inativação (1X NEB buffer 2, 60mM EDTA pH (8.0) a cada reação marcada. Foram adicionados 50 ul da solução “DArT hybridizer” a cada reação e misturadas suavemente. As reações foram desnaturadas a 95°C por 3 minutos; 56°C por 5 minutos; 55°C indefinidamente. Quando as amostras alcançaram os 55°C, foram centrifugadas para eliminar as borbulhas e colocadas no bloco seco pré aquecido a 65°C.

Foram adicionados 60 ul da solução anterior a cada “slide” e colocados nas câmeras de hibridização em banho a 62°C durante 16 horas. As amostras foram lavadas numa solução de 1X SSC, 0,1% SDS, 0,5M DTT por 5 minutos, depois foram enxaguadas numa solução de 1X SSC por 5 minutos, após transferidas para uma solução de 0.2X SSC por 2 minutos e por último, lavadas numa solução 0,02X SSC por 30 segundos. Então, foram centrifugadas a 500g por 7 minutos a 30°C e colocadas num dissecador com sílica gel, ligado a uma bomba a vácuo por 30 minutos. Finalmente, foram escaneadas e a análise dos dados genotípicos foi realizada utilizando o software “DArT soft” ([www.diversityarrays.com](http://www.diversityarrays.com)).



#### **4.5.2 Análise da diversidade genética.**

O software “DArT soft” foi utilizado para analisar as imagens, identificar e fazer a contagem dos marcadores polimórficos. A leitura dos dados foi feita com 1 para presença, 0 para ausência e X para dados perdidos. Os parâmetros escolhidos para a descoberta de novos marcadores polimórficos foi: o parâmetro de qualidade Q, que mede a fração do total da variação através dos indivíduos devido à bi modalidade; a reprodutibilidade derivada das réplicas e que supostamente deveria fornecer os mesmos resultados; o valor de “call rate” que representa o número de leituras versus o número máximo de outras possíveis. O PIC (*polymorphism information content*) foi utilizado para avaliar como os marcadores foram distribuídos entre os dois grupos (avaliados como 0 ou 1) e foi determinado de acordo com Anderson et al. (1993). Este valor é frequentemente utilizado para medir o grau de informação do marcador genético.

A contagem dos marcadores polimórficos foi analisada utilizando o software DARwin (Perrier & Jacquemoud-Collet, 2006) para visualização das relações genéticas entre os genótipos e construção do dendrograma utilizando as estimativas de distância genética de complemento aritmético do coeficiente de Jaccard (1-J). O agrupamento foi realizado utilizando o algoritmo de (*neighbor joining*) (Saitou & Nei, 1987) com um número de 1000 bootstraps.

#### **4.5.3 Análises de coordenadas principais.**

Para visualizar a posição da diversidade dos genótipos no espaço, análises de coordenadas principais foram feitas utilizando as matrizes de distância de Jaccard. A construção do gráfico de dispersão analisando os 89 genótipos incluiu amostras de feijão comum selvagem, genótipos Mesoamericanos e Andinos e amostras dos programas de melhoramento genético. As variedades foram identificadas pela origem nessa análise. As análises de coordenadas principais foram computadas utilizando o software NTSYSpc 2.1 (Rohlf, 2002) e os gráficos tri-dimensionais foram construídos utilizando o software Statistics 7.1 (StatSoft, Inc. 2005).

#### 4.5.4 Genotipagem com marcadores DArTs

Após desenvolvida e testada a metodologia, os DArTs foram disponibilizados para a comunidade do feijão ([www.diversityarrays.com](http://www.diversityarrays.com)) e os DNAs da população AND 277 x SEA 5 foram extraídos e enviados para genotipagem na empresa DArT Pty (Canberra, Australia).

#### 4.6 Genotipagem com marcadores SNPs.

A genotipagem para os 384 SNPs foi conduzida via tecnologia Vera Code® em plataforma BeadXpress (Illumina) no Laboratório de Biotecnologia da Embrapa Arroz e Feijão (Goiás, GO). Um conjunto de 384 marcadores SNPs, validados através do *Prelim file* (<https://icom.illumina.com/Custom/UploadOpaPrelim/>), previamente identificados para *Phaseolus vulgaris* L. (Müller et al., *em elaboração*) e derivados do polimorfismo entre as linhagens BAT477, de origem Mesoamericana, e Jalo EEP558, de origem Andina, foram selecionados para compor o *Oligo Pool Assay* (OPA) de marcadores SNPs. Foi realizado um teste de polimorfismo na população AS. Durante o procedimento de detecção dos SNPs na plataforma *BeadXpress* foram utilizados três oligonucleotídeos, sendo dois alelo-específico (ASO) para cada uma das variações de um mesmo SNP e o terceiro loco-específico (LSO) ligando-se à região 3' do fragmento de DNA contendo o SNP alvo. Após a hibridização, o processo consistiu na extensão das regiões entre o ASO e o LSO, seguido pela fusão das mesmas a partir de uma enzima ligase formando, assim, um único fragmento alelo-específico. Posteriormente, esse fragmento foi amplificado utilizando a enzima *Titanium Taq* DNA polimerase e *primers* complementares à região ASO foram marcados com as fluorescências Cy3 e Cy5. Ao final, os produtos derivados da PCR foram hibridizados através da complementaridade da região LSO com as seqüências presentes na superfície das *beads* holográficas. A genotipagem dos SNPs foi realizada por meio do programa *Genome Studio* versão 1.8.4 (Illumina, EUA) utilizando valores de *Call Rate* variando de 0,80 a 0,90 e *GenTrain*  $\geq$  0,26 para o agrupamento dos SNPs. A clusterização (agrupamento) para os alelos de SNPs referentes a cada linhagem foi feita, *a priori*, de modo automatizado com base na intensidade dos sinais emitidos a partir dos fluoróforos Cy3 e Cy5, os quais foram agrupados em três classes de genótipos representativas dos grupos de homozigotos para os alelos AA ou BB e heterozigoto AB. Para a análise dos

dados, os agrupamentos foram individualmente e manualmente ajustados determinando os melhores *clusters* com base no perfil dos genitores.

#### **4.7 Caracterização e genotipagem com marcadores microssatélites**

Inicialmente, para verificar quais microssatélites eram polimórficos para os genitores SEA5 e AND277, foram utilizados um total de 594 marcadores microssatélites, em sua grande maioria com a temperatura de anelamento já obtida na literatura e alguns re-caracterizados por gradiente.

Os microssatélites utilizados são provenientes da literatura (Benchimol et al., 2007; Blair et al., 2003; Blair et al., 2006; Blair et al., 2008; Blair et al., 2009; Buso et al., 2006; Caixeta et al., 2005; Campos et al., 2007; Campos et al., 2011; Cardoso et al., 2008; Perseguini et al., 2011; Gaitán-Solís et al., 2002; Grisi et al., 2007; Guerra-Sanz, 2004; Hanai et al., 2007).

Cada par de *primer* foi caracterizado quanto a temperatura ideal de anelamento para PCR. As reações de amplificação foram realizadas com 30 ng de DNA, 1U de *Taq*-DNA polimerase, 1,5 mM de cloreto de magnésio, 0,15 mM de cada dNTP, 0,8 pmol/μL de cada *primer* (direto e reverso) e 1x de tampão da enzima, num volume total de 15 μL. As condições de amplificação utilizadas para os SSRs genômicos foram: 1) 94°C, 1 min.; 2) 94°C, 1 min.; 3) temperatura de anelamento (*Ta*) específica para cada SSR por 1 min.; 4) 72°C, 1 min., 5) volta ao passo 2 por 30 vezes; 6) 72°C, 5 min.; 7) 15°C, incubação final. O termociclador MyCycler (BioRAD) foi utilizado. Para os microssatélites que não possuíam temperatura de anelamento definida, as amplificações foram realizadas em gradiente de temperatura, de acordo com as seguintes condições: 1) 94oC, 1 min; 2) 94oC, 1 min; 3) temperatura de anelamento (*Ta*) de 45°C-56°C ou 56°C-65°C por 1 min.; 4) 72oC, 1 min, 5) volta ao passo 2 por 30 vezes; 6) 72°C por 5 min.; 7) 15°C, incubação final.

Para verificar a qualidade das amplificações, primeiramente os produtos das reações de amplificação foram submetidos à eletroforese em gel de agarose 3%, corado com GelRed™ 1X (Biotium, Inc. Hayward, CA) para visualização das bandas. Os locos que apresentaram padrão adequado de amplificação foram submetidos à eletroforese em gel de poliacrilamida denaturante 6%, corados com nitrato de prata, de acordo com Creste et al. (2001).

Os géis de poliacrilamida 6% foram feitos a partir de solução de poliacrilamida 6% (150 mL de acrilamida 40%, 420 g de uréia, 200 mL de TBE 5X, volume completo para 1L) filtrada e polimerizada por aproximadamente 2h com 960 µL de PSA (Persulfato de Amônio) e 60 µL de TEMED. Após aquecimento do gel em cuba de eletroforese (*Nucleic Acid Sequencing Systems – CBS Scientific Company, INC*) por 20 min. à 900 V, 5 ul de reação de cada amostra foram aplicados no gel, preparados a partir de 15 ul de produto de PCR com 5 ul de formamida. Foram feitas duas entradas por placa, sendo que uma entrada continha 54 amostras de DNA mais o *ladder* de 50 pares de base (*Invitrogen*) e a outra, 53 amostras mais o *ladder*, utilizando-se uma placa para genotipar cada microsatélite para as 105 linhagens e os dois genitores. A corrida foi de aproximadamente 4h a 800- 900V.

#### **4.8 Construção do mapa genético AS com marcadores SNPs e SSRs**

As análises de segregação dos SNPs e SSRs das 105 linhagens endogâmicas recombinantes  $F_8$  e dos genitores SEA 5 e AND 277 foram realizadas utilizando o teste de Qui-quadrado, pela equação:  $\chi^2 = \frac{(fo - fe)^2}{fo}$ , em que:  $\chi^2$  = valor calculado do Qui-Quadrado; fo = frequência observada em cada uma das classes; fe = frequência esperada em cada uma das classes. Os valores obtidos do Qui-Quadrado foram convertidos em seus respectivos *p*-valores através do programa estatístico “R” version 2.5.1 “The R Foundation for Statistical Computing, 2007”.

O mapa genético foi construído utilizando o software “OneMap” versão 2.0-1 (Margarido et al. 2007), utilizando as abordagens multiponto e modelos de Markov. Após a identificação dos marcadores redundantes e distorção de segregação, as frações de recombinação foram estimadas entre cada par de marcadores usando a função “twopts”. Em seguida, os marcadores foram atribuídos aos cromossomo usando um LOD de 3,0 e uma distancia genética máxima de 37,5 cM usando a função de mapeamento de Kosambi (1944). Para os marcadores restantes utilizou-se a função “try.seq”. O posicionamento dos marcadores foi refinada usando a função “make.seq”. Para ajudar na desição da posição de cada marcador inserido num grupo de ligação específico, utilizou-se a função “rf.graph.table” e “draw.try=TRUE” para exibir o “heat map”. A nomenclatura dos cromossomo e as posições físicas foram identificadas por meio de comparações de

semelhança das sequências utilizando BLASTN (Altschul et al. 1997) contra o genoma de *Phaseolus vulgaris* (<http://www.phytozome.net/phaseolus>) e o mapa genético integrado para o feijão comum com base em mapeamento de microssatélites de Blair et al. (2011) e Campos et al. (2011). Por isto, os grupos de ligação foram chamados de Pv de 1 a 11 segundo a nomenclatura proposta por Pedrosa et al. (2008)

Os dados fenotípicos da distribuição de normalidade foram avaliados por Skewness, Kurtose e valores Shapiro-Wilk para o experimento dos vasos (Anexo 4) e para o experimento dos rizotron (Anexo 5). A transformação por Box-Cox foi aplicada e o modelo apropriado para normalizar os dados de cada característica foi selecionado usando o parâmetro lambda ( $\lambda$ ) (Osborne, 2010) (Anexos 6 e 7).

#### **4.9 Mapeamento dos locos associados à tolerância ao estresse hídrico no mapa SSR-SNP**

A identificação de QTLs foi realizada utilizando o software QTL Cartographer vs 1.17 (Basten et al., 2005) com a análise de mapeamento por intervalo composto (CIM).

Testes da razão de verossimilhança (LRT) foram utilizados para verificar a presença de QTL. O coeficiente de determinação foi calculado para cada intervalo separadamente ( $R^2$ ) e determinada a variância explicada por um único QTL. Os valores de LOD foram calculados utilizando a fórmula  $LOD = 0,2172 * LRT$ . A regressão linear múltipla para cada posição cromossômica foi aplicada com 5% de nível de significância para a obtenção dos cofatores utilizados na análise. Valores limiares foram identificados para cada característica baseado em 1000 permutações e representado por gráficos usando o Excel 2010 (Anexo 8).

#### **4.10 Construção do mapa genético AS com marcadores SNPs – SSRs – DarTs.**

O mapa genético utilizando os marcadores SNPs, SSRs e DarTs e a identificação dos QTLs nesse mapa foi realizada pela empresa DArT Pty (Canberra, Australia) utilizando o mapeamento por precisão com a metodologia SML (*Statistical machine learning*) descrita por Bedo et al. (2008) e de uso exclusivo da empresa.

## 5 RESULTADOS E DISCUSSÃO

### 5.1 Desenvolvimento da plataforma DArT para feijão comum

#### 5.1.1 Seleção do método de redução da complexidade

Foram testadas oito combinações diferentes de enzimas para escolher a redução da complexidade mais eficiente. Escolheram-se os métodos que apresentaram os melhores perfis de digestão dos DNAs: *PstI/BstNI* e *PstI/TaqI*. O critério para os parâmetros selecionados foi de 100% de reprodutibilidade e de 95% para o “call rate”. O melhor método para a redução da complexidade foi *PstI/BstNI*, mostrando o alto nível de polimorfismo no primeiro teste feito com 2.304 clones (Tabela 3).

Tabela 3. Avaliação do método da redução da complexidade

Método	Nº Marcas	Média	Nº Marcas	Média
	<i>BstNI</i>		<i>TaqI</i>	
<i>PstI/BstNI</i>	42	16,93	24	13,87
<i>PstI/BstNI</i>	46	18,54	18	10,4
<i>PstI/BstNI</i>	47	18,95	27	15,6
<i>PstI/TaqI</i>	41	16,53	28	16,18
<i>PstI/TaqI</i>	39	15,72	44	25,43
<i>PstI/TaqI</i>	33	13,3	32	18,49
TOTAL	248		173	

Reprodutibilidade: 100%

“Call Rate”: 95%

O painel de diversidade foi criado com 89 genótipos. Os recursos genéticos consistiram em amostras dos dois principais “pools” gênicos, Mesoamericano e Andino. Foram identificados 13.824 clones que foram polimórficos nas bibliotecas genômicas de feijão comum, BAT-477, IAC-Alvorada, genótipos Andinos, genótipos Mesoamericanos e genótipos selvagens de *P. vulgaris* (Tabela 4), demonstrando a utilidade dos marcadores DArTs para análises de diversidade genética e mapeamento de características de interesse econômico em feijão comum.

Tabela 4. Bibliotecas disponíveis para *Phaseolus* com tecnologia DArT

Nome	Método	Clones
<i>Phaseolus vulgaris</i> L.	<i>Pst</i> I/ <i>Taq</i> I	1152
<i>Phaseolus vulgaris</i> L.	<i>Pst</i> II/ <i>Bst</i> NI	1152
BAT 477	<i>Pst</i> II/ <i>Bst</i> NI	1536
IAC Alvorada	<i>Pst</i> II/ <i>Bst</i> NI	1536
Andinos	<i>Pst</i> II/ <i>Bst</i> NI	3072
Mesoamericanos	<i>Pst</i> II/ <i>Bst</i> NI	3072
Selvagens <i>P.vulgaris</i>	<i>Pst</i> II/ <i>Bst</i> NI	2304
Selvagens <i>P.dumosus</i>	<i>Pst</i> II/ <i>Bst</i> NI	1536
Selvagens <i>P.coccineus</i>	<i>Pst</i> II/ <i>Bst</i> NI	1536
Selvagens <i>P.tuerkbeimii</i>	<i>Pst</i> II/ <i>Bst</i> NI	768
Selvagens <i>P.acutifolius</i>	<i>Pst</i> II/ <i>Bst</i> NI	768
Selvagens <i>P.anisotrichus</i>	<i>Pst</i> II/ <i>Bst</i> NI	768
Selvagens <i>P.cosaricensis</i>	<i>Pst</i> II/ <i>Bst</i> NI	768
Selvagens <i>P.albescens</i>	<i>Pst</i> II/ <i>Bst</i> NI	768
Selvagens <i>P.lunatus</i>	<i>Pst</i> II/ <i>Bst</i> NI	768
<b>TOTAL</b>		<b>21504</b>

### 5.1.2 Desenvolvimento do Array

As bibliotecas foram expandidas a 21.504 clones empregando o melhor método de redução da complexidade, utilizando a enzima *Pst*I como corte frequente e *Bst*NI como corte secundário. Para otimizar a descoberta de marcadores, a hibridação das representações com todos os genótipos incluiu réplicas aleatórias no desenho do experimento. Foram identificados 2.501 marcadores polimórficos para o painel de 89 genótipos. As bibliotecas de *P. dumosus*, *P. coccineus*, *P. tuerkbeimii*, *P. acutifolius*, *P. anisotrichus*, *P. costaricensis*, *P. albescens* e *P. lunatus* não foram utilizadas na descoberta de marcadores polimórficos nessa análise.

### 5.1.3 Diversidade Genética de Feijão Comum

Um total de 2.501 marcadores polimórficos foram encontrados para o painel de 89 genótipos. A qualidade dos marcadores DArTs foi determinada pela reprodutibilidade de 99,89%, discordância de 0, consenso de 99,8% e *call rate* de 97,11%. Estes dados demonstraram que os arranjos foram bem sucedidos na determinação da diversidade genética para feijão comum.

Os genótipos selecionados representam uma proporção significativa da variação genética em feijão comum nos dois maiores “pools” genéticos. Além disto, os genótipos incluíram linhagens elite dos programas de melhoramento do feijoeiro realizados nos principais centros de pesquisa do mundo.

Utilizando os 2.501 marcadores, foi construído o dendrograma com 89 genótipos incluindo Andinos, Mesoamericanos, acessos selvagens e amostras dos programas de melhoramento genético (Figura 1).

O dendrograma mostrou uma divisão de dois grupos principais. O primeiro grupo teve as amostras do grupo Andino e o segundo grupo agrupou os genótipos do grupo Mesoamericano. Um total de 19 genótipos foram agrupados no grupamento Andinos e 70 no grupo Mesoamericano. Os genótipos usados como controle foram posicionados no grupo Mesoamericano (DOR 364) e no grupo Andino (G19833).

### 5.1.4 Análise de coordenadas principais

A análise de coordenadas principais (PCoA) mostrou, em três dimensões, a representação espacial da variabilidade entre os genótipos. O gráfico mostrou a dispersão diferenciando os dois maiores *pools* gênicos de feijão comum (Andino e Mesoamericano), os quais ficaram bem separados nos agrupamentos (Figura 2). A análise explicou 81,97% da variação contida entre os genótipos.

Acreditamos que esses agrupamentos estão formados pela variabilidade genética que possui cada centro de pesquisa. Essa variabilidade genética pode estar explicada pelos programas de melhoramento genético, as condições dos bancos de germoplasma e os diferentes ambientes onde são plantadas essas variedades.



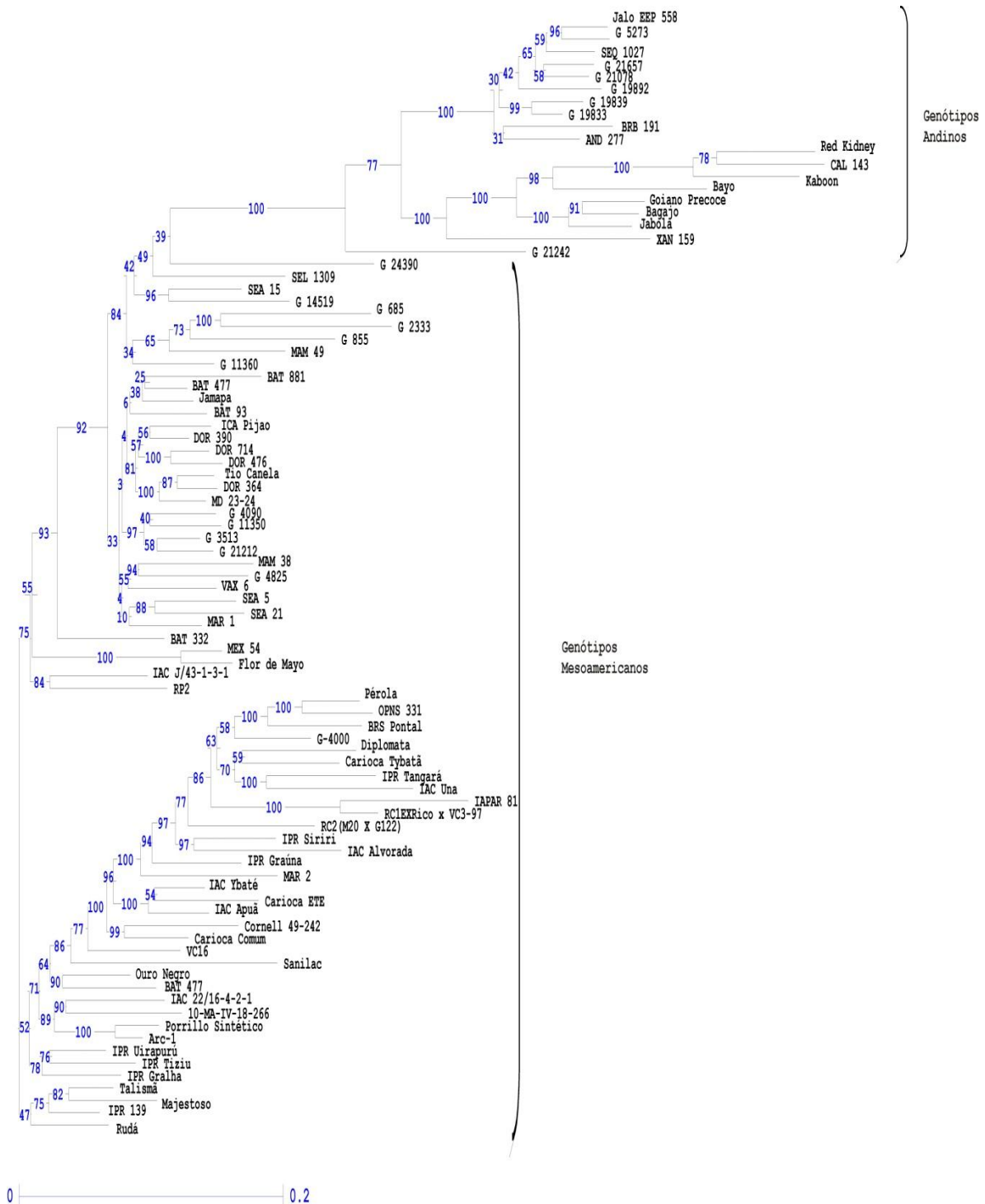


Figura 1. Dendrograma utilizando o método de distância genética do complemento aritmetico de Jaccard e agrupamento com “neighbor-joining” para o perfil de marcadores DArTs gerado para o painel de 89 genótipos.

Este estudo é o primeiro relato e a primeira publicação de feijão comum utilizando a tecnologia DArT. Os resultados mostraram que o desenvolvimento da plataforma DArT foi eficiente para estudar a diversidade genética do feijão comum.

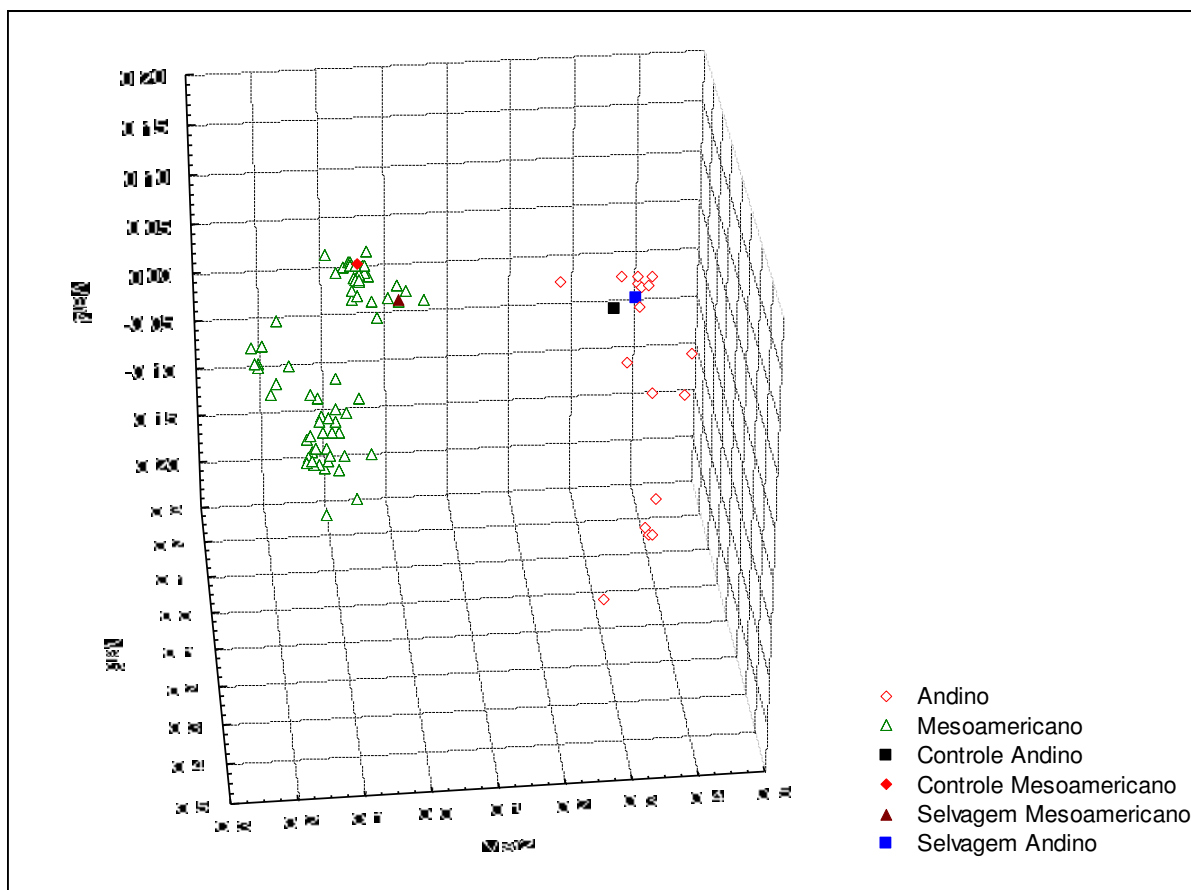


Figura 2. PCoA dos genótipos de feijoeiro dos grupos Andinos, Mesoamericanos, os controles e as amostras selvagens em três dimensões.

Nas análises de diversidade, com 2.501 marcas, os genótipos foram divididos em dois grandes grupos correspondentes aos “pools” gênicos andinos e mesoamericanos, confirmando os resultados encontrados em estudos anteriores com outros marcadores, como os microssatélites (Benchimol *et al.*, 2007; Blair *et al.*, 2009).

A amostra codificada como 1 x RC2G2122 M20-72 que ficou no grupo Mesoamericano, vem do cruzamento original da linhagem andina G122 (sementes beges pintadas de vermelho e cerca de 45 g por 100 sementes) com a linhagem mesoamericana

M20 (sementes tipo carioca com cerca de 23 g por 100 sementes). A F1 foi cruzada novamente com a M20, obtendo-se a progênie RC1, que foi re- cruzada com o genitor M20, obtendo-se a progênie RC2. Na população RC1, ainda, foram selecionadas as plantas mais similares à M20 com base em marcadores RAPD, semelhante ao método usado por Benchimol *et al.* (2005, 2008) proporção de alelos do genitor recorrente. Além disso, foram selecionadas as progênies mais semelhantes agronomicamente ao genitor recorrente, como a 72. Portanto, essa progênie possui alta porcentagem de alelos do M20 o que justifica o seu agrupamento com os genótipos mesoamericano e não apareceu como uma introgressão.

O fato dos dois diferentes acessos de BAT-477 terem sido agrupados em diferentes subgrupos dentro do grupo Mesoamericano pode ser explicado porque os acessos vêm de bancos de germoplasma e programas de melhoramento genético diferentes. Um deles vem do banco de germoplasma do CIAT (Centro Internacional de Agricultura Tropical, Cali, Colômbia) e o outro vem do IAC (Instituto Agrônômico, Campinas). O fato de que os genitores das linhagens, codificadas como BAT terem sido oriundas de cruzamentos de raças e cultivares como "Jamapa" e "Porrillo Sintetico" pode ser evidenciado. No dendrograma, a amostra BAT 477 que vem do CIAT ficou ao lado do genótipo "Jamapa" e a outra amostra de BAT 477, que vem do IAC, ficou perto do genótipo "Porrillo Sintetico". Então, de alguma forma, este acesso evoluiu fixando alelos e/ou genes diferentes em um e outro banco de germoplasma. Vale lembrar, que as sementes de um BAG de tempos em tempos, tem de ser multiplicadas o que poderia aplicar a divergência dos acessos de BAT 477 de diferentes BAGs.

Os "bootstraps" no dendrograma confirmaram que os subgrupos dentro dos grupos foram distintos e confirma a base genética do "pool" gênico Mesoamericano e Andino. No presente estudo, nenhuma classificação anterior de acessos foi realizada, e a tendência de formar grupos por raças distintas não foi detectada. Os grupos tiveram características comuns entre os genótipos como tamanho e cor dos grãos, resistência a patógenos e tolerância à seca.

Os tipos Andinos ficaram amplamente separados dos tipos Mesoamericanos na análise de coordenadas principais (Figura 2), como era esperado. Entre os genótipos selvagens, o genótipo Mesoamericano (G24390), ficou amplamente separado do feijão do

tipo selvagem Andino (G19892). Como controles, foram utilizados os genótipos G19833, para o grupo Andino, e DOR 364, para o grupo Mesoamericano. Esses controles também foram utilizados no estudo de Blair et al. (2009), demonstrando a consistência dos nossos resultados.

Os subgrupos formados em cada *cluster*, foram estruturados também pelos programas de melhoramento de origem e foi nítida a separação de amostras do CIAT em um grupo e do IAC em outro grupo. Este resultado é interessante e pode ser relacionado à pressão de seleção de cada programa de melhoramento genético, por causa da adaptação edafoclimática, entre outros fatores ambientais que podem alterar o genoma das espécies e aumentar a diversidade genética. A maioria dos acessos cultivados formaram um único grupo relativamente compacto (figura 2). Assim sendo, o padrão de separação em subgrupos foi consistente. Isso indica que a tecnologia DArTs, de fato, tem a capacidade de revelar a diversidade genética, com um grau de detalhamento fino conforme o esperado.

O dendrograma foi confirmado pela análise de coordenadas principais (PCoA), a qual explicou 82% da variação (figura 2). Normalmente a porcentagem da variação explicada por marcadores moleculares é inferior a 40% (Benchimol et al., 2005), mas os DArTs explicaram muito mais da variação genética dos dados em três dimensões.

Em relação ao feijão selvagem, os acessos estiveram localizados no grupo das amostras do CIAT. Na maioria das variedades do IAC existe uma pressão de seleção para as características selecionadas no programa. Paralelamente a topologia teria permitido que o feijão fosse domesticado em altitudes determinadas para ser transferido gradualmente para os outros ambientes, resultando em subgrupos genéticos. Beebe (2001) demonstrou que o germoplasma brasileiro de origem andina apresenta diversidade entre feijões andinos para diversas características: hábito de crescimento indeterminado; sensibilidade ao fotoperíodo; moderada tolerância à disponibilidade de fósforo no solo, com as cores de grãos e sementes com padrão chamado de "zebra". Estes acessos podem ser residuais do primeiro germoplasma adaptado a baixas altitudes.

Em conclusão, os resultados da genotipagem em grande escala com a utilização da tecnologia DArT foram bem sucedidos para analisar a diversidade genética do feijoeiro

comum. Este estudo é o primeiro relato usando a tecnologia DArTs para feijão comum e será útil como base para outros estudos utilizando a tecnologia, especialmente na identificação de locos de características quantitativas (Briñez et al., 2012) (Anexo 1).

## **5.2 Construção do Mapa AS**

Foram construídos dois mapas genéticos, um primeiro utilizando SSRs e SNPs e outro que foi feito pela empresa DArT Pty utilizando DArTs, SSRs e SNPs.

### **5.2.1 Mapa com marcadores SSRs – SNPs**

Entre os 594 marcadores microssatélites selecionados para os genitores, 150 (25%) foram polimórficos para a população e 80 SSRs (13% do total) foram mapeados. Um total de 288 marcadores polimórficos foram identificados nos SNPs e 251 foram incluídos no mapa genético. Alguns deles (9,37%) ainda apresentavam heterozigoto os quais não foram utilizados na análise. O mapa de ligação foi construído com um total de 331 marcadores segregantes na população e abrangeu todos os 11 cromossomos com um comprimento total de 1.515,2 cM. Todos os marcadores foram distribuídos entre os cromossomos com uma distancia media entre os marcadores de 4,5 cM. O tamanho dos cromossomos variou de 63,1 cm (cromossomo Pv10), ate 221,2 cM (cromossomo Pv01). O número de marcadores em cada cromossomo variou de 23 para o cromossomo Pv04 ate 40 marcadores no cromossomo Pv03 (Tabela 5). O genoma foi saturado com base na construção do mapa de ligação sem espaços grandes entre os marcadores.

Tabela 5. Distribuição dos SSRs e SNPs, mapeados nos 11 cromossomos de feijoeiro comum, no mapa genético desenvolvido a partir da população AND277 x SEA5 (AS) utilizando o software OneMap. A saturação, comprimento, distância média entre os marcadores para cada cromossomo pode ser evidenciada.

Cromossomo	SSR	SNP	Nº locos ligados	Comprimento (cM)	Distância (cM) entre <i>loci</i> <sup>1</sup>
Pv01	10	26	36	221,2	6,1
Pv02	11	21	32	161,4	5,0
Pv03	10	30	40	159,4	3,9
Pv04	5	18	23	128,4	5,5
Pv05	5	22	27	147,2	5,4
Pv06	8	19	27	148,3	5,4
Pv07	8	22	30	179,6	5,9
Pv08	7	22	29	86,7	2,9
Pv09	7	17	24	112,4	4,6
Pv10	5	23	28	63,1	2,2
Pv11	4	31	35	107,5	3,0
Total	80	251	331	1515,2	4,5

<sup>1</sup> Média Aritmética

Os marcadores SNPs ficaram distribuídos em todos os cromossomo, variando de 17 (Pv09) a 31 SNPs (Pv11) e os marcadores SSRs variaram de 4 (Pv11) a 10 (Pv01, Pv02, Pv03).

O cromossomo Pv03 foi o que apresentou maior saturação, com 40 marcas ligadas, entre elas 30 SNPs e 10 SSRs. O cromossomo Pv01 apresentou maior distância de mapa encontrado, portando 221,2 cM e um total de 36 marcas ligadas.

Para o feijoeiro comum, vários mapas de ligação foram construídos utilizando técnicas de baixa reprodutibilidade como os RAPDs (Blair et al., 2012), ou robustos utilizando marcadores RGAs e EST-SSRs (Campos et al., 2011; Hanai et al., 2010). Mapas oriundos de populações de diferentes “pools” gênicos tem a vantagem de apresentar um nível mais elevado de polimorfismo e, por conseguinte, possibilitar a obtenção de mapas mais saturados. Na população analisada neste estudo, entre os marcadores microssatélites selecionados nos genitores, obteve-se um número elevado de polimorfismo, 80 marcadores foram úteis (13%) para o mapeamento. Blair et al. (2012) usaram um cruzamento de

mesoamericano x mesoamericano, onde cerca de 595 marcadores foram avaliados e 60 foram incluídos no respectivo mapa. Yuste-Lisbona et al (2012) usaram uma população derivada do mesmo “pool” gênico onde foram avaliados 1.035 SSRs e 251 marcadores SNPs, e apenas 10,2% e 7,2%, respectivamente, de polimorfismo foi detectado.

O polimorfismo encontrado na população AS tipo RIL utilizada para a construção do mapa foi de 13% para os microssatélites e de 75% para os marcadores SNPs. Estes resultados refletem a importância de se considerar as propriedades dos marcadores SSRs utilizados, ou seja, o comprimento e o número de unidades de repetição e a posição dos SSRs dentro da sequência transcrita (Garcia et al., 2011). Os SNPs e marcadores SSR foram bem distribuídas (Figura 3). O tamanho total do mapa, a co-localização de marcadores, o tamanho dos cromossomos e as posições mantidas entre os 11 cromossomos, quando comparada com outros mapas (Córdoba et al., 2010; Campos et al., 2011; Garcia et al., 2011; Blair et al., 2012), bem como o genoma do feijão disponível (<http://www.phytozome.net/phaseolus>), confirmam a robustez e confiabilidade do mapa genético gerado neste trabalho a partir dos SNPs e SSRs utilizados. Os marcadores saturaram todos os onze cromossomos, cobrindo todo o genoma, e permitindo a identificação de QTLs em dois sistemas de irrigação.

### **5.2.2 Mapa com marcadores SSRs – SNPs – DArTs.**

O mapa construído pela empresa Diversity Arrays Technology ([www.diversityarrays.com](http://www.diversityarrays.com)), incluiu cinco classes de marcadores SSRs, 91 SNPs e 4468 DArTs. Os microssatélites foram localizados nos cromossomos Pv01, Pv07, Pv09 e Pv11, os SNPs ficaram nos 11 cromossomos com exceção do Pv10 e os DArT ficaram em todos os cromossomos com um número de marcas em média de 406 DArTs por cromossomo, identificando a menor saturação no cromossomo Pv05 com 269 marcas, e a maior saturação no cromossomo Pv07 com 555 marcas (Tabela 6).

Tabela 6. Distribuição dos SSRs, SNPs e DArT, mapeados nos 11 cromossomos de feijão comum, no mapa genético desenvolvido pela empresa DArT Pty

Cromossomo	SSR	SNP	DArT	Nº locos ligados	Comprimento (cM)	Distância (cM) entre loci <sup>1</sup>
Pv01	1	16	489	506	322,95	0,77
Pv02	0	3	439	442	438,48	0,90
Pv03	0	13	395	408	513,15	0,85
Pv04	0	8	432	440	290,21	1,02
Pv05	0	8	261	269	311	0,8
Pv06	0	5	351	356	310,13	0,67
Pv07	1	10	544	555	335,88	0,77
Pv08	0	13	292	305	369,86	0,70
Pv09	2	9	399	410	377,14	0,81
Pv10	0	0	428	428	213,99	0,66
Pv11	1	6	438	445	361,64	0,78
Total	5	91	4468	4564	3844,43	0,79

<sup>1</sup> Média Aritmética

O mapa gerado pela empresa DArT Pty (SSR-SNP-DArT, Tabela 6) apresentou maior saturação do que o anterior (SSR-SNP) com um grande número de marcadores ligados (4.468), um comprimento total de 3.844,43 cM e saturação média entre os marcadores de 0,79 cM. Só foram incluídos cinco microssatélites dos 150 que foram polimórficos nos genitores, e 91 SNPs dos 288 polimórficos na população. A genotipagem da população foi realizada utilizando a tecnologia “DArTseq” que apresenta a vantagem de



gerar todas as sequências dos marcadores inseridos no mapa para posteriores análises. Porém, a tecnologia apresentou várias desvantagens em relação ao aproveitamento da informação com o uso das ferramentas disponíveis (softwares), rotinas de mapeamento e desenho dos mapas, forçando a utilização do “pipeline” da empresa. Quando construído o mapa utilizando os softwares de mapeamento “OneMap” e “Mapmaker”, inserindo os marcadores DArTs, SSR e SNPs, os marcadores DArTs foram ligados em apenas três cromossomos, em uma mesma posição em cada, evidenciando a presença de um artefato, pois na redução da complexidade da técnica, foram utilizadas três enzimas diferentes de restrição, sendo uma de corte raro, e duas de corte frequente que resultam em vários sítios espalhados no genoma. Já quando construído com o “pipeline” da empresa, a maioria dos marcadores SSR não foram ligados ao mapa, indicando que existe um “gap” entre as tecnologias para mapeamento destes diferentes tipos de marcador. Resultados similares foram encontrados por Jing et al. (2009) para trigo (*Triticum monococcum*) onde os DArTs e os SSRs tiveram tendência a formar grupos independentes ao longo dos cromossomos. Vários mapas em outras espécies tem sido construídos utilizando estes três tipos de marcadores além da utilização de outros marcadores como AFLPs e RFLPs, sendo incluídos poucos SSRs no mapa (Akbari et al., 2006). Wenzl et al. (2006) descreveram as necessidades de integrar a informação gerada pelos marcadores DArTs com os dados produzidos a partir das tecnologias baseadas na utilização de gel de agarose e poliacrilamida. O autor descreve que foram encontrados resultados com pouca reprodutibilidade no conjunto de dados, com alta densidade, analisados utilizando o programa JoinMap. A dificuldade de analisar grande quantidade de dados utilizando esse programa também foi encontrado por outros autores como Isidore et al. (2003) e Van et al. (2005).

Em média, cada cromossomo apresentou mais do que 420 marcadores em cada grupo, limitando o desenho do mapa e excedendo a capacidade do programa utilizado, o “Mapchart”, software para apresentação gráfica de mapas de ligação e QTLs (Voorrips et al., 2011). O Mapa apresentou um tamanho maior do que os mapas acima citados por outros autores, porém a distância média entre os marcadores foi menor a qualquer outro mapa publicado, evidenciando vantagem para sua utilização nos programas de

melhoramento genético, além de possuir todas as sequências dos marcadores inseridos no mapa SSR-SNP-DArT.

### **5.3 Fenotipagem das plantas em vasos**

Os padrões de respostas fisiológicas e morfológicas das plantas de feijoeiro submetidas a seca em condições de casa de vegetação foram medidas aos 33 dias após o plantio (DAP) (Tabela 7), nesta data, o estresse hídrico das plantas da casa de vegetação não irrigada após 20 dias de restrição hídrica apresentava potencial de água no solo médio de 160 kPa. Durante o experimento, a média de temperatura e umidade do tratamento irrigado foi de 32,1°C e 52,9% e, do tratamento não irrigado foram de 34,5°C e 45,2%, respectivamente.

A análise de variância dos padrões de resposta morfológica relacionada à tolerância à seca detectou diferenças significativas entre os genitores e as RILs para a maioria dos caracteres (Tabela 7). Na casa de vegetação irrigada, os genitores SEA 5 e AND 277, diferiram para massa seca da folha, temperatura da folha, dias para o florescimento, número de vagens, número de sementes por vagem, produtividade, número de sementes, massa de grãos e massa das vagens. Entre as RILs, todas as características foram estatisticamente significativas. Os dois genitores foram semelhantes em termos de clorofila, área foliar, massa fresca da folha, massa fresca do caule e massa seca do caule. A herdabilidade foi menor para massa das vagens (0,28) e maior para massa fresca da folha (0,93) talvez pela menor influência do ambiente sob a segunda característica somado ao erro experimental.

Na casa de vegetação não irrigada, os genitores diferiram para área foliar, massa fresca da folha, massa seca da folha, temperatura da folha, dias para o florescimento, número de vagens, número de sementes por vagem, produtividade, número de sementes, massa das sementes e massa das vagens. Entre as RILs, a massa fresca das folhas e massa seca das folhas foram características não significativas estatisticamente. A herdabilidade foi menor para massa seca da folha (0,12) e maior para produtividade (0,95). O genitor SEA 5 apresentou maior área foliar, massa fresca da folha, massa seca da folha, produtividade e massa de sementes.

Tabela 7. Estatística descritiva para as medidas de características quantitativas na população AND277 x SEA5 de feijoeiro comum do experimento de tratamentos irrigado e não irrigado nos vasos, em casa de vegetação.

Características	Irigado					Não irrigado				
	Genitores			Média das RILs	$h^2_g$	Genitores			Média das RILs	$h^2_g$
	SEA 5	AND 277	Diff			SEA5	AND 277	Diff		
Clorofila	42,46	43	ns	41,69*	0,54	23,85	26,96	ns	26,81*	0,71
Área foliar	2099	2315,3	ns	1402,59*	0,74	390,25	149	*	303,96*	0,87
Massa fresca da folha	32,83	37,5	ns	26,59*	0,93	4,75	1,83	*	3,17 <sup>ns</sup>	0,21
Massa fresca do caule	23,83	24,66	ns	17,07*	0,89	5,87	4,16	ns	4,92*	0,56
Massa seca da folha	4	6,83	*	3,47*	0,64	1,87	1	*	1,01 <sup>ns</sup>	0,12
Massa seca do caule	3,33	4,33	ns	2,04*	0,36	1,25	1,5	ns	1,18*	0,39
Temperatura da folha	23	30,33	*	28,25*	0,61	33,33	35,5	*	34,63*	0,44
Dias até o florescimento	31	37,75	*	36,77*	0,92	38	36	*	36,37*	0,94
Número de vagens	13	5,25	*	10,56*	0,49	13	5	*	14,53*	0,53
Número de sementes/vagem	4,23	2,32	*	2,8*	0,64	4,5	2,7	*	2,75*	0,73
Produtividade (g/planta)	9,46	4,6	*	6,67*	0,9	11,99	3,52	*	7,74*	0,95
Número de sementes	53,33	12,25	*	27,79*	0,35	48	10,5	*	37,65*	0,46
Massa das sementes (g/100 sem.)	21,88	38,68	*	22,02*	0,84	23,18	12,47	*	23,89*	0,73
Massa de vagem	4,14	1,14	*	2,28*	0,28	4,33	2,12	*	3,87*	0,68

\* Diferenças significativas a probabilidade de 0,05, ns: não significativo, Diff: diferenças entre os genitores,  $h^2_g$ : herdabilidade no sentido amplo

No tratamento irrigado, a produtividade média de todos os genótipos foi de 2,84 g/planta e 1,97 g/planta no tratamento sob estresse hídrico. Este resultado mostrou 30% de redução no rendimento de grãos devido a seca, calculado pelo índice de intensidade de seca ( $DII=1-Xds/Xns$ ).

Tolerância à seca é uma característica quantitativa complexa controlada por vários genes de efeito pequeno ou QTLs (Barnabas et al 2008; Fleury et al 2010). Para lidar com a complexidade das respostas das plantas à seca, é vital compreender a base fisiológica e

genética desta resposta (Passioura, 2010). A escolha adequada e descrição do ambiente é fundamental para qualquer esforço de fenotipagem (Poorter et al., 2012a). Por essa razão, e tentando manter as condições ambientais, os experimentos foram realizados em casa de vegetação, com aplicações adequadas de água em condições experimentais. Especial atenção deve ser dada sobre os efeitos do tamanho do recipiente no crescimento e performance das plantas (Pieruschka & Poorter, 2012). Com o objetivo de reduzir o efeito do tamanho do recipiente, recipientes maiores, com 14 kg de capacidade de solo foram usados e não foram realizadas análise de raiz neste sistema.

De modo geral, os pesquisadores usam um dos três tipos de tratamentos impostos para avaliar o estresse hídrico, conhecido em seus experimentos com dois níveis de abastecimento de água (irrigação para não estresse e de seca para o estresse hídrico), para quantificar os efeitos da intensidade e duração da seca no crescimento da cultura e produtividade de sementes (Blair et al., 2012; Sabadin et al., 2012). Neste estudo, o estresse hídrico terminal foi escolhido porque afeta mais de 60% da produção mundial de feijão (White & Singh, 1991) e, as áreas mais afetadas na América Latina estão no Nordeste do Brasil e na América Central. No entanto, apesar da seca terminal ser uma característica das mais sérias, os resultados desta experiência encontraram 30% de diminuição da produção de grãos e, surpreendentemente, algumas das RILs, mesmo com a redução, apresentaram maiores rendimentos médios sob condições de seca terminal. A análise de variância das características quantitativas mostraram que o genitor SEA 5 foi superior em quase todas as características que foram significativas, medidas sob condições de estresse hídrico, exceto para a temperatura da folha, onde foi inferior do que a temperatura do AND 277 (Tabela 7). Ao comparar os dois tratamentos, o desempenho do SEA 5 foi melhor sob estresse hídrico para o número de sementes/vagem, produtividade, massa de sementes e das vagens e o número de vagens foi mantido. Curiosamente para este genótipo, sob irrigação, o valor para os dias até a floração foi muito inferior ao AND 277, e sob condições de estresse atrasou o florescimento, sendo estatisticamente significativa em ambos os casos.

Gonçalves (2013) estudou linhagens de feijoeiro (*Phaseolus vulgaris* L.) tolerantes à seca e os resultados relataram que dentre as cultivares consideradas tolerantes a seca, a cultivar SEA 5 foi superior para várias das características avaliadas. Essa cultivar exibiu

maior potencial hídrico foliar, corroborando com os resultados obtidos por Pimentel & Perez (2000). Neste estudo, o SEA 5 também apresentou a menor redução da produtividade na fase vegetativa (61,3%), e uma das menores reduções com déficit na pre-floração (9,4%) e pós- floração (34,3%). Além disto, a cultivar SEA 5 destacou-se para massa de mil grãos, produtividade de grãos e parâmetros fisiológicos e morfológicos.

### **5.3.1 Identificação de QTLs em vasos no mapa SSR-SNPs.**

De um total de 22 QTLs identificados no experimento, 14 QTLs foram localizados sob condições de irrigação (Tabela 8) e oito sob condições específicas de estresse hídrico (Tabela 9). Um QTL para massa seca da vagem foi expresso apenas sob tratamento com estresse, explicando 14% da variação fenotípica com alelo negativo para o QTL indicando que o genitor SEA 5 contribuiu para esta característica, com um LOD de 3,48, e em especial o marcador BAR3100 localizado próximo ao pico deste QTL. Outros QTLs foram detectados em ambas as condições, mas em diferentes cromossomos (Figura 3). Os maiores efeitos da variação fenotípica foram associados a clorofila, em ambas as condições, com coeficiente de determinação ( $R^2$ ) de 32,8% e 32,1%, respectivamente. No tratamento não irrigado, massa fresca do caule, massa das sementes e número de sementes apresentaram QTLs de maior efeito, respectivamente com  $R^2$  de 18,25%, 17,31% e 15,5%. Esses QTLs foram encontrados ligados ao marcador BM159 no cromossomo Pv03, marcador BAR 3474 no cromossomo Pv01 e marcador BAR3045 do cromossomo Pv07, com a contribuição do alelo do genitor SEA 5 (Tolerante). Foram encontrados alelos positivos e negativos para os QTLs de clorofila e massa fresca do caule, em ambas as condições, indicando que ambos os genitores contribuíram para estas características. Todos os QTLs detectados no tratamento não irrigado apresentaram a contribuição do alelo do genitor SEA 5, com exceção de um de clorofila, identificado no cromossomo Pv11, e um de temperatura da folha identificado no cromossomo Pv07. A massa fresca e seca das folhas, e temperatura da folha, apresentaram um alelo positivo para o AND 277, sob tratamento irrigado. A temperatura da folha teve alelo positivo do AND 277, em ambos os tratamentos. O número de vagens e dias para o florescimento, sob tratamento irrigado, e número de sementes e massa das sementes, em ambos os tratamentos, apresentaram alelo negativo do SEA 5. Alelos positivos e negativos contribuíram para a produtividade (g/planta) no tratamento

irrigado. A maior parte dos locos de características quantitativas identificadas tiveram a sua contribuição de SEA 5 (13), ao invés do genitor AND 277 (9).

Catorze QTLs foram identificados no tratamento irrigado, com valores de  $R^2$  que variam de 10% até 33%, localizados nos cromossomos Pv01, Pv03, Pv04, Pv05, Pv06, Pv07 e Pv11(Tabela 8).

Tabela 8. Descrição dos QTLs identificados no experimento em vasos no tratamento irrigado, via análise de Mapeamento por Intervalo Composto (CIM), utilizando o mapa genético AS com marcadores SSR-SNPs. Os QTLs identificados foram nomeados conforme a característica do QTL encontrado no experimento de vasos. O intervalo no mapa AS (cM) e o marcador do pico dos gráficos de cada QTL, em seus respectivos cromossomo, o LOD de cada pico, efeitos aditivos e efeitos na variância fenotípica ( $R^2$ ) de cada QTL também são apresentadas

Característica	Cromossomo	QTL	Intervalo (cM)	Marcador*	LOD	Efeito aditivo	R(%)
Clorofila	Pv01	C1.1 <sup>AS</sup>	114,7-140,7	BAR3651(0cM)	9,83	0,36	0,33
Clorofila	Pv01	C1.2 <sup>AS</sup>	96,83-111,08	BAR3454 (2cM)	3,47	-0,21	0,11
Massa fresca da folha	Pv07	LBF7.1 <sup>AS</sup>	65,3-101,92	BAR3122 (0cM)	4,06	0,13	0,14
Massa fresca do caule	Pv01	SBF1.1 <sup>AS</sup>	81,43-102,44	IAC11 (0cM)	3,27	0,07	0,10
Massa fresca do caule	Pv07	SBF7.1 <sup>AS</sup>	65,3-105,92	BAR4417 (0cM)	3,76	0,08	0,13
Massa seca da folha	Pv07	LBD7.1 <sup>AS</sup>	71,3-103,92	BAR4417 (1cM)	3,02	0,04	0,11
Temperatura da folha	Pv04	LT4.1 <sup>AS</sup>	11,41-42,23	BAR4881 (0cM)	3,35	0,61	0,12
NºVagens	Pv07	NP7.1 <sup>AS</sup>	73,3-102,92	BAR3682 (0 cM)	3,21	-0,06	0,13
NºSementes	Pv06	NS6.1 <sup>AS</sup>	77,71-102,54	PVM21 (5 cM)	3,31	-0,38	0,13
Massa de semente (g/100 sem.)	Pv05	SW5.1 <sup>AS</sup>	42,67-116,38	BAR4677 (5 cM)	3,91	-0,29	0,14
Dias até o florescimento	Pv01	DF1.1 <sup>AS</sup>	117,7-140,76	BAR4823 (0 cM)	3,17	-0,01	0,11
Dias até o florescimento	Pv03	DF3.1 <sup>AS</sup>	74,55-159,28	BAR3097 (0 cM)	5,09	-0,02	0,14
Produtividade (g/planta)	Pv01	PRD1.1 <sup>AS</sup>	81,43-98,87	BAR3084 (0 cM)	3,18	0,19	0,10
Produtividade (g/planta)	Pv11	PRD11.1 <sup>AS</sup>	50,03-75,83	BAR4938 (0 cM)	3,49	-0,19	0,12

\* Entre parênteses está a distância do marcador ao pico do QTL;

No experimento dos vasos os resultados mostraram a identificação de oito QTLs no ambiente de seca (Tabela 9) explicando diferentes valores (0,09% a 32%) do porcentual da variabilidade genética, e localizados nos cromossomos Pv01, Pv03, Pv06, Pv07, Pv08 e

Pv11. Todos os QTLs identificados em condições de seca tiveram o alelo do genitor SEA 5 exceto o QTL da temperatura foliar (LT7.1<sup>AS</sup>). QTLs foram identificados nos 11 cromossomos, exceto no cromossomo Pv02, Pv09 e Pv10. QTLs sobrepostos nos dois tratamentos foram identificados nos cromossomos Pv01, Pv3, Pv06, Pv07 e Pv11 para diferentes características. Blair et al (2010) também encontraram alguns QTLs sobrepostas e sugeriram pleiotropia de genes que controlam duas ou mais características. No presente estudo, os QTLs para massa fresca da folha, massa fresca do caule, massa seca da folha e o número de vagens agruparam no cromossomo Pv07 em condição irrigada. Apesar do QTL LBD7.1AS para massa seca da folha apresentar um maior intervalo, mostrou o mesmo marcador (BAR4417) no pico do gráfico do QTL SBF7.1<sup>AS</sup> para massa fresca do caule.

Tabela 9. Descrição dos QTLs identificados no experimento de vasos no tratamento não irrigado relacionados à tolerância à seca encontrados via análise de Mapeamento por Intervalo Composto (CIM), utilizando o mapa genético AS de feijoeiro comum. Os QTLs identificados foram nomeados conforme a característica do QTL encontrado no experimento de vasos. O intervalo no mapa AS (cM) e o marcador do pico dos gráficos de cada QTL, em seus respectivos cromossomo, o LOD de cada pico, efeitos aditivos e efeitos na variância fenotípica ( $R^2$ ) de cada QTL também são apresentadas

Característica	Cromossomo	QTL	Intervalo (cM)	Marcador*	LOD	Efeito aditivo	R (%)
Clorofila	Pv06	C6.1 <sup>AS</sup>	95,54-140,45	BAR5885 (0cM)	10,83	-3,87	0,32
Clorofila	Pv11	C11.1 <sup>AS</sup>	56,23-78,68	BAR3594(0cM)	4,30	2,14	0,11
Massa fresca do caule	Pv03	SBF3.1 <sup>AS</sup>	72,55-98,25	BM159 (2cM)	4,14	-0,11	0,18
Temperatura foliarF	Pv07	LT7.1 <sup>AS</sup>	140,9-171,98	BAR2897(6 cM)	3,04	0,57	0,14
NºSementes	Pv07	NS7.1 <sup>AS</sup>	29-54,34	BAR3045 (3 cM)	4,08	-0,52	0,16
NºSementes	Pv08	NS8.1 <sup>AS</sup>	39,86-63,42	BAR4250 (8 cM)	4,08	-0,04	0,09
Massa de semente (g/100 sem.)	Pv01	SW1.1 <sup>AS</sup>	54,88-76,87	BAR3474 (0 cM)	4,39	-0,12	0,17
Massa seco da vagem	Pv11	PWD11.1 <sup>AS</sup>	56,23-77,52	BAR3100 (0 cM)	3,48	-0,15	0,14

\* Entre parênteses está a distância do marcador ao pico do QTL;

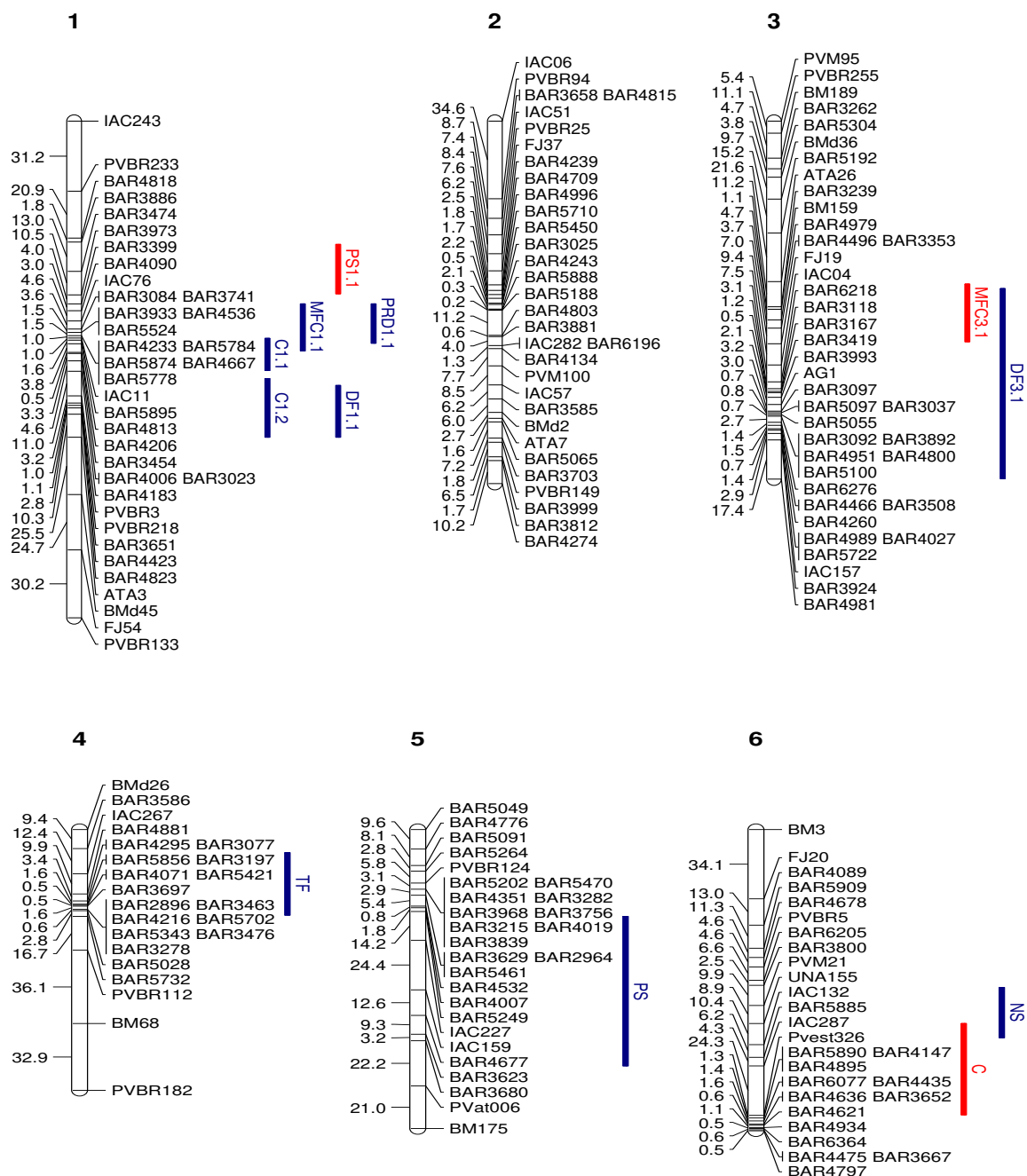


Figura 3. Mapa genético AS de feijão comum mostrando a localização dos 22 QTLs encontrados relacionados às características de tolerância à seca e condições normais de irrigação, no experimento de vasos. Os QTLs em azul são derivados do experimento em casa de vegetação irrigada e os QTLs em vermelho são derivados dos experimentos em casa de vegetação sob estresse. C=clorofila, PSV= Massa seco da vagem, PRD= produtividade, TI= temperatura foliar, MFF=massa fresca da folha, MFC=massa fresca do



caule, MSF= massa seca da folha, NV= número de vagens, NS= número de sementes, PS=massa das sementes, DF= dias até o florescimento.

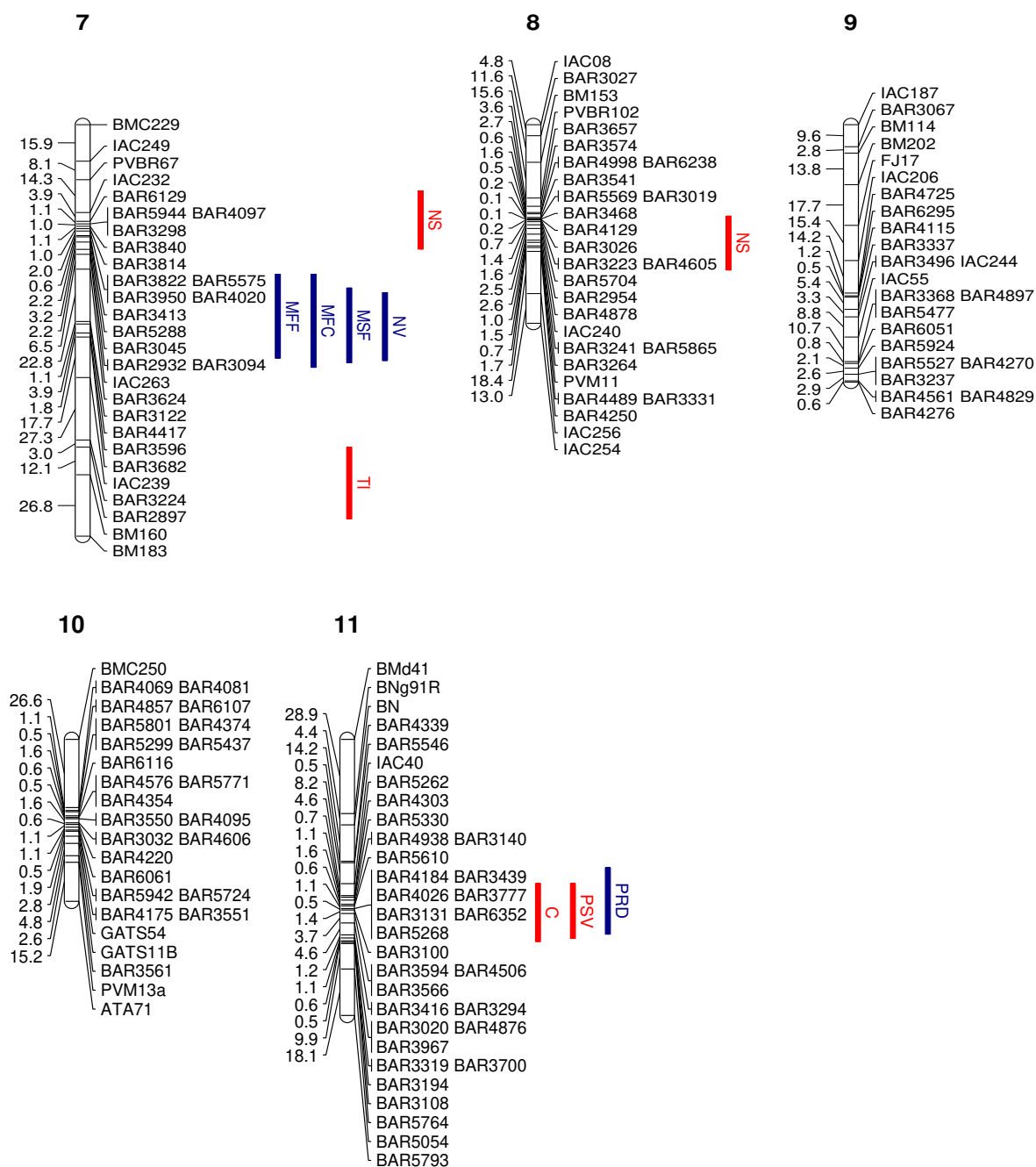


Figura 3. Mapa genético AS de feijão comum mostrando a localização dos 22 QTLs encontrados relacionados às características de tolerância à seca e condições normais de irrigação, no experimento de vasos. Os QTLs em azul são derivados do experimento em casa de vegetação irrigada e os QTLs em vermelho são derivados dos experimentos em

casa de vegetação sob estresse. C=clorofila, PSV= massa seco da vagem, PRD= produtividade, TI= temperatura foliar, MFF=massa fresca da folha, MFC=massa fresca do caule, MSF= massa seca da folha, NV= número de vagens, NS= número de sementes, PS=massa das sementes, DF= dias até o florescimento.

Para clorofila, massa fresca do caule, temperatura da folha, número de sementes e massa de sementes, QTLs foram identificados em ambos os tratamentos. Para os dias até o florescimento, massa fresca da folha, massa seca da folha, número de vagens e produtividade foram detectados QTLs apenas no tratamento irrigado. Para massa seca da vagem, um QTL foi identificado no tratamento não irrigado. A característica de massa seca das sementes foi importante, uma vez que o enchimento das sementes é inibida sob estresse hídrico, pode-se dizer que sementes pesadas podem indicar tolerância à seca e levar a um maior rendimento de grãos e/ou produtividade (Ramírez-Vallejo & Kelly, 1998). Além disso, o alelo para o tamanho de semente, nos dois tratamentos, foi oriundo do genitor tolerante à seca, SEA 5. Para massa de sementes, foram descobertos dois QTLs nas análises, um no cromossomo Pv01 e o outro no cromossomo Pv05. Blair et al. (2012) também encontraram QTLs para massa de sementes no cromossomo Pv05 em condição de seca e Broughton et al. (2003) encontraram QTLs no cromossomo Pv01 nas mesmas condições. QTLs para dias até o florescimento foram encontrados no cromossomo Pv01 e Pv03, enquanto Broughton et al. (2003) localizaram QTLs no cromossomo Pv01 e Pv08.

### **5.3.2 Identificação de QTLs em vasos no mapa SSR-SNPs-DArTs**

Foram identificados 18 QTLs no tratamento não irrigado pela empresa DArT Pty utilizando o próprio “pipeline” da empresa. QTLs para clorofila, área foliar, massa fresca do caule, massa seca da folha, massa seca do caule, temperatura da folha, número de sementes por vagem, massa de 100 sementes e dias até o florescimento foram detectados. Os QTLs de maior efeito foram para clorofila (C6.2<sup>AS</sup>), área foliar (LA3.1<sup>AS</sup>), massa fresca do caule (SBF3.1<sup>AS</sup>), número de sementes por vagem (NSP10.1<sup>AS</sup>, NSP 1.1<sup>AS</sup>, NSP5.1<sup>AS</sup>) e massa de 100 sementes (SW1.1<sup>AS</sup>) (Tabela 10).

Tabela 10. Descrição dos QTLs identificados no experimento em vasos no tratamento não irrigado relacionados à tolerância à seca encontrados via análise de “Statistical Machine Learning” (SML), utilizando o mapa genético AS de feijão comum com os marcadores SSR-SNP-DArT. Os QTLs identificados foram nomeados conforme a característica do QTL encontrado no experimento de vasos não irrigado. O marcador do pico dos gráficos de cada QTL, em seus respectivos cromossomos, o PAVE ou significância e a posição de cada QTL também são apresentados.

Característica	Cromossomo	QTL	Marcador	PAVE	Posição (cM)
Clorofila	Pv06	C6.2 <sup>AS</sup>	100025029 F 0	0,051	106,78
Clorofila	Pv06	C6.2 <sup>AS</sup>	100018482 F 0	0,026	106,78
Área foliar	Pv03	LA3.1 <sup>AS</sup>	100015192 F 0	0,022	118,78
Área foliarF	Pv03	LA3.2 <sup>AS</sup>	100027180 F 0	0,010	122,63
Massa fresca do caule	Pv03	SBF3.1 <sup>AS</sup>	100004453 F 0	0,021	289,49
Massa fresca do caule	Pv03	SBF3.2 <sup>AS</sup>	100028860 F 0	0,018	250,65
Massa seca da folha	Pv08	LBD8.1 <sup>AS</sup>	BARC-PV-0004144	0,016	212,86
Massa seca da folha	Pv08	LBD8.2 <sup>AS</sup>	BARC-PV-0002954	0,011	210,48
Massa seca do caule	Pv02	SBD2.1 <sup>AS</sup>	100027898 F 0	0,012	12,62
Temperatura da folha	Pv08	LT8.1 <sup>AS</sup>	100010997 F 0	0,011	50,94
NºSementes/vagem	Pv10	NSP10.1 <sup>AS</sup>	100041132 F 0	0,028	68,16
NºSementes/vagem	Pv01	NSP1.1 <sup>AS</sup>	100035462 F 0	0,021	295,46
NºSementes/vagem	Pv05	NSP5.1 <sup>AS</sup>	100022390 F 0	0,021	150,97
NºSementes/vagem	Pv04	NSP4.1 <sup>AS</sup>	100043301 F 0	0,011	215,95
NºSementes/vagem	Pv01	NSP1.1 <sup>AS</sup>	100030921 F 0	0,010	217,88
Massa Semente (g/100s)	Pv01	SW1.2 <sup>AS</sup>	100008797 F 0	0,064	38,08
Dias até o florescimento	Pv05	DF5.1 <sup>AS</sup>	100030163 F 0	0,016	6,00
Dias até o florescimento	Pv01	DF1.1 <sup>AS</sup>	100016108 F 0	0,014	205,65

PAVE: estatisticamente significativo  $\geq 0.1$

A metodologia de análise utilizada (SML) na identificação de QTLs a partir do mapa com marcadores SSR-SNP e DArT, evidenciou QTLs diferentes aos identificados pelo método de Mapeamento por Intervalo Composto (CIM), exceto para o QTL de clorofila C6.2<sup>AS</sup>, identificado no cromossomo Pv06, na posição 106,78 cM, no mapeamento SML e que coincide com o QTL C6.1<sup>AS</sup>, no cromossomo Pv06, identificado pela metodologia CIM, que esta localizado no intervalo de 95,54-140,45 cM, no tratamento não

irrigado. A metodologia “Statistical Machine Learning” detectou maior número de QTLs (18), do que o mapeamento por Intervalo Composto (8) nas condições de déficit hídrico.

Bedo et al. (2008) compararam o desempenho das metodologias SML (*Statistical Machine Learning*), CIM (*Composite Interval Mapping*), BIM (*Bayesian Interval Mapping*) e MR (*Single Marker Regression*) em um conjunto de dados de uma progênie proveniente de um cruzamento entre duas cultivares de cevada. Os resultados mostraram que o método SML identificou os QTLs com precisão enquanto o método BIM subestimou o número de QTLs. Os QTLs identificados por SML coincidiram com a localização daqueles conhecidos na literatura. A metodologia SML evidenciou aproximadamente metade dos QTLs relatados pelas análises de CIM e MR, o que é esperado já que estas metodologias não utilizam testes independentes. Estes dois últimos métodos também foram suscetíveis a produzir estimativas excessivamente otimistas dos efeitos dos QTLs. A resolução do QTL (definição do pico) oferecida pelo SML foi consistentemente superior aos métodos MR, CIM e BIM. Os autores concluem que a metodologia SML produz melhores estimativas dos efeitos dos QTLs porque elimina a tendência otimista no desempenho preditivo dos outros métodos de identificação de QTLs, além de produzir picos mais estreitos do que os outros métodos, exceto o método BIM, e, portanto, identifica QTLs com maior precisão. É mais robusto para genotipagem e mapeamento de ligação e identifica marcadores ligados a QTLs na ausência de mapa genético.

#### **5.4 Fenotipagem das plantas em rizotron**

As plantas em condições de estresse foram mantidas 27 dias sem abastecimento de água e as plantas do tratamento controle foram irrigadas duas vezes ao dia com 200 ml de água. O potencial matricial do solo foi mantido entre 0 - 30 kPa, o tempo todo para o tratamento irrigado, e para os primeiros 25 dias após o plantio, no tratamento não irrigado. Depois disso, a água não foi mais adicionada e o potencial matricial do solo começou a aumentar rapidamente até atingir 199 kPa, em quatro dias (sem água), para a maioria do solo das plantas e manteve-se até ao final do experimento em 52 dias após o plantio. A temperatura média das folhas do tratamento não irrigado foi de 22,9 °C e 22,6 °C para o

tratamento irrigado, a temperatura ambiental média da casa de vegetação foi de 25,5 °C e a umidade de 33,8%.

Os resultados das análises de variância aplicados nas 11 características entre os genitores e as RILs, nos dois tratamentos, mostraram diferenças significativas (Tabela 11).

Tabela 11. Análise de variância para características quantitativas dos genitores SEA5, AND277 e das linhagens recombinantes da população AS de feijão comum em casa de vegetação irrigada e não irrigada nos rizotrons.

Característica	Irigado						Não irrigado					
	Genitores			Média das RILs	h <sup>2</sup> <sub>g</sub>		Genitores			Média das RILs	h <sup>2</sup> <sub>g</sub>	
	SEA5	AND 277	Diff				SEA 5	AND277	Diff			
Clorofila	40,55	40,1	ns	39,84*	0.85		42,63	36,25	ns	38,17*	0,61	
Área foliar	485	594,8	ns	391,99*	0.78		410,63	179,1	*	217,47*	0,30	
Massa fresca da folha	5,68	7,25	ns	4,23*	0.75		3,61	1,82	*	1,99 <sup>ns</sup>	0,05	
Massa fresca do caule	4,55	4,48	ns	4,13 <sup>ns</sup>	0.17		2,31	3,23	ns	2,20 <sup>ns</sup>	0,01	
Massa seca da folha	0,81	1,9	*	0,93 <sup>ns</sup>	0.13		1,41	0,78	ns	0,66 <sup>ns</sup>	0,11	
Massa seca do caule	0,41	0,98	ns	0,7 <sup>ns</sup>	0.06		0,73	1,59	*	0,53 <sup>ns</sup>	0,13	
Temperatura da folha	19,5	19	ns	19,08 <sup>ns</sup>	0.2		21,66	19	*	20,86*	0,40	
Comprimento total de raízes (cm planta <sup>-1</sup> )	1248	2222,5	ns	1847,81 <sup>ns</sup>	0.07		2371,4	1079,6	*	1775,72*	0,3	
Área de superfície radicular	135,62	205,33	ns	181,97*	0.45		239,63	93,89	*	198,54*	0,82	
Volume de raízes (cm <sup>3</sup> planta <sup>-1</sup> )	1,17	1,5	ns	1,49*	0.64		1,41	0,59	*	0,95*	0,21	
Diâmetro de raízes (mm)	0,35	0,29	ns	0,3 <sup>ns</sup>	0.01		0,26	0,25	ns	0,26 <sup>ns</sup>	0,07	

\* Diferenças significativas à probabilidade de 0,05, ns: não significativo, Diff: diferenças entre os genitores, h<sup>2</sup><sub>g</sub>: herdabilidade no sentido amplo

Entre os genitores, foram encontradas diferenças para área foliar, massa fresca da folha, massa seca do caule, temperatura da folha, comprimento total de raízes, área superficial da raiz, e para o volume de raiz, e não significativas para clorofila, massa fresca do caule, massa seca da folha e diâmetro de raiz no tratamento não irrigado. Por outro lado, a única característica no tratamento irrigado estatisticamente significativa entre os genitores foi a massa seca da folha. Entre as linhagens endogâmicas recombinantes foram encontradas diferenças estatisticamente significativas para clorofila, área foliar, temperatura

da folha, comprimento total de raízes e volume da raiz. As diferenças foram não significativas para massa fresca da folha, massa fresca do caule, massa seca da folha e do caule e diâmetro de raiz para o tratamento de estresse hídrico. As diferenças para as RILs foram significativas para clorofila, área foliar, massa fresca da folha e volume de raiz e não significativas para massa fresca do caule, massa seca da folha e do caule, temperatura da folha, comprimento total de raiz e diâmetro da raiz no tratamento irrigado.

O genitor SEA 5 foi melhor para todas as características estatisticamente significativas, com exceção da massa seca do caule no tratamento não irrigado. Ele se comportou da mesma forma como quando exposto a condições normais de água, com exceção de comprimento da raiz e área superficial da raiz, em que cresceu muito mais em condições de estresse do que em condições normais. O genitor AND 277 reduziu significativamente a área foliar, massa fresca da folha, comprimento total de raiz, área superficial e volume de raiz no tratamento não irrigado, enquanto o SEA 5 manteve os valores das características tão semelhantes quanto ao tratamento irrigado, apresentando maior comprimento da raiz e área superficial radicular.

A herdabilidade no sentido amplo calculada variou entre as características e os tratamentos e foi relativamente moderadas (Tabela 10). Clorofila no tratamento irrigado e a área superficial da raiz no tratamento não irrigado tiveram os maiores valores de herdabilidade ( $h^2=0,85$  e  $0,82$ ), respectivamente, enquanto que a massa fresca da folha e do caule, e o diâmetro da raiz tiveram os menores valores de herdabilidade ( $h^2=0,05$ ,  $0,01$  e  $0,07$ ), respectivamente, no tratamento não irrigado.

O melhoramento clássico para tolerância a seca é baseado na seleção para produção e seus componentes em ambientes com limitação de água (Blum 1985; Millan et al., 2006). Os germoplasmas são geralmente testados para duas características importantes em seca, sistema radicular cumprido e tamanho da área foliar (Turner et al., 2001; Saxena 2003). O objetivo do experimento em rizotron foi a análise da folha, caule e raiz do sistema como fator importante na planta para adaptação a seca. A maior parte dos métodos usados para avaliar raízes não são só demorados e trabalhosos, mas também são de caráter destrutivo para a planta (Beebe et al., 2006).

A estratégia adotada neste estudo no experimento da avaliação de raiz foi a utilização da linhagem avançada para tolerância a seca SEA 5 cruzada com o genótipo suscetível ao estresse hídrico, AND 277, gerando linhagens endogâmicas recombinantes para detectar os QTLs de maior efeito envolvidos no desenvolvimento da raiz e que possam ser utilizados na seleção indireta em programas de melhoramento genético para a adaptação à seca. Para as características mensuradas em condições irrigadas, os genitores tiveram quase o mesmo desempenho nas características estatisticamente não significativas, exceto para massa seca da folha. Enquanto que no tratamento não irrigado, foi evidente a superioridade do genótipo SEA 5. Por exemplo, a área foliar do AND 277 apresentou uma redução de 70% enquanto que para o SEA 5 foi de apenas 15%, mostrando uma melhor resposta ao estresse hídrico. Para massa fresca da folha, a redução do AND 277 foi de 75% enquanto para o SEA 5 foi 36 %. O comprimento total de raiz para o SEA 5 foi quase o dobro quando comparado na condição irrigada e apresentando uma redução de 50% para o AND 277. O genótipo SEA 5 foi previamente descrito por ter raízes pivotantes profundas (CIAT 2004) o que pode ter um papel fundamental no desempenho da produtividade (White & Castillo, 1992). Estes resultados confirmaram os resultados encontrados por Singh & Teran (2002) e por Asfaw & Blair (2012) demonstrando o nível superior do genótipo SEA 5 para tolerância a seca. Recentemente, o estudo de Gonçalves (2013), também mostrou que a cultivar SEA 5 possui comprimento de raiz diferenciado conforme a fase de aplicação do déficit hídrico. Taiz & Zeiger (2009) relataram que o aumento das raízes nas camadas mais profundas do solo durante o período de restrição hídrica exige alocação de assimilados para seus ápices em crescimento, no entanto, durante este período, há uma exigência de assimilados entre raízes e frutos, o que explica o fato pelo qual as plantas são mais sensíveis ao déficit hídrico no período reprodutivo.

#### **5.4.1 Identificação de QTLs em rizotron no mapa SSR-SNPs.**

O nome do QTL, número, localização e efeitos associados com a tolerância a seca foram resumidos na Tabela 12. A identificação dos QTLs foi realizado com 331 *loci* do mapa genético incluindo SSRs e SNPs, cobrindo uma distância total do genoma do feijão comum de 1.515,2 cM. Os QTLs foram localizados nos cromossomos Pv01, Pv02, Pv03, Pv05 e Pv06. Um total de 19 QTLs foram identificados para 11 características em resposta

ao estresse hídrico e tratamento irrigado utilizando as análises de mapeamento por intervalo composto (CIM). A maioria dos QTLs foram identificados no tratamento irrigado (15) enquanto quatro QTLs foram identificados no tratamento não irrigado (Tabela 13). Entre os 13 QTLs identificados de maior efeito ( $> 15\%$ ), um foi para clorofila (15,66%) no tratamento irrigado, no cromossomo Pv03, um para a área foliar (20,30%), no tratamento não irrigado, três de massa fresca do caule (21,73%, 15,88%, 15,77%), um para massa fresca da folha (26,67%), um para massa seca da folha (16,80%), um para massa seca do caule (22,10%), um para o comprimento das raízes (19,20%), um para volume de raízes (16,60%) e dois para a área superficial da raiz (18,60% e 32,60%). A maioria dos QTLs com maiores efeitos (C3.1<sup>AS</sup>, LA1.1<sup>AS</sup>, SBF1.1<sup>AS</sup>, SBF3.3<sup>AS</sup>, LBF1.1<sup>AS</sup>, LBD1.1<sup>AS</sup>, SBD1.1<sup>AS</sup>, RL2.1<sup>AS</sup> e RSA1.1<sup>AS</sup>) apresentaram a contribuição do alelo do genitor SEA 5. Os QTLs LA1.1<sup>AS</sup>, SBF1.1<sup>AS</sup>, LBF1.1<sup>AS</sup>, SBD1.1<sup>AS</sup> e RSA1.1<sup>AS</sup> se localizaram no mesmo intervalo de marcador e com o mesmo marcador mais próximo ao pico do QTL, o PVBR3 no cromossomo Pv01, o QTL SBF3.3<sup>AS</sup> e SBD3.1<sup>AS</sup> com marcador BM189 no cromossomo Pv03 e os QTLs RL6.1<sup>AS</sup> e RSA6.1<sup>AS</sup> com marcador BM 3, no cromossomo Pv06. QTLs foram identificados para todas as características avaliadas, com exceção da temperatura foliar.



Tabela 12. Identificação de QTLs no experimento em rizotron no tratamento irrigado via análise CIM utilizando o mapa genético AS de feijão comum a partir do mapa com marcadores SSR-SNP

Característica	QTL	Cromossomo	Intervalo	Marcador	LOD	Threshold	Efeito aditivo	R(%)
Clorofila	C3.1 <sup>AS</sup>	Pv03	120,3-138,9	BAR5055	4,52	3,04	-0,25	15,66
Clorofila	C5.1 <sup>AS</sup>	Pv05	46,6-67,8	IAC227	3,26	3,04	-0,15	10,55
Massa fresca do caule	SBF1.1 <sup>AS</sup>	Pv01	111,7-140,7	PVBR3	7,67	3,20	-0,12	21,73
Massa fresca do caule	SBF3.1 <sup>AS</sup>	Pv03	6,3-33	BM189	3,7	3,20	0,08	9,28
Massa fresca do caule	SBF3.2 <sup>AS</sup>	Pv03	50,9-70,9	BAR5192	3,5	3,20	0,12	15,88
Massa fresca do caule	SBF3.3 <sup>AS</sup>	Pv03	82,6-108,2	BAR3353	5,12	3,20	-0,12	15,77
Massa fresca do caule	SBF6.1 <sup>AS</sup>	Pv06	38,1-57	BAR4089	4,14	3,20	0,09	10,99
Massa fresca da folha	LBF1.1 <sup>AS</sup>	Pv01	107-140,7	PVBR3	8,92	3,15	-0,26	26,67
Massa seca da folha	LBD1.1 <sup>AS</sup>	Pv01	111,7-140,7	BAR4423	5,18	3,12	-0,13	16,80
Massa seca do caule	SBD1.1 <sup>AS</sup>	Pv01	113,7-140,7	PVBR3	6,4	3,10	-0,12	22,10
Massa seca do caule	SBD3.1 <sup>AS</sup>	Pv03	5,3-34	BM189	3,1	3,10	0,06	9,80
Comprimento raiz	RL2.1 <sup>AS</sup>	Pv02	50,6-72,7	PVBR25	5,58	3,21	-6,33	19,20
Volume da raiz	RV6.1 <sup>AS</sup>	Pv06	0-49	FJ20	3,31	2,99	0,40	16,60
Area superficial da raiz	RSA1.1 <sup>AS</sup>	Pv01	111-140	PVBR3	5,02	3,14	-1,96	18,60
Diametro da raiz	RD6.1 <sup>AS</sup>	Pv06	47-85	PVBR5	3,49	2,94	0,01	13,64

Os quatro QTLs identificados no tratamento não irrigado foram de maior efeito e interessantes para o melhoramento genético destas características associadas a tolerância a seca (Tabela 13). Os QTLs foram localizados nos cromossomos Pv01 e Pv06 (Figura 4).

Tabela 13. Identificação de QTLs no experimento em rizotron no tratamento não irrigado relacionados a tolerância a seca encontrados via análise CIM, utilizando o mapa genético AS de feijão comum com os marcadores SSR-SNPs

Característica	QTL	Cromossomo	Intervalo (cM)	Marcador	LOD	Threshold	Efeito aditivo	R(%)
Área foliar	LA1.1 <sup>AS</sup>	Pv01	111,7-140,7	PVBR3	6,52	3,15	-0,70	20,30
Comprimento raiz	RL1.1 <sup>AS</sup>	Pv01	131,7-156,1	ATA3	3,19	3,09	-3,12	9,75
Comprimento raiz	RL6.1 <sup>AS</sup>	Pv06	0-34	BM3	3,28	3,09	4,37	19,20
Area superficial raiz	RSA6.1 <sup>AS</sup>	Pv06	0-57	BM3	5,18	3,00	0,33	32,60

Um dos mecanismos mais importantes na respostas a adaptação a seca é a absorção eficiente da água através de um sistema radicular profundo e equilibrado (Kamoshita et al., 2008), Porém a seleção fenotípica para características de raiz é um processo lento e de trabalho intenso o que justifica a busca de estratégias alternativas como a identificação de QTLs em casa de vegetação, seguido pela seleção indireta através do uso de marcadores moleculares (Asfaw & Blair, 2012). Três QTLs identificados no tratamento com déficit hídrico foram associados aos mecanismos de adaptação a seca no experimento dos rizotrons, sendo que dois foram para comprimento de raiz (RL1.1<sup>AS</sup> e RL6.1<sup>AS</sup>) e um para área foliar (LA1.1<sup>AS</sup>).

Os resultados encontrados no experimento dos vasos também identificaram QTLs de maior efeito para área foliar confirmando esta característica como mecanismo de adaptação a seca. Os QTLs identificados para área foliar e comprimento de raiz no tratamento não irrigado tiveram a contribuição do alelo do genitor SEA 5. Estes resultados mostram uma vez mais a superioridade da cultivar SEA 5 para tolerância a seca. Gonçalves (2013) relatou no seu estudo que a cultivar SEA 5 foi a que mais contribuiu para o incremento da produtividade de grãos em condições de deficiência hídrica e se destacou como genitor por apresentar os maiores valores positivos para a capacidade geral de combinação para produtividade de grãos e massa de mil grãos além de ser superior nos parâmetros fisiológicos e morfológicos, mostrando-se como cultivar potencial para compor blocos de cruzamento para tolerância a seca.

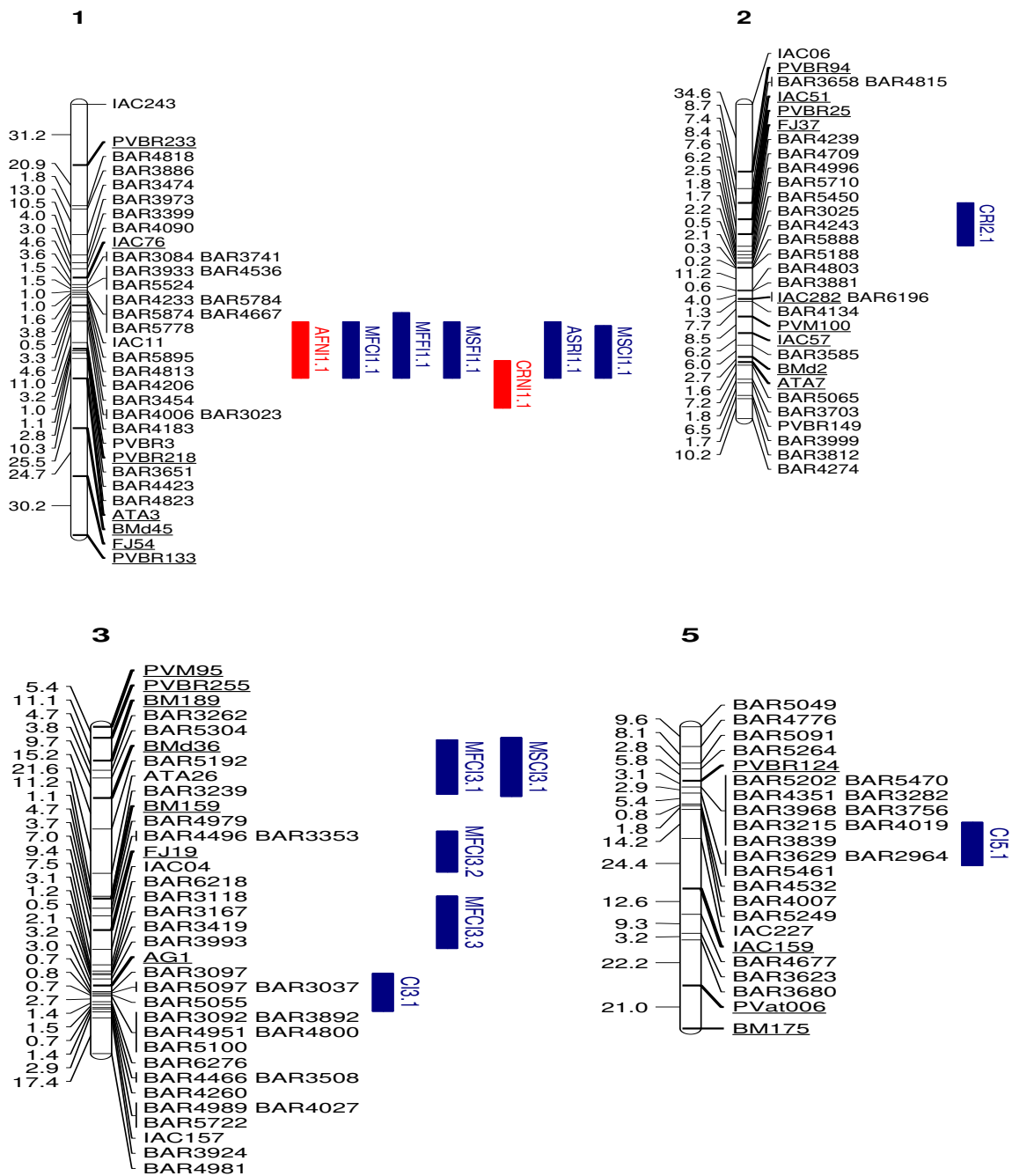


Figura 4. Mapa genético AS de feijão comum mostrando a localização dos 19 QTLs encontrados a partir do mapa com marcadores SSR-SNP relacionados às características de tolerância à seca e condições normais de irrigação no experimento de rizotron. Os QTLs em azul são derivados do experimento em casa de vegetação irrigada e os QTLs em vermelho

são derivados dos experimentos em casa de vegetação com estresse. Cl=clorofila, MFF=massa fresca da folha, MFC=massa fresca do caule, MSF= massa seca da folha, MSC=massa seca do caule, ASR=área superficial da raiz, CR=comprimento de raiz, AF=área foliar, VR=volume da raiz, DR=diâmetro da raiz.

6

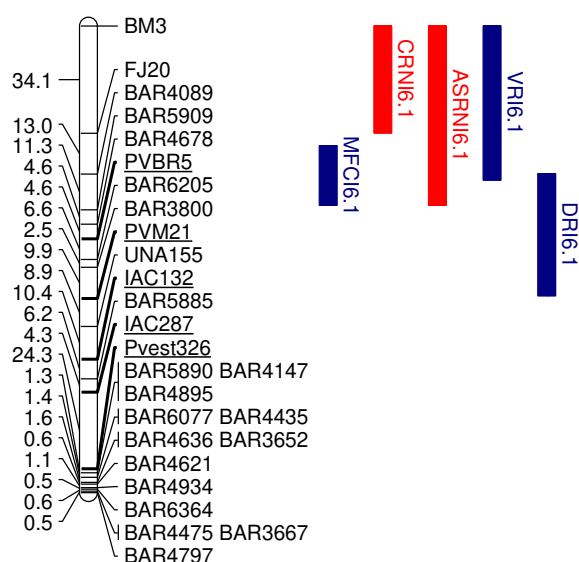


Figura 4. Mapa genético AS de feijão comum mostrando a localização dos 19 QTLs encontrados a partir do mapa com marcadores SSR-SNP relacionados às características de tolerância à seca e condições normais de irrigação no experimento de rizotron. Os QTLs em azul são derivados do experimento em casa de vegetação irrigada e os QTLs em vermelho são derivados dos experimentos em casa de vegetação com estresse. Cl=clorofila, MFF=massa fresca da folha, MFC=massa fresca do caule, MSF= massa seca da folha, MSC=massa seca do caule, ASR=área superficial da raiz, CR=comprimento de raiz, AF=área foliar, VR=volume da raiz, DR=diâmetro da raiz.

#### 5.4.2 Identificação de QTLs em rizotron a partir do mapa SSR-SNPs-DArT.

Foram identificados 10 QTLs pela empresa DArT Pty para o tratamento irrigado para as características clorofila, área foliar, massa fresca da folha, massa fresca do caule, massa seca da folha, comprimento de raiz, área superficial da raiz e temperatura foliar sendo os de maior efeito para área foliar (LA1.1<sup>AS</sup>), massa fresca do caule (SBF3.1<sup>AS</sup>), massa seca da folha (LBD 1.1<sup>AS</sup>), comprimento da raiz (RL1.1<sup>AS</sup>) e área superficial da raiz (RSA10.1<sup>AS</sup>) (Tabela 14).

Tabela 14. Identificação de QTLs no experimento em rizotron no tratamento irrigado encontrados via análise SML a partir do mapa com os marcadores SSR-SNP-DArT

Característica	QTL	Cromossomo	Marcador	PAVE	Posição cM
Clorofila	C9.1 <sup>AS</sup>	Pv09	100025821	0,013	325,12
Área foliar	LA1.1 <sup>AS</sup>	Pv01	100032102	0,022	213,14
Massa fresca da folha	LBF1.1 <sup>AS</sup>	Pv01	100032102	0,011	213,14
Massa fresca do caule	SBF3.1 <sup>AS</sup>	Pv03	100016086	0,020	390,68
Massa fresca do caule	SBF6.1 <sup>AS</sup>	Pv06	100089317	0,011	261,69
Massa seca da folha	LBD1.1 <sup>AS</sup>	Pv01	100024096	0,023	213,14
Massa seca da folha	LBD1.2 <sup>AS</sup>	Pv01	100030811	0,016	233,34
Comprimento de raiz	RL1.1 <sup>AS</sup>	Pv01	100009123	0,020	202,03
Área superficial da raiz	RSA10.1 <sup>AS</sup>	Pv10	100021789	0,023	80,35
Temperatura foliar	LT1.1 <sup>AS</sup>	Pv01	100013136	0,013	213,14
Temperatura foliar	LT9.1 <sup>AS</sup>	Pv09	100015358	0,011	156,38

PAVE: estatisticamente significativo  $\geq 0.1$

Para o tratamento não irrigado foram identificados cinco QTLs, três para área foliar, um para massa fresca do caule e outro para volume da raiz. O QTL de maior efeito foi para área foliar (LA8.1AS) no cromossomo oito (Tabela 15).

Seleção fenotípica para as características de raiz são processos lentos e trabalhosos, o que justifica a busca de estratégias alternativas, como a identificação de QTLs em casas de vegetação seguidos de seleção indireta através do uso de marcadores de alta reprodutibilidade. A associação das características da raiz e marcadores genéticos por meio de uma abordagem de mapeamento de QTLs em uma população de RILs estável avaliadas em experimentos com repetições e devidamente conduzidos, pode contribuir para uma melhor compreensão da variabilidade das características hereditárias de interesse para a raiz

(Asfaw & Blair. 2012). Esta abordagem foi utilizada em estudos de controle genético para características de raiz no melhoramento de tolerância a seca em arroz (*Oryza sativa*; Steel et al. 2006; Courtosis et al. 2009), milho (*Zea mays*; Tuberosa et al. 2007; Ruta et al. 2010) e grão de bico (*Cicer arietinum*; Vadez et al. 2008; Gaur et al. 2008).

Tabela 15. Identificação de QTLs no experimento em rizotron no tratamento não irrigado encontrados via análise SML a partir do mapa com os marcadores SSR-SNP-DArT

Característica	QTL	Cromossomo	Marcador	PAVE	Posição cM
Área foliar	LA8.1 <sup>AS</sup>	Pv08	100023528	0,03	309,76
Área foliar	LA3.1 <sup>AS</sup>	Pv03	100006334	0,01	454,45
Área foliar	LA2.1 <sup>AS</sup>	Pv02	100018180	0,01	31,24
Massa fresca do caule	SBF8.1 <sup>AS</sup>	Pv08	100015514	0,01	326,24
Volume da raiz	RV2.1 <sup>AS</sup>	Pv02	100088280	0,01	401,25

PAVE: estatisticamente significativo  $\geq 0.1$

Em feijão comum, análises de QTLs foram utilizados na análise de pêlos radiculares e acidificação da rizosfera (Yan et al. 2004), desenvolvimento basal de raiz e adaptação a baixos níveis de fosforo (Ochoa et al. 2006), resposta da raiz a auxinas (Remans et al. 2008). O primeiro estudo de identificação de QTLs em raiz em feijão comum com efeito sobre a tolerância a seca foi relatado recentemente por Asfaw & Blair (2012), sendo que foram detectados QTLs para massa seca de raiz, volume e comprimento de raiz.

Estes QTLs foram identificados no Pv11 diferentemente dos QTLs revelados neste estudo que foram localizados para volume de raiz no cromossomo Pv02 e para comprimento de raiz no Pv01 e Pv06, respectivamente, no tratamento não irrigado.

Com o objetivo de validar os QTLs identificados neste estudo e estimar a estabilidade dos mesmos, experimentos futuros em campo avaliando diferentes anos e locais devem ser realizados antes da sua aplicação nos programas de melhoramento genético como método indireto de seleção ou para estudos que possibilitem que marcadores alelo específicos, que co-segrem com locos de tolerância sejam identificados, viabilizando a clonagem posicional dos genes de tolerância. Com isso, os QTLs estáveis serão detectados, para que se possa refinar a região destes QTLs a fim de utilizá-los para seleção assistida por marcadores, visando que o melhorista possa manipular a variação

genética, realizando introgressão e piramidação dos genes de tolerância através de seleção assistida por marcadores ou transgenia.

## CONSIDERAÇÕES FINAIS

A seca é o fator abiótico que mais afeta a produtividade nas culturas e hoje em dia é causada por chuvas insuficientes e padrões de precipitação alterados (Toker et al., 2007). O melhoramento genético do feijoeiro pode ajudar a aumentar a área de cultivo em áreas onde hoje não é possível plantar por falta de água. A tolerância à seca deve ser um tema de prioridade e um objetivo importante para o melhoramento genético, pois pode causar grandes variações na produção de ano para ano, podendo acarretar a falta de alimento em algumas populações (Beebe et al. 2011).

A natureza da seca e sua interação com vários fatores ambientais fazem a identificação de QTLs específicos da seca muito mais complexa. Poucos QTLs têm sido identificados para tolerância à seca, entre eles, Schneider et al (1997b), estudaram a genética da resistência à seca, utilizando locos controladores de características quantitativas (QTLs) detectados com marcadores RAPD. Quatro marcadores em uma população e cinco em uma segunda população de RILs foram relatados como importantes para a resistência à seca. Beebe et al (2006a) relataram a identificação de marcadores de QTLs sob seca e, em condições irrigadas, em uma população de RILs derivada do cruzamento do SEA 5 x MD 23-24, descobrindo um QTL em comum nas duas temporadas de seca, um QTL específico para cada temporada e alguns QTLs comuns ao tratamento irrigado.

A população desenvolvida proveniente de cruzamento entre genótipos Andino x Mesoamericano permite explorar adicionalmente a diversidade alélica dos QTLs para tolerância à seca. A utilização de SNPs, SSRs e marcadores DArTs em conjunto com as características contrastantes da população geraram uma maior quantidade de marcadores polimórficos.

Embora os tipos de adaptações ao déficit hídrico não sejam sempre específicos e muitas plantas apresentem uma ou mais adaptações que aumentam a tolerância ao déficit hídrico, neste estudo QTLs de maior efeito foram identificados em condições de seca, sendo os mais interessantes àqueles associados às três linhas principais de defesa contra a seca, sendo que para a característica da área foliar foram identificados três QTLs (LA3.1<sup>AS</sup>, LA3.2<sup>AS</sup>, LA1.1<sup>AS</sup>), um para massa fresca do caule (SBF3.1<sup>AS</sup>), um para massa seca do



caule (SBD2.1<sup>AS</sup>) e dois para comprimento de raiz (RL1.1<sup>AS</sup>, RL6.1<sup>AS</sup>). Para os componentes de produtividade, que são as características mais importantes para o melhorista, foram identificados dois QTLs para massa de semente (SW1.1<sup>AS</sup>, SW1.2<sup>AS</sup>), cinco para número de sementes por vagem (NSP10.1<sup>AS</sup>, NSP1.1<sup>AS</sup>, NSP1.1<sup>AS</sup>, NSP5.1<sup>AS</sup>, NSP4.1<sup>AS</sup>), dois para dias até o florescimento (DF5.1<sup>AS</sup>, DF1.1<sup>AS</sup>), dois para número de sementes (NS7.1<sup>AS</sup>, NS8.1<sup>AS</sup>). Estes QTLs podem ser utilizados para direcionar o melhoramento genético afim de validar cultivares potencialmente tolerantes a seca e para o mapeamento fino de QTLs sob condições de déficit hídrico.

Os avanços na biotecnologia abrem uma nova perspectiva para compreender a complexa regulação de respostas fisiológicas e mecanismos que controlam a adaptação de cultivares a estresses abióticos. Métodos de marcadores moleculares poderosos como os DArTs, geram um grande volume de informações potencialmente úteis para assessorar os programas de melhoramento. O sucesso na geração de novas cultivares vai depender do uso destas tecnologias e informações geradas, aliadas a uma boa triagem para medir importantes características fenotípicas. Para isto, equipes multidisciplinares envolvendo áreas de fisiologia de plantas, biologia molecular e melhoramento genético devem ser integradas.

Os programas de melhoramento genético não parecem estar priorizando a tolerância a seca entre os fatores de seleção, ou seja, ainda não incorporaram a deficiência hídrica como uma característica fundamental dos seus ambientes alvo. Ao invés disso, os melhoristas estão buscando testar seus materiais elites em ensaios específicos, sob estresse, ou aproveitando oportunidades em que o estresse ocorre naturalmente em ensaios de valor de cultivo e uso, para detectar diferenças quanto a esta característica. Embora estas informações possam auxiliar no posicionamento de mercado das cultivares, tal estratégia dificilmente irá resultar em ganhos cumulativos para tolerância a seca. Por outro lado, os melhoristas parecem aguardar passivamente que alguma inovação significativa surja das pesquisas moleculares. Seria importante que os melhoristas buscassem uma maior aproximação com a biotecnologia, e vice-versa, para que as pesquisas em tolerância à seca obtivessem resultados concretos em um tempo mais curto.

Isto demonstra a importância deste trabalho já que existe pouca informação e poucos grupos envolvidos em desvendar toda a complexidade de resposta a este tipo de estresse. Nosso grupo é um grupo multidisciplinar que está se fortalecendo cada dia trocando informações e experiências com os outros grupos que trabalham com tolerância a seca no Brasil e no mundo. Grandes esforços e investimentos estão sendo feitos em infraestrutura, equipamentos e capacitação do pessoal no IAC, com o objetivo de fazer investigação de ponta e com qualidade, visando que nossa contribuição à pesquisa seja de grande ajuda ao futuro dos estudos relacionados a deficiência hídrica.

## CONCLUSÕES

- ✓ As análises mostraram que os genitores foram contrastantes para a tolerância a seca.
- ✓ O genótipo SEA 5 foi superior quanto a tolerância à seca para as características de área foliar, comprimento e volume de raiz, massa de 100 sementes e produtividade.
- ✓ A genotipagem com SSRs, SNPs e DArTs evidenciou um alto nível de polimorfismo na população AS.
- ✓ Os dois mapas genéticos construídos (SSR-SNP, SSR-SNP-DArT) são os mais saturados até o momento, e apresentam robustez nos marcadores ordenados nos respectivos cromossomos e quanto aos QTLs detectados.
- ✓ Marcadores DArTs são extremamente polimórficos e abundantes no genoma do feijão e sua utilização é viável desde que se tenham os *softwares* adequados para análise dos dados com grande quantidade de marcadores.
- ✓ Dos QTLs associados à seca, 75% tiveram a contribuição do genitor SEA 5.
- ✓ QTLs interessantes para o melhoramento para tolerância à seca foram identificados para área foliar, massa fresca do caule e massa de 100 sementes.
- ✓ A característica de temperatura foliar não mostrou resultados favoráveis para estudos futuros de identificação de QTLs relacionados ao déficit hídrico, sendo um descritor que não deve ser considerado.

## REFERÊNCIAS

- Aerts R, Berecha G, Gijbels P, Hundera K, Glabeke SV, Vandepitte K, Muys B, Ruiz IR, Honnay O (2013) Genetic variation and risk of introgression in the wild *Coffea arabica* gene pool in South-western Ethiopian montane rainforests. *Evolutionary Applications* 6:243-253.
- Aguiar SR, Moda-Cirino V, Faria RT, Vidal LHI (2008) Avaliação de linhagens promissoras de feijoeiro (*Phaseolus vulgaris* L.) tolerantes ao déficit hídrico. *Semina: Ciências Agrárias* 29:1-14.
- Akbari M, Wenzl P, Vanessa C, Carling J, Xia L, Yang S, Uszynski G, Mohler V, Lehmensiek A, Kuchel H, Hayden MJ, Howes N, Sharp P, Rathmell B, Vaughan P, Huttner E, Kilian A (2006) Diversity Arrays Technology (DART) for high-throughput profiling of the hexaploid wheat genome. *Theoretical and Applied Genetics* 113: 1409-1420.
- Akhunov E. et al. (2009) Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina Golden Gate assay. *Theor. Appl. Genet* 119: 507–517.
- Alheit KV, Reif JC, Maurer HP, Hahn V, Weissmann EA, Miedaner T, Wurschum T (2011) Detection of segregation distortion loci in triticales (x *Triticosecale* Wittmack) based on high-density DART marker consensus genetic linkage map. *BMC genomics* 12:380
- Altschul SF, Madden TL, Schaffer A, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25, 3389-3402.
- Allard, R.W. Principles of Plant Breeding. 2 ed. New York: John Wiley & Sons, 1999. 254p.
- Ariyaratne HM, Coyne DP, Jung G, Skroch PW, Vidaver JR, Steadman AK, Miklas PN, Basset MJ (1999) Molecular mapping of disease resistance genes for halo blight, common bacterial blight, bean common mosaic virus in a segregating population of common bean. *Journal of the American Society of Horticultural Science* 124:654-662.
- Armelin JM, Canniatti-Brazaca SG, Piedade SM, Machado FMVF, Spoto MHF (2007) Avaliação física de feijão carioca irradiado. *Ciência e Tecnologia de Alimentos* 27:498-502.
- Asfaw A & Blair MW (2012) Quantitative trait loci for rooting pattern traits of common beans grown under drought stress versus non-stress conditions. *Molecular Breeding* 30:681-695.
- Barnabas B, Jager K, Feher A (2008) The effect of drought and heat stress on reproductive processes in cereals. *Plant Cell Environ* 31:11–38
- Basset MJ (1991) A revised linkage map of common bean. *HortScience* 26:834-836.
- Basten CJ, Weir BS, Zeng ZB. QTL cartographer version 1.17: a reference manual and tutorial for QTL mapping. Raleigh, USA: North Carolina State University, 2005.
- Bedo J, Wenzl P, Kowalczyk A, Kilian A (2008) Precision-mapping and statistical validation of quantitative trait loci by machine learning. *BMC Genetics* 9: 35
- Beebe S (2012) Plant breeding reviews vol 36. In: Jules Janick, eds. Common bean breeding in the tropics. Wiley-blackwell. Oxford, UK.
- Beebe S, Ramírez J, Jarvis A, Rao IM, Mosquera G, Bueno JM, Blair MW (2011a) Genetic improvement of common beans and the challenges of climate change. In: Yadav SS;

- Redden R; Hatfield JL; Lotze-Campen H; Hall A, eds. Crop adaptation to climate change. Wiley-Blackwell, Oxford, UK.
- Beebe S, Rao IM, Cajiao I, Grajales M (2008) Selection for drought resistance in common bean also improves yield in phosphorus limited and favorable environments. *Crop Science* 48:582–592.
- Beebe S, Rengifo J, Gaitan E, Duque MC, Tohme J (2001) Diversity and origin of Andean landraces of common bean. *Crop Sci* 41:854–862.
- Beebe SE, Rao IM, Blair M, Tovar E, Grajales M and Cajiao C (2006a). Identificación de QTL para resistencia a sequía en líneas recombinantes (RILs) de la cruz MD 23- 24 x SEA 5. Paper presented at the LII Annual Meeting of the PCCMCA (Program Cooperativo Centroamericano para el Mejoramiento de Cultivos y Animales), Montelimar, Nicaragua. 24–28 April, 2006.
- Beebe SE, Rao IM, Blair MW, Acosta-Gallegos JA (2010) Phenotyping common beans for adaptation to drought. In: JM Ribaut and P Monneveux (eds.) Drought phenotyping in crops: from theory to practice. Generation Challenge Program Special Issue on Phenotyping, pp. 311-334
- Benchimol LL, Campos T, Carbonell SAM, Colombo CA, Chioratto AF, Formighieri EF, Gouve LRL, Pereira de Souza A (2007) Structure of genetic diversity among common bean (*Phaseolus vulgaris* L.) varieties of Mesoamerican and Andean origins using new developed microsatellite markers. *Genet Resour Crop Evol* 54:1747–1762
- Benchimol LL, Souza CL, Souza AP (2005) Microsatellite assisted backcross selection in maize. *Genet Mol Biol* 28:789–797.
- Benjamini Y & Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society* 7:289-300.
- Bernardo R (2008) Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Science* 48:1649-1664.
- Bitocchi E, Nanni L, Bellucci E, Rossi M, Giardini A, Zeuli PS, Logozzo G, Stougaard J, McClean P, Attene G, Papa R (2011) Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. [www.pnas.org/cgi/doi/10.1073/pnas.1108973109](http://www.pnas.org/cgi/doi/10.1073/pnas.1108973109).
- Blair MW, Buendía HF, Giraldo MC, Métais I, Peltier D (2008) Characterization of AT-rich microsatellites in common bean (*Phaseolus vulgaris* L.). *Theoretical and Applied Genetics* ., 118:91-103.
- Blair MW, Chavarro CM, Muñoz-Torres MC, Giraldo MC, Pedraza F (2009) Development and diversity of Andean-derived, gene-based microsatellites for common bean (*Phaseolus vulgaris* L.). *BMC Plant Biology* 9:1-14.
- Blair, M.W. et al. (2009) Development of microsatellite markers for common bean (*Phaseolus vulgaris* L.) based on screening of nonenriched, small-insert genomic libraries. *Genome* 52: 772–782.
- Blair MW, Galeano CH, Tovar E, Torres MCM, Castrillón AV, Beebe SE, Rao IM (2012) Development of a Mesoamerican intra-genepool genetic map for quantitative trait loci detection in a drought tolerant susceptible common bean (*Phaseolus vulgaris* L.) cross. *Mol Breeding* 29:71–88.
- Blair MW, Giraldo MC, Buendía HF, Tovar E, Duque MC, Beebe SE (2006) Microsatellite marker diversity in common bean (*Phaseolus vulgaris* L.) *Theoretical and Applied Genetics* 113:100-109.

- Blair MW, Hurtado N, Chavarro CM, Muñoz-Torres MC, Giraldo MC, Pedraza F, Tomkins J, Wing R (2011) Gene-based SSR markers for common bean (*Phaseolus vulgaris* L.) derived from root and leaf tissue ESTs: an integration of the BMc series. *BMC Plant Biology* 11:50.
- Blair MW, Pedraza F, Buedia HF, Gaitán-Solís E, Beebe SE, Gepts P, Tohme J (2003) Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.). *Theoretical and Applied Genetics* 107:1362-1374.
- Blum A (1985) Breeding crop varieties for stress environments. *Critical Rev Plant Sci* 2:199-238.
- Bonierbale MW, Plaisted RL, Tanksley SD (1988) RFLP maps base on a common set of clones reveal modes of chromosomal evolution in potato and tomato. *Genetics*, 120:1095-1103.
- Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic map in man using restriction length polymorphism. *Am J Hum Genet* 32:314-331.
- Briñez B, Blair MW, Kilian A, Carbonell SAM, Chiorato AF, Benchimol LR (2012) A whole genome DArT assay to assess germplasm collection diversity in common beans. *Mol Breeding* 30:181-193
- Broughton WJ, Hernandez G, Blair M, Beebe S, Gepts P, Vanderleyden J (2003) Beans (*Phaseolus* spp.) – model food legumes. *Plant and Soil* 252:55-128.
- Budic M, Saboti J, Megli V, Kos J, Kidri M (2013) Characterization of two novel subtilases from common bean (*Phaseolus vulgaris* L.) and their responses to drought. *Plant Physiology and Biochemistry* 62:79-87.
- Burle ML, Fonseca JR, Kami JA, Gepts P (2010) Microsatellite diversity and genetic structure among common bean (*Phaseolus vulgaris* L.) landraces in Brazil, a secondary center of diversity. *Theoretical and Applied Genetics* 121:801-813.
- Buso GSC, Amaral ZPS, Brondani RPV, Ferreira ME (2006) Microsatellite markers for the common bean *Phaseolus vulgaris*. *Molecular Ecology Notes* 6:252-254.
- Butcher PA, Williams ER, Whitaker D, Ling S, Speed TP, Moran GF (2002) Improving linkage analysis in outcrossed forest trees: an example from *Acacia mangium*. *Theoretical and Applied Genetics* 104:1185-1191.
- Caixeta ET, Borém A, Kelly JD (2005) Development of microsatellite markers based on BAC common bean clones. *Crop Breeding and Applied Biotechnology* 5:125-133.
- Campos T, Benchimol LL, Carbonell SAM, Chioratto AF, Formigueri EF, Souza AP (2007) Microsatellites for genetic studies and breeding programs in common bean. *Pesq. Agropec. Bras* 42:401-408.
- Campos T, Oblessuc PR, Sforça DA, Cardoso JMK, Baroni RM, Sousa ACB, Carbonell SAM, Chioratto AF, Rubiano LLB, Souza AP (2011) Inheritance of growth habit detected by genetic linkage analysis using microsatellites in the common bean (*Phaseolus vulgaris* L.). *Molecular Breeding* 27:549-560.
- Cardoso JMK, Oblessuc PR, Campos T, Sforça DA, Carbonell SAM, Chioratto AF, Formighieri EF, Souza AP, Benchimol LL (2008) New microsatellite markers developed from an enriched microsatellite common bean library. *Pesq. Agrop. Brasil*.
- Carneiro MS & Vieira MLC (2002) Mapas genéticos em plantas. *Bragantia* 61:89-100.
- Chase CD, Ortega VM, Vallejos CE (1991) DNA restriction fragment length polymorphisms correlate with isozyme diversity in *Phaseolus vulgaris* L. *Theoretical and Applied Genetics* 81:806-811.

- Chen M, Wang QY, Cheng XG, Xu ZS, Li LC, Ye XG, Xia LQ, Ma YZ (2007) GmDREB2, a soybean DRE-binding transcription factor, conferred drought and high – salt tolerance in transgenic plants. *Biochemical and Biophysical Research Communication*, 353 (2):299-305.
- CIAT, 2004: Annual Report. [http://www.ciat.cgiar.org/beans/pdfs/output\\_1\\_04.pdf](http://www.ciat.cgiar.org/beans/pdfs/output_1_04.pdf). Acesso 10 de Outubro de 2012.
- CIAT, 2012: Phenotyping Common Beans for Adaptation to Drought: Protocol for Field Evaluation, <http://www.ciat.cgiar.org/beans/default.aspx>. Acesso 10 de Novembro de 2012.
- Close, T.J. et al. (2009) Development and implementation of highthroughput SNP genotyping in barley. *BMC Genomics* 10, 582
- Coelho ASG. Considerações gerais sobre a análise de QTL's. In: Pinheiro, J.B.; Carneiro, I.F. (Eds.). *Análise de QTL no melhoramento de plantas*. Goiânia: FUNAPE, 2000. p.1-36.
- CONAB – Centro Nacional de Abastecimento, 2012. Disponível em: <http://www.conab.gov.br>. Acesso em 14.Dez.2012.
- Córdoba JM, Chavarro C, Rojas F, Muñoz C, Blair MW (2010) Identification and Mapping of Simple Sequence Repeat Markers from Common Bean (*Phaseolus vulgaris* L.) Bacterial Artificial Chromosome End Sequences for Genome Characterization and Genetic–Physical Map Integration. *The Plant Genome* 3:154–165.
- Cortés AJ, Chavarro MC, Blair MW (2011) SNP marker diversity in common bean (*Phaseolus vulgaris* L.). *Theoretical and Applied Genetics* 123:827-845.
- Courtosis B, Ahmadi N, Khowaja F, Price AH, Rami JF, Frouin J, Hamelin C, Ruiz M (2009) Rice root genetic architecture: meta-analysis from drought QTL database. *Rice* 2:115-128.
- Creste S, Tulmann A, Figueira A (2001) Detection of single sequence repeat polymorphism in denaturing polyacrylamide sequencing gels by silver staining. *Plant Molecular Biology Reporter* 19:299-306.
- Debouck, DG, Toro O, Paredes OM, Johnson WC, Gepts P (1993) Genetic diversity and ecological distribution of *Phaseolus vulgaris* in northwestern South America. *Economic Botany* 47:408-423.
- Didonet AD, Silva SC (2004) Elementos climáticos e produtividade do feijoeiro. *Informe Agropecuário*, EPAMIG, Belo Horizonte 25:13-19.
- Doorenbos J, Kassam AH. Yield response to water. Roma: FAO, 193p. (Irrigation and Drainage Paper, 33), 1979.
- Dourado N, Fancelli AL. Produção de feijão. Guaíba: Agropecuária, 385p. 2000.
- Edwards MD, Helentjaris T, Wright S, Stuber CW (1992). Molecular-marker-facilitated investigations of quantitative-trait loci in maize. Analysis based on genome saturation with isozyme and restriction fragment length polymorphism markers. *Theoretical and Applied Genetics* 83:765-774.
- Fageria NK, Baligar VC, Jones CA. Common bean and cowpea. In: FAGERIA, N. K.; BALIGAR, V. C.; JONES, C. A. Growth and mineral nutrition of field crops. New York: Maecel Dekker, p. 280-318, 1991.
- Faleiro FG, Schuster I, Ragagnin VA, Cruz CD, Corrêa RX, Moreira MA, Barros EG (2003) Caracterização de linhagens endogâmicas recombinantes e mapeamento de locos de características quantitativas associados a ciclo e produtividade do feijoeiro-comum. *Pesquisa Agropecuária Brasileira* 38:1387-1397.

- Faleiro FG. Melhoramento e mapeamento genético do feijoeiro-comum: análise de características quantitativas, morfológicas, moleculares e de resistência a doenças. Tese de Doutorado, Universidade Federal de Viçosa, 2000.
- Ferreira ME, Gattapaglia D. Introdução ao uso de marcadores moleculares em análise genética. 2 ed. Brasília: EMBRAPA – CENARGEN, 220 p, 1995.
- Fischer KS, Lafitte R, Fukai S, Atlin G, Hardy B (eds) (2003) Breeding rice for drought-prone environments International Rice Research Institute, Los Baños, Philippines.
- Fleury D, Jefferies S, Kuchel H, Langridge P (2010) Genetic and genomic tools to improve drought tolerance in wheat. *J Exp Bot* 61:3211–3222
- Flint-Garcia SA & Thornsberry JM. (2003) Structure of linkage disequilibrium in plants. *Annual Reviews in Plant Biology* 54:357-374.
- Freyre R, Skroch PW, Geffroy V, Adam-Blondon AF, Shirmohamadali A, Johnson WC, Llaca V, Nodari RO, Pereira PA, Tsai SM, Tohme J, Dron M, Nienhuis J, Vallejos CE, Gepts P (1998) Towards an integrated linkage map of common bean. 4. Development of a core linkage map and alignment of RFLP maps. *Theoretical and Applied Genetics* 97:847-856.
- Gaitán-Solís E, Duque MC, Edwards KJ, Tohme J (2002) Microsatellite in Common Bean (*Phaseolus vulgaris*): Isolation, Characterization, and Cross-Species Amplification in- *Phaseolus* ssp. *Crop Science* 42:2128-2136.
- Gaitán-Solís E, Choi IY, Quigley C, Cregan P, Tohme J (2008) Single nucleotide polymorphisms in common bean: their discovery and genotyping using a multiplex detection system. *Plant Genome* 1:125–134.
- Ganal MW, Altmann T, Roder MS (2009) SNP identification in crop plants. *Current Opinion in Plant Biology* 2:211-7
- Garcia RAV, Rangel PN, Brondani C, Martins WS, Melo LC, Carneiro MS, Borba TCO, Brondani RPV (2011) The characterization of a new set of EST-derived simple sequence repeat (SSR) markers as a resource for the genetic analysis of *Phaseolus vulgaris*. *BMC Genetics* 12:41-55.
- Gaur PM, Krishnamurthy L, Krishiwagi J (2008) Improving drought-avoidance root traits in chickpea (*Cicer arietinum* L.): current status of research at ICRISAT. *Plant Production Science* 11:3-11.
- Gelderman H (1975) Investigations on inheritance of quantitative characters in animals by gene markers. I.Method. *Theoretical and Applied Genetics* 46:319-330.
- Gepts P, Bliss FA (1984) Enhanced available methionine concentration associated with higher phaseolin levels in common bean seeds. *Theoretical and Applied Genetics* 69: 47-53.
- Gepts PR, Nodari R, Tsai R, Koinange EMK, Llaca V, Gilbertson, Guzman P (1993) Linkage Mapping in Common Bean, *Annual Report Bean Improvement Cooperative* 36:24-38.
- Gonçalves JGR. Identificação de linhagens de feijoeiro (*Phaseolus vulgaris* L.) tolerantes à seca. Tese de Doutorado, Instituto Agrônomo (IAC), Campinas/SP, 2013.
- Gore M.A. et al. (2009) A first-generation haplotype map of maize. *Science* 326, 1115–117.
- Grewal TS. et al. (2008) Mapping quantitative trait loci associated with barley net blotch resistance. *Theor. Appl. Genet* 116: 529–539.
- Grisi MCM, Blair MW, Gepts P, Brondani C, Pereira PAA, Brondani RPV (2007) Genetic mapping of a new set of microsatellite markers in a reference common bean (*Phaseolus vulgaris*) population BAT93 x Jalo EEP558. *Genet. Mol. Res* 6:691-706.



- Guerra-Sanz JM (2004) New SSR markers of *Phaseolus vulgaris* from sequence databases. *Plant Breeding* 123:87-89.
- Guimarães CM. Características morfo-fisiológicas do feijoeiro (*Phaseolus vulgaris* L.) relacionadas com a resistência à seca. 131 p. Tese (Doutorado) – Universidade Estadual de Campinas-UNICAMP, Campinas, 1992.
- Gupta PK, Varshney RK (2000) The development and use of microsatellite markers for genetics and plant breeding with emphasis on bread wheat. *Euphytica* 113:163-185
- Gupta PK, Roy JK, Prasad M (2001) Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. *Current Science* 80:524-535
- Haldane JBS (1919) The combination of linkage values, and the calculation of distance between the loci of linked factors. *Journal of Genetics* 8:299-309.
- Hanai LR, Campos T, Camargo LEA, Benchimol LL, Souza AP, Melotto M, Carbonell SAM, Chioratto AF, Consoli L, Formigueri EF, Siqueira MVBM, Mui Tsai S, Vieira MLC (2007) Development, characterization and comparative analysis of polymorphism at common bean-SSR loci isolated from genic and genomic sources. *Genome* 50:266-277.
- Hanai LR, Santini L, Camargo LEA, Fungaro MHP, Gepts P, Tsai SM, Vieira MLC (2010) Extension of the core map of common bean with EST-SSR, RGA, AFLP, and putative functional markers. *Mol Breeding* 25:25-45.
- Hanai LR. Desenvolvimento de marcadores SSR-EST e construção de mapas genéticos em feijão-comum (*Phaseolus vulgaris* L.). Tese de Doutorado, Universidade de São Paulo Escola Superior de Agricultura “Luiz de Queiroz”, 2008.
- Herrera JC, Combes MC, Cortina H, Alvarado G, Lashermes P (2002) Gene introgression into *Coffea arabica* by way of triploid hybrids (*C. arabica* X *C. canephora*). *Heredity* 89:488-494.
- Hippolyte I. et al., (2010) A saturated SSR/DaRT linkage map of *Musa acuminata* addressing genome rearrangement among bananas. *BMC Plant Biology* 10:65.
- Hiremath PJ, Kumar A, Penmetsa RV, Farmer A, Schlueter J, Chamarthi SK, Whaley AM, Garcia NC, Gaur P, Upadhyaya HD, Kishor PVK, Shah TM, Cook DR, Varshney RK (2012) Large-scale development of cost-effective SNP marker assays for diversity assessment and genetic mapping in chickpea and comparative mapping in legumes. *Plant Biotechnol J* 6:716-732.
- Hoisington D, Khairallah M, Gonzalez-de-Leon D. Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory. Mexico DF: CIMMYT, 1994.
- Hyten, D.L. et al. (2008) High-throughput genotyping with the GoldenGate assay in the complex genome of soybean. *Theor. Appl. Genet.* 116, 945–952
- Hyten DL, Song Q, Fickus EW, Quigley CV, Lim JS, Choi IY, Hwang EY, Pastor-Corrales M, Cregan PB (2010) High-throughput SNP discovery and assay development in common bean. *BMC Genomics* 11:475-490.
- IBGE – Instituto Brasileiro de Geografia e Estatística, 2010. Disponível em: [http://www.ibge.gov.br/home/estatistica/indicadores/agropecuaria/lspa/lspa\\_201007\\_7.shtm](http://www.ibge.gov.br/home/estatistica/indicadores/agropecuaria/lspa/lspa_201007_7.shtm) Acesso em: 20 dez. 2012.
- Isidore E, van Os H, Andrzejewski S, Bakker J, Barrena I, Bryan GJ, Caromel B, van Eck H, Ghareeb B, de Jong W, van Koert P, Lefebvre V, Milbourne D, Ritter E, van der Voort JR, Rousselle-Bourgeois F, van Vliet J, Waugh R (2003) Toward a marker-

- dense meiotic map of the potato genome: lessons from linkage group I. *Genetics* 165:2107-2116.
- Iwata H, Ninomiya S (2006) AntMap: Constructing Genetic Linkage Maps Using an Ant Colony Optimization Algorithm. *Breeding Science* 56:371-377.
- Jaccoud D, Peng K, Feinstein D, Kilian, A (2001) Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Res* 29: E25.
- Jansen RC (1993) Interval mapping of multiple quantitative trait loci. *Genetics*, 135:205-211.
- Jing HC, Bayon C, Kanyuka K, Berry S, Wenzl P, Huttner E, Kilian A, Hammond-Kosack KE (2009) DArT markers: diversity analyses, genomes comparison, mapping and integration with SSR markers in *Triticum monococcum*. *BMC Genomics* 10:458.
- Jin X, Xue Y, Wang R, Xu R, Bian L, Zhu B, Han H, Peng R, Yao Q (2013) Transcription factor *OsAP21* gene increases salt/drought tolerance in transgenic *Arabidopsis thaliana*. *Molecular Biology Reports* 40:1743-1752
- Kasuga M, Miura S, Shinosaki K, Yamaguchi-Shinosaki K (2004) A combination of the *Arabidopsis* DREB1A gene and stress-inducible rd29A promoter improved drought and low temperature stress tolerance in tobacco by gene transfer. *Plant Cell Physiology*, 45(3):346-350.
- Kao CH, Zeng ZB, Teasdale RD (1999) Multiple interval mapping for quantitative trait loci. *Genetics* 152:1203-1216.
- Kelly JD, Gepts P, Miklas PN, Coyne DP (2003) Tagging and mapping of genes and QTL and molecular marker-assisted selection for traits of economic importance in bean and cowpea. *Field Crops Research* 82:135–154.
- Kilian A. et al. (2005) The fast and the cheap, SNP and DArT based whole genome profiling for crop improvement. In: *Proceedings of the International Congress: In the Wake of the Double Helix, from the Green Revolution to the Gene Revolution*. Tuberosa, R., et al., eds. Pp443-461
- Kosambi DD (1944) The estimation of map distances from recombinant values. *Annals of Eugenics* 12:172-175.
- Kwok P-Y. Single nucleotide polymorphisms: methods and protocols. Methods in molecular biology, Humana Press, Totowa, N.J., 2003.
- Lamprecht H (1961) Weitere kopplungsstudien an *Phaseolus vulgaris* mitt einer übersicht über die Koppelungsgruppen. *Agri Hortique Genetica* 19:319-332
- Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199.
- Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln EE, Newburg L (1987) Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174-181.
- Liu BH, Knapp SJ (1992) Gmandel: a program for Mendelian segregation and linkage analysis of individual or multiple progeny populations using log-likelihood ratios. *Journal of Heredity* 81:407-418.
- Liu BH. Statistical Genomics. CRC Press, Boca Raton, Florida, 1998, p.630.
- Lu, XP, Gui Y, Xiao B, Li Y, Tong Z, Liu Y, Bai X, Wu W, Ling X, Huttner E, et al (2013) Development of DArT markers for a linkage map of flue-cured tobacco. *Chinese Science Bulletin* 58:641-648.

- Lynch M, Walsh B. Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA. 980p, 1998.
- Mace ES, Rami JF, Bouchet S, Klein PE, Klein RR, Kilian A, Wenzl P, Xia L, Halloran K, Jordan DR (2008) A consensus genetic map of sorghum that integrates multiple component maps and high-throughput Diversity Array Technology (DArT) markers. *BMC Plant Biology* 9:13.
- Maciel FL, Echeverrigaray S, Gerald LTS, Grazziotin FB (2003) Genetic relationships and diversity among Brazilian cultivars and landraces of common beans (*Phaseolus vulgaris* L.) revealed by AFLP markers. *Genet Res Crop Evol* 50:887–893
- Magalhães AA, Millar AA (1978) Efeito do déficit de água no período reprodutivo sobre a produção de feijão. *Pesquisada Agropecuária Brasileira* 13:55-60.
- Margarido GRA, Souza AP, Garcia AAF (2007) OneMap: software for genetic mapping in outcrossing species. *Hereditas* 144:78-79.
- Marone D, Panio G, Ficco DB, Russo MA, De Vita P, Papa R, Rubiales D, Cattivelli L, Mastrangelo AM (2012) Characterization of wheat DArT markers: genetic and functional features. *Mol Genet Genomics* 9:741-53.
- Matise TC, Perlin M, Chakravarti A (1994) Automated construction of genetic-linkage maps using an expert-system (MULTIMAP) - a human genome linkage map. *Nature Genetics* 6:384-390.
- McNally, K.L. et al. (2009) Genome wide SNP variation reveals relationships among landraces and modern varieties of rice. *Proc.Natl. Acad. Sci* 106 : 12273–12278.
- Métais I, Hamon B, Jalouzet R, Peltier D (2002) Structure and level of genetic diversity in various bean types evidenced with microsatellite markers isolated from a genomic enriched library. *Theoretical and Applied Genetics* 104:1346-1352.
- Mifflin BJ (2000) Crop biotechnology. Where now? *Plant Physiol* 123:17-28.
- Miklas PN, Kelly JD, Beebe SE, Blair MW (2006) Common bean breeding for resistance against biotic and abiotic stresses: From classical to MAS breeding. *Euphytica* 147:105-131.
- Millan T, Clarke HJ, Siddique KHM, Buhariwalla HK, Gaur PM, Kumar J, Gil J, Kahl G, Winter P (2006) Chickpea molecular breeding: New tools and concepts. *Euphytica* 147, 81-103
- Moreira JAA, Azevedo JA, Stone LF, Caixeta TJ. Irrigação. In: Zimmermann, M.I.O.; Rocha, M.; Yamada, T. (Ed.). Cultura do feijoeiro: fatores que afetam a produtividade. Piracicaba: Associação Brasileira para Pesquisa da Potassa e do Fosfato, p.317-340, 1988.
- Morran S, Eini O, Pyvovarenko T, Parent B, Singh R, et al. (2011) Improvement of stress tolerance of wheat and barley by modulation of expression of DREB/CBF factors. *Plant Biotechnol J* 9: 230–249.
- Nayak SN, Zhu H, Varghese N, Datta S, Choi HK, Horres R, Jungling R, Singh J, Kishor PB, Sivaramakrishnan S, Hoisington DA, Kahl G, Winter P, Cook DR, Varshney RK (2010) Integration of novel SSR and gene-based SNP marker loci in the chickpea genetic map and establishment of new anchor points with *Medicago truncatula* genome. *Theoretical and Applied Genetics* 7:1415-41.
- Nelson JC (1997) QGENE: Software for marker-based genome analysis and breeding. *Molecular Breeding* 3:239-245.

- Nepomuceno AL, Neumaier N, Farias JRB, Oya T (2001) Tolerância à seca em plantas: Mecanismos fisiológicos e moleculares. *Biotecnologia, Ciência e Desenvolvimento* 23:12-18.
- Nodari RO, Tsai SM, Gilbertson RL, Gepts P (1993) Towards an integrated linkage of common bean. Development of a RFLP- based linkage map. *Theoretical and Applied Genetics* 85:513-520.
- Norman MJT, Pearson CJ, Searle PGE. Common bean (*Phaseolus vulgaris* L.). In: Norman, M.J.T.; Pearson, C.J.; Searle, P.G.E. (Eds.). The ecology of tropical food crops. Cambridge: CAMBRIDGE University Press, p.208-224, 1995.
- Oblessuc PR, Baroni RM, Garcia AAF, Chioratto AF, Carbonell SAM, Camargo LEA, Benchimol LL (2012) Mapping of angular leaf spot resistance QTL in common bean (*Phaseolus vulgaris* L.) under different environments. *BMC Genetics* 13:50.
- Osborne JW. Improving your data transformations: Applying the Box-Cox transformation. Practical Assessment, Research & Evaluation, v.15, n.12, 9p. 2010
- Passioura JB (2010) Scaling up: the essence of effective agricultural research. *Funct Plant Biol* 37:585–591
- Paterson AH, Damon S, Hewitt JD, Zamir D, Rabinowitch HD, Lincoln SE, Lander ES, Tanksley SD (1991) Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. *Genetics* 127:181-197.
- Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature* 335:721-726.
- Pedrosa-Harand A, Porch T, Gepts P (2008) Standard nomenclature for common bean chromosomes and linkage groups. Annual Report of the Bean Improvement Cooperative 51:106-107.
- Pellegrineschi A, Ribaut JM, Trethowan R, Yamaguchi-Shinozaki K, Hoisintong D (2002) Progress in the genetic engineering of wheat for water-limited conditions. JIRCAS Working Report. p.55-60.
- Perseguini JMKC, Chioratto AF, Zucchi MI, Colombo CA, Carbonell SAM, Mondego JMC, Gazaffi R, Garcia AAF, Campos T, Souza AP, Rubiano LB (2011) Genetic diversity in cultivated carioca common beans based on molecular marker analysis. *Genetics and Molecular Biology* 34:88-102.
- Perrier X, Jacquemoud-Collet JP (2006) DARwin software. [http:// darwin.cirad.fr/Darwin](http://darwin.cirad.fr/Darwin)
- Petes TD, Greenwell PW, Dominska M (1997) Stabilization of microsatellite sequence by variant repeats in the yeast *Saccharomyces cerevisiae*. *Genetics* 146:491-498.
- Petroli CD, Sansaloni CP, Carling J, Steane DA, Vaillancourt RE, et al. (2012) Genomic Characterization of DArT Markers Based on High-Density Linkage Analysis and Physical Mapping to the *Eucalyptus* Genome. *PLoS ONE* 7(9): e44684.
- Pieruschka R & Poorter H (2012) Phenotyping Plants: genes, phenes and machines. *Functional plant biology* 39:813 – 820.
- Pimentel C, Perez AJC (2000) Estabelecimento de parâmetros para avaliação de tolerância a seca em genótipos de feijoeiro. *Pesquisa Agropecuaria Brasileira* 35:31-39.
- Poorter H, Fiorani F, Stitt M, Schurr U, Finck A, Gibon Y, Usadel B, Munns R, Atkin OK, Tardieu F, Pons TL (2012a) The art of growing plants for experimental purposes; a practical guide for the plant biologist. *Functional Plant Biology* 39:821–838.

- Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. *Trends in Plant Science* 1:215-222.
- Priyanka B, Sekhar K, Reddy VD, Rao KV (2010) Expression of pigeonpea hybrid-proline-rich protein encoding gene (CcHyPRP) in yeast and Arabidopsis affords multiple abiotic stress tolerance. *Plant Biotechnol J* 8: 76–87.
- Qin F, Kakimoto M, Maruyama K, Osakabe Y, Tran LS, Shinosaki K, Yamaguchi-Shinosaki K (2007) Regulation and functional analysis of ZMDREB2A in response to drought and heat stresses in ZEA mays L. *The Plant Journal*, 50:54-69.
- Quan R, Hu S, Zhang Z, Zhang H, Huang R (2010) Overexpression of an ERF transcription factor TSRF1 improves rice drought tolerance. *Plant Biotechnol J* 8: 476–488.
- Rafalski A (2002) Applications of single nucleotide polymorphism in crop genetics. *Curr Opin Plant Biol* 5:94-100.
- Raman H, Raman R, Nelson M, Aslam MN, Rajasekaran R, Wratten N, Cowling WA, Kilian A, Sharpe AG, Schondelmaier J (2012) Diversity Array Technology Markers: Genetic Diversity Analyses and Linkage Map construction in Rapessed (*Brassica napus* L.) *DNA Research* 19:51-65.
- Ramirez Vallejo P & Kelly JD (1998) Traits related to drought resistance in common bean. *Euphytica* 99:127-136
- Rao IM (2001) Role of physiology in improving crop adaptation to abiotic stresses in the tropics: the case of common bean and tropical forages. *Handbook of Plant and Crop Physiology*, Marcel Dekker, p. 583–613. New York, USA.
- Reddy OUK, Pepper AE, Abdurakhmonov I, Saha S, Jenkins JN, Brooks T, Bolek Y, El-zik KM (2001) New dinucleotide and trinucleotide microsatellite marker resources for cotton genome research. *Journal of Cotton Science* 5:103-113.
- Ribeiro HJSS, Prudencio SH, Myiagui DT, Ribeiro ELA (2009) Caracterização de concentrado proteico de feijão comum preto, cultivar Iapar 44, novo e envelhecido. *Ciência e Tecnologia de Alimentos* 29:571-580.
- Ruta N, Stamp P, Liedgens M, Fracheboud Y, Hund A (2010) Co-locations of QTL for seedling traits and yield components of tropical maize under water stress conditions. *Crop Science* 50:1385-1392.
- Sabadin PK, Malosetti M, Boer MP, Tardin FD, Santos FG, Guimaraes CT, Gomide RL, Andrade CLT, Albuquerque PEP, Caniato FF, Mollinari M, Margarido GRA, Oliveira BF, Schaffert RE, Garcia AAF, van Eeuwijk FA, Magalhaes JV (2012) Studying the genetic basis of drought tolerance in sorghum by managed stress trials and adjustments for phenological and plant height differences. *Theoretical and Applied Genetics* 124:1389-1402
- Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL, Hunt SE, Cole CG, Coggill PC, Rice CM, Ning Z, Rogers J, Bentley DR, Kwok PY, Mardis ER, Yeh RT, Schultz B, Cook L, Davenport R, Dante M, Fulton L, Hillier L, Waterston RH, McPherson JD, Gilman B, Schaffner S, Van Etten WJ, Reich D, Higgins J, Daly MJ, Blumenstiel B, Baldwin J, Stange-Thomann N, Zody MC, Linton L, Lander ES, Altshuler D. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409:928-933.
- Saitou N, Nei M (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425

- Santalla M, De Ron AM, De La Fuente M (2010) Integration of genome and phenotypic scanning gives evidence of genetic structure in Mesoamerican common bean (*Phaseolus vulgaris* L.) landraces from the southwest of Europe. *Theoretical and Applied Genetics* 120: 1635–1651.
- Saxena NP (2003) Management of drought in chickpea-holistic approach. In: Saxena NP (eds) Management of Agricultural Drought-Agronomic and Genetic options. Oxford and IBH publishing Co.Pvt Ltd. New Delhi, India pp. 103-122
- Shinozaki K, Yamaguchi-Shinosaki K (1996) Molecular responses to drought and cold stress. *Curr Opin Biotechnol* 7(2):161-167
- Schneider KA, Brothers ME, Kelly JF (1997b) Marker-assisted selection to improve drought resistance in common bean. *Crop Sci.* 37: 51–60.
- Schuster I & Cruz CD. Estatística genômica – Aplicada a Populações Derivadas de Cruzamentos Controlados. Universidade Federal de Viçosa, 2008. 568p.
- Schuster I, Vieira SNV, Hamilton S, Sinhorati D, Silva RB, Oliveira MAR (2007) Fluxo gênico em soja na região oeste do Paraná. *Pesquisa Agropecuária Brasileira* 42: 515-520.
- Sgarbieri VC (1980) Estudo do conteúdo e de algumas características das proteínas e sementes de plantas leguminosas. *Ciência e Cultura* 32:78-84.
- Silveira PM, Guimarães CM, Stone LF, Kluthcouski J (1981) Avaliação de cultivares de feijão para resistência a seca, baseada em dias de estresse de água no solo. *Pesquisa Agropecuária Brasileira* 16:693-699.
- Singh SP (1995) Selection for water-stress tolerance in interracial populations of common bean. *Crop Science* 35:118-124.
- Singh SP, Gepts P, Debouck DG (1991) Races of common bean (*Phaseolus vulgaris*, Fabriceae) *Economic Botany* 45:379-396.
- Singh SP, Teran H, Gutierrez JA (2001) Registration of SEA 5 and SEA 13 drought tolerant dry bean germplasm. *Crop Sci* 41:276–7.
- Souza TLPO, Barros EG, Bellato CM, Hwang EY, Cregan PB, Pastor-Corrales MA (2012) Single nucleotide polymorphism discovery in common bean. *Molecular Breeding* 30:419-428.
- Stam P (1993) Construction of integrated genetic maps by means of a new computer package: JoinMap. *Plant Journal* 3:739-744.
- Staub JE, Serquen FC, Gupta M (1996) Genetic markers, map construction, and their application in plant breeding. *HortScience* 31:729-740.
- Steel KA, Prince AH, Shashidhar HE, Witcombe JR (2006) Marker assisted selection to introgress rice QTL controlling root traits into Indian unland variety. *Theoretical and Applied Genetics* 109:1247-1260.
- Suiter KA, Wendell JF, Case JS (1983) Linkage 1: a pascal computer-program for the detection and analysis of genetic-linkage. *Journal of Heredity* 74:203-204.
- Szilagyi L, Tayyar Ş, Ciuca M (2011) Evaluation of genetic diversity in common bean (*Phaseolus vulgaris* l.) using RAPD markers and morpho-agronomic traits. *Romanian Biotechnological Letters* 16:1.
- Taiz L & Zeiger E. Fisiologia do Estresse. In: Fisiologia Vegetal. 4. Ed. Porto Alegre: Artmed, 2009.
- Tanksley SD, Bernatzky R, Lapitan NL, Prince JP (1988) Conservation of gene repertoire but not gene order in pepper and tomato. *Proceedings of the National Academy of Science* 85:6419-6423.

- Tanksley SD, Bernatzky R, Lapitan NL, Prince JP (1988b) Conservation of gene repertoire but not gene order in pepper and tomato. *Proceedings of the National Academic of Sciences* 85:6419-6423
- Tanksley SD, McCouch SR (1997) Seed banks and molecular maps: Unlocking genetic potential from the wild. *Science* 277:1063-1066.
- Tautz D & Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryote genomes. *Nucleic Acids Research* 12:4127-4138.
- Tautz D. (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Resources* 17:6463-6471.
- Teles FL. Mapa genético do feijoeiro comum (*Phaseolus vulgaris* L.) baseado em marcadores microssatélites e RAPD utilizando populações de melhoramento para teor de fibra. 106 p. Dissertação (Mestrado) apresentada à Universidade Federal de Goiás, UFG, Goiânia, 2007.
- Terán H & Singh SP (2002) Selection for drought resistance in early generations of common bean populations. *Canadian Journal of Plant Science* 82:491–497.
- Thoday JM. Effects of specific genes. In: Pollak E, Kempthorne O, Bailey TB. In: International Quantitative Genetics Conference. Proceedings... Ames: Iowa State University, p.141-159, 1977.
- Toker C, Canci H, Yildirim T (2007) Evaluation of perennial wild Cicer species for drought resistance. *Genet Resour Crop Evol* 54:1781–1786
- Trujillo LE, Sotolongo M, Menendez C, Ochogavia ME, Coll Y, et al. (2008) SodERF3, a novel sugarcane ethylene responsive factor (ERF), enhances salt and drought tolerance when overexpressed in tobacco plants. *Plant Cell Physiol* 49: 512–525.
- Tuberosa R, Salvi S, Giuliani S, Sanguineti MC, Bellotti M, Conti S, Landi P (2007) Genome wide approach to investigative and improve maize response to drought. *Crop Science* 47:2120-5141.
- Turner NC, Wright GC, Siddique KHM (2001) Adaptation of grain legumes (pulses) to water-limited environments. *Adv Agron* 71:193-231
- Vadez V, Rao S, Kholova J, Krishnamurthy L, Kashiwagi J, Ratnakumar P, Sharma KK, Bhatnagarmathur P, Basu PS (2008) Root research for drought tolerance in legumes: Quo vadis?. *Journal of Food Legumes* 21:77-85.
- Vallejos CE, Sakyama NS, Chase CD (1992) A molecular marker-based linkage map of *Phaseolus vulgaris* L. *Genetics* 131:733-740.
- Van Os H, Stam P, Visser RGF, van Eck HJ (2005) RECORD: a novel method for ordering loci on a genetic linkage map. *Theoretical and Applied Genetics* 112:30-40.
- Varshney RV, Glaszmann JC, Leung H, Ribaut JM (2010) More genomic resources for less-studied crops. *Trends in Biotechnology* 28:452-460.
- Varshney RV (2010) Gene based marker systems in plants, high throughput approaches for marker Discovery and genotyping. In *Molecular Techniques in Crop Improvement*. Jain, SM and Brar DS, eds. Springer, pp.119-140
- Vasconcelos MJ, Barros EG, Moreira MA, Vieira C (1996) Genetic diversity of the common bean *Phaseolus vulgaris* L. determined by DNA-based molecular markers. *Brazilian Journal of Genetics* 19:447-451.
- Vidigal MCG, Cruz AS, Garcia A, Kami J, Filho PSV, Sousa LL, McClean P, Gepts P, Pastor-Corrales MA (2011) Linkage mapping of the Phg-1 and Co-14 genes for resistance to angular leaf spot and anthracnose in the common bean cultivar AND 277. *Theoretical and Applied Genetics* 122:893–903.

- Voorrips RE (2002) MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *The Journal of Heredity* 93:77-78.
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* 23:4407-4414.
- Wang S, Basten CJ, Zeng ZB. Windows QTL Cartographer ver. 2.5. Department of Statistics, North Carolina State University, Raleigh, NC, 2005.
- Weir B. Genetic data analysis. 2.ed. Sunderland: Sinauer Associates, 1996. 447p.
- Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res* 18:7213-7218.
- Wenzl P, Carling J, Kudrna D, Jaccoud D, Huttner E, Kleinhofs A, Kilian A (2004) Diversity arrays technology (DArT) for whole-genome profiling of barley. *Proc Natl Acad Sci* 101:9915–9920.
- Wenzl P, Li H, Carling J, Zhou M, Raman H, Paul E, Hearnden P, Maier C, Xia L, Caig V, Ovesná J, Cakir M, Poulsen D, Wang J, Raman R, Smith KP, Muehlbauer GJ, Chalmers KJ, Kleinhofs A, Huttner E, Kilian A (2006) A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. *BMC Genomics* 7:206.
- White JW & Castillo JA (1992) Evaluation of diverse shoot genotypes on selected root genotypes of common bean under soil water deficits. *Crop Sci* 32:762-765.
- White JW & Singh SP (1991). Sources and inheritance of earliness in tropically adapted indeterminate common bean. *Euphytica* 55:15–19.
- Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary *primers* are useful as genetic markers. *Nucleic Acids Res* 18:6531-6535.
- Wortmann CS, Kirkby RA, Elude CA, Allen DJ (1998) Atlas of common bean (*Phaseolus vulgaris* L.) production in Africa. CIAT publication No.297. 133pp.
- Xu S (2008) Quantitative Trait Locus Mapping Can Benefit From Segregation Distortion. *Genetics* 180:2201-2208.
- Xu Y, Lu Y, Xie SG, Wan J, Prasanna BM (2012) Whole genome strategies for marker-assisted plant breeding. *Molecular Breeding* 29:833-854.
- Yan, J. et al. (2009) Genetic characterization and linkage disequilibrium estimation of a Global maize collection using SNP markers. *PLoS ONE* 4:8451.
- Yang, S. et al. (2006) Low level of genetic diversity in cultivated pigeonpea compared to its wild relatives is revealed by diversity arrays technology. *Theor. Appl. Genet* 113:585–595.
- Yano M, Sasaki T (1997) Genetic and molecular dissection of quantitative traits in rice. *Plant Mol Biol* 35:145-153.
- Yin SW, Tang CH, Wen QB, Yang XQ, Yuan DB (2010) The relationships between physicochemical properties and conformational features of succinylated and acetylated kidney bean (*Phaseolus vulgaris* L.) protein isolates. *Food Research International, Essex* 43:730-738.
- Yu K, Park SJ, Poysa V (1999) Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus vulgaris* and *Vigna*). *Genome* 42:27-34.
- Yu K, Park SJ, Poysa V, Gepts P (2000) Integration of simple sequence repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). *The Journal of Heredity* 91:429-434.



- Yu MA & Bliss FA (1978) Seed proteins of common bean. *Crop Science* 18:431-437.
- Yuste-Lisbona FJ, Santalla M, Capel C, García-Alcázar M, De La Fuente M, Capel J, De Ron AM, Lozano R (2012) Marker-based linkage map of Andean common bean (*Phaseolus vulgaris* L.) and mapping of QTLs underlying popping ability traits. *BMC Plant Biology* 12:136.
- Zeng ZB (1993) Theoretical basis for separation of multiple linked gene effects in mapping of quantitative trait loci. *Proceedings of the National Academic of Sciences* 90:10972-10976.
- Zeng ZB (1994) Precision mapping of quantitative trait loci. *Genetics* 136:492-496.
- Zhou ML, Ma JT, Zhao YM, Wei YH, Tang YX, Wu YM (2012) Improvement of drought and salt tolerance in *Arabidopsis* and *Lotus corniculatus* by overexpression of a novel *DREB* transcription factor from *Populus euphratica*. *Gene* 506:10-17.

## ANEXOS

### Anexo 1. Desenvolvimento da plataforma DArT para feijão comum (*Phaseolus vulgaris* L.)

Author's personal copy

Mol Breeding (2012) 30:181–193  
DOI 10.1007/s11032-011-9609-3

#### A whole genome DArT assay to assess germplasm collection diversity in common beans

Boris Briñez · Matthew W. Blair · Andrzej Kilian ·  
Sérgio Augusto Morais Carbonell · Allison Fernando Chiorato ·  
Luciana Benchimol Rubiano

Received: 11 February 2011 / Accepted: 24 June 2011 / Published online: 7 July 2011  
© Springer Science+Business Media B.V. 2011

**Abstract** A Diversity Arrays Technology (DArT) marker system was applied for the first time to common beans (*Phaseolus vulgaris* L.) and tested on 89 accessions from the IAC and CIAT common bean breeding programs, as well as on landraces from the FAO collection. Seven frequently used restriction endonucleases were tested in combination with a

rare-cutting restriction enzyme, *Pst*I, to evaluate their suitability for DArT technology. Two restriction enzyme combinations (*Pst*I/*Bst*NI and *Pst*I/*Taq*I) were selected to evaluate polymorphisms and *Pst*I/*Bst*NI, which yielded the most polymorphisms, was used to construct the final array. Genotyping was done by labelling the genomic representations with the fluorescent nucleotides cy3-dUTP and cy5-dUTP. The poly-linker fragment was labelled with 6-FAM and used as a control treatment and standard to determine the amount of DNA spotted on the array for each clone. DArTsoft version 7.3 software was used to analyse, identify and score polymorphic markers. Arrays containing individual fragments from these representations generated DArT fingerprints with a genotype call rate of 97.1% and a scoring reproducibility of at least 99.9%. A total of 2,501 polymorphic markers were found. Neighbour-joining distance matrices were used to create dispersion graphs that distinguished the two major gene pools of common beans and classified the accessions as either Andean or Mesoamerican. A principal coordinate analysis of the DArT marker results explained 82% of the total data variation. These results show that the DArT platform was accurate for studying the genetic diversity of the common bean and efficient for the large-scale detection of polymorphisms. These properties make marker technology a choice for future experiments. This is the first report to describe the use of DArT technology for genotyping the common bean.

**Electronic supplementary material** The online version of this article (doi:10.1007/s11032-011-9609-3) contains supplementary material, which is available to authorized users.

B. Briñez (✉)  
Department of Genetics, Institute of Biology, State  
University of Campinas (UNICAMP), PO Box 6109,  
Campinas, SP, Brazil  
e-mail: borisbrinez@hotmail.com

M. W. Blair  
Department of Plant Breeding, Cornell University, Ithaca,  
NY 14853, USA

A. Kilian  
Diversity Arrays Technology P/L, PO Box 7141,  
Yarralumla, ACT 2600, Australia

S. A. M. Carbonell · A. F. Chiorato · L. B. Rubiano  
The Agronomic Institute (IAC), PO Box 28, Campinas,  
SP, Brazil

M. W. Blair  
International Center for Tropical Agriculture (CIAT),  
A.A. 6713 Cali, Colombia

**Keywords** Andean and Mesoamerican gene pools · Diversity array technology · Genetic diversity · *Phaseolus vulgaris* L.

## Introduction

The common bean (*Phaseolus vulgaris* L.) is one of the most ancient crops of the Americas and also the most important grain legume for human consumption world-wide (Broughton et al. 2003; McClean et al. 2004). This species was domesticated by Middle American and South American Andean cultures (Gepts et al. 1986; Gepts 1998) and progressively dispersed worldwide from these two regions (Angioi et al. 2010; Asfaw et al. 2009; Blair et al. 2009a, 2010b; Santalla et al. 2010; Zhang et al. 2008). Cultivated common beans originated in Latin America from these two recognized centers of domestication about 7,000–8,000 years ago (Gepts and Deboucq 1991). Total bean production worldwide currently exceeds 23 million metric tonnes (MT), of which 7 million MT are produced in Latin America and Africa. Social factors and ecological constraints determine where, whether, by whom and how beans are grown in a particular region. Apart from their high protein content, beans contain large quantities of complex carbohydrates, fibres, oligosaccharides and phytochemicals, such as polyphenols and isoflavones (Anderson et al. 1999).

Molecular markers have been used in common bean breeding programs in a variety of ways (Santalla et al. 2010; Campos et al. 2011), including studies on the origin and diversity of current cultivars (Blair et al. 2009a; Kwak and Gepts 2009; Rossi et al. 2009), domestication of *P. vulgaris* L. (Vega et al. 2010), and the genetics of resistance to important diseases (Miklas et al. 2003, 2006a; Shi et al. 2011) and abiotic constraints (Chen et al. 2010). Several types of molecular markers have been important tools in the characterization and detection of genetic diversity among common bean landraces (Angioi et al. 2010; Asfaw et al. 2009; Blair et al. 2010b). Random amplification of polymorphic DNA (RAPD) markers have been used principally to study the structure of genetic diversity among common bean landraces of Middle American origin (Beebe et al. 2000) while amplified fragment length polymorphism (AFLP) markers have been used to evaluate a variety

of wild beans (Tohme et al. 1996) and the origin of Andean landraces, but not their diversity because of low polymorphism (Beebe et al. 2001). Fingerprinting with AFLP has also been used to characterize yellow bean germplasm from both gene pools (Pallotini et al. 2004). Microsatellites have been widely used because of their ability to detect polymorphisms in a range of germplasms (Blair et al. 2006, 2009b; Benchimol et al. 2007). However, these markers vary in the polymorphism that they detect, depending on the length and sequence of the repeat motif they contain and the microsatellite location in gene-coding or non-coding segments of the genome (Buso et al. 2006; Marotti et al. 2007; Blair et al. 2010a).

All of the molecular markers described above show limitations in their genome coverage, discriminatory ability, reproducibility, technical demands and time required for execution. A number of these limitations can be overcome by using specialized techniques such as high-throughput capillary electrophoresis. However, most of the limitations are related to the sequential nature and high assay costs of these marker technologies, in addition to their reliance on DNA sequence information (Mace et al. 2008). Diversity Arrays Technology (DArT) can overcome these limitations and has been developed as a hybridization-based alternative for the majority of gel-based marker technologies currently in use. The DArT genotyping method was originally developed for rice (Jaccoud et al. 2001) and has subsequently been applied to many other plant species, including barley (Wenzl et al. 2004), cassava (Xia et al. 2005), *Arabidopsis* (Wittenberg et al. 2005), pigeonpea (Yang et al. 2006), wheat (Akbari et al. 2006) and other species. DArT offers a high level of multiplexing and can simultaneously type several thousand *loci* per assay while remaining independent of sequence information. DArT assays generate whole genome fingerprints by scoring the presence versus absence of DNA fragments in representations generated from genomic DNA samples through the process of complexity reduction (Mace et al. 2008).

The aim of this study was to develop a DArT system for common beans and analyze the genetic diversity in a germplasm panel that included groups of cultivated Andean and Mesoamerican beans from two breeding programs, as well as landraces of each gene pool and wild bean accessions. The specific

objectives involved testing different representations based on different endonuclease enzymes and comparing the genetic diversity found with the resulting DArT markers in several breeding accessions derived from different breeding programs and from the two major common bean gene pools.

## Materials and methods

### Sample collection and sources of DNA

Diversity panels were generated using DNA from 89 accessions (Table 1) that consisted mainly of samples from the Agronomic Institute (Instituto Agronômico—IAC, Campinas, SP, Brazil), the Agronomic Institute of Paraná (Instituto Agronômico do Paraná—IAPAR, Londrina, PR, Brazil) and the International Center for Tropical Agriculture (Centro Internacional de Agricultura Tropical—CIAT, Cali, Colombia). In addition, a few samples from the Federal University of Lavras (Universidade Federal de Lavras—UFLA, Lavras, MG, Brazil), the University of São Paulo (Escola Superior de Agricultura “Luiz de Queiroz”—ESALQ/USP, Universidade de São Paulo, Piracicaba, SP, Brazil) and the Federal University of Viçosa (Universidade Federal de Viçosa—UFV, Viçosa, MG, Brazil) were also used. The DNA quality and quantity were checked for each genotype by incubating 1 µl of DNA with 5 µl of 1 × loading dye at 37°C for 2 h. After incubation, the samples were loaded onto 0.8% agarose gels and run in 1 × TAE buffer at 80 V. Short or degraded DNA (< ~4 kb) was eliminated and DNA concentrations of 10–100 ng ml<sup>-1</sup> were used for the DArT construction step.

### Enzyme digestions for the development of a common bean DArT system

Combinations of the rare-cutting restriction enzyme *Pst*I with several restriction endonucleases (*Pst*I/*Bst*NI, *Pst*I/*Taq*I, *Pst*I/*Alu*I, *Pst*I/*Bam*II, *Pst*I/*Mse*I, *Pst*I/*Hae*III and *Pst*I/*Msp*I) (New England Biolabs, Beverly, MA, USA) that cut frequently were initially tested on DNA from three accessions (CAL 143, IAC Alvorada and BAT 477) to determine the restriction enzyme combination that provided the best digestion of bean DNA. All enzymes were tested at a

concentration of 20 U/µl, except for *Bst*NI that was used at 10 U/µl. DNA digestion and adaptor ligation were performed as described by Jaccoud et al. (2001). Briefly, 100 ng of DNA was digested with 9 µl of the digestion/ligation mix containing 0.2 µl of restriction enzymes (except for *Taq*I and *Msp*I for which 0.1 µl was used). After analysing the digestion profile and identifying the best combination for beans, the restriction product was then ligated to *Pst*I adapters (5'-CACGATGGATCCAGTGCA-3' annealed with 5'-CTGGATCCATCGTGCA-3') with 2 U of T4 DNA ligase (New England Biolabs). Subsequently, 1 µl of the restriction/ligation mixture was amplified in 50 µl of amplification mixture using a primer (5'-GATGGATCCAGTGCAG-3') complementary to the adapter sequence and the cycling conditions described by Wenzl et al. (2004): 94°C for 1 min, followed by 30 cycles at 94°C for 20 s, 58°C for 40 s, 72°C for 1 min and 72°C for 7 min.

### Library construction for the DArT array

Once the two best combinations were established, enzyme digestion, library construction and bacterial insert amplification were performed as described by Xie et al. (2006) using a mixture of the 89 accessions. Briefly, the PCR products of each sample were combined and ligated into the pCR<sup>®</sup> 2.1 TOPO vector using a TOPO cloning kit (Invitrogen Life Technologies Corporation, Carlsbad, CA, USA). The ligated insert-vector constructs were transformed into electrocompetent TOP10 *Escherichia coli* cells (Invitrogen), according to the manufacturer's instructions. Transformed cells were screened on medium containing ampicillin (100 µg/ml) and X-gal (40 µg/ml). The white colonies were picked and transferred to 96-well plates containing 60 µl of LB medium (50 µg/mg ampicillin) per well followed by incubation at 37°C for 22 h. The inserts were amplified directly from 1 µl of the LB stock using the M13 forward and reverse primers, and the PCR products were precipitated with isopropanol, washed with ethanol and resuspended in spotting buffer 2 (DArT Spotter solution). The PCR products were subsequently printed onto polylysine-coated slides using a MicroGrid II arrayer (Genomic Solutions, Lincoln, NE, USA). After arraying, the slides were placed in a water bath at 95°C for 2 min to denature the DNA and then immersed in a solution containing 0.1 mM



**Table 1** The common bean accessions, including genotype, source, gene pool and origin, used to develop the DArT markers

No.	Genotype	Source	Gene pool	Cultivar origin
1	AND 277	UFV	Andean	CIAT
2	Arc-1	IAC	Mesoamerican	Mexico
3	Bagajo	IAC	Andean	Brazil
4	BAT 332	UFV	Mesoamerican	CIAT
5	BAT 477	IAC	Mesoamerican	CIAT
6	BAT 477	CIAT	Mesoamerican	CIAT
7	BAT 881	CIAT	Mesoamerican	CIAT
8	BAT 93	CIAT	Mesoamerican	CIAT
9	Bayo	IAC	Andean	Mexico
10	BRB 191	CIAT	Andean	CIAT
11	CAL 143	IAC	Andean	CIAT
12	Carioca Comum	IAC	Mesoamerican	Brazil
13	Carioca ETÉ	IAC	Mesoamerican	Brazil
14	Carioca Tybatã	IAC	Mesoamerican	Brazil
15	Comell 49-242	IAC	Mesoamerican	USA
16	Diplomata	IAC	Mesoamerican	Brazil
17	DOR 364	CIAT	Mesoamerican	CIAT
18	DOR 390	CIAT	Mesoamerican	CIAT
19	DOR 476	CIAT	Mesoamerican	CIAT
20	DOR 714	CIAT	Mesoamerican	CIAT
21	Flor de Mayo	IAC	Mesoamerican	Mexico
22	G 11350	CIAT	Mesoamerican	Mexico
23	G 11360	CIAT	Mesoamerican	Mexico
24	G 14519	CIAT	Mesoamerican	USA
25	G 19833	CIAT	Andean	Peru
26	G 19839	CIAT	Andean	Peru
27	G 19892	CIAT	Andean	Argentina
28	G 21078	CIAT	Andean	Argentina
29	G 21212	CIAT	Mesoamerican	Colombia
30	G 21242	CIAT	Andean	Colombia
31	G 21657	CIAT	Andean	Bulgaria
32	G 2333	CIAT	Mesoamerican	Mexico
33	G 24390	CIAT	Mesoamerican	Mexico
34	G 3513	CIAT	Mesoamerican	Mexico
35	G 4090	CIAT	Mesoamerican	El Salvador
36	G 4825	CIAT	Mesoamerican	Brazil
37	G 5273	CIAT	Andean	Mexico
38	G 685	CIAT	Mesoamerican	Guatemala
39	G 855	CIAT	Mesoamerican	Mexico
40	G-4000	IAC	Mesoamerican	Mexico
41	Goiano Precocoe	IAC	Andean	Brazil
42	IAC 22/16-4-2-1	IAC	Mesoamerican	Brazil
43	IAC Apuã	IAC	Mesoamerican	Brazil
44	IAC J/43-1-3-1	IAC	Mesoamerican	Brazil
45	IAC Ybaté	IAC	Mesoamerican	Brazil

**Table 1** continued

No.	Genotype	Source	Gene pool	Cultivar origin
46	IAC-Alvorada	IAC	Mesoamerican	Brazil
47	IAC-UNA	IAC	Mesoamerican	Brazil
48	IPR 139	IAPAR	Mesoamerican	Brazil
49	IAPAR 81	IAC	Mesoamerican	Brazil
50	IPR Gralha	IAPAR	Mesoamerican	Brazil
51	IPR Grauna	IAPAR	Mesoamerican	Brazil
52	IPR Siriri	IAPAR	Mesoamerican	Brazil
53	IPR Tangará	IAC	Mesoamerican	Brazil
54	IPR Tiziu	IAPAR	Mesoamerican	Brazil
55	IPR Uirapurú	IAPAR	Mesoamerican	Brazil
56	ICA Pijao	CIAT	Mesoamerican	Colombia
57	Jabola	IAC	Andean	Brazil
58	Jalo EEP558	CIAT	Andean	Brazil
59	Jamapa	CIAT	Mesoamerican	Mexico
60	Kaboon	IAC	Andean	Netherlands
61	MA-IV-18-266	UFLA	Mesoamerican	Brazil
62	Majestoso	UFLA	Mesoamerican	Brazil
63	MAM 38	CIAT	Mesoamerican	CIAT
64	MAM 49	CIAT	Mesoamerican	CIAT
65	MAR 1	CIAT	Mesoamerican	CIAT
66	MAR 2	IAC	Mesoamerican	CIAT
67	MD 23-24	CIAT	Mesoamerican	EAP
68	MEX 54	UFV	Mesoamerican	CIAT
69	OPNS 331	IAC	Mesoamerican	Brazil
70	Ouro Negro	UFV	Mesoamerican	Brazil
71	Pérola	IAC	Mesoamerican	Brazil
72	BRS Pontal	IAC	Mesoamerican	Brazil
73	Porrillo Sintético	IAC	Mesoamerican	Mexico
74	BC1 (EXRico × VC3)	UFLA	Mesoamerican	Brazil
75	BC2 (M20 × G122)	UFLA	Mesoamerican	Brazil
76	Red Kidney	IAC	Andean	USA
77	RP2	UFLA	Mesoamerican	Brazil
78	Rudá	UFV	Mesoamerican	Brazil
79	Sanilac	IAC	Mesoamerican	USA
80	SEA 15	CIAT	Mesoamerican	CIAT
81	SEA 21	CIAT	Mesoamerican	CIAT
82	SEA 5	CIAT	Mesoamerican	CIAT
83	SEL 1309	CIAT	Mesoamerican	CIAT
84	SEQ 1027	CIAT	Andean	CIAT
85	Talismã	UFLA	Mesoamerican	Brazil
86	Tio Canela	CIAT	Mesoamerican	EAP
87	VAX 6	CIAT	Mesoamerican	CIAT
88	VC 16	UFLA	Mesoamerican	Brazil
89	XAN 159	USP	Andean	CIAT

DTT and 0.1 mM EDTA and dried by centrifugation (500g, 7 min, at room temperature).

#### DART genotyping analysis

Genotyping was done by labelling the two best genomic representations with the fluorescent nucleotides cy3-dUTP and cy5-dUTP, in a primer extension reaction with random decamers and the exo-Klenow fragment of *E. coli* DNA polymerase, respectively. The poly-linker fragment was labelled with 6-FAM and used as a reference to determine the amount of DNA from each clone spotted on the array (Jaccoud et al. 2001). The hybridization mix was denatured on a Corbett PCR machine at 95°C for 3 min, followed by 56°C for 5 min and 55°C indefinitely until hybridization with the microarray printed with the DART clones. The hybridization reaction was then incubated overnight at 65°C, after which the slides were washed and scanned using a Tecan LS300 confocal laser scanner (Grödig, Salzburg, Austria). The resulting TIF images were analysed using DARTsoft version 7.3, a software package developed by DART P/L and available for DART network members. This software identifies and scores polymorphic markers, with each marker being scored as 0 (absence), 1 (presence) or X (unable to score). The three principal parameters used for marker selection were: the quality parameter (Q value) that measured the fraction of the total variation across all individuals attributable to bimodality, the reproducibility (R) that derived from replicate individuals that were supposed to give identical results, and the call rate (CR) that represented the number of scored spots versus the maximum number of potential scores. The polymorphism information content (PIC) was used to assess how the marker scores were distributed between the two clusters (scored as 0 or 1) and was determined according to Anderson et al. (1993). This value is often used to measure the informativeness of a genetic marker.

#### Data analysis

In the analysis of genetic diversity, the marker scores were analysed with DARwin software (Perrier and Jacquemoud-Collet 2006) using the Jaccard index. The diversity tree was built using the neighbour-joining (NJ) algorithm (Saitou and Nei 1987) that

relaxes the assumption of equal mutation rate in space and time and produces an unrooted tree. The confidence interval of the genetic relationships among the accessions was determined by performing 1,000 bootstraps, with the results being expressed as percentages at the main nodes of each branch. Analysis of molecular variance (AMOVA) was used to partition the genetic variation into within- and among-gene pool diversity based on Arlequin v3.5 software (Excoffier and Lischer 2010). This analysis was used to identify and separate the samples into Mesoamerican and Andean groups based on the neighbour-joining dendrogram and the results were compared with those in the literature. A group with 19 accessions was obtained for the Andean gene pool and one with 70 accessions for the Mesoamerican gene pool. Statistical significance of the difference between mean genetic distances was assessed by Student's *t* test. The same software was used to calculate the average gene diversity over loci (*pi*) and the global  $F_{ST}$  and pairwise  $F_{ST}$  values that were the basis for inferring the pattern of population structure. In the latter analysis we used the Q coefficient from Weir and Cockerham (1984). Principal component analysis (PCA) based on genetic dissimilarity was used to visualize the position of the accessions in three-dimensional space (as an indication of the diversity of each pair of accessions). This analysis was done using the software NTSYSpc 2.1 (Rohlf 2002). The three-dimensional graphs representing the dispersion of the accessions were drawn using Statistica software 7.1 (StatSoft Inc. 2005) which allowed identification of the varieties by origin or type class.

#### Results

##### Selection of the complexity reduction method and array development

As described above, the rare-cutting restriction enzyme *Pst*I was used in combination with a selection of frequently cutting restriction enzymes in order to determine the most effective combination for reducing complexity. The best digestion profiles from the first 51 accessions were generated with the enzyme combinations *Pst*II/*Bst*NI and *Pst*II/*Taq*I (Table 1). Two libraries were produced from

**Table 2** Evaluation of methods for reducing complexity

Method	Spot plate barcode	Number of <i>Pst</i> I/ <i>Bst</i> NI markers	Average	Number of <i>Pst</i> I/ <i>Taq</i> I markers	Average
<i>Pst</i> I/ <i>Bst</i> NI	1	42	16.93	24	13.87
<i>Pst</i> I/ <i>Bst</i> NI	2	46	18.54	18	10.45
<i>Pst</i> I/ <i>Bst</i> NI	3	47	18.95	27	15.62
<i>Pst</i> I/ <i>Taq</i> I	4	41	16.53	28	16.18
<i>Pst</i> I/ <i>Taq</i> I	5	39	15.72	44	25.43
<i>Pst</i> I/ <i>Taq</i> I	6	33	13.34	32	18.49
Total	6	248	99.97	173	99.97

Reproducibility: 100%, Call rate: 95%

genomic representations, one containing 1,152 clones generated with *Pst*I/*Taq*I and the other with 1,152 clones generated with *Pst*I/*Bst*NI. The polymorphism information generated by each restriction enzyme pair was analysed and the criteria for selection of a given restriction enzyme pair was 100% reproducibility and a 95% call rate, indicating the usefulness of DArT markers for analysing genetic diversity in common bean accessions. Based on the number of polymorphic markers, the best combination for complexity reduction was the restriction enzyme pair *Pst*I/*Bst*NI (Table 2). To develop the full array, the libraries were extended to a complete set of 15,360 clones from all of the accessions, again using *Pst*I as the selective primary enzyme and *Bst*NI as the secondary enzyme. To optimize marker discovery, the representations were hybridized with all of the accessions. For this, random replicates were run in a double-spotted experimental design to ensure few false positives.

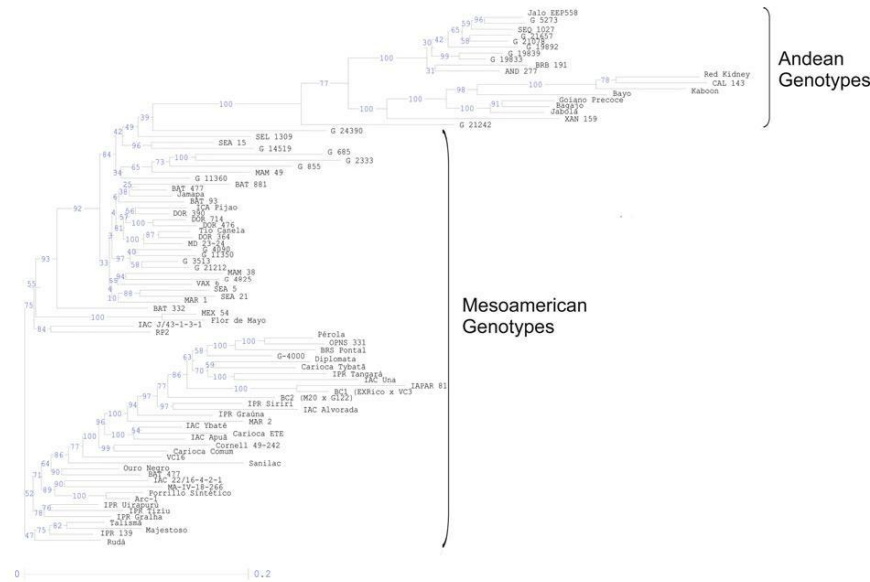
#### Common bean genetic diversity

A total of 2,501 polymorphic markers were identified for common beans among the 89 accessions used to create the array. These accessions were selected as representative of worldwide (from CIAT) and Brazilian (from the IAC and other institutes) bean germplasms of the Andean and Mesoamerican gene pools. The germplasm set included elite lines from different breeding institutes, as well as a number of wild accessions and the cultivated germplasm from both major gene pools. The overall polymorphism information content rate was 0.31, based on the full set of 15,360 clones isolated for the array. The average gene diversity over *loci* for the

Mesoamerican gene pool was  $0.2137 \pm 0.1024$ , and  $0.1810 \pm 0.0905$  for the Andean gene pool. The quality of the DArT markers based on the scoring reproducibility computed from the 67 replicated analyses was 99.9%; the discordance value, which is complementary to the reproducibility and expresses the overall variation in scores among replicates, was 0.00 with a consensus of 99.8%. The average call rate, which expresses the reliability of the final scores and represents the number of scored slides versus the maximum number of potential scores, was 97.1%. This percentage agreed with the values obtained in other studies, e.g., 97.6% in hop (Howard et al. 2011), 95.0% in barley (Wenzl et al. 2004), 99.2% in wheat (Akbari et al. 2006), 92.5% in sugarcane (Heller-Uszynska et al. 2011), 91.6% in banana (Risterucci et al. 2009) and 96.0% in pigeonpea (Yang et al. 2006).

Using the 2,501 polymorphic markers, a dendrogram was generated with the 89 accessions (Fig. 1). The neighbor-joining dendrogram resulted in a tree with two major clusters, one containing the Andean samples and the other containing the Mesoamerican samples. Overall, 19 accessions were classified as Andean and 70 as Mesoamerican. One of the control accessions was found in the Mesoamerican group (DOR 364) and the other in the Andean side of the tree (G19833). AMOVA was used to determine the variation explained by the Mesoamerican and Andean gene pools and to compare the variability between versus within gene pools. Inter-gene pool differences were significant ( $P < 0.001$ ) and accounted for 54.7% of variability, equivalent to an  $F_{ST}$  value of 0.5. The remaining variability (45.3%) was attributable to individual variation within each of the two gene pools (Table 3).





**Fig. 1** Neighbour-joining analysis of diverse common bean accessions based on 2,501 DArT markers using Jaccard's dissimilarity coefficients. The numbers on the branches indicate bootstrap values (expressed in percentages; based on 1,000 replications)

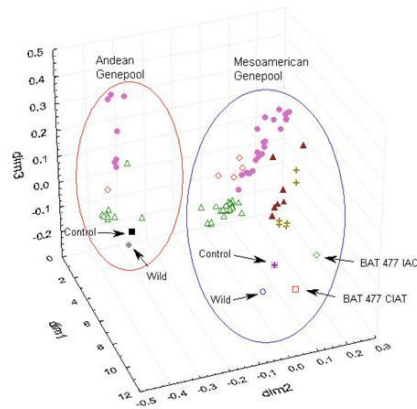
PCA clearly distinguished the two major gene pools represented by the Andean and Mesoamerican groups (Fig. 2). This analysis explained 82% of the variation across the three axes of the graph. Two subgroups were found within each cluster separating CIAT and Brazilian/IAC accessions. Finally, each gene pool subgroup was structured by cultivar source: one group consisted mainly of CIAT accessions (Fig. 2) while the other consisted mostly of IAC accessions (Fig. 2).

## Discussion

This is the first complete report on the development of DArT technology for common beans. As shown here, the DArT platform was efficient in assessing the genetic diversity of common beans based on the 2,501 polymorphic markers developed in this work. The analysis of diversity showed that the accessions formed two major groups corresponding to the Andean and Mesoamerican gene pools, thus

**Table 3** Analysis of molecular variance (AMOVA) for Andean and Mesoamerican common bean gene pools

Source of variation	Degrees of freedom	Sum of squares	Variance components	Percentage of variation	Level of significance, <i>P</i>
Between Andean and Mesoamerican accessions	1	7,933.403	258.287	54.71	<0.001
Among individuals within gene pools	87	18,601.136	213.806	45.29	<0.001
Total	88	26,534.539	472.093		



**Fig. 2** Principal coordinate analysis based on the analysis of 89 common bean accessions. This analysis explained 82% of the variation and showed the position of control and wild accessions for each gene pool. filled diamond UFLV; empty triangle CIAT; filled circle IAC; filled triangle UFLA; plus LAPAR

confirming the results of earlier studies with simple sequence repeat (SSR) markers (Benchimol et al. 2007; Blair et al. 2006, 2009a, b, 2010a) and showing that DArT libraries were biologically representative of the species. In addition, the DArT array results also separated the accessions according to their origin, seed colour, disease resistance and drought tolerance, as well as by their breeding institution.

Backcross breeding has been extensively used to transfer favourable alleles of monogenic traits from donor genotypes (elite or non-elite) to elite inbred lines (Openshaw et al. 1994); high heritability polygenic traits have also been transferred by this method, although to a lesser extent (Rinke and Sentz 1961; Shaver 1976). For monogenic traits, 6–8 generations of backcrossing followed by 1–2 generations of selfing have been recommended to simultaneously introgress the favourable allele and recover the genotype of the recurrent inbred parent (Fehr, 1987). Molecular markers could be used to improve the effectiveness of backcross breeding by monitoring the genetic similarity between the genotypes of the backcrosses and their recurrent inbred parent, known as “background selection”. DArT technology

proved useful for this kind of study in common beans since it was capable of clustering introgressed accessions derived from a recurrent backcrossing program to reveal common genomic backgrounds. For example, the origin of the genotype identified as BC2 (M20 × G122) and located in the Mesoamerican group was confirmed as being a backcross with the Mesoamerican accession M20 used as a recurrent parent that had previously been crossed to BC1 generated from the M20 × G122 cross. When the BC2 (M20 × G122) accession was in the first backcross generation (BC1), plants that were more similar to the M20 recurrent parent were selected based on RAPD markers and the proportion of alleles of the recurrent parent, as proposed by Benchimol et al. (2005). The backcross progeny were also selected for the similarity of their agronomic characteristics such as grain yield, grain type, upright plants and resistance to angular leaf spot similar to the recurrent parent M20 (Cameiro et al. 2010). M20 is an advanced line from the UFLA breeding program; this line has a ‘carioca’ grain type, a type II growth habit and shows anthracnose (*Colletotrichum lindemuthianum* (Sacc. & Magnus) Briosi & Cava) and angular leaf spot (*Pseudocercospora griseola* (Sacc.) Crous & U. Braun) resistance. G122 is an Andean line with white mold (*Sclerotinia sclerotiorum* (Lib.) de Bary) resistance. The resultant BC2 progeny have a high percentage of M20 parental alleles, and DArT markers revealed this similarity by correctly grouping these plants in the Mesoamerican gene pool.

DArT was also effective in predicting accessions of mixed ancestry, such as the inter-specific line XAN159 that comes from the cross [(G4449 × G10022) × G40020] × G4509 with a high amount of introgression from tepary beans (Muñoz et al. 2004). This genotype is more similar to the Andean group than to the Mesoamerican group, as described by Johns et al. (1997) and Svetleva et al. (2006).

As expected, the Andean types were efficiently separated from Mesoamerican types in PCA. The division of the accessions into two major groups indicated agreement between this study and the results of previous reports using molecular markers (Singh et al. 1991; Becerra and Gepts 1994; Blair et al. 2006; Benchimol et al. 2007). The Andean accessions included a few Jalo-type beans and other accessions while most of the accessions were of

Mesoamerican origin. Among wild beans, the Mesoamerican genotype (G24390) clustered distant from the wild Andean bean (G19892). The controls (G19833 for the Andean group and DOR 364 for the Mesoamerican group) also clustered separately. These samples were also used as controls by Blair et al. (2009a, 2010a, b), who confirmed the consistency of the DArT results with SSR analysis.

Another interesting finding was that two BAT 477 accessions occurred in different subgroups within the Mesoamerican group. This could be explained by the fact that these accessions came from different breeding institutes (CIAT and IAC) and from different germplasm banks, respectively. Therefore, some selection, genetic drift or open fertilization may have occurred, leading one accession to diverge from the other at some stage of advanced line development (CIAT, unpublished). Pedigree analysis showed that BAT 477 was derived from a multiple cross that included well-known landraces and successful cultivars such as Jamapa and Porrillo Sintético. Indeed, the BAT 477 sample that came from CIAT grouped closer to the Jamapa genotype whereas the other BAT 477 sample, which came from IAC, was located closer to the Porrillo Sintético genotype.

Among Mesoamerican beans, the cluster formed by accessions IAC Ybaté, Carioca ETÉ, IAC Apuã, Cornell 49-242 and Carioca Comum was derived from the IAC breeding program. The Cornell 49-242 genotype is resistant to anthracnose (Gonçalves-Vidigal et al. 2009) and has been widely used to incorporate anthracnose resistance into lines from the IAC breeding program. The Carioca Comum genotype was used as a common ancestor of IAC Ybaté, Carioca ETÉ and IAC Apuã to obtain beans with the carioca type seed coat, which is a cream-striped Mesoamerican bean preferred in most Brazilian states. All of the carioca accessions were closely grouped, showed the typical carioca grain type, and had high tolerance to diseases such as angular leaf spot (*Phaeoisariopsis griseola*), anthracnose caused by the fungus *C. lindemuthianum*, bean wilting caused by *Fusarium oxysporum* formae specialis (f. sp.) phaseoli (Fop) and bean common mosaic virus (BCMV). These diseases can reduce crop productivity by up to 95%.

Another Mesoamerican subgroup consisted of the accessions IPR Uirapuru, IPR Tiziu and IPR Gralha from the IAPAR breeding program. The IPR Gralha

genotype has large grains common to Andean gene pool accessions. However, this variety clustered with the Mesoamerican group and may have introgression from this gene pool. The Talismã and Majestoso accessions, from the Federal University of Lavras, clustered together with other varieties from the Lavras breeding program. Majestoso was derived from a blend of varieties and was positioned close to Talismã (Fig. 1). This genotype was subsequently used to improve OPNS 331 and was later released as BRSMG-Majestoso. For this reason, these two accessions may share a high proportion of their genetic background.

Three other small-seeded carioca-type beans, including Pérola, OPNS 331 and BRS-Pontal, grouped together as expected. The latter two accessions came from crosses using the Pérola genotype as a common ancestor. Their plant architecture is semi-erect to prostrate, the seed appearance is carioca-type and the seed size is relatively small. These accessions are tolerant to angular leaf spot and have become the most popular grain in Brazil, hence their importance as a baseline for this analysis. The IAC Diplomata and Tybatã varieties clustered in the same group and shared resistance to bean golden mosaic virus (BGMV) disease. Both of these varieties both have G4000 as a common ancestor that was used as a source of disease resistance.

Another group of Mesoamerican beans was formed by accessions MAM 38, G4825, VAX 6, SEA 5, SEA 21 and MAR 1, all of which share the same characteristics (cream-colored grain and resistance to low phosphorus availability). Some of these advanced lines or accessions are also known for their drought resistance and have been used in studies of abiotic stress (Terán and Singh 2002; Kelly 2008; CIAT 2008). Notably, these CIAT lines were developed mainly for the tropics rather than for subtropical regions. The CIAT and IAC Mesoamerican accessions were distinct, possibly because of different selection pressures in each breeding program and/or the presence of differing agro-ecological factors in tropical versus subtropical environments; these factors probably increased the genetic distance between the germplasm of each location. The subgroup clustering pattern was consistent with the experiments, indicating that DArT technology faithfully detected small differences among accessions, with excellent bootstrap support.



The separation of the CIAT and Brazilian (IAC) accessions in the Andean and Mesoamerican gene pools probably reflects the selection pressure that would have permitted beans that were domesticated in highland or tropical environments to gradually be moved into lowland or sub-tropical environments, resulting in genetically distinct subgroups. Beebe et al. (2001) found that Brazilian germplasm of Andean origin has several distinct traits such as indeterminate bush growth, photoperiod insensitivity, moderate tolerance to a low phosphorus availability, unique grain colours and a seed striping pattern known as 'zebra'. These accessions may be remnants of the first germplasm adapted to low altitudes that resulted from exchange between the Andes and Brazil. More recently, Blair et al. (2009a, b) found that Jalo-type beans of light yellow colour often have a mixed or inter-gene pool origin, suggesting another mechanism for the adaptation of Andean beans to the lowland environments more typical of Mesoamerican beans. Finally, carioca types appear to have been selected from the Mesoamerican gene pool only in Brazil and are not common in other tropical environments. Asfaw et al. (2009) found similar results in African accessions, where farmers' preferences differed in Ethiopia and Kenya and resulted in the grouping of accessions by country, in accordance with the ecological adaptation, cooking value and marketability of each bean type.

AMOVA indicated that the two major gene pools were strongly separated ( $F_{ST} = 0.5$ ,  $P < 0.001$ ) and that individuals within gene pools showed high variability, similar to that observed for BAT477 accessions derived from different germplasm banks. Asfaw et al. (2009) reported that the genetic differentiation ( $F_{ST}$ ) between two Andean control accessions versus two Mesoamerican control accessions was  $F_{ST} = 0.6$ .

In summary, DArT profiling proved to be useful in elucidating the pattern of relatedness among a wide range of modern bean accessions from the most important breeding programs and common bean pools. Though SSR and DArT marker systems are characterized by different information content on a per-locus basis, it must be emphasized that the use of biallelic dominant markers such as DArT markers to characterize the genetic stocks usually deployed in genetic analyses does not imply losses of genetic information. The high number of available DArT

markers, their cost-effectiveness and relatively high polymorphism content are ideal characteristics for extensive genome-wide screening for genetic diversity purposes. Additionally, DArT markers in beans were able to detect similar levels of diversity to SSRs, mainly because of the possibility of genotyping thousands of *loci* without any previous sequence information. The results obtained with the DArT markers are in good agreement with those obtained with highly informative genomic SSR *loci*, which up to now have represented the markers of choice to investigate genetic relationships and to carry out genetic and/or association mapping studies in common beans (Benchimol et al. 2007; Campos et al. 2011; Blair et al. 2010c).

In conclusion, a high-throughput genotyping platform based on DArT technology proved very useful for analysing the genetic diversity in common beans and may be applied to bean cultivar fingerprinting and/or germplasm analysis. This approach may be particularly relevant for breeding programs that use marker-assisted selection systems to monitor breeding achievements. Since this is the first complete report of a DArT platform for common beans, the results described here provide a useful starting point for future genetic, physical and association mapping studies.

**Acknowledgments** The authors thank Prof. Dr. João Bosco (UFLA, Lavras, MG, Brazil) for donating germplasm and for assistance during this work, and Drs. Nelson da Silva (IAPAR, Londrina, PR., Brazil), Everaldo Barros (UFV, Viçosa, MG, Brazil) and Maria Lucia Carneiro Vieira (ESALQ/USP, Piracicaba, SP, Brazil) for donating germplasms. This work was sponsored by Fundag (Fundação de Apoio à Pesquisa Agrícola) and FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo, grants numbers 2009/50275-0 and 2010/51186-9).

## References

- Akbari M, Wenzl P, Caig V, Carling J, Xia L, Yang S, Uszynski G, Mohler V, Lehmensiek A, Kuchel H, Hayden MJ, Howes N, Sharp P, Vaughan P, Rathnell B, Huttner E, Kilian A (2006) Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *Theor Appl Genet* 113:1409–1420
- Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic-linkage maps. *Genome* 36:181–186
- Anderson JW, Smith BM, Washnock CS (1999) Cardiovascular and renal benefits of dry bean and soybean intake. *Am J Clin Nutr* 70:464S–474S

- Angioi SA, Rau D, Attene G, Nanni L, Bellucci E, Logozzo G, Negri V, Zeuli PLS, Papa R (2010) Beans in Europe: origin and structure of the European landraces of *Phaseolus vulgaris* L. Theor Appl Genet 121:829–843. doi: 10.1007/s00122-010-1353-2
- Asfaw A, Blair M, Almekinders C (2009) Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) landraces from the East African highlands. Theor Appl Genet 120:1–12
- Becerra V, Gepts P (1994) RFLP diversity of common bean (*Phaseolus vulgaris*) in its centres of origin. Genome 37:256–263
- Beebe S, Skroch PW, Tohme J, Duque MC, Pedraza F, Nienhuis J (2000) Structure of genetic diversity among common bean landraces of Middle American origin based on correspondence analysis of RAPD. Crop Sci 40:264–273
- Beebe S, Rengifo J, Gaitan E, Duque MC, Tohme J (2001) Diversity and origin of Andean landraces of common bean. Crop Sci 41:854–862
- Benchimol LL, Souza CL, Souza AP (2005) Microsatellite-assisted backcross selection in maize. Genet Mol Biol 28:789–797
- Benchimol LL, Campos T, Carbonell SAM, Colombo CA, Chioratto AF, Formighieri EF, Gouve LRL, Pereira de Souza A (2007) Structure of genetic diversity among common bean (*Phaseolus vulgaris* L.) varieties of Mesoamerican and Andean origins using new developed microsatellite markers. Genet Resour Crop Evol 54:1747–1762
- Blair MW, Giraldo MC, Buendía HF, Tovar E, Duque MC, Beebe SE (2006) Microsatellite marker diversity in common bean (*Phaseolus vulgaris* L.). Theor Appl Genet 113:100–109
- Blair MW, Díaz LM, Buendía HF, Duque MC (2009a) Genetic diversity, seed size associations and population structure of a core collection of common beans (*Phaseolus vulgaris* L.). Theor Appl Genet 119:955–972
- Blair MW, Muñoz-Torres M, Giraldo MC, Pedraza F (2009b) Development and diversity assessment of Andean-derived, gene-based microsatellites for common bean (*Phaseolus vulgaris* L.). BMC Plant Biol 9:100
- Blair MW, Chaves A, Tofiño A, Calderón JF, Palacio JD (2010a) Extensive diversity and inter-gene pool introgression in a world-wide collection of indeterminate snap bean accessions. Theor Appl Genet 120:1381–1391
- Blair MW, Gonzalez LF, Kimani PM, Butare L (2010b) Genetic diversity, inter-gene pool introgression and nutritional quality of common beans (*Phaseolus vulgaris* L.) from Central Africa. Theor Appl Genet 121:237–248
- Blair MW, Prieto S, Díaz LM, Buendía HF, Cardona C (2010c) Linkage disequilibrium at the APA insecticidal seed locus of common bean (*Phaseolus vulgaris* L.). BMC Plant Biol 10:79. doi:10.1186/1471-2229-10-79
- Broughton WJ, Hernández G, Blair M, Beebe S, Gepts P, Vanderleyden J (2003) Beans (*Phaseolus* spp.) model food legumes. Plant Soil 252:55–128
- Buso GSC, Amaral ZPS, Brondani RPV, Ferreira ME (2006) Microsatellite markers for the common bean *Phaseolus vulgaris*. Mol Ecol Notes 6:252–254
- Campos T, Oblessuc PR, Sforc DA, Cardoso JMK, Baroni RM, De Sousa ACB, Carbonell SAM, Chioratto AF, Garcia AAF, Rubiano LB, De Souza AP (2011) Inheritance of growth habit detected by genetic linkage analysis using microsatellites in the common bean (*Phaseolus vulgaris* L.). Mol Breed 27:549–560
- Carneiro FF, Bosco JS, Leite ME (2010) Marker-assisted backcrossing using microsatellites and validation of SCAR *Phs* marker for resistance to white mold in common bean. Electron J Biotechnol 13:9–10
- Chen J, Zhang X, Jing R, Blair MW, Mao X, Wang S (2010) Cloning and genetic diversity analysis of a new *P5CS* gene from common bean (*Phaseolus vulgaris* L.). Theor Appl Genet 120:1393–1404
- CIAT (2008) Improved beans for the developing world. Annual Report I-P2. CIAT, Cali
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567
- Fehr WR (1987) Principles of cultivars development. Collier Macmillan Publishers, London, pp 172–192
- Gepts P (1998) Origin and evolution of common bean: past events and recent trends. Hort Sci 33:1124–1130
- Gepts P, Debouck D (1991) Origin, domestication and evolution of the common beans (*Phaseolus vulgaris* L.). In: Schoonhoven A, Voysey O (eds) Common beans: research for crop improvement. CAB International, Wallingford, pp 7–54
- Gepts P, Osborne TC, Rashka K, Bliss FA (1986) Phaseolin protein variability in wild forms and landraces of the common bean (*Phaseolus vulgaris* L.): evidence for multiple centers of domestication. Econ Bot 40:451–468
- Gonçalves-Vidigal MC, Vidigal Filho PS, Medeiros AF, Pastor-Corrales MA (2009) Common bean landrace Jalo Listras Pretas as the source of a new Andean anthracnose resistance gene. Crop Sci 49:133–138
- Heller-Uszyska K, Uszyski G, Huttner E, Evers M, Carlig J, Caig V, Aitken K, Jackson P, Piperidis G, Cox M, Gilmour R, D'Hont A, Butterfield M, Glaszmann JC, Kilian A (2011) Diversity arrays technology effectively reveals DNA polymorphism in a large and complex genome of sugarcane. Mol Breed. doi:10.1007/s11032-010-9460-y
- Howard EL, Whittok SP, Jakse J, Carling J, Matthews PD, Probasco G, Henning JA, Darby P, Cerenak A, Javomik B, Killian A, Koutoulis A (2011) High-throughput genotyping of hop (*Humulus lupulus* L.) utilising diversity arrays technology (DART). Theor Appl Genet 122:1265–1280
- Jaccoud D, Peng K, Feinstein D, Kilian A (2001) Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res 29:4 e25
- Johns MA, Skroch PW, Nienhuis J, Hinrichsen P, Bascur G, Muñoz-Schick C (1997) Gene pool classification of common bean landraces from Chile based on RAPD and morphological data. Crop Sci 37:605–613
- Kelly J (2008) Annual report of the bean improvement cooperative. <http://www.css.msu.edu/bic/reports.cfm>. Accessed 08 Nov 2010
- Kwak M, Gepts P (2009) Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., Fabaceae). Theor Appl Genet 118:979–992

- Mace E, Xia L, Jordan Dr, Halloran K, Parh DK, Huttner E, Wenzl P, Kilian A (2008) DArT markers: diversity analyses and mapping in *Sorghum bicolor*. BMC Genomics 9:26
- Marotti L, Bonetti A, Minelli M, Catizone P, Dinelli G (2007) Characterization of some Italian common bean (*Phaseolus vulgaris* L.) landraces by RAPD, semi-random and ISSR molecular markers. Genet Resour Crop Evol 54:175–188
- McClellan P, Kami J, Gepts P (2004) Genomic and genetic diversity in common bean. In: Wilson RF, Stalker HT, Brummer EC (eds) Legume crop genomics. AOCS Press, Champaign, pp 60–82
- Miklas PN, Coyne DP, Grafton KF, Mutlu N, Reiser J, Lindgren DT, Singh SP (2003) A major QTL for common bacterial blight resistance derives from the common bean great northern landrace cultivar Montana No. 5. Euphytica 131:137–146
- Miklas PN, Kelly J, Beebe S, Blair M (2006) Common bean breeding for resistance against biotic and abiotic stresses: from classical to MAS breeding. Euphytica 147:105–131
- Muñoz LC, Blair MW, Duque MC, Tohme J, Roca W (2004) Introgression in common bean x tepary bean interspecific congruity-backcross lines as measured by AFLP markers. Crop Sci 44:637–645
- Openshaw SJ, Jarboe SG, Beavis WD (1994) Marker-assisted selection in backcross breeding. In: Proceedings of the symposium “Analysis of Molecular Marker Data”, joint plant breeding symposia series. American Society for Horticultural Science/Crop Science of America, Corvallis, Oregon, pp 41–43
- Pallotini L, Garcia E, Kami E, Barcaccia G, Gepts P (2004) The genetic anatomy of a patented yellow bean. Crop Sci 44:968–977
- Perrier X, Jacquemoud-Collet JP (2006) DARwin software. (<http://darwin.cirad.fr/Darwin>)
- Rinke EH, Sentz JC (1961) Moving corn-belt germplasm northward. Ann Hybrid Corn Ind Conf 16:53–56
- Risterucci AM, Hippolyte I, Perrier X, Xia L, Caig V, Evers M, Huttner E, Kilian A, Glaszmann JC (2009) Development and assessment of diversity arrays technology for high-throughput DNA analyses in musa. Theor Appl Genet 119:1093–1103
- Rohlf F (2002) NTSYS pc. Numerical taxonomy system. Exeter Publishing, Setauket
- Rossi M, Bitocchi E, Bellucci E, Nanni L, Rau D, Attene G, Papa R (2009) Linkage disequilibrium and population structure in wild and domesticated populations of *Phaseolus vulgaris* L. Evol Appl 2:504–522
- Saitou N, Nei M (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425
- Santalla M, De Ron AM, De La Fuente M (2010) Integration of genome and phenotypic scanning gives evidence of genetic structure in Mesoamerican common bean (*Phaseolus vulgaris* L.) landraces from the southwest of Europe. Theor Appl Genet 120:1635–1651
- Shaver DL (1976) Conversations for earliness in maize inbreds. Maize Genet Coop Newslett 50:20–23
- Shi C, Chaudhary S, Yu K, Park SJ, Navabi A, McClellan PE (2011) Identification of candidate genes associated with CBB resistance in common bean HR45 (*Phaseolus vulgaris* L.) using cDNA-AFLP. Mol Biol Rep 38:75–81
- Singh SP, Nodari R, Gepts P (1991) Genetic diversity in cultivated common bean: I. Allozymes. Crop Sci 31:19–23
- StatSoft, Inc (2005) STATISTICA (data analysis software system) version 7.1. ([www.statsoft.com](http://www.statsoft.com))
- Svetleva D, Pereira G, Carlier J, Cabrita L, Leitão J, Genchev D (2006) Molecular characterization of *Phaseolus vulgaris* L. Accessions included in Bulgarian collection by ISSR and AFLP analyses. Sci Hort 109:198–206
- Terán H, Singh SP (2002) Comparison of sources and lines selected for drought resistance in common bean. Crop Sci 42:64–70
- Tohme J, González DO, Beebe S, Duque MC (1996) AFLP analysis of gene pools of a wild bean core collection. Crop Sci 36:1375–1384
- Vega EP, Pañeda A, Rodríguez-Suárez C, Campa A, Giraldez R, Ferreira JJ (2010) Mapping of QTLs for morpho-agronomic and seed quality traits in a RIL population of common bean (*Phaseolus vulgaris* L.). Theor Appl Genet 120:1367–1380
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370
- Wenzl P, Carling J, Kudma D, Jaccoud D, Huttner E, Kleinhofs A, Kilian A (2004) Diversity arrays technology (DArT) for whole-genome profiling of barley. Proc Natl Acad Sci USA 101:9915–9920
- Wittenberg AHJ, Van der Lee T, Cayla C, Kilian A, Visser RGF, Schouten HJ (2005) Validation of the high-throughput marker technology DArT using the model plant *Arabidopsis thaliana*. Mol Genet Genomics 274:30–39
- Xia L, Peng K, Yang S, Wenzl P, Carmen de Vicente M, Fregene M, Kilian A (2005) DArT for high-throughput genotyping of cassava (*Manihot esculenta*) and its wild relatives. Theor Appl Genet 110:1092–1098
- Xie Y, McNally K, Li CY, Leung H, Zhu YY (2006) A high-throughput genomic tool: diversity array technology complementary for rice genotyping. J Integr Plant Biol 48:1069–1076
- Yang S, Pang W, Ash G, Harper J, Carling J, Wenzl P, Huttner E, Zong X, Kilian A (2006) Low level of genetic diversity in cultivated pigeonpea compared to its wild relatives is revealed by diversity arrays technology. Theor Appl Genet 113:585–595
- Zhang X, Blair MW, Wang S (2008) Genetic diversity of Chinese common bean (*Phaseolus vulgaris* L.) landraces assessed with simple sequence repeat markers. Theor Appl Genet 117:629–640

**Anexo 2.** Amostras da placa número um de feijão comum das espécies *P. tuerckheimii*, *P. dumosus*, *P. coccineus* L, *P. acutifolius*, *P. anisotrichos* procedentes da Universidade de Cork na Irlanda.

Posição no gel	Espécies	Genótipo	Comentários
1	<i>Phaseolus tuerckheimii</i> Donn.Sm.	G40578	wild
3	<i>Phaseolus dumosus</i> Macfad.	G36286	wild G36286
5	<i>Phaseolus dumosus</i> Macfad.	G36286	wild G36286
7	<i>Phaseolus dumosus</i> Macfad.	G36286	wild G36286
9	<i>Phaseolus dumosus</i> Macfad.	G35877	wild G35877
11	<i>Phaseolus dumosus</i> Macfad.	G35877	wild G35877
13	<i>Phaseolus dumosus</i> Macfad.	G35877	wild G35877
15	<i>Phaseolus dumosus</i> Macfad.	G35877	wild G35877
2	<i>Phaseolus dumosus</i> Macfad.	G35908	wild G35908
4	<i>Phaseolus dumosus</i> Macfad.	G35908	wild G35908
6	<i>Phaseolus dumosus</i> Macfad.	G35908	wild G35908
8	<i>Phaseolus dumosus</i> Macfad.	G35946	wild G35946
10	<i>Phaseolus dumosus</i> Macfad.	G35946	wild G35946
12	<i>Phaseolus dumosus</i> Macfad.	G35946	wild G35946
14	<i>Phaseolus dumosus</i> Macfad.	G35758	wild G35758
16	<i>Phaseolus dumosus</i> Macfad.	G35758	wild G35758
17	<i>Phaseolus dumosus</i> Macfad.	G35758	wild G35758
19	<i>Phaseolus dumosus</i> Macfad.	G35758	wild G35758
21	<i>Phaseolus dumosus</i> Macfad.	G36045	wild G36045
23	<i>Phaseolus dumosus</i> Macfad.	G36045	wild G36045
25	<i>Phaseolus dumosus</i> Macfad.	G36045	wild G36045
27	<i>Phaseolus dumosus</i> Macfad.	PI311165	Landr PI311165
29	<i>Phaseolus dumosus</i> Macfad.	PI311165	Landr PI311165
31	<i>Phaseolus dumosus</i> Macfad.	PI311165	Landr PI311165



Posição no gel	Espécies	Genótipo	Comentários
18	<i>Phaseolus dumosus</i> Macfad.	PI190074	Landr PI190074
20	<i>Phaseolus dumosus</i> Macfad.	PI190074	Landr PI190074
22	<i>Phaseolus dumosus</i> Macfad.	PI190074	Landr PI190074
24	<i>Phaseolus dumosus</i> Macfad.	PI194585	Landr PI194585
26	<i>Phaseolus dumosus</i> Macfad.	PI194585	Landr PI194585
28	<i>Phaseolus dumosus</i> Macfad.	PI194585	Landr PI194585
30	<i>Phaseolus dumosus</i> Macfad.	PI195389	Landr PI195389
32	<i>Phaseolus dumosus</i> Macfad.	PI195389	Landr PI195389
33	<i>Phaseolus dumosus</i> Macfad.	PI195389	Landr PI195389
35	<i>Phaseolus dumosus</i> Macfad.	PI311183	Landr PI311183
37	<i>Phaseolus dumosus</i> Macfad.	PI311183	Landr PI311183
39	<i>Phaseolus dumosus</i> Macfad.	PI311183	Landr PI311183
34	<i>Phaseolus dumosus</i> Macfad.	PI317563	Landr PI317563
36	<i>Phaseolus dumosus</i> Macfad.	PI317563	Landr PI317563
38	<i>Phaseolus dumosus</i> Macfad.	PI317563	Landr PI317563
41	<i>Phaseolus dumosus</i> Macfad.	PI311194	Landr PI311194
43	<i>Phaseolus dumosus</i> Macfad.	PI311194	Landr PI311194
45	<i>Phaseolus dumosus</i> Macfad.	PI195399	Landr PI195399
47	<i>Phaseolus coccineus</i> L.	G35757	wild G35757
49	<i>Phaseolus coccineus</i> L.	G35757	wild G35757
51	<i>Phaseolus coccineus</i> L.	G35757	wild G35757
53	<i>Phaseolus coccineus</i> L.	G36002	wild G36002
55	<i>Phaseolus coccineus</i> L.	G36002	wild G36002
42	<i>Phaseolus coccineus</i> L.	G36002	wild G36002
44	<i>Phaseolus coccineus</i> L.	G36353	wild G36353
46	<i>Phaseolus coccineus</i> L.	G36353	wild G36353
48	<i>Phaseolus coccineus</i> L.	G36006	wild G36006
50	<i>Phaseolus coccineus</i> L.	G36006	wild G36006



Posição no gel	Espécies	Genótipo	Comentários
52	<i>Phaseolus coccineus</i> L.	G36285	wild G36285
54	<i>Phaseolus coccineus</i> L.	G36285	wild G36285
56	<i>Phaseolus coccineus</i> L.	G35837A	wild G35837A
57	<i>Phaseolus coccineus</i> L.	G35837A	wild G35837A
59	<i>Phaseolus coccineus</i> L.	G35837A	wild G35837A
61	<i>Phaseolus coccineus</i> L.	G36206	wild G36206
63	<i>Phaseolus coccineus</i> L.	G36206	wild G36206
65	<i>Phaseolus coccineus</i> L.	PI205360	Landr PI205360
67	<i>Phaseolus coccineus</i> L.	PI205360	Landr PI205360
69	<i>Phaseolus coccineus</i> L.	PI205360	Landr PI205360
71	<i>Phaseolus coccineus</i> L.	PI317572	Landr PI317572
58	<i>Phaseolus coccineus</i> L.	PI317572	Landr PI317572
60	<i>Phaseolus coccineus</i> L.	PI317572	Landr PI317572
62	<i>Phaseolus coccineus</i> L.	PI433237	cult PI433237
64	<i>Phaseolus coccineus</i> L.	PI433237	cult PI433237
66	<i>Phaseolus coccineus</i> L.	PI451863	cult PI451863
68	<i>Phaseolus coccineus</i> L.	PI451863	cult PI451863
70	<i>Phaseolus coccineus</i> L.	PI451863	cultPI451863
72	<i>Phaseolus coccineus</i> L.	PI451868	cult PI451868
73	<i>Phaseolus coccineus</i> L.	PI451868	cult PI451868
75	<i>Phaseolus coccineus</i> L.	PI451868	cult PI451868
77	<i>Phaseolus coccineus</i> L.	PI451870	cult PI451870
79	<i>Phaseolus coccineus</i> L.	PI451870	cult PI451870
74	<i>Phaseolus coccineus</i> L.	PI451872	cult PI451872
76	<i>Phaseolus coccineus</i> L.	PI451872	cult PI451872
78	<i>Phaseolus dumosus</i> Macfad.	G35641	cult G35641
81	<i>Phaseolus dumosus</i> Macfad.	G35641	cult G35641
83	<i>Phaseolus dumosus</i> Macfad.	G35641	cult G35641

Posição no gel	Espécies	Genótipo	Comentários
85	<i>Phaseolus dumosus</i> Macfad.	G36127	cult G36127
87	<i>Phaseolus dumosus</i> Macfad.	G36127	cult G36127
89	<i>Phaseolus dumosus</i> Macfad.	G36127	cult G36127
91	<i>Phaseolus dumosus</i> Macfad.	PI355423	cult PI355423
93	<i>Phaseolus dumosus</i> Macfad.	PI355423	cult PI355423
95	<i>Phaseolus dumosus</i> Macfad.	PI355423	cult PI355423
82	<i>Phaseolus coccineus</i> L.	G36256	cult G36256
84	<i>Phaseolus coccineus</i> L.	G36256	cult G36256
86	<i>Phaseolus coccineus</i> L.	G36256	cult G36256
88		nn1	sp
90		nn2	sp
92	<i>Phaseolus acutifolius</i> A.Gray	G40286	wild G40286
94	<i>Phaseolus acutifolius</i> A.Gray	G40286	wild G40286
96	<i>Phaseolus anisotrichos</i> Schltl.	G40553	wild G40553

**Anexo 3.** Amostras da placa número dois de feijão comum das espécies, *P. coccineus* L., *P. acutifolius*, *P. costaricensis*, *P. albescens*, *P. vulgaris* L., híbridos *vulgaris* x *dumosus*, *P. lunatus* L., e amostras de feijão comum sp., Procedentes da Universidade de Cork na Irlanda.

Posição no gel	Espécies	Genótipo	Comentários
1	<i>Phaseolus costaricensis</i> Freytag & Debouck	G36286	wild G36286
3	<i>Phaseolus costaricensis</i> Freytag & Debouck	G36286	wild G36286
5	<i>Phaseolus costaricensis</i> Freytag & Debouck	G40604	wild G40604
7	<i>Phaseolus coccineus</i> L.	G36334	wild G36334
9	<i>Phaseolus coccineus</i> L.	G36334	wild G36334
11	<i>Phaseolus coccineus</i> L.	G36334	wild G36334
13	<i>Phaseolus coccineus</i> L.	G35859	wild G35859
15	<i>Phaseolus coccineus</i> L.	G35859	wild G35859
2	<i>Phaseolus coccineus</i> L.	G35859	wild G35859

Posição no gel	Espécies	Genótipo	Comentários
4	<i>Phaseolus coccineus</i> L.	G35650	wild G35650
6	<i>Phaseolus coccineus</i> L.	G35650	wild G35650
8	<i>Phaseolus coccineus</i> L.	PI201336	Landr PI201336
10	<i>Phaseolus coccineus</i> L.	PI201336	Landr PI201336
12	<i>Phaseolus coccineus</i> L.	PI313503	Landr PI313503
14	<i>Phaseolus coccineus</i> L.	PI313503	Landr PI313503
16	<i>Phaseolus coccineus</i> L.	PI325603	Landr PI325603
17	<i>Phaseolus coccineus</i> L.	PI325603	Landr PI325603
19	<i>Phaseolus coccineus</i> L.	PI346951	cult PI346951
21	<i>Phaseolus coccineus</i> L.	PI346951	cult PI346951
23	<i>Phaseolus coccineus</i> L.	PI449381	cult PI449381
25	<i>Phaseolus coccineus</i> L.	PI449381	cult PI449381
27	<i>Phaseolus albescens</i> McVaugh ex R.Delgad. & A.Delgado	PL3592	wild PL3592
29	<i>Phaseolus albescens</i> McVaugh ex R.Delgad. & A.Delgado	PL3592	wild PL3592
31	<i>Phaseolus vulgaris</i> L.	G12853	wild G12853
18	<i>Phaseolus vulgaris</i> L.	G12853	wild G12853
20	<i>Phaseolus vulgaris</i> L.	G12853	wild G12853
22	<i>Phaseolus vulgaris</i> L.	G12853	wild G12853
24	<i>Phaseolus vulgaris</i> L.	G12853	wild G12853
26	<i>Phaseolus vulgaris</i> L.	G50387	wild G50387
28	<i>Phaseolus vulgaris</i> L.	G50387	wild G50387
30	<i>Phaseolus vulgaris</i> L.	G50387	wild G50387
32	<i>Phaseolus vulgaris</i> L.	G50387	wild G50387
33	<i>Phaseolus vulgaris</i> L.	G50387	wild G50387
35	<i>Phaseolus vulgaris</i> L.	G50388B	wild G50388B
37	<i>Phaseolus vulgaris</i> L.	G50388B	wild G50388B
39	<i>Phaseolus vulgaris</i> L.	G50388B	wild G50388B
34	<i>Phaseolus vulgaris</i> L.	G50388B	wild G50388B

Posição no gel	Espécies	Genótipo	Comentários
36	Phaseolus vulgaris L.	G19908	wild G19908
38	Phaseolus vulgaris L.	G19908	wild G19908
41	Phaseolus vulgaris L.	G19908	wild G19908
43	Phaseolus vulgaris L.	G19908	wild G19908
45	Phaseolus vulgaris L.	G19908	wild G19908
47	Phaseolus vulgaris L.	G19909	wild G19909
49	Phaseolus vulgaris L.	G19909	wild G19909
51	Phaseolus vulgaris L.	G19909	wild G19909
53	Phaseolus vulgaris L.	G19909	wild G19909
55	Phaseolus vulgaris L.	G19909	wild G19909
42	Phaseolus vulgaris L.	G23463	wild G23463
44	Phaseolus vulgaris L.	G23463	wild G23463
46	Phaseolus vulgaris L.	G23463	wild G23463
48	Phaseolus vulgaris L.	G23463	wild G23463
50	Phaseolus vulgaris L.	G23463	wild G23463
52	Phaseolus vulgaris L.	G7479	wild G7479
54	Phaseolus vulgaris L.	G7479	wild G7479
56	Phaseolus vulgaris L.	G7479	wild G7479
57	Phaseolus vulgaris L.	G7479	wild G7479
59	Phaseolus vulgaris L.	G7479	wild G7479
61	Phaseolus vulgaris L.	G24358	wild G24358
63	Phaseolus vulgaris L.	G24358	wild G24358
65	Phaseolus vulgaris L.	G24358	wild G24358
67	Phaseolus vulgaris L.	G21244	wild G21244
69	Phaseolus vulgaris L.	G21244	wild G21244
71	Phaseolus vulgaris L.	G21244	wild G21244
58	Phaseolus vulgaris L.	G21244	wild G21244
60	Phaseolus vulgaris L.	G21244	wild G21244

Posição no gel	Espécies	Genótipo	Comentários
62		Hyb G35980F	Hyb G35980F vulgaxdumosus
64		Hyb G35980F	Hyb G35980F vulgaxdumosus
66		Hyb G35980F	Hyb G35980F vulgaxdumosus
68		Hyb G35980F	Hyb G35980F vulgaxdumosus
70		Hyb G35980F	Hyb G35980F vulgaxdumosus
72		Hyb G50998C	Hyb G50998C vulgaxdumosus
73		Hyb G50998C	Hyb G50998C vulgaxdumosus
75		Hyb G35980E	Hyb G35980E vulgaxdumosus
77		Hyb G35980E	Hyb G35980E vulgaxdumosus
79	Phaseolus vulgaris L.	PI109863	cult PI109863
74	Phaseolus vulgaris L.	PI109863	cult PI109863
76	Phaseolus vulgaris L.	PI109860	cult PI109860
78	Phaseolus vulgaris L.	PI109860	cult PI109860
81	Phaseolus vulgaris L.	PI109861	cult PI109861
83	Phaseolus vulgaris L.	PI109861	cult PI109861
82		nn3	sp
84		nn4	sp
86	Phaseolus lunatus L.	G25224	wild G25224
88	Phaseolus lunatus L.	G25224	wild G25224
90	Phaseolus acutifolius A.Gray	G40286	wild G40286

**Anexo 4.** Estatísticas descritivas para características quantitativas avaliadas na população AND 277 X SEA 5 do experimento realizado em vasos em casa de vegetação.

Características	Skewness	Kurtosis	Shapiro-wilk	<i>p</i> valor
CI	0,2521	2,3155	0,9743	<0,0001*
CNI	0,3265227	0,5741205	0,987999	0,0056*
AFI	3,2192017	51,74894	0,668798	<0,0001*
AFNI	0,2524645	1,2333302	0,981124	0,0003*
MFFI	0,6048856	3,3423528	0,958267	<0,0001*
MFFNI	0,3104951	0,8344974	0,985677	0,0045*
MFCI	0,5038602	2,0778767	0,968003	<0,0001*
MFCNI	0,5471411	0,3567925	0,980573	0,0002*
MSFI	0,2278624	2,1416397	0,976136	<0,0001*
MSFNI	0,2124753	-0,434425	0,989543	0,0216 <sup>ns</sup>
MSCI	0,385813	1,2699492	0,984967	0,0021*
MSCNI	0,0430564	-0,292912	0,993814	0,1802 <sup>ns</sup>
TII	0,0003301	-0,602149	0,989967	0,0281 <sup>ns</sup>
TINI	0,162045	-0,248152	0,994755	0,2955 <sup>ns</sup>
NVI	0,4427387	3,2771693	0,962555	<0,0001*
NVNI	0,5490551	1,7796124	0,974562	0,0001*
SVI	1,2567547	17,002439	0,865663	<0,0001*
SVNI	0,0109032	2,5221149	0,961073	<0,0001*
NSI	0,8686761	5,4055397	0,950637	<0,0001*
NSNI	0,5657035	1,3434045	0,978579	0,0005*
PSI	1,0914729	6,0948002	0,901571	<0,0001*
PSNI	0,1329638	17,375184	0,830788	<0,0001*
FI	-1,014628	4,6333071	0,879631	<0,0001*
FNI	-1,71401	11,76586	0,645045	<0,0001*
PSVI	0,4873849	1,977694	0,9745	<0,0001*
PSVNI	0,4259785	2,526472	0,950539	<0,0001*
PSSI	2,5860714	20,334402	0,810543	<0,0001*
PSSNI	-0,224504	0,483523	0,989632	0,0504 <sup>ns</sup>

CI/CNI: Clorofila Irrigado/Não irrigado, AFI/AFNI: Área foliar Irrigado/Não irrigado, MFFI/MFFNI: Massa fresca folha irrigado/Não irrigado, MFCI/MFCNI: Massa fresca caule irrigado/Não irrigado, MSFI/MCFNI: Massa seca folha irrigado/Não irrigado, MSCI/MSCNI: Massa seca caule irrigado/Não irrigado, TII/TINI: temperatura da folha irrigado/Não irrigado, NVI/NVNI: Número de vagens irrigado/Não irrigado, SVI/SVNI: Número de sementes por vagem irrigado/Não irrigado, NSI/NSNI: Número de sementes irrigado/Não irrigado, PSI/PSNI: Massa de 100 sementes Irrigado/Não irrigado, FI/FNI: Dias até florescimento Irrigado/Não irrigado, PSVI/PSVNI: Massa seco de vagens Irrigado/Não irrigado, PSSI/PSSNI: Produtividade Irrigado/Não irrigado.

\* Nível de significância corresponde a  $p < 0,01$ ; <sup>ns</sup>: não significativo.

**Anexo 5.** Estatísticas descritivas para características quantitativas medidas na população AND277 X SEA5 do experimento em casa de vegetação em rizotrons.

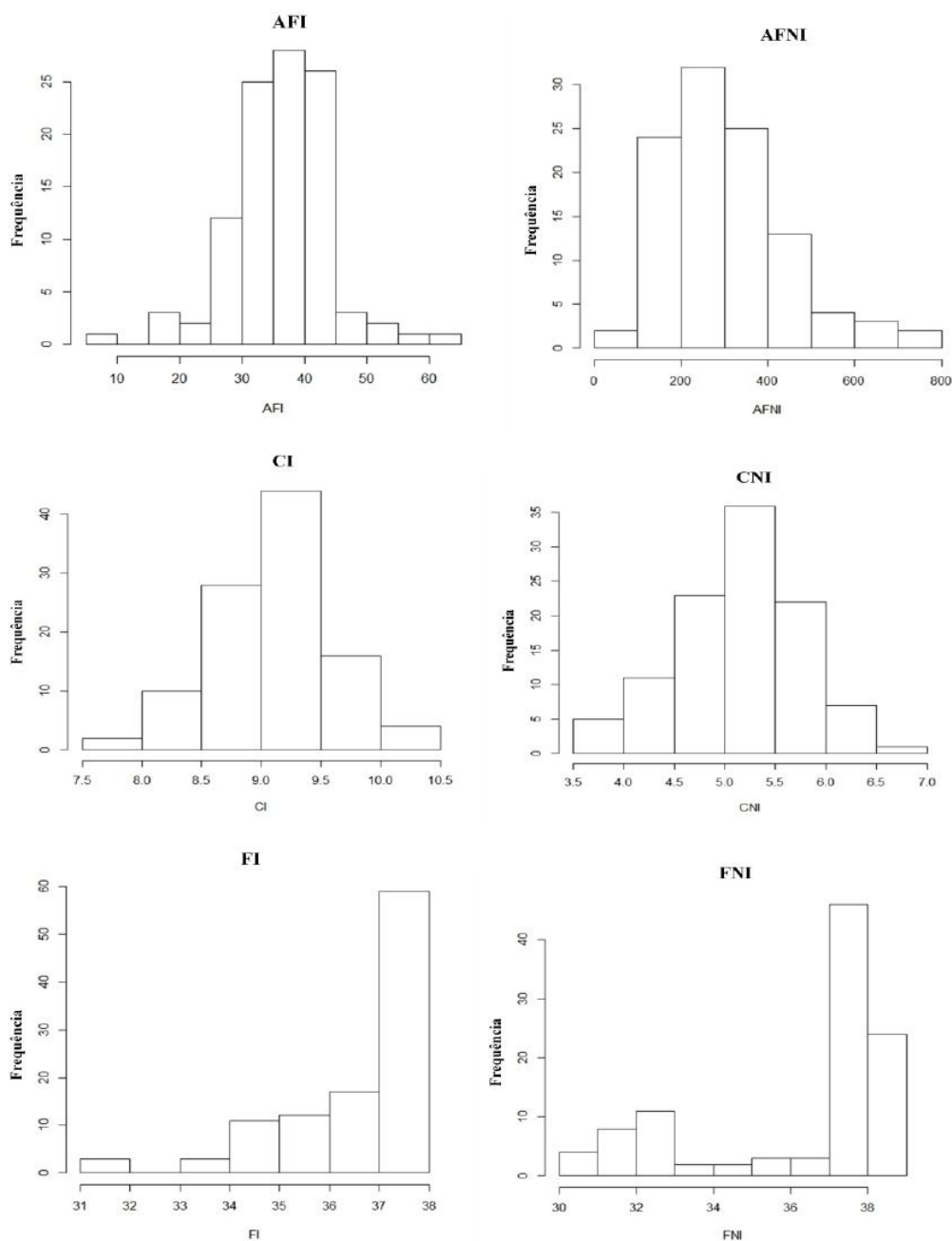
Característica	Skewness	Kurtosis	Shapiro-wilk	<i>p</i> valor
CI	-0,107757	0,7549668	0,98015	0,0077*
CNI	0,0423237	-0,035853	0,994268	0,5673 <sup>ns</sup>
AFI	-0,006149	1,2081877	0,978912	0,0059*
AFNI	0,0667326	1,0597483	0,98107	0,004*
MFFI	0,0711364	0,2302449	0,97863	0,0056*
MFFNI	0,1064899	27,584026	0,759573	<0,0001*
MFCI	0,043923	28,895651	0,721645	<0,0001*
MFCNI	0,0460647	0,4398309	0,991738	0,2241 <sup>ns</sup>
MSFI	0,1792994	1,6944576	0,92579	<0,0001*
MSFNI	0,2537405	0,5475163	0,982651	0,0183 <sup>ns</sup>
MSCI	0,258925	3,6340647	0,898345	<0,0001*
MSCNI	0,3612703	4,1237391	0,940278	<0,0001*
TII	0,5038891	2,4018606	0,937107	<0,0001*
TINI	-0,412564	1,6829834	0,929508	<0,0001*
CRI	0,437326	1,5561357	0,977555	0,0026*
CRNI	0,1341793	0,5587391	0,991498	0,2126 <sup>ns</sup>
ASRI	0,137251	3,3291797	0,960749	<0,0001*
ASRNI	0,4091746	2,2150666	0,96551	<0,0001*
VRI	0,0282132	1,2924399	0,982619	0,0142 <sup>ns</sup>
VRNI	0,1383549	1,8750819	0,964718	<0,0001*
DRI	0,0973758	2,6351945	0,966574	0,0001*
DRNI	0,9989349	6,0990852	0,870029	<0,0001*

CI/CNI: Clorofila Irrigado/Não irrigado, AFI/AFNI: Área foliar Irrigado/Não irrigado, MFFI/MFFNI: Massa fresca folha irrigado/Não irrigado, MFCI/MFCNI: Massa fresca caule irrigado/Não irrigado, MSFI/MCFNI: Massa seca folha irrigado/Não irrigado, MSCI/MSCNI: Massa seca caule irrigado/Não irrigado, TII/TINI: temperatura da folha irrigado/Não irrigado, CRI/CRNI: Comprimento de raízes Irrigado/Não irrigado, ASRI/ASRNI: Área superficial de raízes Irrigado/Não irrigado, APRI/APRNI: Área projetada de raízes Irrigado/Não irrigado, VRI/VRNI: Volume de raízes Irrigado/Não irrigado, DRI/DRNI: Diâmetro de raízes Irrigado/Não irrigado.

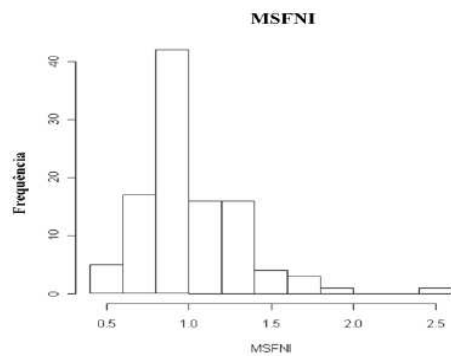
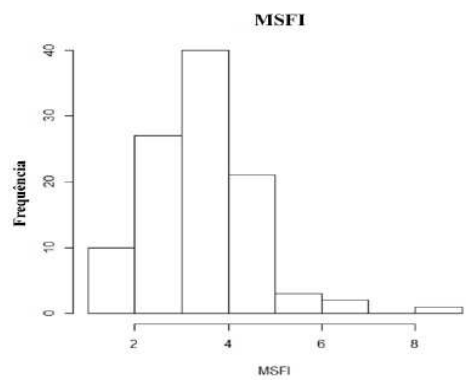
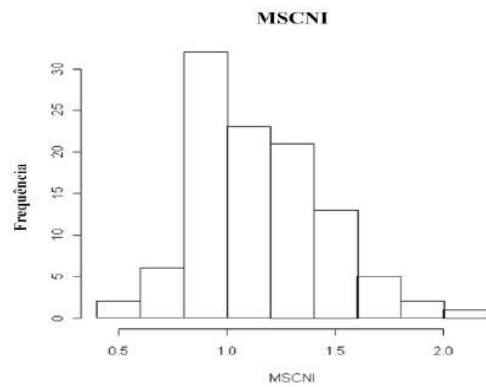
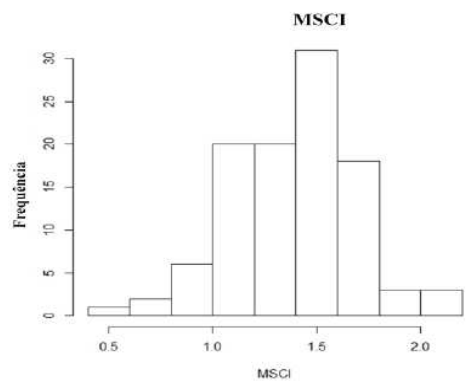
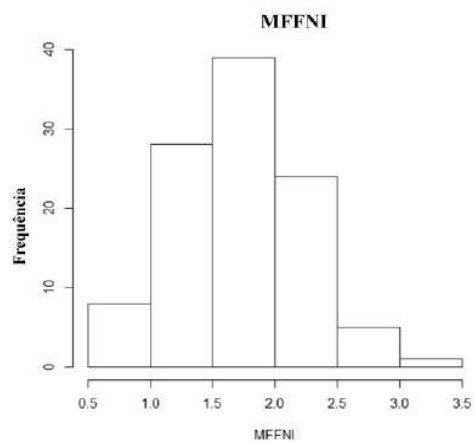
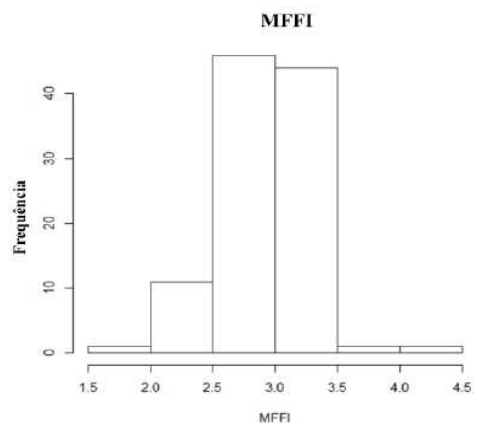
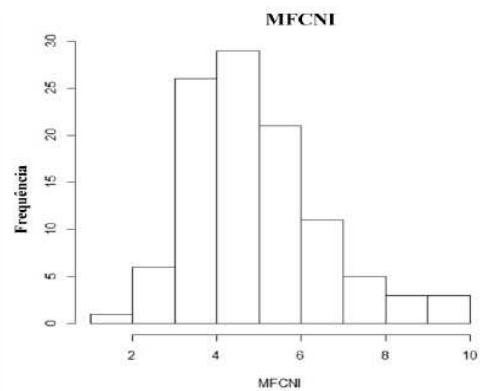
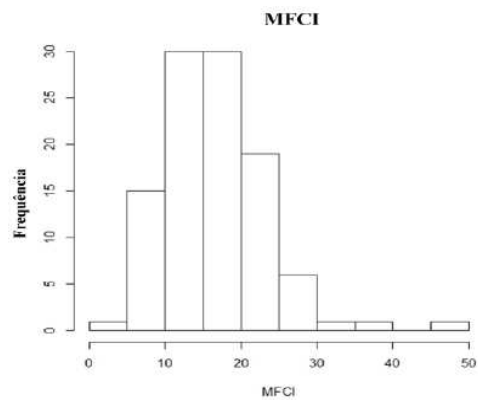
\* Nível de significância corresponde a  $p < 0,01$ .

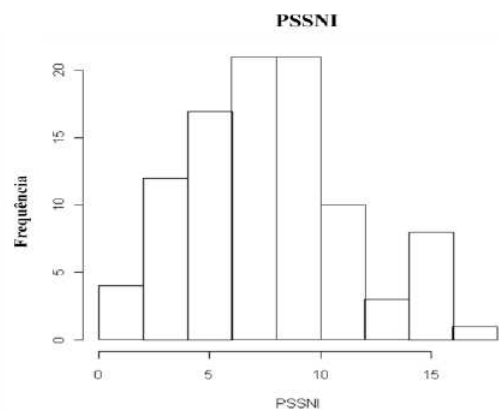
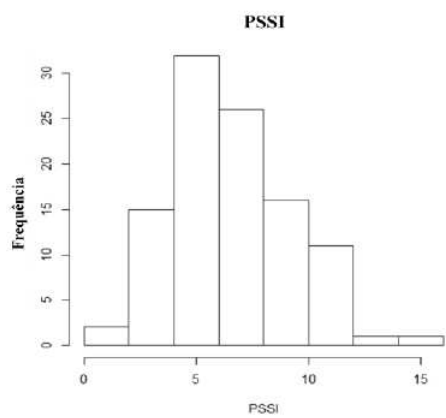
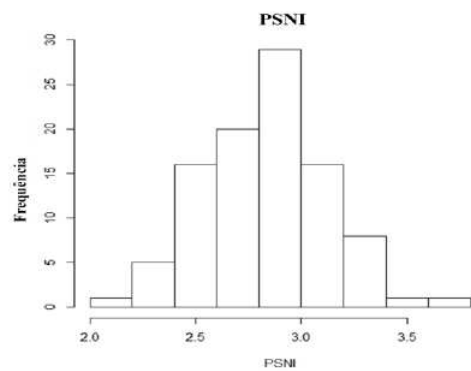
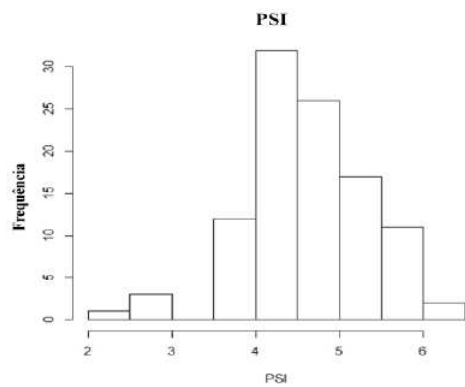
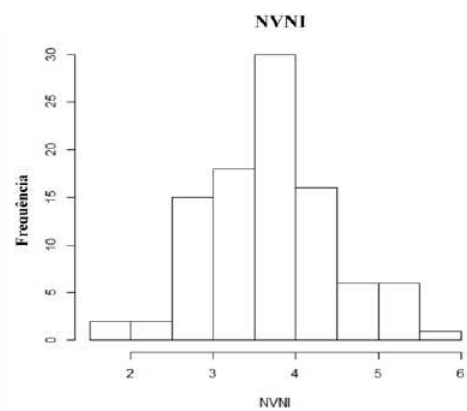
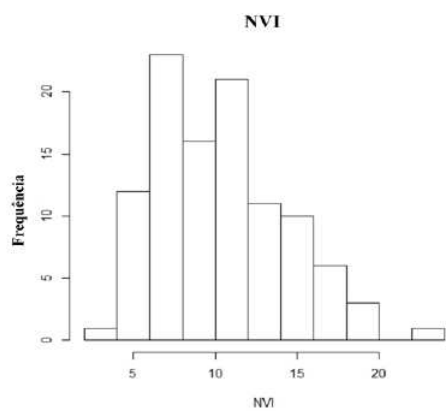
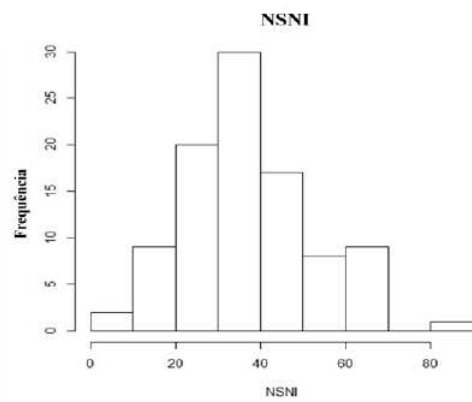
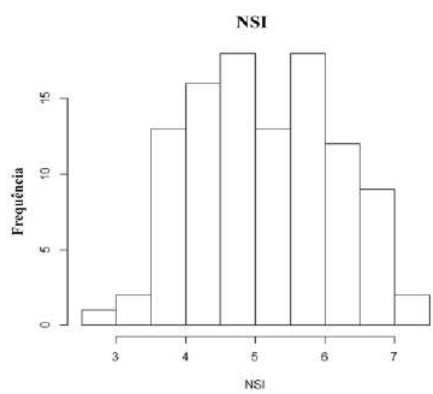
<sup>ns</sup>: não significativo

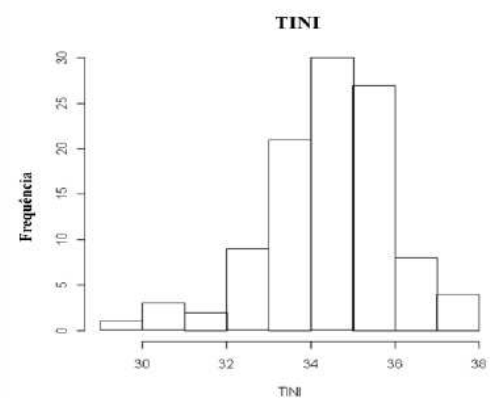
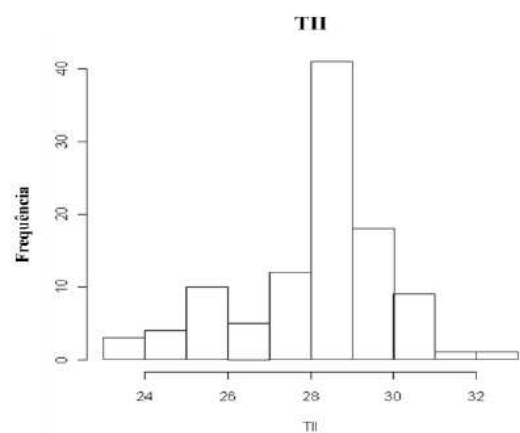
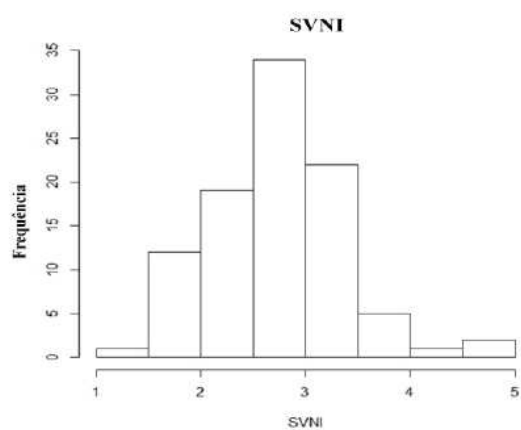
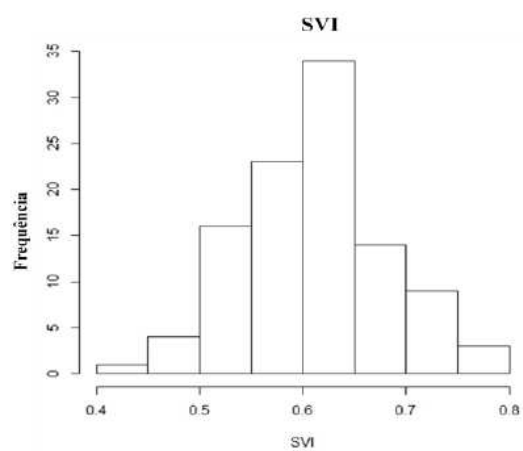
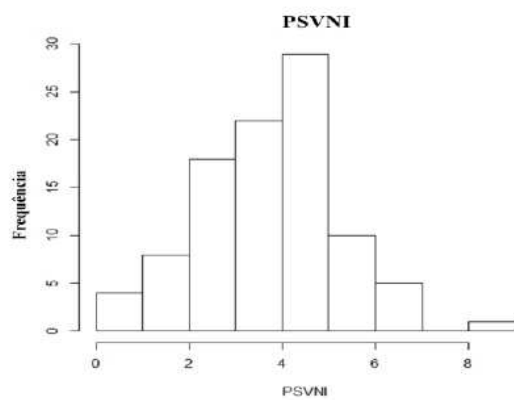
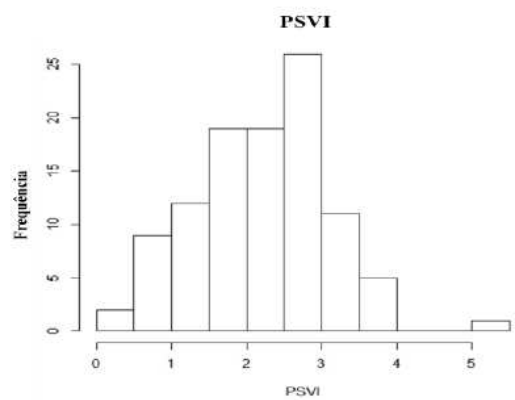
**Anexo 6.** Histogramas dos valores normalizados das características avaliadas na população AS durante o experimento em casa de vegetação em vasos em ambiente irrigado e não irrigado.



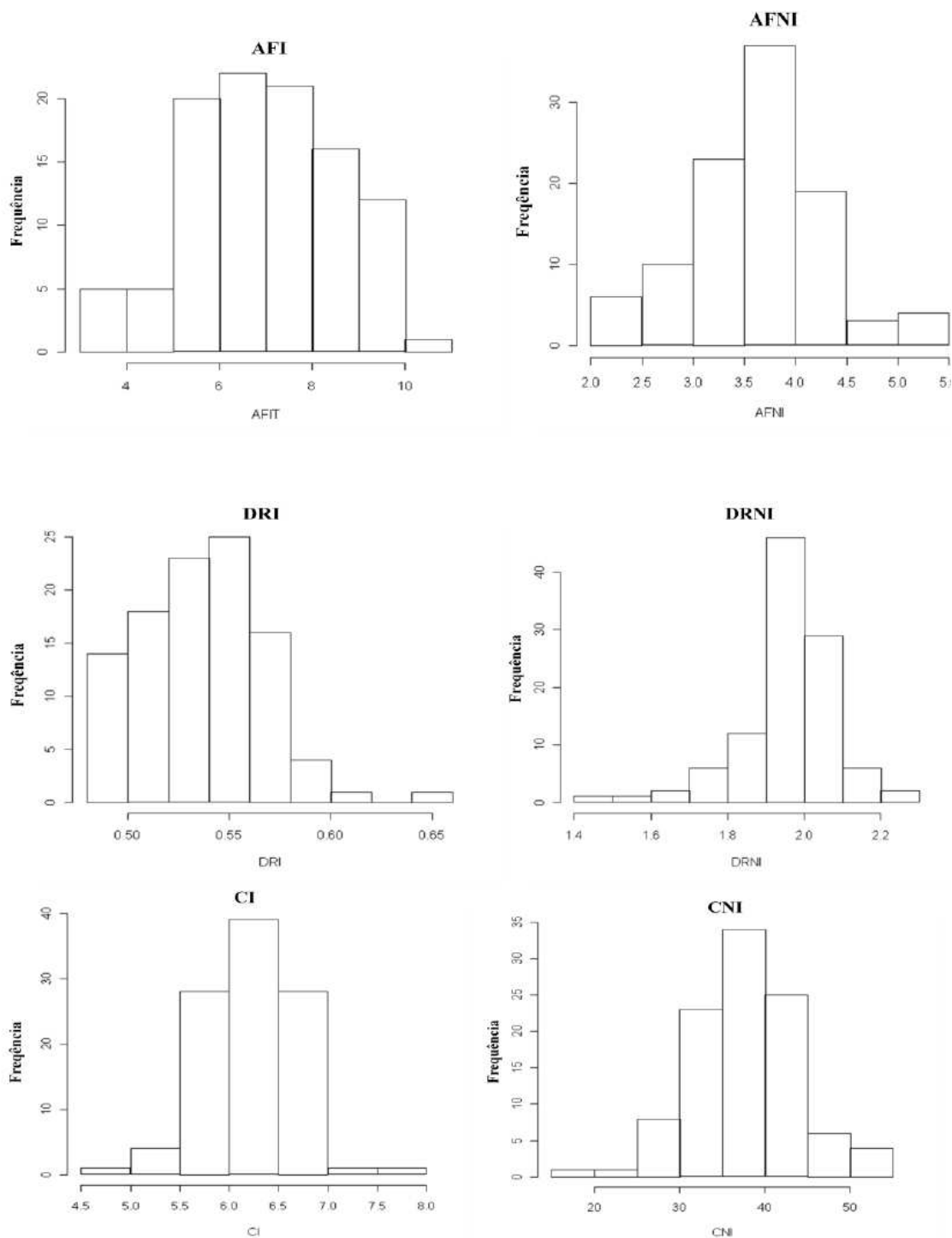


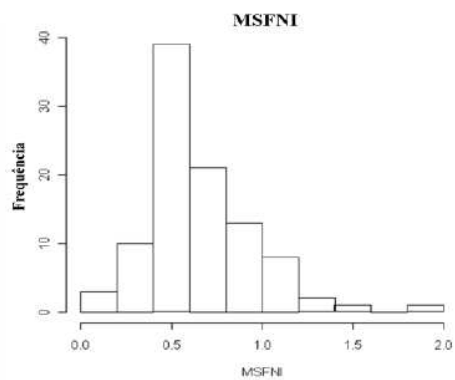
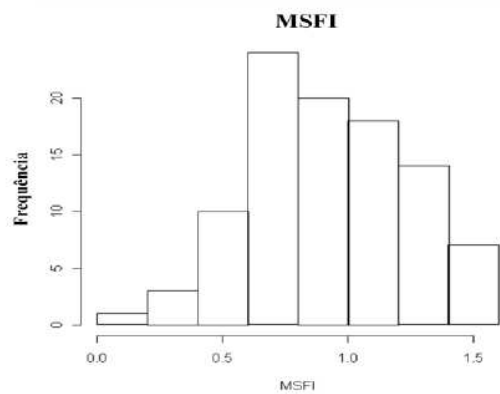
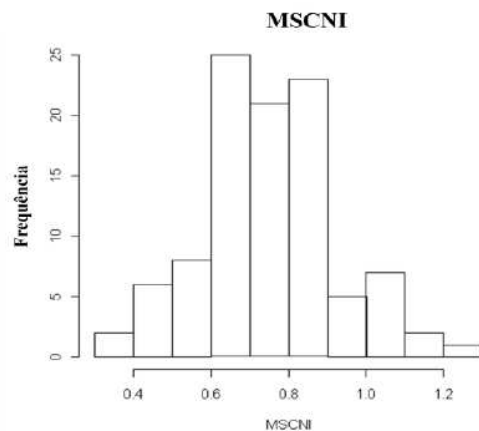
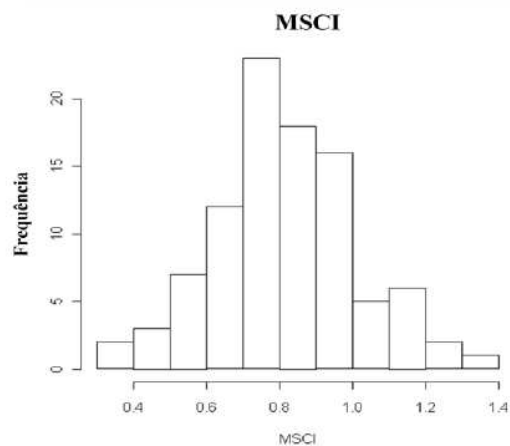
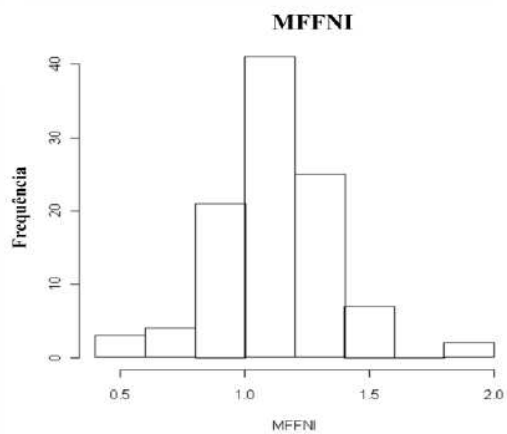
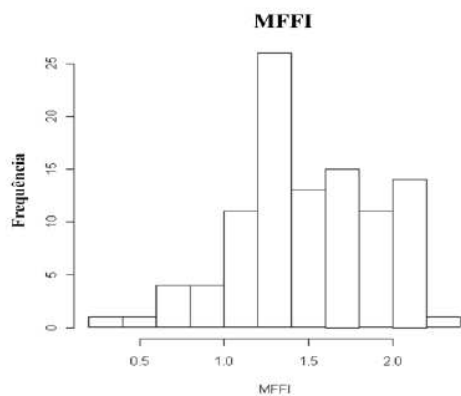
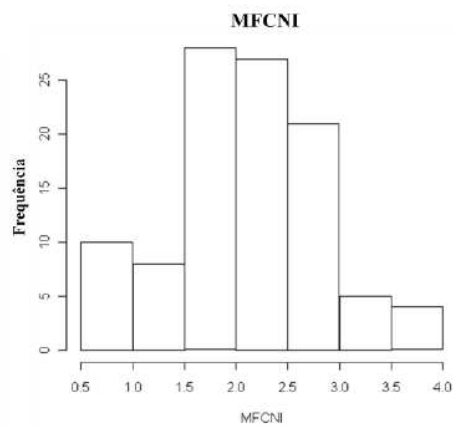
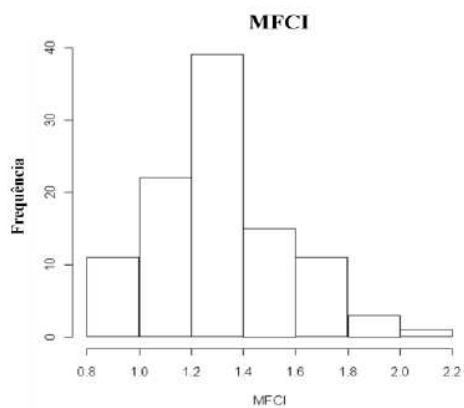


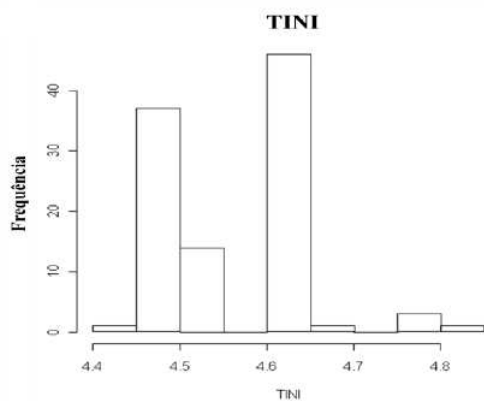
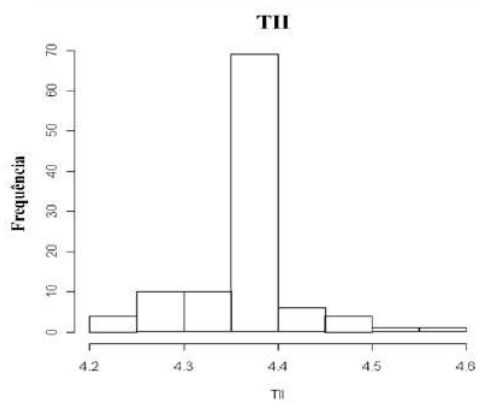
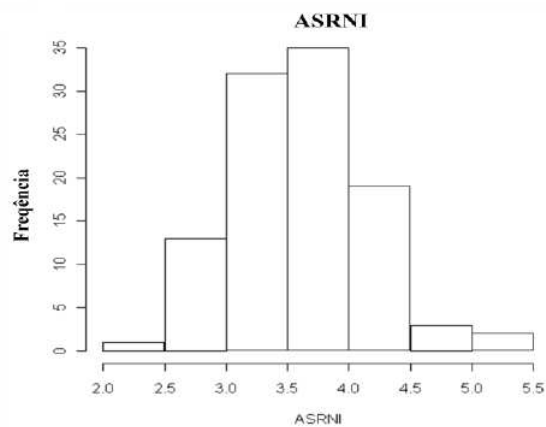
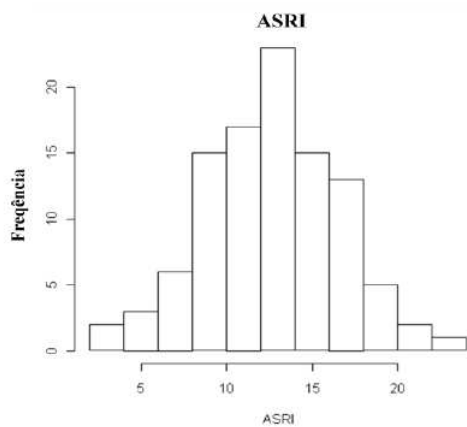
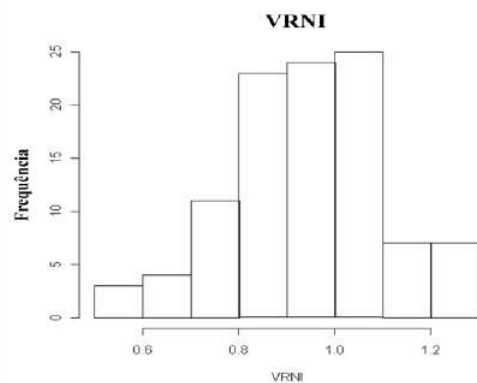
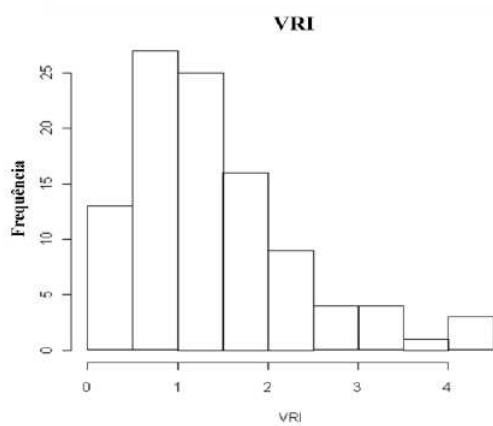
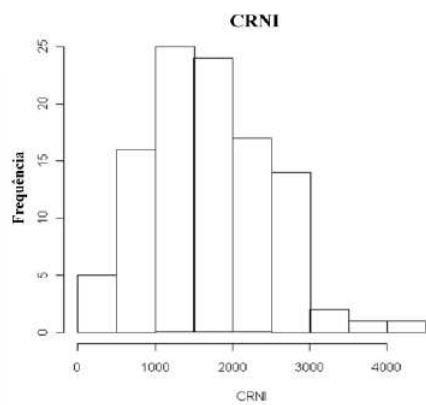
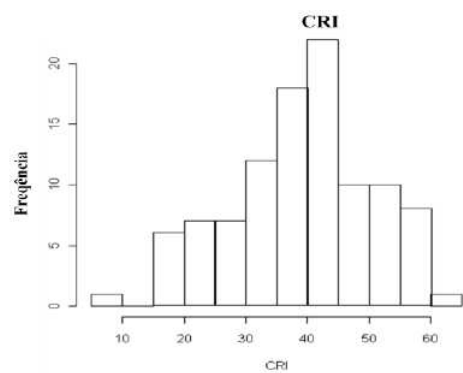




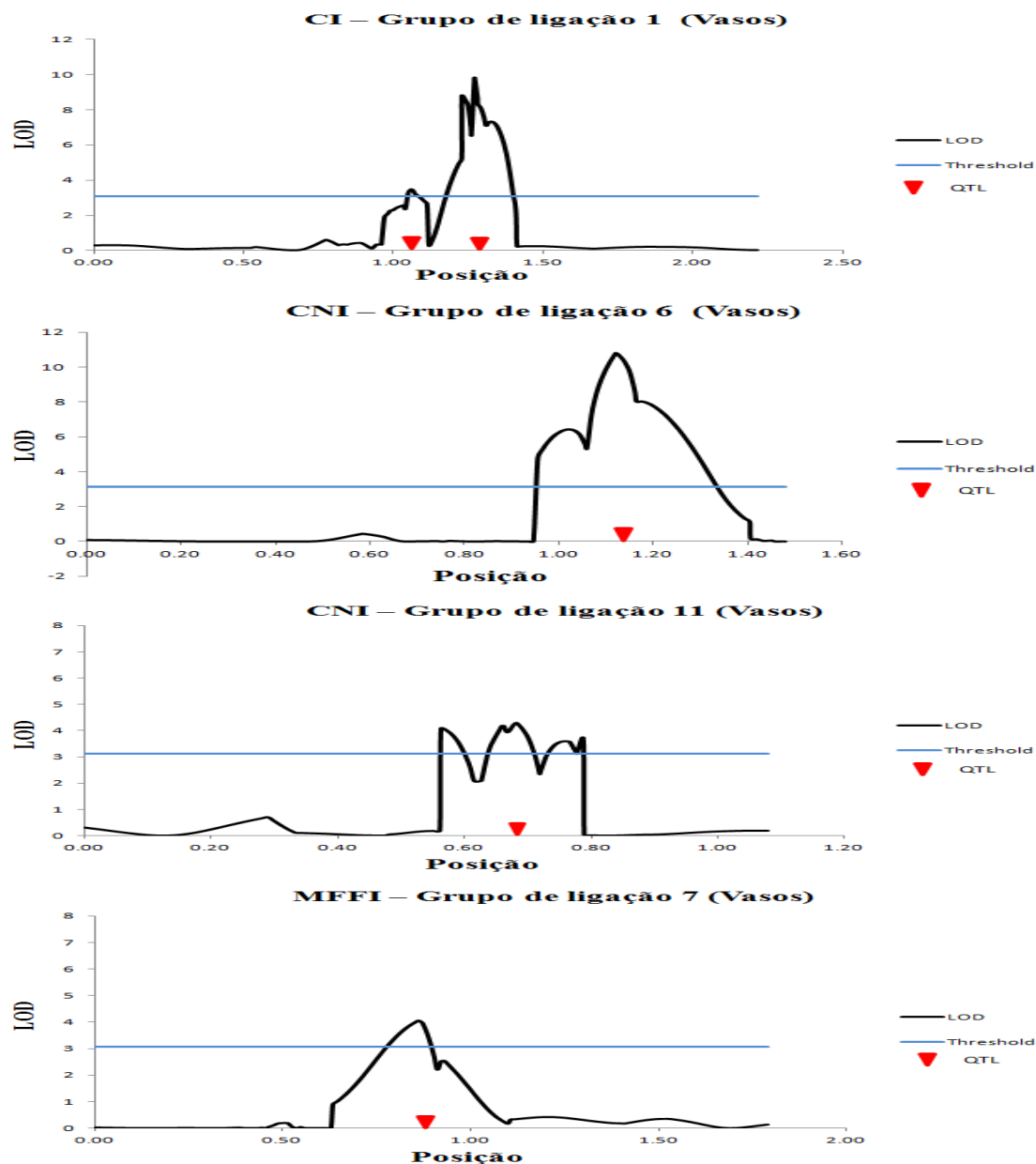
**Anexo 7.** Histogramas dos valores normalizados das características avaliadas na população AS durante o experimento em casa de vegetação em rizotrons em ambiente irrigado e não irrigado.

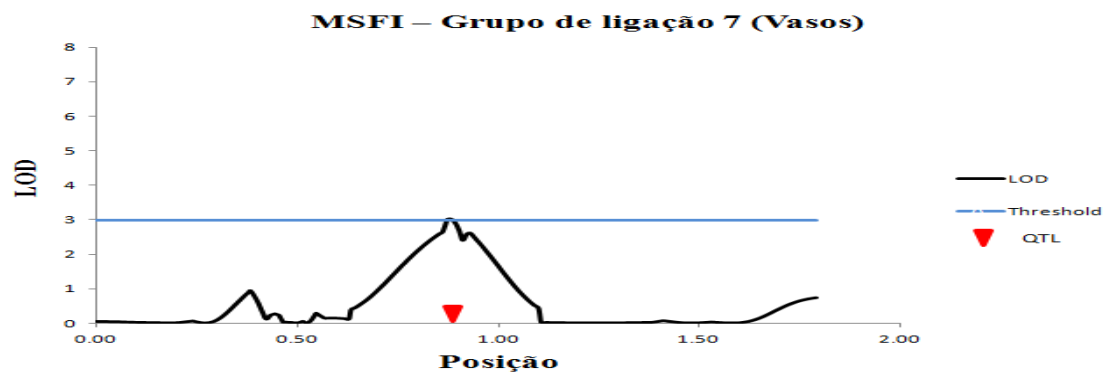
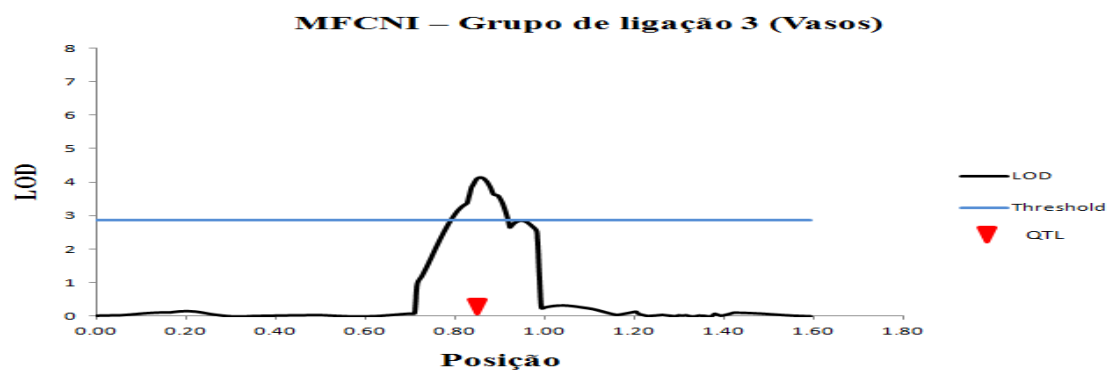
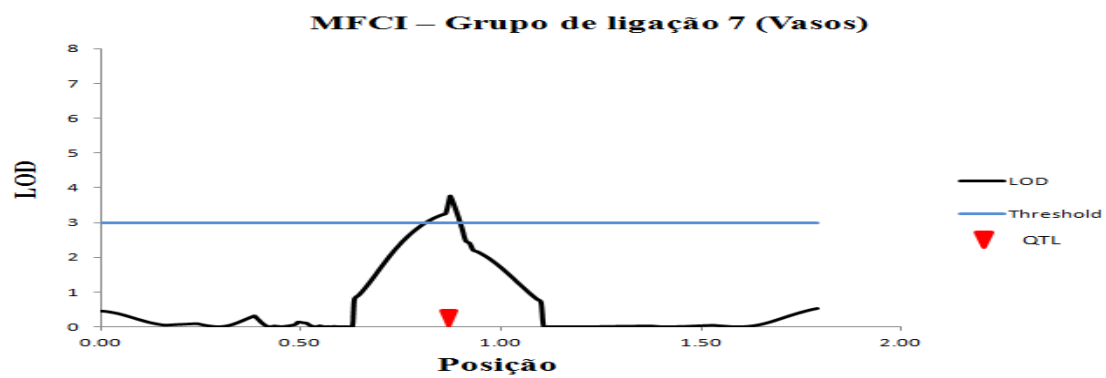
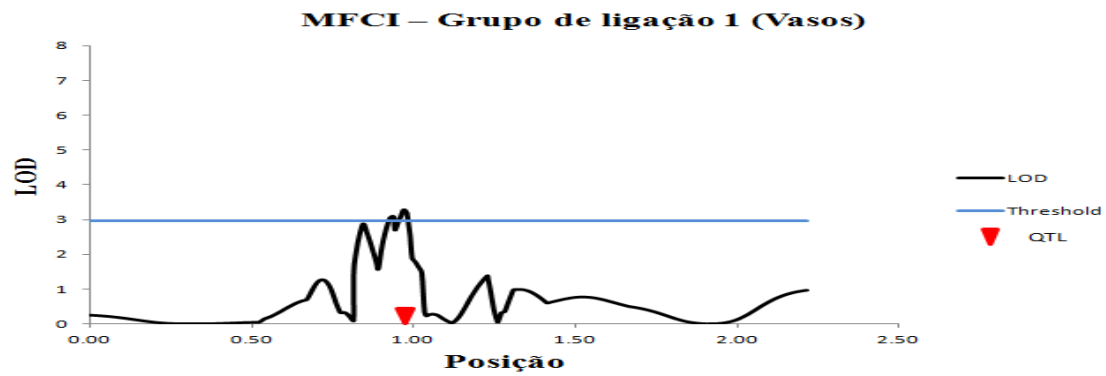






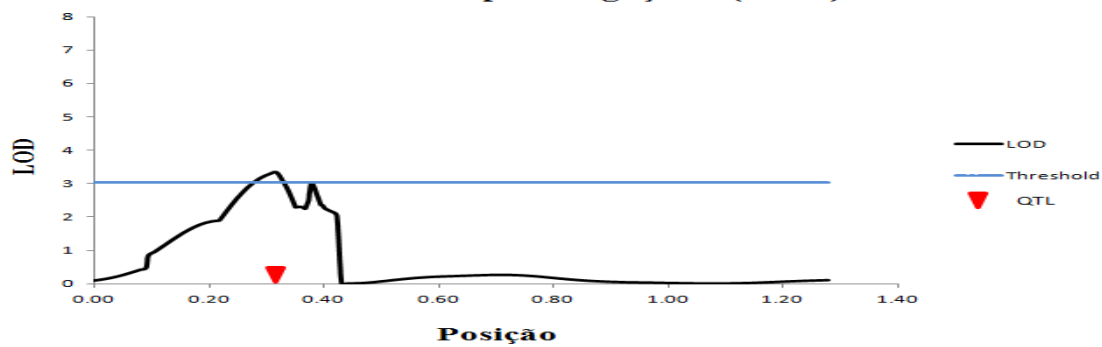
**Anexo 8.** Gráficos dos valores de LOD encontrados pela análise de CIM para a identificação dos QTLs relacionados à seca em experimento realizados em vasos e em rizotrons.



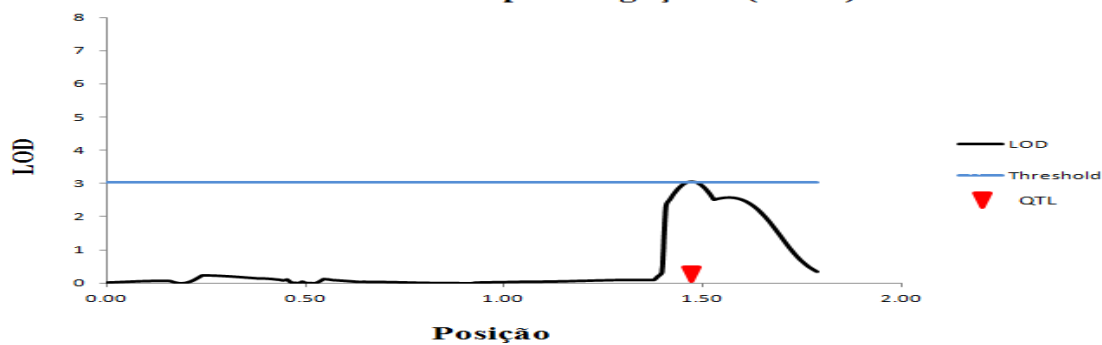




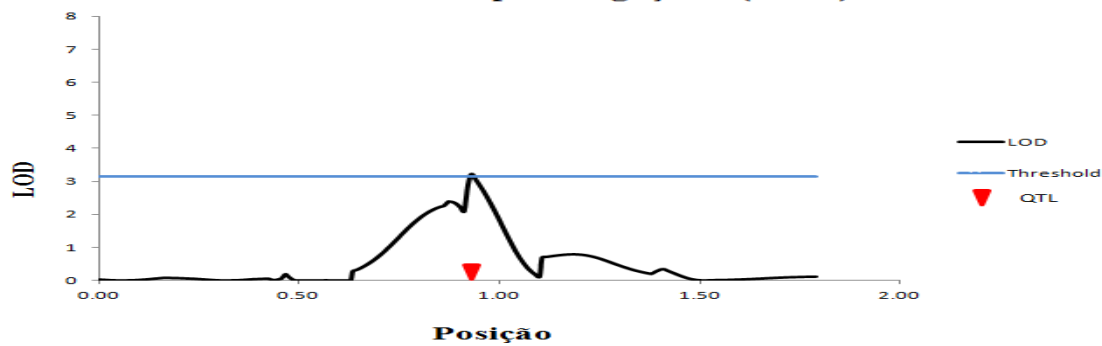
**TII – Grupo de ligação 4 (Vasos)**



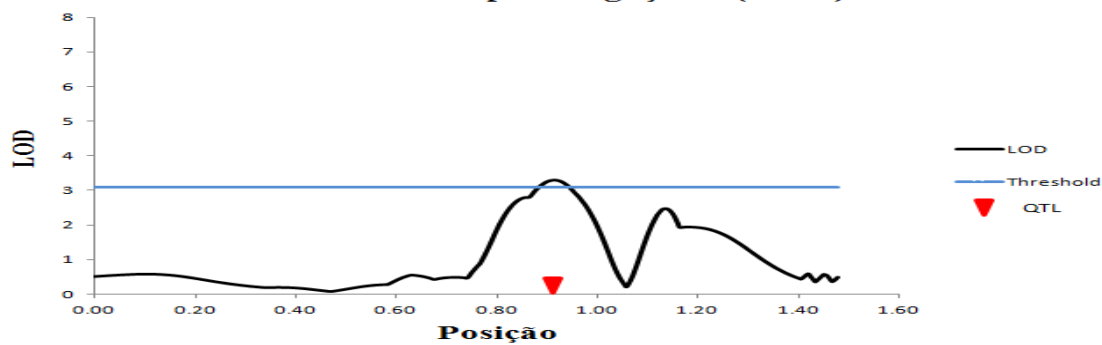
**TINI – Grupo de ligação 7 (Vasos)**

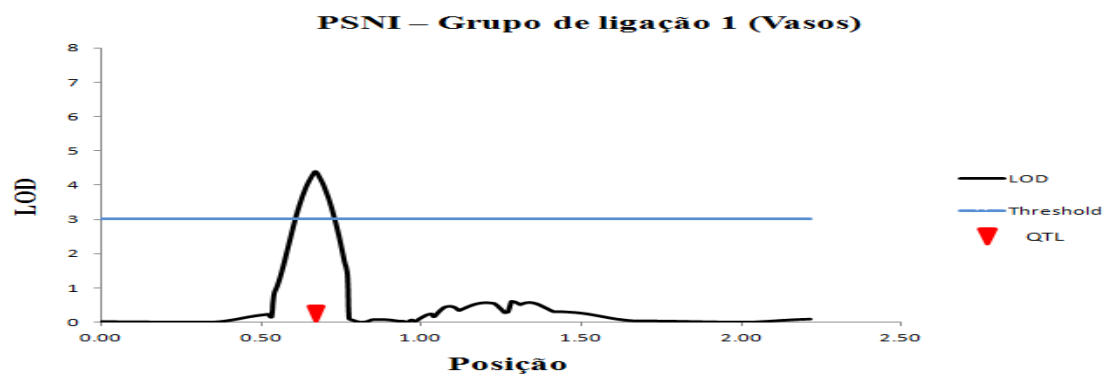
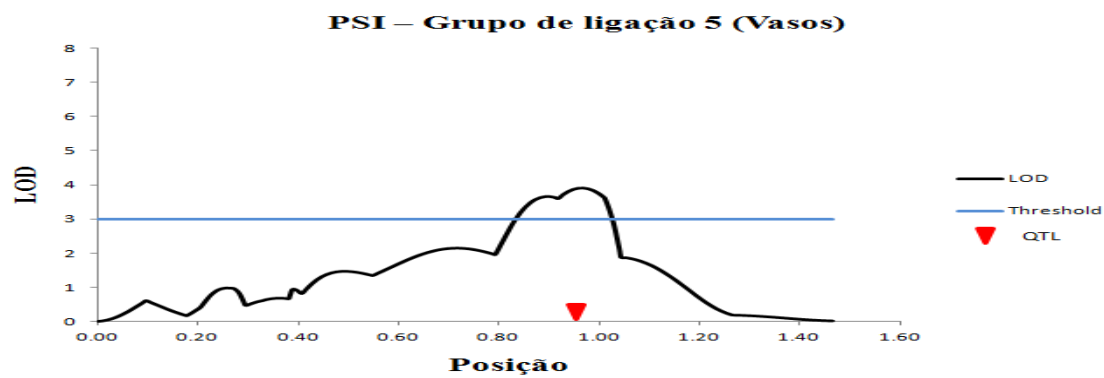
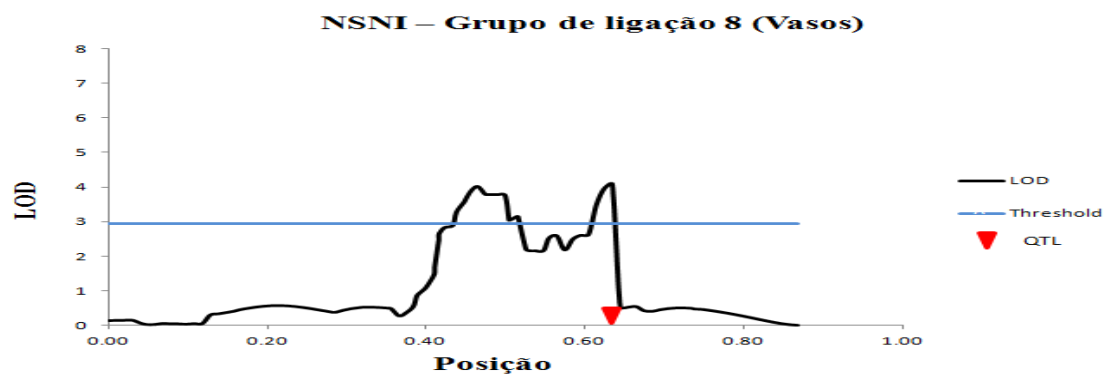
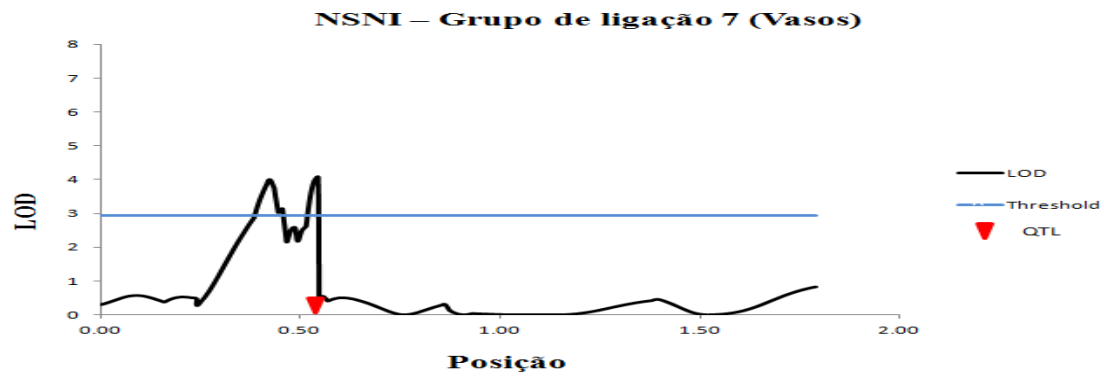


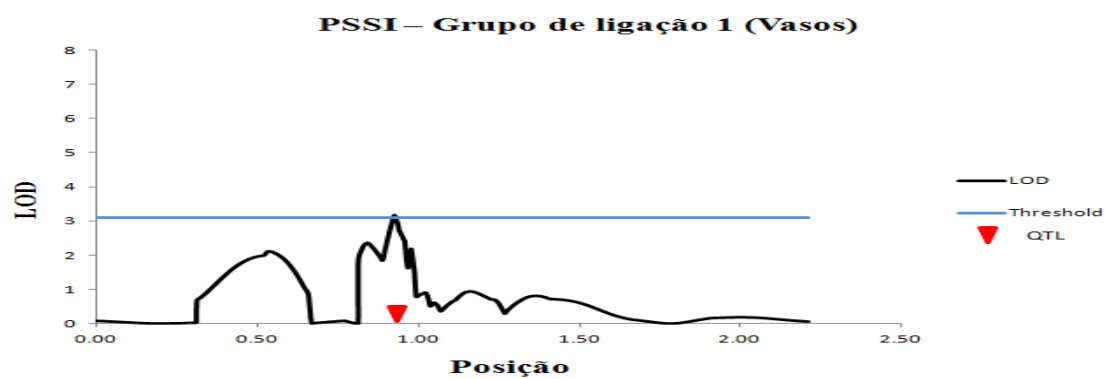
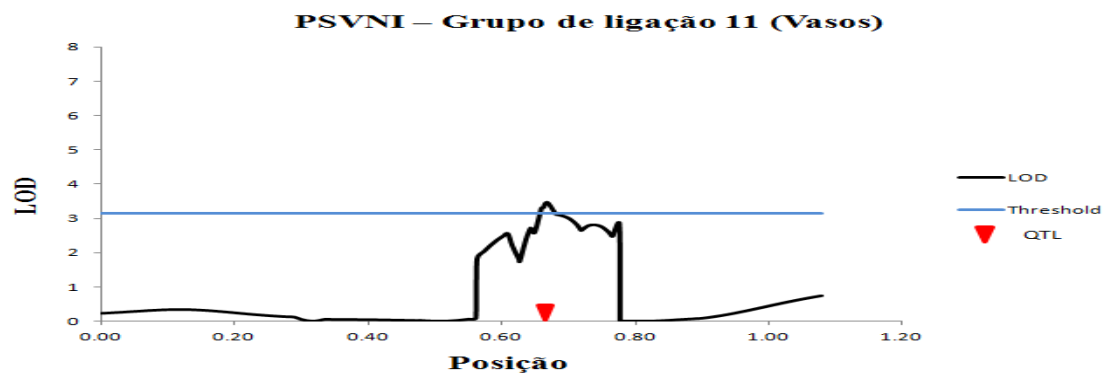
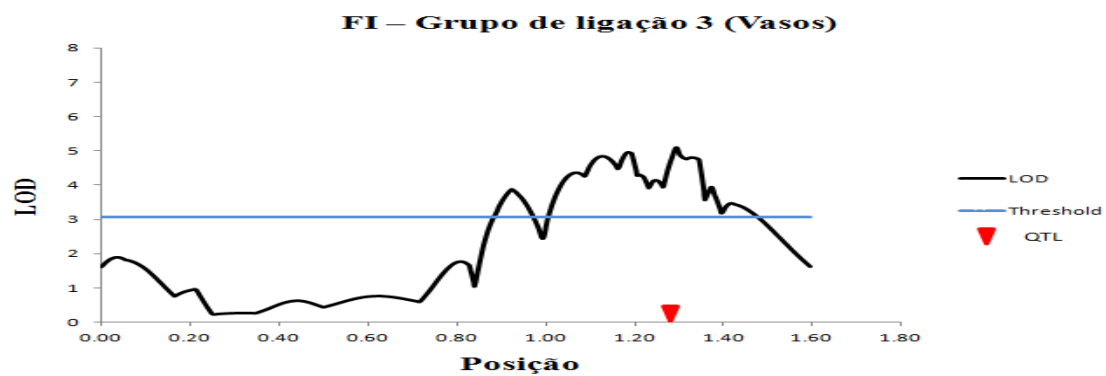
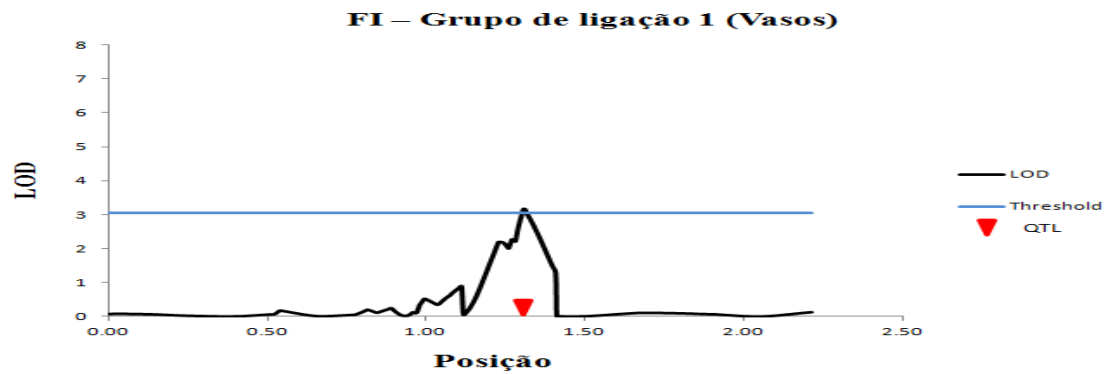
**NVI – Grupo de ligação 7 (Vasos)**

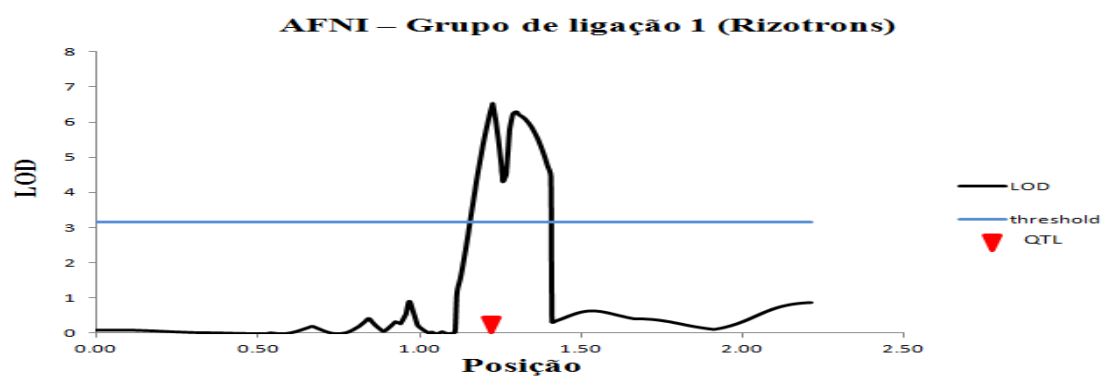
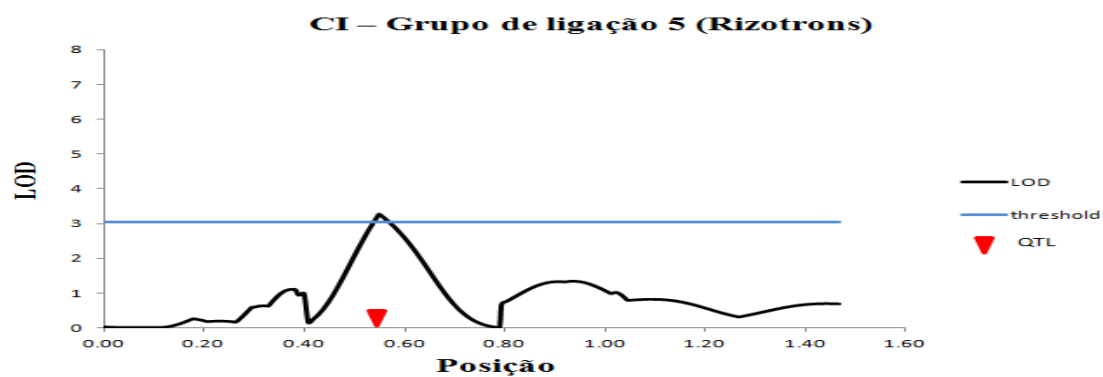
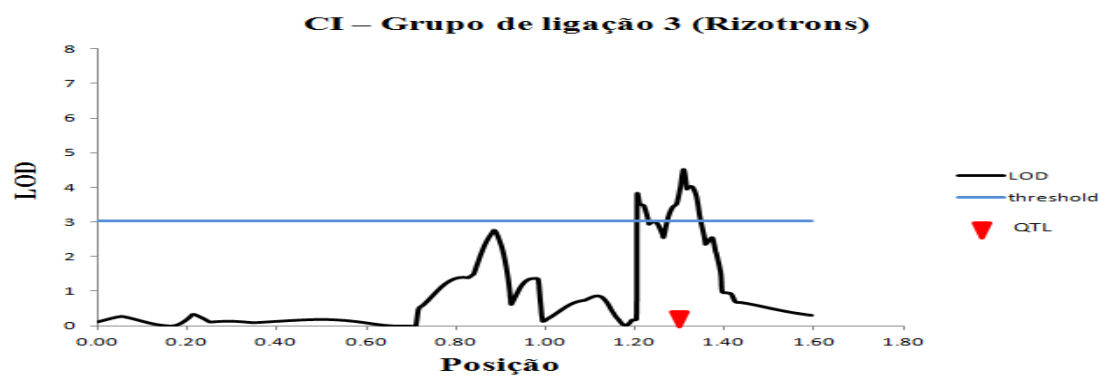
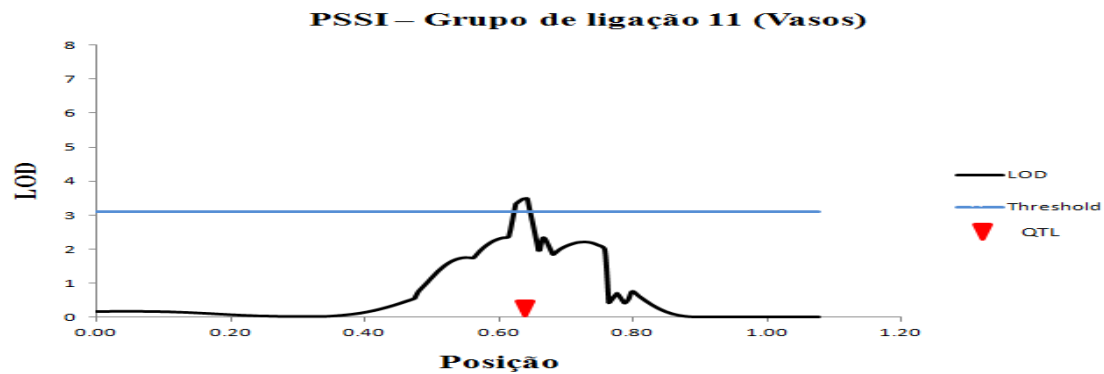


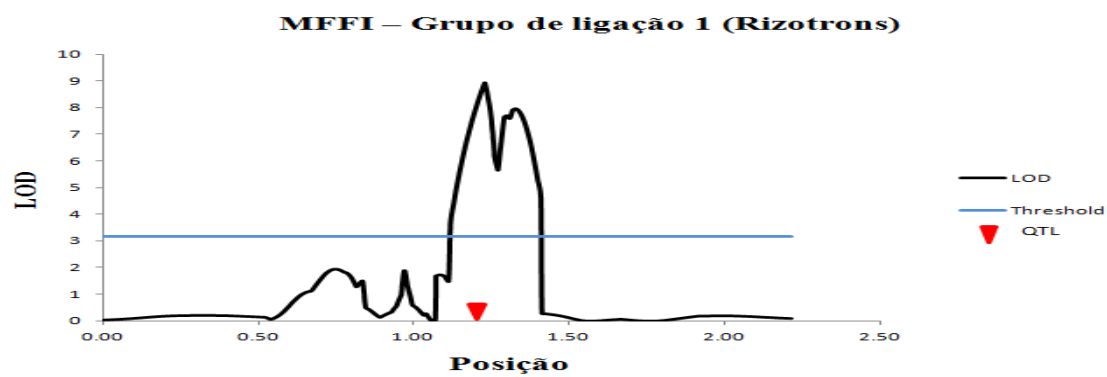
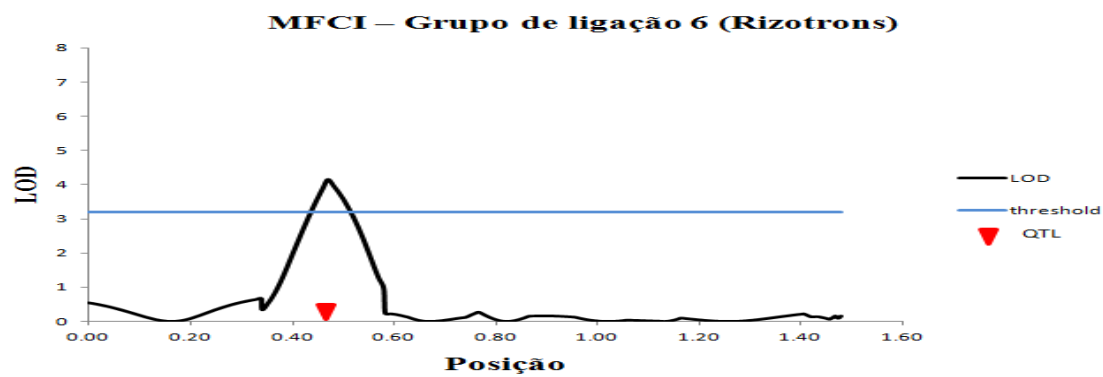
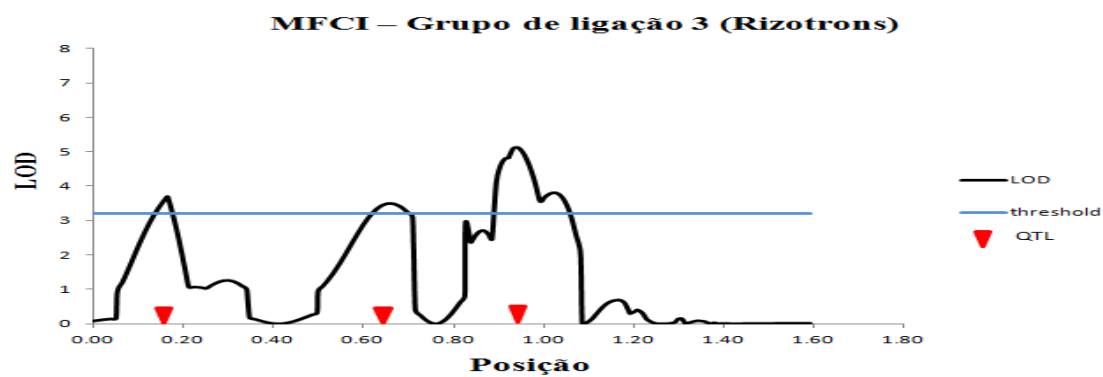
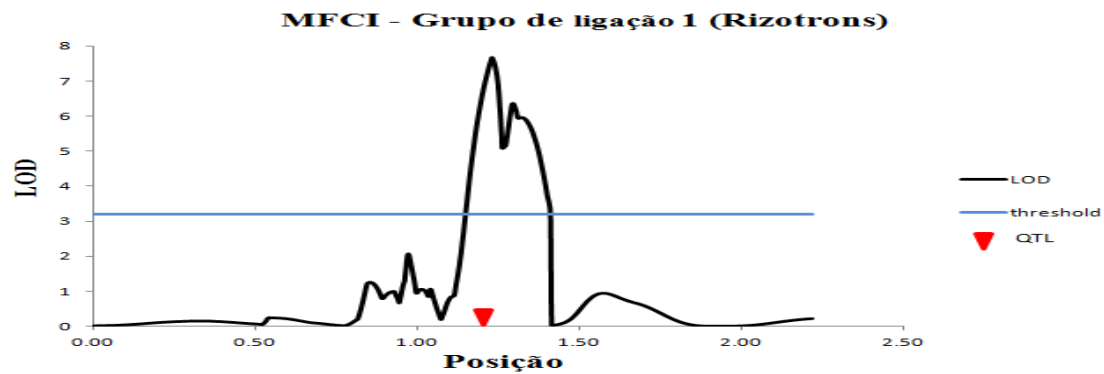
**NSI – Grupo de ligação 6 (Vasos)**

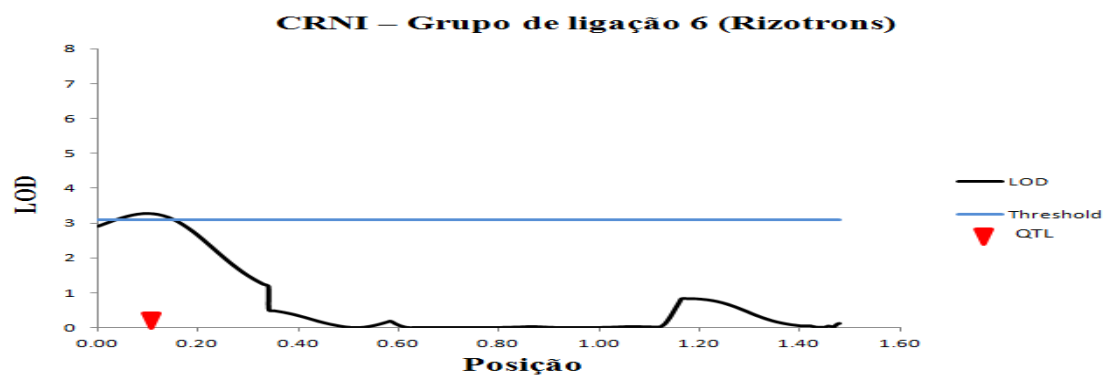
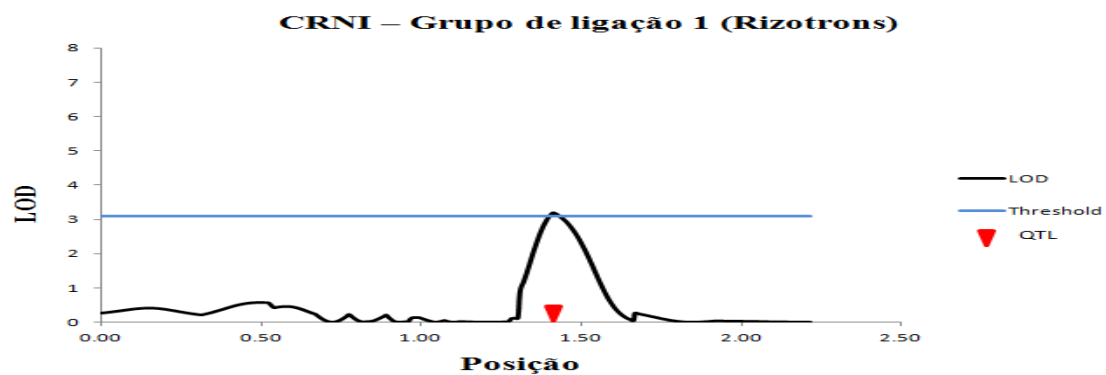
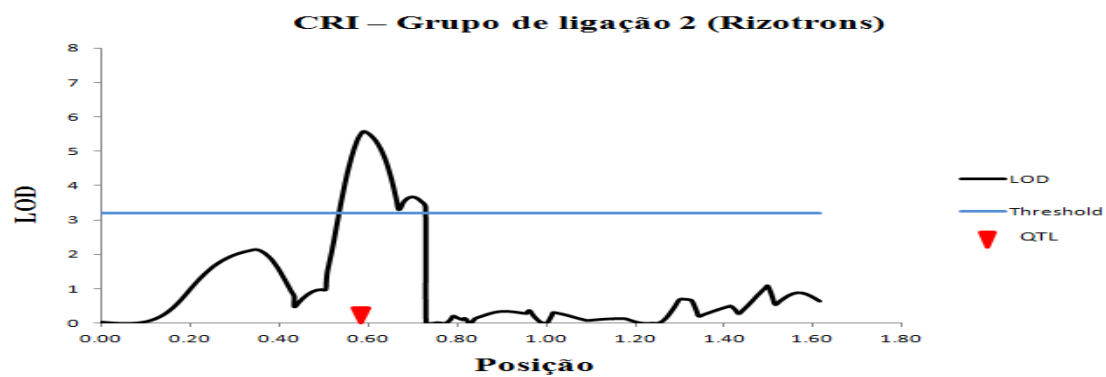
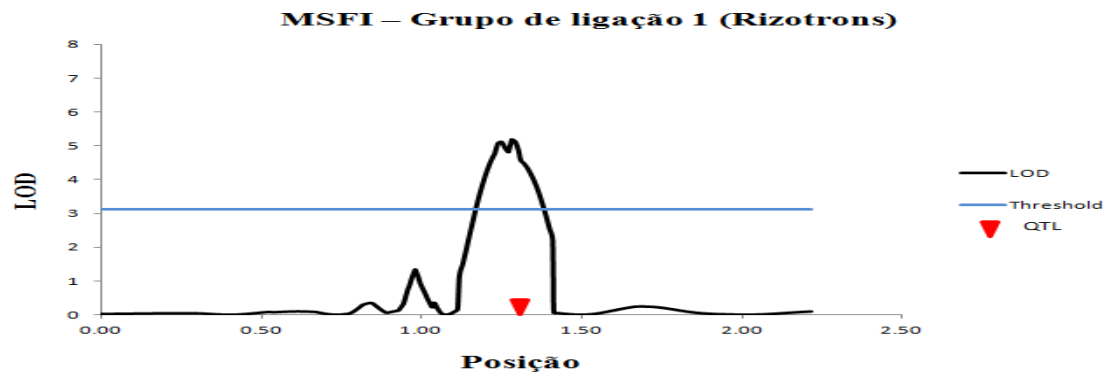


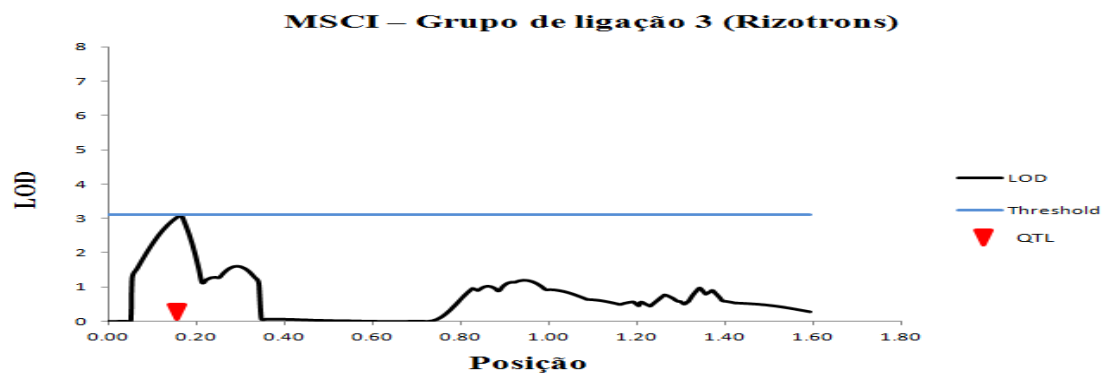
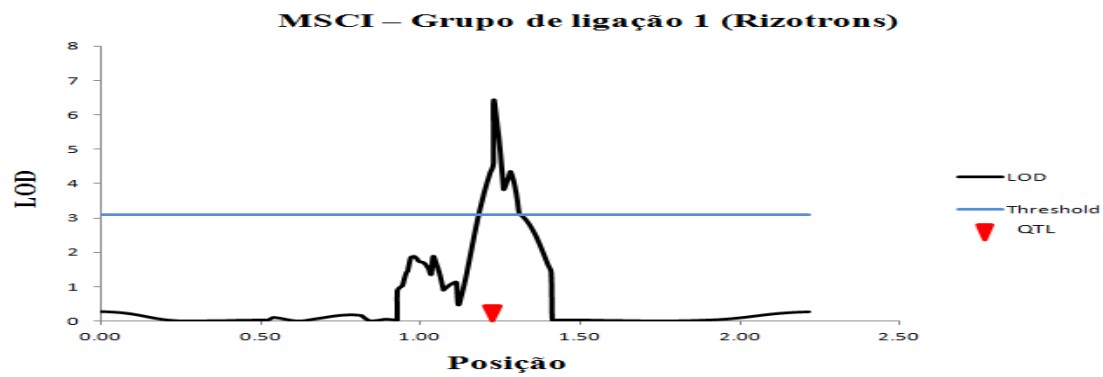
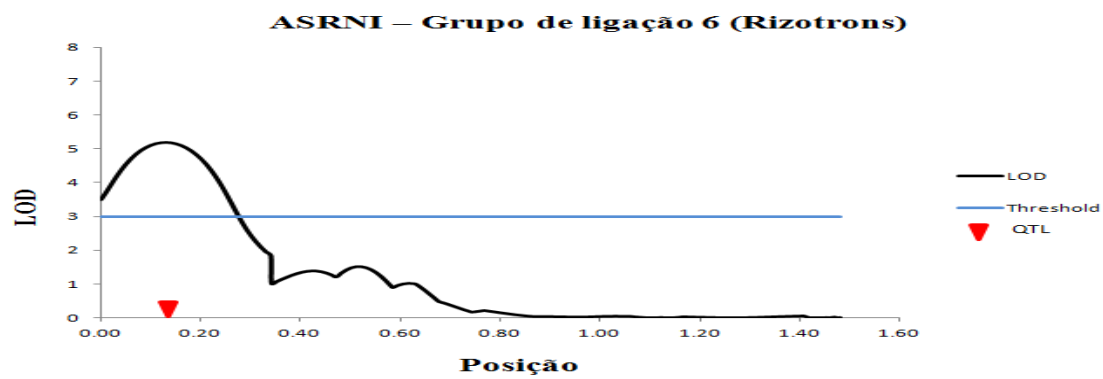
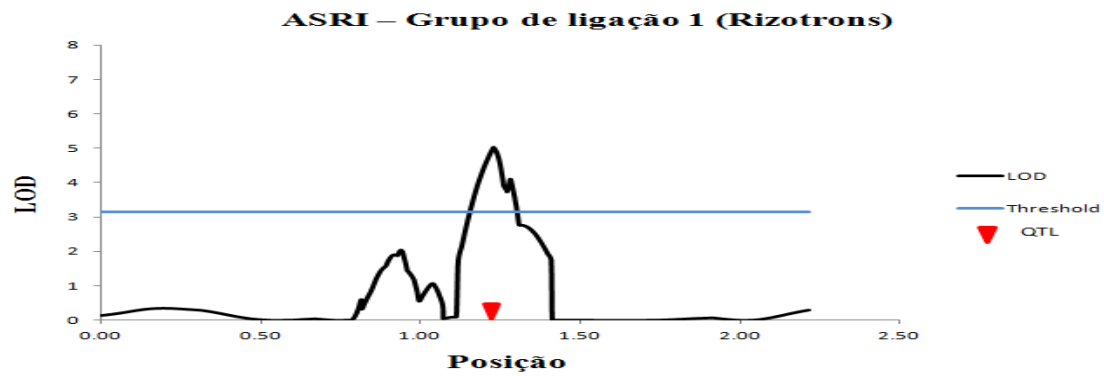




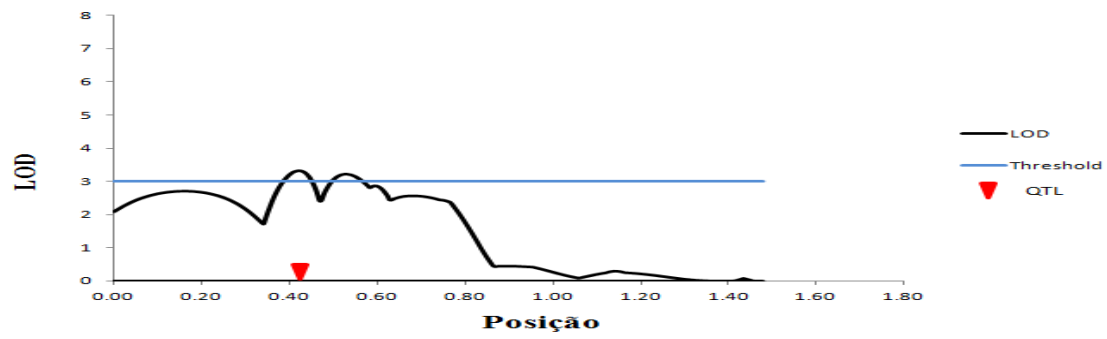








### VRI – Grupo de ligação 6 (Rizotrons)



### DRI – Grupo de ligação 6 (Rizotrons)

