



UNIVERSIDADE
ESTADUAL de LONDRINA

LUAN GONÇALVES CRUZ

**DESENVOLVIMENTO DE FERRAMENTAS MOLECULARES
E FENOTÍPICAS PARA AVALIAÇÃO DOS TIPOS DE
CRESCIMENTO DA SOJA (G. MAX)**

Londrina
2017

LUAN GONÇALVES CRUZ

**DESENVOLVIMENTO DE FERRAMENTAS MOLECULARES
E FENOTÍPICAS PARA AVALIAÇÃO DOS TIPOS DE
CRESCIMENTO DA SOJA (G. MAX)**

Dissertação apresentada ao programa de Pós Graduação em Agronomia, da Universidade Estadual de Londrina, como requisito parcial à obtenção do título de Mestre em Agronomia.

Orientador: Prof. Dr. Cássio Egidio Cavenaghi Prete.

Londrina
2017

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

Cruz, Luan Gonçalves.

DESENVOLVIMENTO DE FERRAMENTAS MOLECULARES E FENOTÍPICAS PARA AVALIAÇÃO DOS TIPOS DE CRESCIMENTO DA SOJA (G. MAX) / Luan Gonçalves Cruz. - Londrina, 2017.

45 f.: il.

Orientador: Cássio Egidio Cavenaghi Prete.

Dissertação (Mestrado em Agronomia) - Universidade Estadual de Londrina, Centro de Ciências Agrárias, Programa de Pós-Graduação em Agronomia, 2017.

Inclui bibliografia.

1. Mapeamento associativo - Tese. 2. Polimorfismo de Base Única - Tese. 3. Melhoramento de soja - Tese. I. Prete, Cássio Egidio Cavenaghi. II. Universidade Estadual de Londrina. Centro de Ciências Agrárias. Programa de Pós-Graduação em Agronomia. III. Título.

LUAN GONÇALVES CRUZ

**DESENVOLVIMENTO DE FERRAMENTAS MOLECULARES E
FENOTÍPICAS PARA AVALIAÇÃO DOS TIPOS DE CRESCIMENTO
DA SOJA (G. MAX)**

Dissertação apresentada ao programa de Pós Graduação em Agronomia, da Universidade Estadual de Londrina, como requisito parcial à obtenção do título de Mestre em Agronomia.

BANCA EXAMINADORA

Orientador: Prof. Dr. Cássio Egidio Cavenaghi Prete
Universidade Estadual de Londrina - UEL

PhD. Anderson Rotter Meda
Tropical Melhoramento & Genética - TMG

Dr. Ricardo Gonçalves
Universidade Estadual de Londrina - UEL

Dr. Leandro Simões Azeredo Gonçalves
Universidade Estadual de Londrina - UEL

PhD. Adriana Maria Polizel Podanosqui
Tropical Melhoramento & Genética - TMG

Londrina, 22 de março de 2017

AGRADECIMENTOS

Primeiramente a Deus, pelo dom da vida, e sempre presente em momentos difíceis.

A meus pais, Jair e Célia e minha irmã Michele por minha formação de caráter e valores.

A minha esposa Bruna e meu filho Eduardo pelo amor, carinho e compreensão.

A todos os colaboradores da Tropical Melhoramento & Genética em especial, Dr. Alexandre Garcia, PhD. Anderson Meda, Sr. Claudinei Rios, PhD. João Maldonado, Dr. Romeu Kiihl, pelo suporte, compreensão e todo apoio para o desenvolvimento do trabalho.

Ao Programa de Pós-Graduação em Agronomia da Universidade Estadual de Londrina pela oportunidade de realização do Mestrado.

Ao orientador Dr. Cássio pelo auxílio e direcionamento dos trabalhos.

A todos que de alguma forma contribuíram com este trabalho.

Muito obrigado.

CRUZ, Luan Gonçalves. **Desenvolvimento de ferramentas moleculares e fenotípicas para avaliação dos tipos de crescimento da soja (G. Max)**. 2017. 45 f. Dissertação (Mestrado em Agronomia) – Universidade Estadual de Londrina, Londrina, 2017.

RESUMO

As interações de diversos fatores determinam a duração do ciclo vegetativo na cultura da soja, dentre eles podem ser destacados alguns fatores externos tais como; temperatura, duração do período do dia, qualidade e intensidade da luz, e os internos como; plantas insensíveis ao comprimento do dia (período juvenil longo), e os tipos de crescimento da haste da soja que podem ser distribuídos em três sendo eles: indeterminado, determinado e semideterminado. A classificação em três tipos se deve ao fato de apresentarem características distintas entre eles; arquitetura, duração do período de florescimento, altura final, distância entre os dois últimos nós reprodutivos, produção de racemo terminal. A expressão desses fenótipos pode ser influenciada por condições edafoclimáticas e a interação epistática entre *Dt1* e *Dt2*, fato que dificulta o processo de caracterização. Os loci *Dt1* e *Dt2* já foram mapeados, no entanto são necessários esclarecimentos sobre a região do locus *Dt2* e sua interação com *Dt1*. Foram analisados polimorfismos no genoma da soja (tipo SNPs) tendo como objetivo validar marcadores moleculares associados ao loci *Dt1* e *Dt2* visando a fácil caracterização e distinção dos tipos de crescimento terminal da haste da soja aplicável a programas de melhoramento. Se utilizou duas abordagens para identificar marcadores ligados a polimorfismos de base única do tipo SNPs que possuem a capacidade de distinguir os tipos de crescimento da haste da soja. A primeira com um painel associativo contendo 384 materiais representativos distribuídos igualmente entre os três tipos de crescimento, e uma segunda abordagem se utilizou uma população segregante com 903 indivíduos oriunda do cruzamento entre Will (*Dt1Dt1;Dt2Dt2*) semideterminado, TMG 2517 (*Dt1Dt1;dt2dt2*) indeterminado, onde foram realizadas avaliações fenotípicas mensurando; altura de plantas (APT), número de vagens no racemo terminal (NVR), distância entre as duas últimas gemas reprodutivas da haste principal (DGH). Dos quatro marcadores propostos para *Dt2*, três conseguiram distinguir plantas Semideterminadas de Indeterminadas; TMG_161815, TMG161817 e TMG161818 com a respectiva acurácia 84,9%, 87,8% e 85,6%. Os dois marcadores sintetizados para *Dt1* foram eficientes para diferenciar plantas Determinadas dos demais tipos, TMG_161902 obteve acurácia de 86,5% e TMG_161910 com 86,1%. A utilização da análise conjunta dos três fatores fenotípicos mensurados na população segregante, possibilitou a validação dos marcadores moleculares desenvolvidos para *Dt2*.

Palavras-chave: Mapeamento associativo. Polimorfismo de base única. Melhoramento de soja.

CRUZ, Luan Gonçalves. **Development of molecular and phenotypic tools for evaluation of soybean growth types (G. Max)**. 2017. 45 p. Dissertação (Master's in Agronomy) – University of Londrina, Londrina, 2017.

ABSTRACT

The interactions of several factors determine the duration of the vegetative cycle in the soybean crop, among them can be highlighted some external factors such as; Temperature, duration of the period of the day, quality and intensity of light, and the internal as; Day insensitive plants (long juvenile period), and the types of soybean stem growth that can be distributed in three being: indeterminate, determinate and half-determined. The classification in three types is due to the fact of presenting distinct characteristics between them; Architecture, duration of flowering period, final height, distance between the last two reproductive nodes, terminal racemic production. The expression of these phenotypes can be influenced by edaphoclimatic conditions and the epistatic interaction between Dt1 and Dt2, which complicates the characterization process. The locus Dt1 and Dt2 have already been mapped, however clarification is required on the Dt2 locus region and its interaction with Dt1. We analyzed polymorphisms in the soybean genome (type SNPs) to validate molecular markers associated with the Dt1 and Dt2 loci aiming at the easy characterization and distinction of the types of terminal growth of the soybean stem applicable to breeding programs. Two approaches have been used to identify markers linked to single base polymorphisms of the SNPs type that have the ability to distinguish the growth types from the soybean stem. The first one with an associative panel containing 384 representative materials equally distributed among the three growth types, and a second approach was used a segregating population with 903 individuals from the intersection between semidetermined Will (Dt1Dt1; Dt2Dt2), TMG 2517 (Dt1Dt1; dt2dt2) Undetermined, where phenotypic evaluations were carried out; (APT), number of pods in the terminal racemus (NVR), distance between the last two reproductive buds of the main stem (DGH). Of the four markers proposed for Dt2, three were able to distinguish Semi-determined plants from Indeterminate; TMG_161815, TMG161817 and TMG161818 with the respective accuracy 84.9%, 87.8% and 85.6%. The two markers synthesized for Dt1 were efficient to differentiate plants. Determined from the other types, TMG_161902 obtained accuracy of 86.5% and TMG_161910 with 86.1%. The use of the joint analysis of the three phenotypic factors measured in the segregating population allowed the validation of the molecular markers developed for Dt2.

Keywords: Associative mapping. Single base polymorphism. Soybean breeding.

LISTA DE FIGURAS

Figura 3.1 - Detalhe da gema apical de uma planta “típica indeterminada”.....	31
Figura 3.2 - Detalhe da formação de um racemo apical de uma planta “típica Semideterminada”.....	32
Figura 3.3 - Destaque em vermelho para uma planta que genotipicamente é indeterminada que pode gerar confusão na hora da fenotipagem.	33
Figura 3.4 - Gráfico Manhattan Plot demonstrando qual SNP foi mais significativo para Dt1.....	34
Figura 3.5 - Gráfico Manhattan Plot demonstrando qual SNP foi mais significativo para Dt2.....	35
Figura 3.6 - Gráfico da leitura da fluorescência resultante da PCR, demonstrando a baixa capacidade do marcador TMG_161816 em separar plantas que possuem alelos recessivos ou dominantes de nulos ou erros, exibindo os 384 materiais genotipados no painel associativo	36
Figura 3.7 - Gráfico da leitura da fluorescência resultante da PCR, neste caso demonstra a elevada capacidade do marcador TMG_161902 em separar plantas que possuem alelos recessivos ou dominantes de nulos ou erros, 384 materiais do painel associativo.....	37

LISTA DE TABELAS

- Tabela 3.1** - Acurácia e porcentagem dos tipos de erros que ocorreram para os marcadores propostos em no painel associativo com 384 acessos para os três tipos de crescimento..... 35
- Tabela 3.2** - Variedades comerciais e Acessos com a descrição de registro e o genótipo identificado pelos 5 marcadores propostos para os tipos de crescimento..... 38
- Tabela 3.3** - Altura total de plantas, ciclo, comprimento entre os dois últimos internódios e número de vagens no racemo terminal em 903 indivíduos da população segregante para Dt2..... 39
- Tabela 3.4** - Acurácia na distinção dos tipos de crescimento, porcentagens dos tipos de erros ocorridos I e II e proporção de segregação na população segregante de 903 para Dt2. 41

SUMÁRIO

1	INTRODUÇÃO	10
2	REVISÃO BIBLIOGRÁFICA	12
2.1	ASPECTOS GERAIS DA CULTURA DA SOJA.....	12
2.2	TIPOS DE CRESCIMENTO DA HASTE DA SOJA.....	12
2.3	DESCRITORES MORFOLÓGICOS	15
2.4	MARCADORES MOLECULARES DE DNA.....	15
2.5	MARCADORES SNP	16
2.6	MAPEAMENTO ASSOCIATIVO.....	17
	REFERÊNCIAS BIBLIOGRÁFICAS	19
3	ARTIGO A: VALIDAÇÃO DE MARCADORES MOLECULARES PARA OS TIPOS DE CRESCIMENTO TERMINAL DA HASTE DA SOJA	23
3.1	RESUMO	23
3.2	ABSTRACT	24
3.3	INTRODUÇÃO	25
3.4	MATERIAL E MÉTODOS	27
3.4.1	ESTUDOS <i>IN-SILICO</i> PARA LOCALIZAÇÃO DE MARCADORES PARA OS GENES <i>Dt1</i> E <i>Dt2</i>	27
3.4.2	<i>IDENTIFICAÇÃO DE MARCADORES PARA O GENES</i>	28
3.4.3	PAINEL DE MATERIAIS CONTRASTANTES PARA MAPEAMENTO ASSOCIATIVO.....	28
3.4.4	EXTRAÇÃO DE DNA, PCR E LEITURA DE FLUORESCÊNCIA	28
3.4.5	CONDIÇÕES DO LOCAL.....	29
3.4.6	GERAÇÃO DA POPULAÇÃO SEGREGANTE PARA MAPEAMENTO	30
3.4.7	AVALIAÇÕES MORFOLÓGICAS	30
3.4.8	ANÁLISES ESTATÍSTICAS	33
3.5	RESULTADOS E DISCUSSÃO	34
3.5.1	IDENTIFICAÇÃO DE MARCADORES PARA OS GENES <i>Dt1</i> E <i>Dt2</i>	34
3.5.2	PCR DO PAINEL ASSOCIATIVO.....	35

3.5.3	VALIDAÇÃO DE MARCADORES CANDIDATOS PARA Dt_2	39
3.5.4	SEGREGAÇÃO.....	41
3.6	CONCLUSÕES.....	42
	REFERÊNCIAS BIBLIOGRÁFICAS	43

1 INTRODUÇÃO

A soja (*Glycine max* (L.) Merrill), é a principal leguminosa comercializada como *commodity* no Mundo sendo utilizada principalmente como fonte proteica na alimentação animal e humana, além de servir para produção de biocombustíveis.

O agronegócio é responsável pela maioria de todo o valor exportado pelo Brasil, e a soja é o principal produto deste total exportado possuindo, portanto, uma enorme representação econômica para o país.

Os desafios climáticos, ambientais e sociais para produção do grão são crescentes e para viabilizar a atividade os produtores devem buscar todas ferramentas disponíveis. Um dos principais fatores para se atingir sucesso na atividade é a utilização de variedades produtivas e adaptadas aos variados fatores e sistemas de produção.

A importância econômica da cultura vem estimulando o investimento de muitos programas de melhoramento, que tem contribuído para o desenvolvimento de um mercado de sementes de soja competitivo e atualizando de acordo com a necessidade dos produtores, anualmente são lançadas no mercado diversas cultivares cada vez mais produtivas, adaptadas às doenças, pragas e particularidades de cada sistema de cultivo e região.

Para se registrar uma nova cultivar de soja em países pertencentes à União Internacional para Proteção das Obtenções Vegetais (UPOV) como é o caso do Brasil, a cultivar deve ser distinta homogênea e estável, para essa caracterização são obrigatórios utilizar descritores morfológicos.

Um grande desafio para os programas de melhoramento é que muitas dessas características são determinadas pela interação de vários genes e estes descritores podem estar sujeitos a variações devido a efeitos do ambiente, sendo ainda que muitos deles exigem ser avaliados em fases avançadas da cultura, demandando tempo, mão de obra, e espaço físico.

Uma das características morfológicas diferenciadoras de cultivares é o tipo de crescimento terminal da haste, que é afetado por dois loci, Dt_1 e Dt_2 . Através da interação epistática destes genes são produzidos três tipos de crescimento distintos; indeterminado, determinado e semideterminado (BERNARD, 1972).

Os tipos de crescimento podem interferir em diversas características morfológicas sendo elas; período de desenvolvimento vegetativo e duração do florescimento, tamanho e arquitetura da haste principal, distância entre os últimos nós reprodutivos e desenvolvimento de racemo terminal.

Com a evolução de tecnologias em genotipagem e conhecimento de variações existentes no genoma da soja diversos marcadores moleculares têm sido desenvolvidos dentre os principais estão os marcadores de DNA do tipo SNP (single nucleotide polymorphisms) por serem muito práticos e acessíveis em termos de custos.

A grande vantagem é que o diagnóstico obtido por marcadores moleculares não sofre interferência do ambiente possibilitando repetibilidade e precisão nas avaliações realizadas. Esses marcadores são compostos por sequências de DNA dentro do gene, ligadas ou muito próximas aos loci responsáveis por determinar uma característica de interesse.

O locus *Dt1* da soja, é homólogo ao gene de *Arabidopsis*, e já foi mapeado por (LIU, et al., 2010). No caso do locus *Dt2* apresenta funcionalidade exclusiva para soja se comparado a outras espécies com genoma conhecido, indicando uma mutação com ganho de função (PING et al., 2014; LUI et al., 2016), segundo os autores os marcadores desenvolvidos até o momento apresentam baixa predição e isso se deve ao fato de existirem poucos materiais com o fenótipo semideterminado disponíveis para serem analisados.

Devido à importância agrônômica da característica em estudo, à necessidade de esclarecimentos em torno das interações entre os loci surgiu a necessidade de validar a variação alélica no genoma para os loci *Dt1* e *Dt2*, tendo o presente trabalho o objetivo de validar marcadores moleculares com eficiência prática para os tipos de crescimento da haste e produzir ferramentas para distinguir os tipos de crescimento que seja aplicável a programas de melhoramento de soja.

2 REVISÃO BIBLIOGRÁFICA

2.1 ASPECTOS GERAIS DA CULTURA DA SOJA

A espécie da soja (*Glycine max*, 2n=40) (L.) Merr. é essencialmente autógama com taxa de polinização cruzada inferior a 1%, pertence à ordem *Fabales*, família *Fabaceae*, subfamília *Papilionoideae*, tribo *Phaseoleae*, gênero *Glycine*. (BORÉM & MIRANDA, 2005).

Possui centro de origem na China e foi domesticada há cerca de 5 mil anos de sua espécie ancestral *G. soja* (HYTEN et al., 2006; WANG et al., 2016). No Brasil o cultivo teve importância comercial a partir da década de 1960, sendo cultivada principalmente na região Sul devido a adaptação natural a altas latitudes, para a expansão da cultura foi necessário o desenvolvimento de variedades com período juvenil longo que são adaptadas para regiões de baixa latitude, fato que possibilitou o Brasil se tornar um dos maiores produtores Mundiais de soja (ALMEIDA et al., 1999).

A soja é a principal commodity agrícola da balança comercial brasileira representando 87,3% do total de vendas externas de produtos agrícolas, sendo que o agronegócio é responsável por 52,5% de todo valor exportado pelo Brasil (MAPA, 2016) é a cultura que mais se expandiu nas últimas três décadas compreendendo 57% da área plantada em grãos no país. O grão é uma importante fonte de proteína destinada principalmente à produção de ração animal e óleo vegetal, apresentando uma crescente utilização na alimentação humana (BRASIL, 2015a).

2.2 TIPOS DE CRESCIMENTO DA HASTE DA SOJA

Os três tipos de crescimento da soja indeterminado, determinado e semideterminado são responsáveis por distinguir diversas características fenotípicas que são agronomicamente interessantes; incluindo porte, período de crescimento vegetativo e reprodutivo, distância entre os últimos nós reprodutivos e desenvolvimento de racemo terminal (BERNARD, 1972; VICENTE et al., 2016).

Análises genéticas clássicas demonstram que os tipos de crescimento da soja são regulados pela interação epistática entre dois loci: *Dt1* e *Dt2*. Em genótipos *Dt1*__; *Dt2*__ produzem fenótipos semideterminados e em *Dt1Dt1*; *dt2dt2* o fenótipo é indeterminado, pois *Dt2* é completamente dominante sobre *dt2* já na presença de (*dt1dt1*; *Dt2*__) o fenótipo se torna do

tipo de crescimento determinado indicando efeito epistático recessivo de *dt1* sobre a expressão do locus *Dt2* (BERNARD, 1972).

Os processos bioquímicos que desencadeiam o desenvolvimento dos tipos de crescimento são modulados principalmente pela conversão de meristemas apicais vegetativos para reprodutivos (WONG et al., 2008).

Estudos destes processos foram melhor explorados em *Arabidopsis*, onde se demonstrou que após a indução floral por estímulos externos como: comprimento do dia, qualidade de luz, temperatura e fatores internos como: idade e balanço hormonal, são expressos os genes como: *LEAFY (LFY)* e *APETALA1 (API)* responsáveis por desencadear processos que resultam na conversão das gemas apicais vegetativas em reprodutivas (WEIGEL, et al., 1992; GUSTAFSON-BROWN et al., 1994)

Por outro lado, pode ser que não ocorra a floração se estiverem presentes genes repressores florais tais como *Terminal Flower1 (TFL1)* (SHANNON & MEEKS-WAGNER, 1991; BRADLEY et al., 1997), o *TFL1* inibe a expressão de *LFY* e *API* mantendo o crescimento vegetativo em gemas apicais, fato que resulta no tipo selvagem de *Arabidopsis* onde as plantas crescem indefinidamente devido a presença de *TFL1* originando o tipo de crescimento indeterminado das plantas (LILJEGREN et al., 1999).

Já em mutantes para perda de função de *tfl1* os genes *LFY* e *API* são expressos permitindo o pleno florescimento resultando no tipo de crescimento determinado (WAGNER et al., 1999), além disso ocorrem casos onde *LFY* e *API* podem se ligar a *TFL1* inibindo sua expressão, fato que também favorece a conversão de meristemas apicais vegetativos para reprodutivos (WINTER et al., 2011).

Apesar de serem necessárias elucidacões acerca de *TFL1* ortólogos purativos de *Arabidopsis* são conservados em diversas espécies sustentando seu papel de repressor floral tal como *Dt1* em soja (*Glycine max*) (LIU et al., 2010; TIAN et al., 2010), *Det* em ervilha (*Pisum sativum*) (FOUCHER et al., 2003), *PvTFL1y* em feijão comum (*Phaseolus vulgaris*) (REPINSKI et al., 2012).

Na cultura da soja a alteração da forma do alelo *dt2* para *Dt2* condiciona a expressão de *Dt1*. O alelo *dt2* na forma recessiva possibilita a expressão de *Dt1* conferindo às plantas o tipo indeterminado, na presença da forma dominante *Dt2* o fenótipo se torna semideterminado e a expressão de *Dt1* é reprimida (PING et al., 2014; LIU et al., 2016).

Encontrado naturalmente em maior frequência em plantas selvagens o tipo de crescimento indeterminado possui plantas que apresentam prolongado período de florescimento, que é

simultâneo ao desenvolvimento vegetativo, geralmente com hastes principais longas e curvas (BERNARD, 1972), que apresentam inflorescência axilar mantendo o crescimento vegetativo mesmo após o início do florescimento (HEATHELERLY & SMITH, 2004). As plantas com tipo de crescimento indeterminados comparados às do tipo determinados em experimentos conduzidos por dois anos a campo exibiram níveis superiores significativos para atributos como: condutância estomática, densidade estomática, número de células epidérmicas por área e comprimento de nervura total por área. Já o comprimento das células guarda, massa foliar e tamanho da folha tende a ser maior nos tipos determinados (TANAKA & SHIRAIWA, 2009). A soja *G. max* foi selecionada do seu ancestral conhecido a espécie *G. soja* (HYMOWITZ, 2004), onde se encontra pouca frequência de plantas com o tipo de crescimento determinado indicando que esta característica foi obtida através de domesticação para ambientes específicos (TIAN et al., 2010). O tipo de crescimento determinado é controlado por um alelo recessivo *dt1* que praticamente finaliza o desenvolvimento vegetativo logo após iniciar o processo de florescimento, concentra a duração do período de floração e as plantas possuem haste principal ereta com inflorescência racemosa apical (BERNARD, 1972; TIAN et al., 2010). O locus *Dt1* foi identificado no cromossomo 19 (Glyma19g37890.1), através da introdução da região genômica GmTFL1b que foi confirmada por um silenciamento de genes induzida por vírus. O alelo *dt1* pode condicionar o hábito determinado através da alteração de GmTFL1b (LIU et al., 2010).

Já o tipo de crescimento semideterminado partilha as características fenotípicas de ambos os tipos, sendo um tipo de crescimento “intermediário”, pois nele as plantas possuem longo período de florescimento que é simultâneo ao desenvolvimento vegetativo, porém com haste principal ereta e inflorescência racemosa apical e axilar ao mesmo tempo (BERNARD, 1972; 1982; VICENTE et al., 2016), as plantas possuem altura inferior às do tipo indeterminado conferindo melhor resistência a acamamento (CHANG et al., 1982), em ambientes irrigados foi possível observar produtividades superiores em plantas com o tipo de crescimento semideterminado comparado as indeterminadas (SETIYONO et al., 2007).

O locus *Dt2* apresenta funcionalidade exclusiva para soja e está localizado no cromossomo 18 (Glyma18g50910) (PING et al., 2014; LIU et al., 2016) se comparado a culturas já estudadas para tal atributo como; tomate (*Solanum lycopersicum*) (ELKIND et al., 1991; PNUELI et al., 1998; FRIDAMAN et al., 2002), grão de bico (*Cicer arietinum*) (HEGDE, 2011) e ervilha (*Pisum sativum*) (BERBEL et al., 2012). Na cultura do tomate ao contrário da soja o tipo de crescimento semideterminado é controlado por um alelo recessivo *sdt* que exibe efeito epistático dominante *Sdt/sdt* (PNUELI et al., 1998), no caso do grão de bico um parente

próximo da soja também apresenta um padrão epistático distinto e muito similar ao encontrado no tomate (HEGDE, 2011), para a cultura da ervilha também se verifica um padrão inverso ao observado em soja (BERBEL et al., 2012).

2.3 DESCRITORES MORFOLÓGICOS

Descritores morfológicos são utilizados para caracterizar a soja e diversas culturas, sendo necessários para proteção e registro de novas variedades em países pertencentes à União Internacional para Proteção das Obtenções Vegetais (UPOV, 2015). Os descritores mínimos exigidos pelo Sistema Nacional de Proteção de Cultivares (SNPC) para um registro abrangem características fenotípicas das variedades, reação a doenças, nematoides e do tegumento à peroxidase, além do teor de óleo e proteína, totalizando 38 itens. Dentre os descritores morfológicos estão os tipos de crescimento, que na cultura da soja é dividido em três: indeterminado, determinado e semideterminado (BRASIL, 2015b).

A caracterização de descritores através de marcadores moleculares de DNA possibilita a distinção de cultivares muito similares devido ao elevado nível de resolução genética e confiabilidade do método (TOPPA & JADOSK., 2013).

A UPOV possui um grupo de trabalho composto por especialistas bioquímicos que estudam a aplicação de novas técnicas bioquímicas e moleculares, para distinção e registro de novas cultivares (UPOV, 2015), de maneira que marcadores moleculares se tornam reconhecidos e aceitos para caracterização de cultivares. Geralmente poucos *loci* proporcionam a diferenciação completa de genótipos avaliados. Apesar do avanço, no Brasil, o uso de marcadores como descritores ainda não é aceito pelo Serviço Nacional de Proteção de Cultivares (SNPC, 2015). Entretanto, a ferramenta já é amplamente utilizada em programas de melhoramento genético (BORÉM & CAIXETA, 2009).

2.4 MARCADORES MOLECULARES DE DNA

O genoma de *G. max* contém 56.044 genes codificadores de proteínas distribuídos em 20 cromossomos, com 85% de parálogos oriundos de eventos de duplicação a cerca de 13 a 59 milhões de anos atrás (SCHMUTZ, et al., 2010, LEHMANN et al., 2011).

A disponibilidade da sequência do genoma da soja possibilita o mapeamento de uma ampla gama de genes fornecendo a base para estudos e diversas descobertas (SCHUMUTZ et al.,

2010). Essas informações podem ser acessadas em bancos de dados abertos (<http://www.phytozome.net/soybean>, <http://soybase.org>, <http://soykb.org>).

As técnicas que utilizam marcadores de sequências de DNA possibilitam analisar características intrínsecas do genótipo dos materiais sem a necessidade de avaliações fenotípicas e a consequente interferência de fatores externos. Esses marcadores são ferramentas auxiliares muito importantes em diversas etapas do melhoramento genético, tais como a caracterização de germoplasma e até na seleção e desenvolvimento de materiais elite (FERREIRA & GRATTAPAGLIA, 1998). A utilização de marcadores moleculares para caracterização e seleção de tipos de crescimento da haste da soja pode ser uma excelente ferramenta para programas de melhoramento (VICENTE et al., 2016).

Cada cromossomo contém moléculas únicas de DNA onde estão localizadas as informações genéticas. As técnicas de análises moleculares permitem definir pontos específicos no cromossomo e estes são denominados “marcadores moleculares”. Um marcador molecular é definido como um segmento específico de DNA ligado aos loci responsáveis pela expressão de um fenótipo. Dentre os tipos de marcadores moleculares que utilizam segmentos de DNA estão o; *Random Fragment Polymorphic* (RAPD), o *Restriction Fragment Length Polymorphism* (RFLP), o *Amplified Fragment Length Polymorphism* (AFLP) e os *Simple Sequence Repeat* (SSR) ou microssatélites (FERREIRA & GRATTAPAGLIA, 1998), e os *Single Nucleotide Polymorphism* (SNP) (CHO et al., 1999).

2.5 MARCADORES SNP

Marcadores de polimorfismo de nucleotídeo único (*Single Nucleotide Polymorphism*, SNP) são definidos como regiões do DNA onde ocorreu a substituição de apenas uma base dentre amostras realizadas em uma população (RISCH & MERIKANGAS, 1996). Estes marcadores são amplamente utilizados devido ao fato de que são a forma de variação mais abundante e de ampla cobertura no genoma, além de apresentarem baixo custo, elevada reprodutibilidade, confiabilidade e possibilidade de automação (RESENDE et al., 2008).

As tecnologias desenvolvidas para detectar o polimorfismo em SNPs são de natureza bialélica detectando apenas dois alelos em cada loco, isso se deve ao fato de que esses marcadores exibem esse padrão em maior frequência de ocorrência. Os SNPs são uma forma de polimorfismo encontrada abundantemente no genoma das plantas, e no caso da soja a

frequência se dá em 1 SNP para cada 200 pares de base, variando de acordo com a diversidade genética das amostras comparadas (GRAEF et al., 2004).

Geralmente estudos com esses tipos de marcadores utilizam milhares SNPs em uma ampla amostragem de plantas com elevado polimorfismo entre elas, de forma a possibilitar a associação de marcadores com alta densidade a cerca de uma característica quantitativa baseando estudos de genética associativa GWAS (Genome-Wide Association Studies).

2.6 MAPEAMENTO ASSOCIATIVO

O mapeamento associativo é baseado no desequilíbrio de ligação (DL) de alelos correspondentes em uma população, esse modelo visa associar caracteres fenotípicos de interesse com o genótipo utilizando marcadores moleculares (ZIEMS et al., 2014). Podemos entender DL como alteração não randômica de alelos em diferentes lócus que ocorre dentro de uma população (FLINT-GARCIA et al., 2003). Os estudos envolvendo DL são utilizados principalmente em duas situações; para analisar uma característica contrastante que afeta uma população como indivíduos resistentes e suscetíveis a uma doença, e uma segunda opção seria analisar a correlação da transmissão do DL através da família oriunda de um parental heterozigoto (ORAGUZIE et al., 2007).

O delineamento experimental e a densidade dos marcadores podem ser determinados pelo tamanho do DL. A extensão e distribuição do DL deve ser mensurada para se determinar a densidade dos marcadores e a resolução do mapeamento. Quando o DL se estende por uma curta distância serão necessários vários marcadores, já em casos onde essa distância é longa podendo ser medida até em cM poucos marcadores são suficientes, porém a resolução de mapeamento é baixa (ZHU et al., 2008).

Diversos modelos estatísticos podem ser utilizados para se estimar o DL, dentre os índices estatísticos mais utilizados estão o índice D' que consiste em parametrizar o desequilíbrio de ligações, mais utilizado entre lócus com maiores frequências, e o r^2 (coeficiente de determinação), mais empregado para alelos raros ou com pouca frequência na população, que permite ser um indicativo de como os marcadores estão associados a região do genoma responsável pela característica de interesse (FLINT-GARCIA et al., 2003).

Com o auxílio de dados produzidos por trabalhos de ressequenciamento podem ser encontrados elevado número de variações alélicas (MALDONADO et al., 2016) que podem ser utilizadas para encontrar SNPs. Variações alélicas que estão dentro ou próximos ao gene

podem regular modificações drásticas de aminoácidos resultantes. (PATIL et al., 2015; MALDONADO et al., 2016).

REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, L.A. de; KIIHL, R.A. de S.; MIRANDA, M.A.C. de; CAMPELO, G.J. de A. Melhoramento da soja para regiões de baixas latitudes. In: QUEIROZ, M.A. de; GOEDERT, C.O.; RAMOS, S.R.R., ed. **Recursos genéticos e melhoramento de plantas para o Nordeste brasileiro**. Petrolina: Embrapa Semi-Arido, 1999. não paginado.
- BERBEL, A.; FERRADIZ, C.; HECHT, V.; DALMAIS, M.; LUND, O.S.; SUSSIMILCH, F.C. *VEGETATIVE1* is a essential for development of the compound inflorescence in pea. **Nat Commun.** 3:797, 2012.
- BERNARD, R.L. Two genes affecting stem termination in soybeans. **Crop Sci.** 12: 235–239. 1972.
- BORÉM, A.; MIRANDA, G.V. **Melhoramento de plantas**. Viçosa: UFV, 2005. 525 p.
- BORÉM, A.; CAIXETA, E. T. **Marcadores moleculares**. 2. ed. Viçosa, MG: UFV, 2009. 532 p.
- BRASIL (A) - Ministério da Agricultura, Pecuária e Abastecimento. **Soja**. Disponível em: <<http://www.agricultura.gov.br/vegetal/culturas/soja>> Acesso em: 15/08/2015
- BRASIL (B) - Ministério da Agricultura, Pecuária e Abastecimento. **Requisitos Mínimos para Determinação do Valor de Cultivo e Uso de Soja (*Glycine Max*) para Inscrição no Registro Nacional de Cultivares**. Disponível em: <<http://www.agricultura.gov.br/vegetal/registros-autorizacoes/registro/registro-nacional-cultivares/formularios-registro-cultivares-requisitos>> Acesso em: 15/08/2015
- BRADLEY, D.; RATCLIFFE, O.; VINCENT, C.; CARPENTER, R.; COEN, E. Inflorescence commitment and architecture in Arabidopsis. **Science** 275: 80–83, 1997.
- CHANG, J.F; Green, D.E; Shibles R. Yield and agronomic performance of semi-determinate and indeterminate soybeans estem types. **Crop Sci.** 1982;22(1):97-101.
- CHO, R. J.; MINDRINOS, M.; RICHARDS, D. R.; SAPOLSKY, R. J.; ANDERSON, M.; DRENKARD, E.; DEWDNEY, J.; REUBER, T. L.; STAMMERS, M.; FEDERSPIEL, N.; THEOLOGIS, A.; YANG, W. H.; HUBBELL, E.; AU, M.; CHUNG, E. Y.; LASHKARI, D.; LEMIEUX, B.; DEAN, C.; LIPSHUTZ, R. J.; AUSUBEL, F. M.; DAVIS, R. W.; OEFNER, P. J. Genome-wide mapping with biallelic markers in Arabidopsis thaliana. **Nature Genetics**, New York, v. 23, p. 203-207, 1999.
- ELKIND, Y.; GURNICK, A.; KEDAR, N. Genetics of semideterminate growth habit in tomato. **HortScience**, 1991 26: 1074–1075.
- FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores RAPD e RFLP em análise genética**. 3 ed. Brasília: EMBRAPA-CENARGEN, 1998. 220 p.
- FLINT-GARCIA, S. A.; S. A.; THORNSBERRY, J. M.; BUCKLER, E. S. Structure of linkage disequilibrium in plants. **Annual Review Of Plant Biology**, Palo Alto, v. 54, n.1, p.357-374, jun. 2003. Annual Reviews.

- FOUCHER, F.; MORIN, J.; COUTIADE, J.; CADIOUX, S.; ELLIS, N.; BANFIELD, M.J. *DETERMINATE* and *LATE FLOWERING* are two *TERMINAL FLOWER1/CENTRORADIALIS* homologs that control two distinct phases of flowering initiation and development in pea. **Plant Cell**. 2003;15(11):2742-54
- FRIDAMAN, E.; LIU, Y.S.; CARMEL-GOREN, L.; GUR, A.; SHORESH, M.; PLEBAN, T.; ESHED, Y.; ZAMIR, D. Two tightly linked QTLs modify tomato sugar content via different physiological pathways. **Mol. Genet. Genomics** 266: 821–826, 2002.
- GRAEF, D.W.G.L.; DIERS, A.M.P.B.W. Identification of putative QTL that underlie yield in interspecific soybean backcross populations. **Theoretical and Applied Genetics (TAG)**, 458-467, 2004.
- HEATHELERLY, L.G; SMITH, J.R. Effect of soybean stem growth habit on height and node number after beginning bloom in the midsouthem USA. **Crop Sci**. 2004;44(5):1855-8.
- HEGDE, V.S. Morphology and genetics of a new found determinate genotype in chickpea. **Euphytica**. 2011. 182: 35–42.
- HYMOWITZ, T. Speciation and cytogenetics. In: BOERMA, H.R.; SPECHT, J.E. (Ed.). **Soybeans: improvement, production and uses**. 4th ed. Madison: American Society of Agronomy, Crop Science Society of America and Soil Science Society of America, 2004. chap. 4, p. 97-136.
- HYTEN, D.L.; SONG, Q.; ZHU, Y.; CHOI, I.Y.; NELSON, R.L.; COSTA, J.M.; SPECHT, J.E.; SHOEMAKER, R.C.; CREGAN, P.B. **Impacts of genetic bottlenecks on soybean genome diversity**. Proceedings of the National Academy of Sciences of the USA, Washington, v. 103, p. 16666–16671, 2006.
- LEHMANN, S.; GUMY, C., BLATTER, E.; BOEFFEL, S.; FRICKE, W., RENTSCH, D. In planta function of compatible solute transporters of the AtProT family. **J. Exp. Bot**. 2011. 62, 787–796.
- LILJEGREN, S.J.; GUSTAFSON-BROW, C.; PINYOPICH, A.; DITTA, G.S.; YANOFSKY, M.F. Interactions among *APETALA1*, *LEAFY*, and *TERMINAL FLOWER1* specify meristem fate. **Plant Cell**. 1999;11(6):1007-18.
- LIU, B.; WATANABE, S.; UCHIYAMA, T.; KONG, F.; KANAZAWA, A.; XIA, Z.; NAGAMATSU, A.; ARAI, M.; YAMADA, T.; KITAMURA, K.; MASUTA, C.; HARADA, K.; ABE, J. The soybean stem growth habit gene *Dt1* is an ortholog of *Arabidopsis TERMINAL FLOWER1*. **Plant Physiol**. 153:198–210, 2010.
- LIU Y.; ZHANG D.; PING J.; LI S.; CHEN Z.; MA J. Innovation of a Regulatory Mechanism Modulating Semideterminate Stem Growth through Artificial Selection in Soybean. **PLoS Genet**. 2016, 12(1)
- MALDONADO S.V.J.; VALLIYODAN B.; JOSHI T.; KHAN S.M.; LIU Y.; WANG J.; ABDELNOOR R.V. Avaliação da variação genética entre cultivares brasileiras de soja por resequenciamento do genoma. **BMC Genomics**. 2016, 17:110.
- PATIL G.; VALLIYODAN B.; DESHMUKHM R.; PRINCE S.; NICANDER B.; ZHAO M.; SONAH H.; SONG L.; LIN L.; CHAUDHARY J.; LIU Y.; JOSHI T.; XU D.; HENRY T.

Soybean (*Glycine max*) SWEET gene family: insights through comparative genomics, transcriptome profiling and whole genome re-sequence analysis. **BMC Genomics**. 2015,16:520

PNUELI L.; CARMEL-GOREN L.; HAREVEN D.; GUTFINGER T.; ALVAREZ J.; GANAL M.; ZAMIR D.; LIFSCHITZ E. (1998). The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. **Development** 125: 1979–1989, 1998.

PING, J.; LIU, Y.; SUN, L.; ZHAO, M.; LI, Y.; SHE, M.; SUI, Y.; LIN, F.; LIU, X.; TANG, Z.; NGUYEN, H.; TIAN, Z.; QIU, L.; NELSON, L. R.; CLEMENTE E. T.; SPECHT, E. J.; MA, J. Dt2 is a gain-of-function MADS-domain factor gene that specifies semideterminacy in soybean. **Plant Cell** 26: 2831–2842, 2014.

RESENDE, M.D.V.; LOPES, P.S.; SILVA, R.L. da; PIRES, I.E. Seleção genômica ampla (GWS) e maximização da eficiência do melhoramento genético. **Pesquisa Florestal Brasileira**, Colombo, n. 56, p. 63, 2008.

REPINSKI, S.L.; KWAK, M.; GEPTS, P. The common bean growth habit gene *PvTFL1y* is a functional homolog of *Arabidopsis TFL1*. **Theor Appl Genet**.2012;124(8):1539-47.

RISCH, N.; MERIKANGAS, K. The future of genetic studies of complex human diseases. **Science** 273:1516-1517, 1996.

SETIYONO, T.D.; WEISS, A.; SPECHT, J.; BASTIDAS, A.M.; CASSMAN, K.G.; DOBERMANN, A. Understanding and modeling the effect of temperature and daylength on soybean phenology under high-yield conditions. **Field Crops Res**. 100: 257–271, 2007.

SCHUMUTZ, J.; CANNON, S.B.; SCHULUETER, J.; MA, J.; MITROS, T.; NELSON, W.; HYTEN, D.L.; SONG, Q.; THELEN, J.J.; CHENG, J.; XU, D.; HELLSTEN, U.; MAY, G.D.; YU, Y.; SAKURA, T.; UMEZAWA, T.; BHATTACHARYA, M.K.; SANDHU, D.; VALLIYODAN, B.; LINDQUIST, E.; PETO, M.; GRANT, D.; SHU, S.; GOODSTEIN, D.; BARRY, K.; FUTRELL, G.M.; ABERNATHY, B.; DU, J.; THIAN, Z.; ZHU, L.; GILL, N.; JOSHI, T.; LIBAULT, M.; SETHURAMAN, A.; ZANG, X.C.; SHINOZAKI, K.; NQUYEN, H.T.; WING, R.A.; CREGAN, P.; SPECHT, J.; GRIMWOOD, J.; ROKHSAR, D.; STACEY, G.; SHOEMAKER, R.C.; JACKSON, S.A. Genome sequence of the palaeopolyploid soybean. **Nature** 463:178–183, 2010.

SHANNON S.; MEEKS-WAGNER D.R.; A mutation in the *Arabidopsis TFL1* gene affects inflorescence meristem development. **Plant Cell**. 1991;3(9):877-92.

GUSTAFSON-BROWN C.; SAVIDG B.; YANOFSKY M.F. Regulation of the *Arabidopsis* floral homeotic gene *APETALA1*. **Cell**. 1994;76(1):131-43.

ORAGUZIE, N.C; RIKKERINK, E.H.A; GARDINER, S.E; SILVA, H.N. Association mapping in plants. New York, **Springer**, 2007. 277 p.

TANAKA, Y. & SHIRAIWA, T., Stem growth habit affects leaf morphology and gas exchange traits in soybean. **Annals of Botany** 104: 1293–1299, 2009.

TIAN, Z.; WANG, X.; LEE, R.; LI, Y.; SPECHT, J.E.; NELSON, R.L. Artificial selection for determinate growth habit in soybean. **Proc Natl Acad Sci USA**. 2010;107(19):8563-8.

TOPPA, B.V.E.; JADOSKI, J.C. O uso dos marcadores moleculares no melhoramento genético de plantas. **Scientia Agraria Paranaensis**. Volume 12, n. 1, p.1-5, 2013.

UPOV - International Union for the Protection of New Varieties of Plants Disponível em: <<http://www.upov.int/about/en/organigram.html#other>>. **UPOV**. Acesso em 10/08/2015.

WAGNER, D.;SABLOWSKI, R.W.M.; MEYEROWITZ, E.M. Transcriptional activation of *APETALA1* by *LEAFY*. **Science**. 1999;285(5427):582-4.

WEIGEL, D.; ALVAREZ, J.; SMYTH, D.R.; YANOFSKY, M.F.; MEYEROWITZ, E.M. *LEAFY* controls floral meristem identity in Arabidopsis. **Cell**, 1992;69(5):843-59.

WINTER, C.M.; AUSTIN, R.S.; BLAVILLAIN-BAUFUME, S.; REBACK, M.A.; MONNIAUX, M.; WU, M.F. *LEAFY* target genes reveal floral regulation logic, cis motifs, and a link to biotic stimulus response. **Dev. Cell**. 2011;20(4):430-43.

WONG, C.E.; SINGH, M.B; BHALLA, P.L. Molecular processes underlying the floral transition in the soybean shoot apical meristem. **Plant Journal**.2009;57(5):832-45.

VICENTE D.; SCHUSTER I.; LAZZARI F.; PARANZINI D.P.J.; OLIVEIRA R.A.M.; PRETE C.E.C.; Mapping and validation of molecular markers of genes Dt1 and Dt2 to determine the type of stem growth in soybean. **Acta Scientiarum**. 2016; 1679-9275.

ZIEMS, L.A.; HICKEY, L. T.; HUNT, C. H.; MACE, E. S.; PLATZ, G. J.; FRANCKOWIAK, J. D.; JORDAN, D. R. Association mapping of resistance to *Puccinia hordei* in Australian barley breeding germoplasm. **Theoretical and Applied Genetics**, Heidelberg, v. 127, p. 1199-1212, 2014.

ZHU, C. et al. Status and prospects of association mapping in plants. **The Plant Genome**, Madison, v, 1, p. 5-20, 2008.

3 ARTIGO A: VALIDAÇÃO DE MARCADORES MOLECULARES PARA OS TIPOS DE CRESCIMENTO TERMINAL DA HASTE DA SOJA

3.1 RESUMO

A planta de soja pode ser classificada em três tipos de crescimento terminal da haste: indeterminado, determinado e semideterminado. A expressão desses fenótipos pode ser influenciada pelo ambiente de cultivo e a interação epistática de dois loci (*Dt1* e *Dt2*), fato que pode dificultar o processo de caracterização. Foram analisados polimorfismos no genoma da soja (tipo SNPs) tendo como objetivo validar marcadores moleculares associados ao loci *Dt1* e *Dt2* visando a fácil caracterização e distinção dos tipos de crescimento terminal da haste da soja aplicável a programas de melhoramento. Se utilizou duas abordagens para identificar marcadores ligados a polimorfismos de base única do tipo SNPs que possuem a capacidade de distinguir os tipos de crescimento da haste da soja. A primeira com um painel associativo contendo 384 materiais representativos distribuídos igualmente entre os três tipos de crescimento, e uma segunda abordagem se utilizou uma população segregante com 903 indivíduos oriunda do cruzamento entre Will (*Dt1Dt1;Dt2Dt2*) semideterminado, TMG 2517 (*Dt1Dt1;dt2dt2*) indeterminado, onde foram realizadas avaliações fenotípicas mensurando; altura de plantas (APT), número de vagens no racemo terminal (NVR), distância entre as duas últimas gemas reprodutivas da haste principal (DGH).

Dos quatro marcadores propostos para *Dt2*, três conseguiram distinguir plantas Semideterminadas de Indeterminadas; TMG_161815, TMG161817 e TMG161818 com a respectiva acurácia 84,9%, 87,8% e 85,6%. Os dois marcadores sintetizados para *Dt1* foram eficientes para diferenciar plantas Determinadas dos demais tipos, TMG_161902 obteve acurácia de 86,5% e TMG_161910 com 86,1%.

A utilização da análise conjunta dos três fatores fenotípicos mensurados na população segregante, possibilitou a validação dos marcadores moleculares desenvolvidos para *Dt2*.

Palavras-chave: Interação entre *Dt1* e *Dt2*, painel associativo, população segregante.

3.2 ABSTRACT

The soybean plant can be classified into three types of terminal stem growth: indeterminate, determinate and semi-determined. (Dt1 and Dt2), which may hamper the characterization process. Non-genomic polymorphisms of soybean (type SNPs) were analyzed aiming to validate molecular markers associated with the Dt1 and Dt2 loci aiming at the easy characterization and distinction of the types of terminal growth of the soybean stem applicable to breeding programs. Two approaches have been used to identify markers linked to single base polymorphisms of the SNPs type that have a capability to distinguish the growth types from the soybean stem. (Dt1Dt1; Dt2Dt2) semi-determined, undetermined, unidentified, unidentified, undetermined TMG 2517 (Dt1Dt1; dt2dt2), where phenotypic results were measured; Plant height (APT), number of pods without terminal racemes (NVR), distance between the last two reproductive buds of the main stem (DGH).

Of the four markers proposed for Dt2, three were able to distinguish Semi-determined plants from Indeterminate; TMG_161815, TMG161817 and TMG161818 with a respective 84.9%, 87.8% and 85.6% accuracy. The two markers synthesized for Dt1 were efficient for differentiating plants. For other types, TMG_161902 obtained 86.5% accuracy and 86.1% TMG_161910.

The use of the joint analysis of the three phenotypic factors measured in the segregating population, allowed a validation of the molecular markers developed for Dt2.

Keywords: Interaction between *Dt1* and *Dt2*, associative panel, segregating population.

3.3 INTRODUÇÃO

Os três tipos de crescimento da soja indeterminado, determinado e semideterminado são responsáveis por distinguir diversas características fenotípicas que são agronomicamente interessantes; incluindo porte, período de crescimento vegetativo e reprodutivo, distância entre os últimos nós reprodutivos e desenvolvimento de racemo terminal (BERNARD, 1972; VICENTE, 2016).

Análises genéticas clássicas demonstram que os tipos de crescimento da soja são regulados pela interação epistática entre dois loci: *Dt1* e *Dt2*. Em genótipos *Dt1--*; *Dt2--* produzem fenótipos semideterminados e em *Dt1;Dt1*; *dt2;dt2* o fenótipo é indeterminado, pois *Dt2* é completamente dominante sobre *dt2* já na presença de (*dt1dt1*; *Dt2--*) o fenótipo se torna do tipo de crescimento determinado indicando efeito epistático de *dt1* sobre a expressão do locus *Dt2* (BERNARD, 1972).

As relações dos processos bioquímicos que desencadeiam o desenvolvimento dos tipos de crescimento são modulados principalmente pela conversão de meristemas apicais vegetativos para reprodutivos (WONG et al., 2008).

Os estudos destes processos foram melhor explorados em *Arabidopsis*, onde se demonstrou que após a indução floral por estímulos externos como: comprimento do dia, qualidade de luz, temperatura e fatores internos como: idade e balanço hormonal, são expressos os genes *LEAFY (LFY)* e *APETALA1 (API)* responsáveis por desencadear processos que resultam na conversão das gemas apicais vegetativas em reprodutivas (WEIGEL, et al., 1992; GUSTAFSON-BROWN et al., 1994).

A floração pode não ocorrer se estiverem presentes os genes repressores florais tais como *Terminal Flower1 (TFL1)* (SHANNON & MEEKS-WAGNER, 1991; BRADLEY et al., 1997). O *TFL1* inibe a expressão de *LFY* e *API* mantendo o crescimento vegetativo em gemas apicais, fato que resulta no tipo selvagem de *Arabidopsis* onde as plantas crescem indefinidamente devido a presença de *TFL1*, originando o tipo de crescimento indeterminado das plantas (LILJEGREN et al., 1999).

Já em mutantes para perda de função de *Tfl1* os genes *LFY* e *API* são expressos permitindo o pleno florescimento resultando no tipo de crescimento determinado (WAGNER et al., 1999), além disso ocorrem casos onde *LFY* e *API* podem se ligar a *TFL1* inibindo sua expressão, fato que também favorece a conversão de meristemas apicais vegetativos para reprodutivos (WINTER et al., 2011).

Apesar de serem necessárias elucidações acerca de *TFL1* ortólogos purativos de *Arabidopsis* são conservados em diversas espécies sustentando seu papel de repressor floral tal como *Dt1* em soja (*Glycine max*) (LIU et al., 2010; TIAN et al., 2010), *Det* em ervilha (*Pisum sativum*) (FOUCHER et al., 2003), *PvTFL1y* em feijão comum (*Phaseolus vulgaris*) (REPINSKI et al., 2012).

Na cultura da soja a alteração da forma do alelo *dt2* ou *Dt2* condiciona a expressão de *Dt1*. O alelo *dt2* na forma recessiva possibilita a expressão de *Dt1* conferindo às plantas o tipo indeterminado, já na presença da forma dominante *Dt2* o fenótipo se torna semideterminado e a expressão de *Dt1* é reprimida (PING et al., 2014; LIU et al., 2016).

Encontrado naturalmente em maior frequência em plantas selvagens o tipo de crescimento indeterminado possui plantas que apresentam prolongado período de florescimento, que é simultâneo ao desenvolvimento vegetativo, geralmente com hastes principais longas e curvas (BERNARD, 1972), que apresentam inflorescência axilar mantendo o crescimento vegetativo mesmo após o início do florescimento (HEATHELERLY & SMITH, 2004).

A soja *G. max* foi selecionada do seu ancestral conhecido a espécie *G. soja* (HYMOWITZ, 2004), onde se encontra pouca frequência de plantas com o tipo de crescimento determinado indicando que esta característica foi obtida através de domesticação para ambientes específicos (TIAN et al., 2010). O tipo de crescimento determinado é controlado por um alelo recessivo *dt1* que praticamente finaliza o desenvolvimento vegetativo logo após iniciar o processo de florescimento, concentra a duração do período de floração e as plantas possuem haste principal ereta com inflorescência racemosa apical (BERNARD, 1972; TIAN et al., 2010). O locus *Dt1* foi identificado no cromossomo 19 (Glyma19g37890.1), através da introdução da região genômica GmTFL1b que foi confirmada por um silenciamento de genes induzida por vírus. O alelo *dt1* pode condicionar o hábito determinado através da alteração de GmTFL1b (LIU et al., 2010).

Já o tipo de crescimento semideterminado partilha as características fenotípicas de ambos os tipos, sendo um tipo de crescimento “intermediário”, pois nele as plantas possuem longo período de florescimento que é simultâneo ao desenvolvimento vegetativo, porém com haste principal ereta e inflorescência racemosa apical e axilar ao mesmo tempo (BERNARD, 1972; VICENTE et al., 2016), as plantas possuem altura inferior às do tipo indeterminado conferindo melhor resistência a acamamento (CHANG et al., 1982), em ambientes irrigados foi possível observar produtividades superiores em plantas com o tipo de crescimento semideterminado comparado as indeterminadas (SETIYONO et al., 2007).

O locus *Dt2* apresenta funcionalidade exclusiva para soja e está localizado no cromossomo 18 (Glyma18g50910) (PING et al., 2014; LIU et al., 2016) se comparado a culturas já estudadas para tal atributo como; tomate (*Solanum lycopersicum*) (ELKIND et al., 1991; PNUELI et al., 1998; FRIDAMAN et al., 2002), grão de bico (*Cicer arietinum*) (HEGDE, 2011) e ervilha (*Pisum sativum*) (BERBEL et al., 2012). Na cultura do tomate ao contrário da soja o tipo de crescimento semideterminado é controlado por um alelo recessivo *sdt* que exibe efeito epistático dominante *Sdt/sdt* (PNUELI et al., 1998), no caso do grão de bico e ervilha apresentam padrão epistático similar ao encontrado no tomate (HEGDE, 2011), para a cultura da ervilha também se verifica um padrão inverso ao observado em soja (BERBEL et al., 2012).

De acordo com a importância agrônômica da característica em estudo, à necessidade de esclarecimentos em torno das interações entre os loci *Dt1* e *Dt2*, surgiu a necessidade de validar a variação alélica no genoma. Objetivou-se validar marcadores moleculares com eficiência prática para os tipos de crescimento da haste e produzir ferramentas para distinguir os tipos de crescimento da haste da soja que seja aplicável a programas de melhoramento.

3.4 MATERIAL E MÉTODOS

3.4.1 ESTUDOS *IN-SILICO* PARA LOCALIZAÇÃO DE MARCADORES PARA OS GENES *DT1* E *DT2*

Para identificação dos SNPs mais representativos para os genes em estudo, foram realizadas análises *in-silico* utilizando como base informações da literatura (LIU et al., 2010; PING et al., 2014) que podem ser melhor explorados com auxílio do banco de dados da SoyKB disponível em (<http://soykb.org/>).

Os resultados fenotípicos obtidos no site do grin (<http://www.ars-grin.gov/>) foram correlacionados com os dados genotípicos oriundos do soybase e foi realizado um estudo de análise de genética associativa (GWAS) para determinação dos principais SNPs relacionados com a característica estudada.

Como abordagem adicional foram prospectados SNPs das regiões dos genes *Dt1* e *Dt2* correspondendo respectivamente ao Glyma19g37890.1 (LIU et al., 2010), e Glyma18g50910 (PING et al., 2014). Essas localizações foram inseridas no site Soykb.org, onde é

disponibilizado um banco de dados com todos os SNPs relacionados com essas localizações, posteriormente foi aplicado um filtro para os SNPs que causam mutação não sinônima dentro do gene.

Após análises interpretativas dos resultados gerados pelo SoyKB se definiu as regiões mais representativas para distinguir os dois lócus em estudo. Foram selecionados dois SNPs para *Dt1* e quatro para *Dt2*. A sequência de aproximadamente 150 pb ao redor de cada SNP foi utilizada para síntese de ensaios tipo KASP (LGC).

3.4.2 IDENTIFICAÇÃO DE MARCADORES PARA O GENES

Com o auxílio de dados produzidos por trabalhos de ressequenciamento e disponibilizados em rede pelo site da soykb.org podem ser encontrados um elevado número de variações alélicas (MALDONADO et al., 2016) que foram utilizadas para encontrar os SNPs. Variações alélicas que poderiam estar dentro ou próximos ao gene que possam regular modificações drásticas de aminoácidos resultantes foram validados por dados de ressequenciamento de dois trabalhos, o primeiro que utilizou 28 cultivares brasileiras (MALDONADO et al., 2016) e o segundo com 106 acessos disponíveis por (PATIL et al., 2015).

3.4.3 PAINEL DE MATERIAIS CONTRASTANTES PARA MAPEAMENTO ASSOCIATIVO

O painel de mapeamento associativo foi montado com variedades registradas pelo Ministério de Agricultura e Pecuária (MAPA) e acessos descritos pelo Germoplasm Resources Information Network (GRIN) totalizando 384 materiais descritos como sendo contrastantes para os três tipos de crescimento, foram explorados ainda materiais já estudados e descritos como derivados de um parental recorrente como linhagens isogênicas utilizadas em outros estudos como o de TANAKA & SHIRAIWA (2009), que utilizou quatro pares de cultivares com suas respectivas linhas isogênicas que diferem apenas para tipos de crescimento.

3.4.4 EXTRAÇÃO DE DNA, PCR e LEITURA DE FLUORESCÊNCIA

O tecido foliar utilizado para extração de DNA em ambos os ensaios tanto do painel associativo quanto o da população segregante, foi coletado retirando-se discos de 5mm de diâmetro das folhas jovens, no estágio fenológico V2, onde os níveis de expressão dos genes

candidatos são superiores se comparados a V0, V1 e V3 (PING et al., 2014). Uma amostra de aproximadamente 100 mg (4 discos foliares) foi coletada e armazenada em ultra freezer a -80 °C até o momento da extração do DNA.

O DNA foi extraído de acordo com protocolo que é fundamentado na utilização de detergente dodecil sulfato de sódio (SDS), que foi adaptado do descrito por KEIM et al. (1988).

Em ambos os marcadores utilizados, a reação de PCR (Polymerase Chain Reaction) foi realizada via tecnologia KASP™ genotyping, que apresenta como vantagens o baixo custo e praticidade, conduzida em placas de 384 células em volume final de 5,02µL, com 2,44µL de DNA e 2,58µL de KASP máster mix da LGC, que contém todas as substâncias necessárias para a síntese de novas cópias da sequência desejada. O equipamento utilizado foi do tipo hidrociclador HYDROCYCLER-16 da empresa LGC, segundo a configuração “Touchdown”, que realiza vários ciclos em diversas temperaturas decrescentes divididas em três etapas; sendo a primeira a 94°C por 15 minutos que é responsável pela desnaturação (quebra das pontes de hidrogênio da fita dupla de DNA), na segunda etapa são realizados 10 ciclos de 94°C por 20 segundos e 55-61°C por 1 minuto, onde ocorrem o anelamento e se finaliza com a terceira etapa com 26 ciclos de 94°C por 20 segundos e 61°C por 1 minuto necessários para amplificação da sequência de interesse específica para o PRIMER utilizado, obtendo assim múltiplas cópias da sequência alvo.

Para a leitura da fluorescência das placas, foi utilizado o aparelho PHERAstarPlus SNP, e para a interpretação e visualização dos dados, os softwares Kraken e PHERAstar.

Uma análise de correlação entre os dados fenotípicos e genotípicos foi realizada para a validação dos marcadores.

3.4.5 CONDIÇÕES DO LOCAL

O experimento foi desenvolvido em casa de vegetação, na estação experimental da empresa Tropical Melhoramento e Genética situada no município de Cambé – PR. As condições de temperatura e umidade são controladas por equipamentos automatizados mantendo-se; (28°C ±2°), e (60 UR% ±10%).

O substrato utilizado é uma mistura de solo do tipo latossolo vermelho eutrófico, esterco, areia, na respectiva proporção de 3:1:1 em vasos de 8 litros.

3.4.6 GERAÇÃO DA POPULAÇÃO SEGREGANTE PARA MAPEAMENTO

Os parentais utilizados para cruzamentos foi Will ($Dt1Dt1; Dt2Dt2$) que apresenta fenótipo do tipo de crescimento semideterminado (SMD) e o parental masculino foi TMG 2517 ($Dt1Dt1; dt2dt2$) com fenótipo do tipo de crescimento indeterminado (IND). As sementes dos materiais Will foram obtidas do Departamento de Agricultura dos Estados Unidos (USDA), enquanto que as sementes da linhagem TMG 2517 foram obtidas da empresa Tropical Melhoramento & Genética (TMG), Brasil.

A cultivar Will é oriunda dos cruzamentos de (T117 ($Dt1Dt1; Dt2Dt2$) retrocruzada com as cultivares mais produtivas da época; Clark ($Dt1Dt1; dt2dt2$) 6 vezes) que foi retrocruzado mais uma vez com Williams ($Dt1Dt1; dt2dt2$) por mais 6 vezes) originado a Will um material muito similar a Williams, porém mais precoce e semideterminado que foi registrada por BERNARD & CREMEENS (1988). Já a linhagem TMG 2517 é uma indeterminada elite típica, oriunda do programa de melhoramento genético da TMG.

Os cruzamentos foram realizados no período de fevereiro a abril de 2015. Foram obtidas 70 progênies em F_1 , semeadas em vasos individuais na data de 20 de julho de 2015. A geração F_2 foi semeada em janeiro de 2016 em casa de vegetação e a semeadura de F_3 foi em setembro de 2016. Os estudos do padrão de segregação da herança genética foram realizados em $F_2:3$ com no mínimo 15 sementes de cada planta F_2 totalizando em uma população com 903 indivíduos analisados.

Os parentais e as plantas F_1 , F_2 e F_3 receberam 14,5 horas de luz diárias, por um período de 20 dias, com o objetivo de atrasar o início do florescimento, possibilitando o desenvolvimento vegetativo das plantas, já que no período de cultivo (dias curtos) as plantas recebem estímulo para florescimento desde o início de seu desenvolvimento.

3.4.7 AVALIAÇÕES MORFOLÓGICAS

Características fenotípicas foram mensuradas em F_3 utilizando 15 plantas originárias de cada planta de geração F_2 para inferir o tipo de segregação. Foram avaliadas somente as características morfológicas que podem apresentar diferenças significativas segundo a literatura (BERNARD, 1972) que são; altura de plantas medidas do solo até a última vagem medidas em centímetros (APT), contagem do número de vagens no racemo terminal (NVR),

distância entre as duas últimas gemas reprodutivas da haste principal medidas em milímetros (DGH).

De acordo com as dificuldades para se diferenciar as plantas para os tipos de crescimento (BERNARD, 1972; VICENTE et al., 2016) foram realizadas avaliações visuais fenotípicas complementares para se distinguir os tipos de crescimento em dois estádios fenotípicos distintos, uma no estágio R6 (pleno enchimento de vagens) por notas em uma escala de 0 a 100, onde a nota 0 indica o tipo indeterminado observado na figura 3.1 e nota 100 para o tipo semideterminado (figura 3.2).



Figura 3.1 - Detalhe da gema apical de uma planta "típica indeterminada".



Figura 3.2 - Detalhe da formação de um racemo apical de uma planta “típica Semideterminada”.

A escala utilizada foi adaptada do original de Ping et al. (2014) para se ampliar a sensibilidade, além de uma segunda avaliação em R8 (maturação completa) que definiu os tipos de crescimento por uma “determinação visual” no momento da colheita.

Em todas as avaliações se considerou a interação de algumas características que para o tipo indeterminado “típico” podem ser observadas uma arquitetura da haste principal com segmentos em inclinações sequencias com ângulo curvos em torno de 45° resultando em um tipo de “Zig-Zag”, ocorre também o prolongamento da distância entre os dois últimos nós reprodutivos que na maioria dos casos é superior a 2 cm como são observados na figura 3.1.

Ocorreram também casos onde essa distância atingiu 0,5 cm, fato que pode dificultar a diferenciação visual e a formação de racemo terminal que geralmente ocorre em plantas com habito de crescimento semideterminado devido à redução brusca do desenvolvimento vegetativo que resulta no agrupamento das gemas reprodutivas terminais, mas em alguns casos também pode ser encontrado plantas com genótipo indeterminado, mas com haste terminal totalmente em “Zig-Zag” e com 3 vagens agrupadas na gema terminal, dificultando o processo de distinção visual destacado na figura 3.3.



Figura 3.3 - Destaque em vermelho para uma planta que genotipicamente é indeterminada que pode gerar confusão na hora da fenotipagem.

3.4.8 ANÁLISES ESTATÍSTICAS

As análises para as populações de mapeamento foram expressas por média e desvio padrão, comparadas por teste estatístico de qui-quadrado. O delineamento experimental utilizado foi inteiramente casualizado, utilizando o teste TUKEY a 5% para avaliação de outros atributos como: altura de plantas (APT), número de vagens no racemo terminal (NVR), distância entre as duas últimas gemas reprodutivas da haste principal (DGH), os dados foram processados por meio do programa SISVAR.

A utilização de modelos estatísticos é a forma de se conseguir agrupar todas as informações disponíveis de genótipos, esses métodos são capazes de analisar concomitantemente características quantitativas e qualitativas. Esse tipo de análise é possível com o modelo de Ward- MLM (modified Location Model) (FRANCO et al., 2001).

Foi utilizado o método de análise de correlação estatística Ward-MLM (Modified Location Model) RR_Blup que permite o agrupamento concomitante de vários tipos de características dentre elas; dados de avaliações morfológicas, informações genóticas, resultados de análise de principais componentes (PCA) e dados do sistema. Esse método foi executado no programa GAPIT que agrupa todos esses fatores com base em tabelas genética, a análise conjunta de todos esses fatores gera uma forte matriz de correção que foi programada para corte de valores com mais de 5% de erro reduzindo a maior parte dos falsos positivos que possam estar presentes (LIPKA et al., 2012). Os SNPs com frequência de alelo mínimo

(minor allele frequency – MAF) menor que 0,1 foram removidos da análise e o algoritmo VanRaden foi utilizado para produção de uma matriz de parentesco a fim de diminuir a presença de falsos positivos. SNPs com valores de p-value acima de $7,19e-4$ foram considerados significativos. De acordo com os resultados obtidos o programa gerou um gráfico exibindo qual marca genética seria mais indicada para explicar as características analisadas.

3.5 RESULTADOS E DISCUSSÃO

3.5.1 IDENTIFICAÇÃO DE MARCADORES PARA OS GENES *Dt1* e *Dt2*

A mesma abordagem foi utilizada para identificar os SNPs mais significativos para ambos os genes em estudo. Foi realizado um mapeamento associativo utilizando informações disponíveis no GRIN inseridas no programa GAPIT, onde foram rodadas as análises selecionando os SNPs com o melhor valor de p-value, que é a medida estatística que determina o corte. Os SNPs que romperam essa linha de corte foram os mais significativos para a característica estudada (em anexo os gráficos Manhattan Plot nas figuras 3.4 e 3.5).

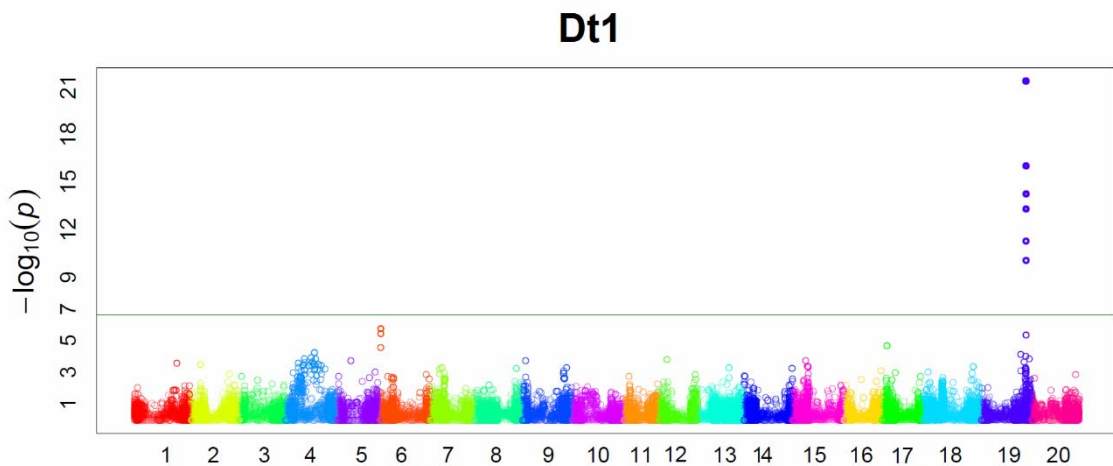


Figura 3.4 - Gráfico Manhattan Plot demonstrando qual SNP foi mais significativo para Dt1.

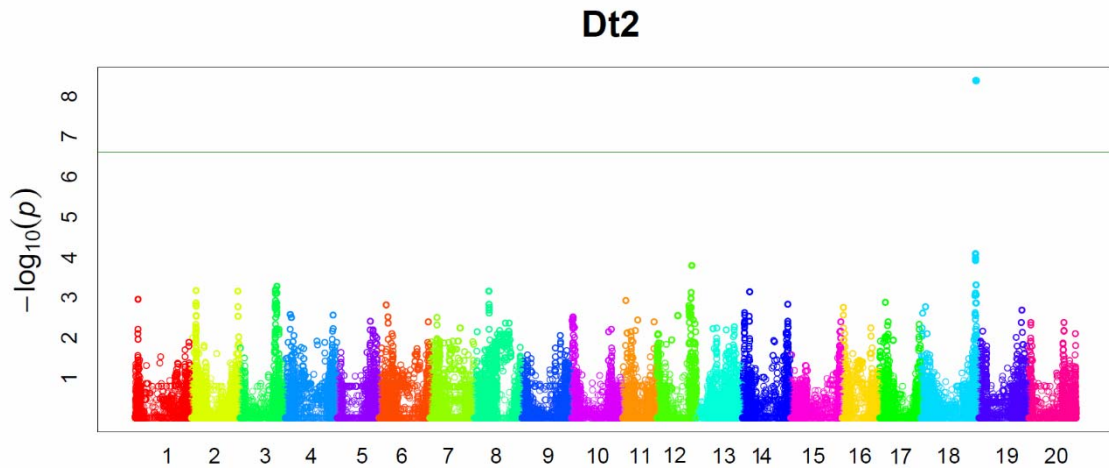


Figura 3.5 - Gráfico Manhattan Plot demonstrando qual SNP foi mais significativo para Dt2.

No estudo foram sintetizados dois SNPs para *Dt1* e quatro para *Dt2*, a primeira estratégia foi obter SNPs por mapeamento associativo onde se identificou os SNPs mais significativos para as características estudadas, já o segundo passo foi realizar uma análise utilizando um banco de dados do site da soykb.com filtrado por região de mutação não sinônima dentro do gene, nesse site foi realizada uma análise comparativa com base em 106 acessos e 28 variedades brasileiras resequenciadas (MALDONADO et al., 2016; PATIL et al., 2015) onde foi possível identificar os SNPs utilizados neste estudo, um para *Dt1* e três para *Dt2*.

3.5.2 PCR DO PAINEL ASSOCIATIVO

Após finalizar a PCR dos 384 materiais propostos para o painel associativo, os marcadores moleculares sintetizados foram avaliados através de análise gráfica, que determina se o marcador amplificou corretamente para os materiais sabidamente caracterizados, possibilitando a separação de bandas dos materiais com fenótipos conhecidos para os alelos positivos e negativos dos loci de interesse. Dos quatro marcadores moleculares sintetizados para *Dt2* apenas um deles foi descartado TMG_161816, pois não conseguiu separar os alelos para os materiais com os tipos de crescimento distintos que pode ser observado na figura 3.6 e sua acurácia foi bem inferior aos outros marcadores propostos (tabela 3.1).

Tabela 3.1 - Acurácia e porcentagem dos tipos de erros que ocorreram para os marcadores propostos em no painel associativo com 384 acessos para os três tipos de crescimento.

Marcadores	Gene	Acurácia	Erro I	Erro II
TMG_161902	Dt1	86,5%	7,3%	24,5%
TMG_161910	Dt1	86,1%	7,4%	25,2%
TMG_161815	Dt2	59,7%	67,2%	56,0%
TMG_161816	Dt2	47,5%	50,0%	47%

TMG_161817	Dt2	67,6%	24,6%	35,7%
TMG_161818	Dt2	61,2%	88,2%	55,7%

Os dois marcadores sintetizados para *Dt1* foram eficientes para distinção dos alelos para os materiais utilizados no painel associativo, como podem ser observados na figura 3.7 e a acurácia para ambos foi elevada em torno de 86% (tabela 3.1).

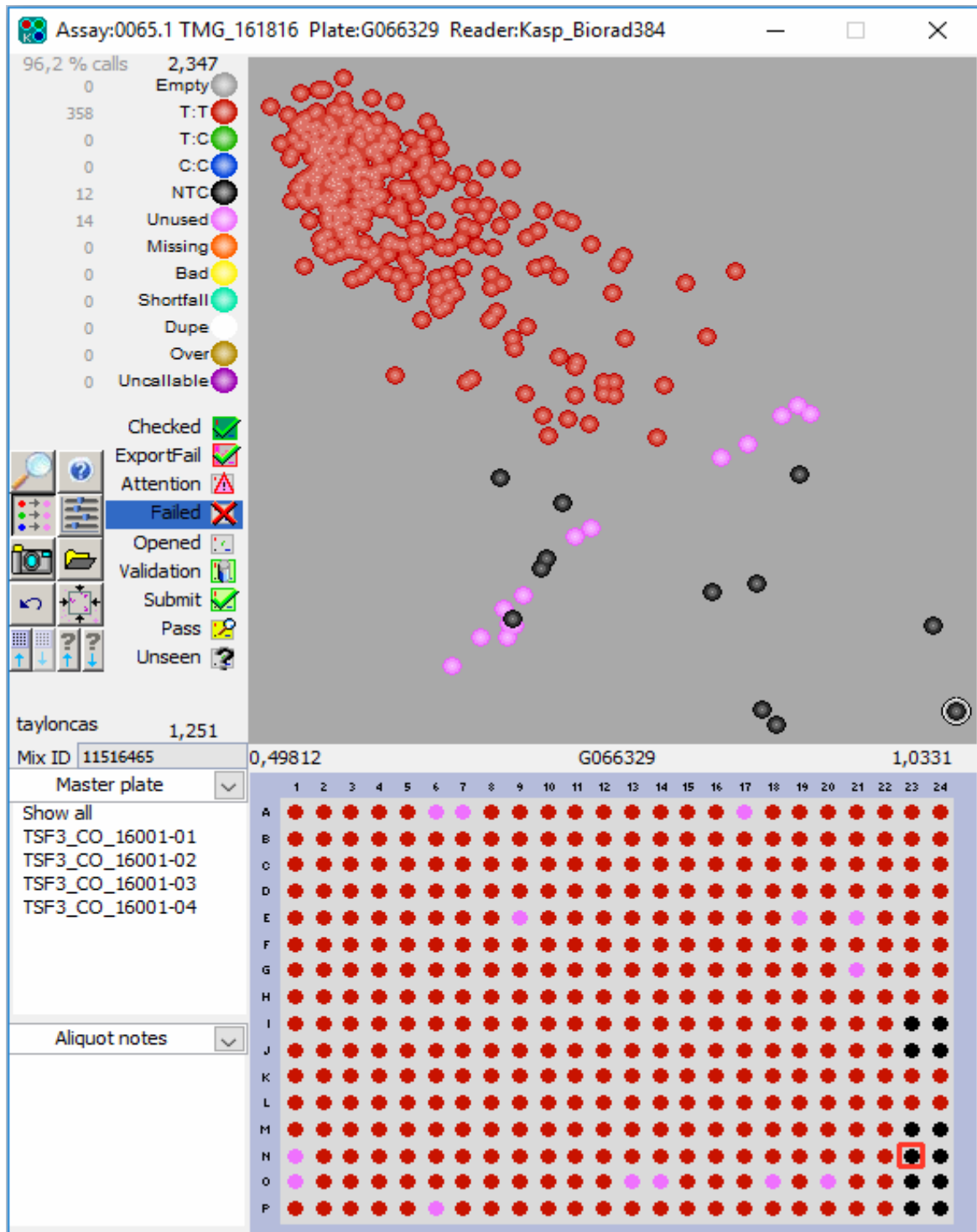


Figura 3.6 - Gráfico da leitura da fluorescência resultante da PCR, demonstrando a baixa capacidade do marcador TMG_161816 em separar plantas que possuem alelos recessivos ou dominantes de nulos ou erros, exibindo os 384 materiais genotipados no painel associativo.

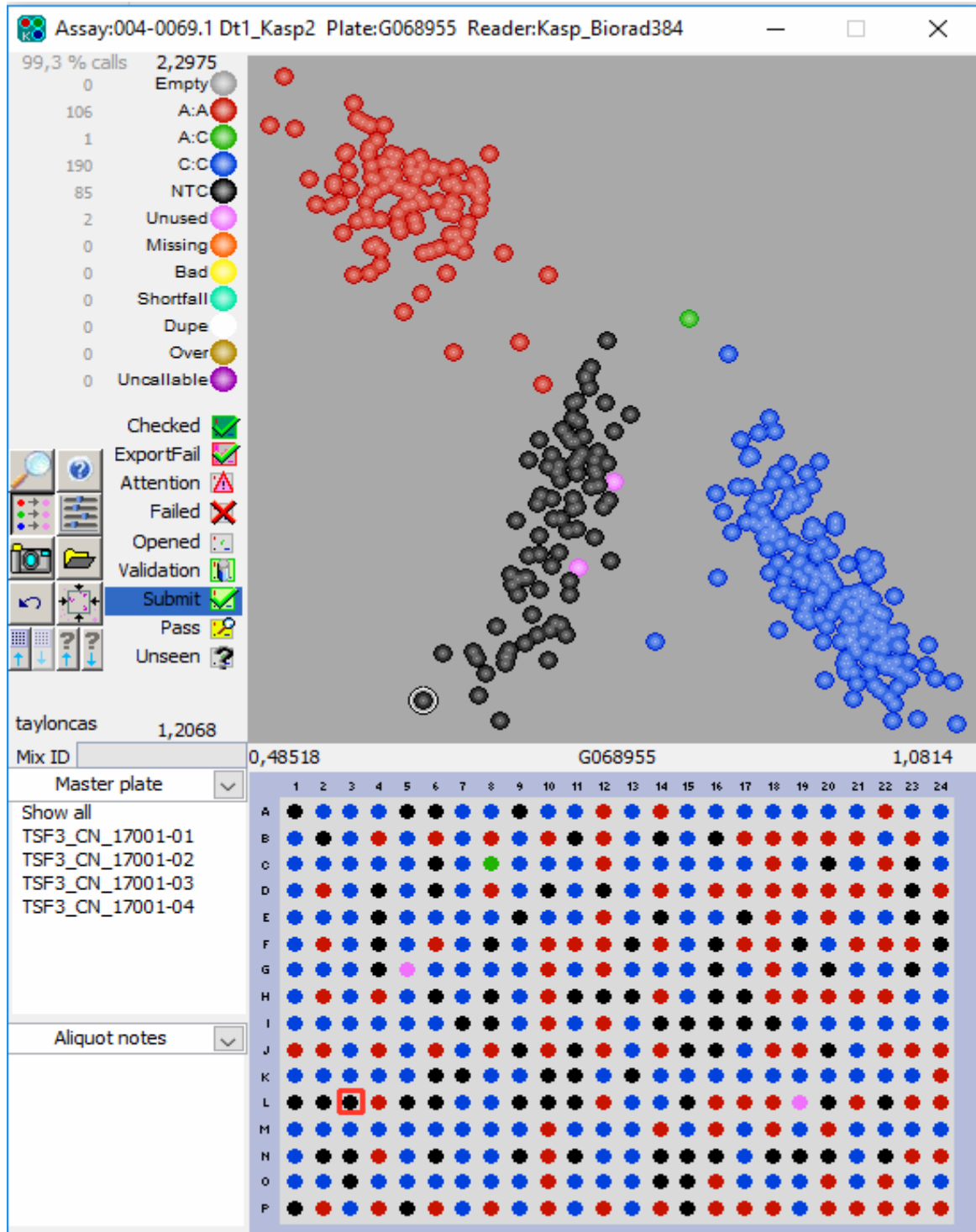


Figura 3.7 - Gráfico da leitura da fluorescência resultante da PCR, neste caso demonstra a elevada capacidade do marcador TMG_161902 em separar plantas que possuem alelos recessivos ou dominantes de nulos ou erros, 384 materiais do painel associativo.

O erro do tipo I ocorre quando rejeitamos a hipótese mesmo quando ela é verdadeira, ou seja, se a planta possui o alelo para os locos *Dt2*, mas por qualquer que seja o motivo o marcador não amplifica a sequência para *Dt2*. Nesta parte do estudo somente os dois marcadores propostos para *Dt1* poderiam ser aceitos, pois tiveram em ambos apenas 7% de erros do tipo I, já os quatro marcadores propostos para *Dt2* seriam descartados devido ao elevado número de

erros do tipo I, que foram de 24% no marcador TMG_161817 até 88% no marcador TMG_161818 (tabela 3.1).

O erro do tipo II acontece quando se aceita a hipótese quando ela é falsa, ou seja, as amostras da planta não contem os loci *Dt2*, mas o marcador amplifica mesmo assim gerando os chamados “falsos positivos”. Aqui também seriam aceitos os dois marcadores para *Dt1* que tiveram em torno de 25% de erros do tipo II ficando em uma média aceitável, mas novamente os 4 marcadores para *Dt2* nesse painel tiveram um resultado insatisfatório que foi de 36% para TMG_161817 chegando até 56% para TMG_161815, (tabela 3.1).

Esses resultados não possibilitam tirar conclusões a cerca destes marcadores para *Dt2* evidenciando a necessidade de estudos complementares como o de uma população segregante realizado neste estudo para se definir a qualidade e utilidade desses marcadores, o elevado número de erros encontrado neste painel associativo se deve ao fato de que vários materiais são registrados ou descritos erroneamente para os tipos de crescimento (tabela 3.2), corroborando com a conclusão de outros trabalhos (BERNARD, 1972; VICENTE et al, 2016).

Tabela 3.2 - Variedades comerciais e Acessos com a descrição de registro e o genotipo identificado pelos 5 marcadores propostos para os tipos de crescimento.

Variedade ou Acesso	Tipo de crescimento descrito (MAPA ou GRIN)	TMG_161 902	TMG_16 1910	TMG_16 1815	TMG_16 1817	TMG_16181 8
TMG1066RR	DETERMINADO	A:A (dt1dt1)	A:A (dt1dt1)	A:A (Dt2Dt2)	A:A (Dt2Dt2)	G:G (Dt2Dt2)
TMG4001	DETERMINADO	A:A (dt1dt1)	A:A (dt1dt1)	A:A (Dt2Dt2)	A:A (Dt2Dt2)	G:G (Dt2Dt2)
PI538781	DETERMINADO	C:C (Dt1Dt1)	T:T (Dt1Dt1)	A:A (Dt2Dt2)	A:A (Dt2Dt2)	G:G (Dt2Dt2)
PI153229	DETERMINADO	C:C (Dt1Dt1)	T:T (Dt1Dt1)	A:A (Dt2Dt2)	A:A (Dt2Dt2)	G:G (Dt2Dt2)
TMG7161RR	INDETERMINADO	C:C (Dt1Dt1)	T:T (Dt1Dt1)	T:T (dt2dt2)	G:G (dt2dt2)	A:A (dt2dt2)
TMG7363RR	INDETERMINADO	C:C (Dt1Dt1)	T:T (Dt1Dt1)	T:T (dt2dt2)	G:G (dt2dt2)	A:A (dt2dt2)
PI540883	INDETERMINADO	C:C (Dt1Dt1)	T:T (Dt1Dt1)	A:A (Dt2Dt2)	A:A (Dt2Dt2)	G:G (Dt2Dt2)
PI542057	INDETERMINADO	C:C (Dt1Dt1)	T:T (Dt1Dt1)	A:A (Dt2Dt2)	A:A (Dt2Dt2)	G:G (Dt2Dt2)
Will	SEMI	C:C	T:T	A:A	A:A	G:G
TIBAGI RR	DETERMINADO	(Dt1Dt1)	(Dt1Dt1)	(Dt2Dt2)	(Dt2Dt2)	(Dt2Dt2)
	SEMI	C:C	T:T	A:A	A:A	G:G
AS3570IPRO	DETERMINADO	(Dt1Dt1)	(Dt1Dt1)	(Dt2Dt2)	(Dt2Dt2)	(Dt2Dt2)
	SEMI	A:A	A:A	T:T	G:G	A:A (dt2dt2)
PI416963	DETERMINADO	(dt1dt1)	(dt1dt1)	(dt2dt2)	(dt2dt2)	
	SEMI	A:A	A:A	T:T	G:G	A:A (dt2dt2)
	DETERMINADO	(dt1dt1)	(dt1dt1)	(dt2dt2)	(dt2dt2)	

Destaque em negrito para os materiais com registro ou descrição divergente da identificada pelos marcadores propostos.

3.5.3 VALIDAÇÃO DE MARCADORES CANDIDATOS PARA *Dt2*

Após a análise gráfica dos resultados da PCR do painel associativo, se selecionou três marcadores moleculares candidatos para *Dt2*, que foram utilizados para PCR na população segregante, os parentais utilizados para estudo desta população foram confirmados por genotipagem no painel associativo sendo homocigotos para ambos os alelos estudados no caso de Will (*Dt1Dt1;Dt2Dt2*), e para a linhagem TMG2517 com o genótipo (*Dt1Dt1;dt2dt2*), portando a única segregação esperada e confirmada desse cruzamento foi apenas para *Dt2*.

Na análise para escolha do melhor marcador dentre os três se utilizou alguns critérios discutidos a seguir; o marcador TMG_161815 obteve maior número de erro do tipo I, mas foi o que resultou em menores números de falsos positivos ou erros do tipo II, já o marcador TMG_161818 resultou em intermediário número de erros tanto do tipo I quanto do tipo II, logo o marcador mais indicado seria o TMG_161817 que apresentou menor número de erros do tipo I e menor número de falsos positivos, o marcador TMG_161817 é uma mutação não sinônima que está situada em uma região dentro do gene, a probabilidade de causar uma alteração da proteína funcional é elevada, fato que pode ter possibilitado menor número de recombinações melhorando a confiabilidade dos resultados.

Em relação altura total de plantas (APT) ambos os marcadores foram eficientes para separar os materiais para os tipos de crescimento, sendo que as plantas do tipo indeterminado (IND) apresentaram (APT) superior as do tipo semideterminado (SMD) estatisticamente falando, um fato interessante é que plantas heterocigotas ficaram em uma faixa intermediária de (APT) indicando um efeito de dose de *Dt2* sobre a expressão de *Dt1*, esses dados podem ser observados na tabela 3.3.

Tabela 3.3 - Altura total de plantas, ciclo, comprimento entre os dois últimos internódios e número de vagens no racemo terminal em 903 indivíduos da população segregante para *Dt2*.

Genótipo	Altura (cm)	Ciclo (dias)	Compr.2 nós	Nº Vag. Rac.
dt2:dt2	90,81 a	101,22 a	2,5 a	1,27 a
Dt2:dt2	73,89 b	99,09 b	1,74 b	1,89 b
Dt2:Dt2	70,17 c	99,10 b	1,78 b	2,08 b
CV%	17,98	7,5	93,27	62,5

Após análises de normalidade foi realizado o Teste Tukey a 5% o genótipo foi validado pelo marcador TMG_161817.

Nos trabalhos de (PING et al., 2014; LIU et al., 2016) são descritos que *Dt2* determina a expressão de *Dt1*, nos resultados deste estudo as plantas com genótipo *Dt2;dt2* inibiram gradativamente a expressão de *Dt1* indicando um efeito de dose do alelo *Dt2* sobre *Dt1;Dt1*.

O gene *TFL1* em *Arabidopsis* é um ortólogo putativo de *Dt1* em soja (LIU et al., 2010; TIAN et al., 2010) que é conhecido como repressor floral *Terminal Flower1 (TFL1)* (SHANNON & MEEKS-WAGNER, 1991; BRADLEY et al., 1997), responsável por desencadear os processos bioquímicos que mesmo após as plantas receberem estímulos florais tanto externos como internos pode inibir a conversão de gemas apicais vegetativas em reprodutivas fazendo com que as plantas mantenham-se com o tipo crescimento indeterminado (LILJEGREN et al., 1999).

Trabalhos mais recentes puderam comprovar que na cultura da soja a alteração da forma do alelo *Dt2* para *dt2* condiciona a expressão de *Dt1* ou seja o alelo *dt2* na forma recessiva possibilita a expressão de *Dt1* conferindo às plantas o tipo indeterminado, na presença da forma dominante *Dt2* o fenótipo se torna semideterminado e a expressão de *Dt1* é reprimida (PING et al., 2014; LIU et al., 2016) corroborando com esses trabalhos os dados apresentados aqui podemos observar que plantas heterozigotas com genótipo *Dt2;dt2* induziram a produção parcial de *Dt1;Dt1* resultando na altura intermediária dessas plantas se comparadas a plantas homozigotas para as duas formas do alelo *Dt2;Dt2*.

O ciclo de desenvolvimento das plantas também foi influenciado pela interação dos alelos *Dt1* e *Dt2*, plantas homozigotas com tipo de crescimento indeterminado tiveram ciclo mais longo se comparadas com as dos genótipos homozigoto e heterozigotos semideterminado, podendo ser observados na tabela 3.3, sendo que *Dt2; dt2* foi estatisticamente igual a *Dt2; Dt2* para o ciclo de desenvolvimento, pelo teste de Tukey ao nível de 5% de probabilidade.

A interação dos alelos *Dt2* com *Dt1* pode interferir na distância entre os dois últimos internódios, as plantas (SMD) e (DET) são conhecidas por produzir um agrupamento e maior número de vagens nas gemas terminais, já genótipos com tipo de crescimento (IND) não produzem este “racemo terminal” (BERNARD, 1972; VICENTE et al., 2016), neste estudo foi observado que plantas com genótipo para o tipo (IND) tiveram um expressivo prolongamento da distância entre as duas últimas gemas reprodutivas de plantas (SMD) homozigotas e ou heterozigotas para o locus (*Dt2;__*). Para a característica do número de vagens no racemo terminal, também foi observado que onde o genótipo *dt2;dt2* foi expresso, as plantas (IND) produziram número inferior de vagens no racemo terminal comparadas às (SMD) com os genótipos (*Dt2;__*).

As avaliações por notas visuais foram validadas pelos resultados da genotipagem dos três marcadores moleculares utilizados, onde foram identificadas plantas com o genótipo *dt2;dt2* se obteve 97% de correspondência com as avaliações visuais realizadas, indicando elevada eficiência na distinção visual para o tipo de crescimento indeterminado e possibilidade de validação dos marcadores propostos. Já no caso das plantas genotipadas para o tipo de crescimento semideterminado, onde os marcadores amplificaram para *Dt2;Dt2* se obteve 43% de correspondência com avaliações visuais e para plantas heterozigotas *Dt2;dt2* onde o fenótipo expresso também deveria ser para o tipo semideterminado, resultou em uma correspondência de apenas 40% muito similar a obtida para os genótipos homizigotos, indicando que podem ocorrer um elevado número de erros na fenotipagem desta característica. Esses resultados reforçam o fato de que alguns materiais já foram registrados erroneamente para os tipos de crescimento terminal da haste (BERNARD, 1972; VICENTE et al., 2016), pois apenas as avaliações fenotípicas visuais principalmente em plantas para o tipo de crescimento semideterminado (*Dt2;__*) não são suficientes para se distinguir entre os dois tipos (IND) de (SMD) na tabela 3.2 podem ser observadas alguns cultivares com genótipo divergente do registrado.

Quando se analisou conjuntamente os dados mensurados; altura de plantas, distância entre os últimos internódios e número de vagens no rácemo terminal, foi possível distinguir as plantas para os tipos indeterminado e semideterminado validados pelos três marcadores selecionados para *Dt2* que podem ser observados na tabela 3.4.

Tabela 3.4 - Acurácia na distinção dos tipos de crescimento, porcentagens dos tipos de erros ocorridos I e II e proporção de segregação na população segregante de 903 para *Dt2*.

Marcadores	Gene	Acurácia %	Erro I %	Erro II %	Proporção SMD/IND
TMG_161815	<i>Dt2</i>	84,9	29,4	0,6	2,81:1
TMG_161817	<i>Dt2</i>	87,8	25,7	1	3,05:1
TMG_161818	<i>Dt2</i>	85,6	29,1	1,5	3,06:1
Avaliação Visual	<i>Dt2</i>	98,4	3,9	0	1,90:1

Os tipos de crescimento foram distintos através da combinando três fatores mensurados; altura de plantas > ou < que 80cm, distância entre os dois últimos internódios > ou < que 1 cm e número de vagens no racemo terminal > ou < que 1 que resultou em um filtro contendo 303 indivíduos do total de 903 da população estudada.

3.5.4 SEGREGAÇÃO

No trabalho de (PING et al., 2014) também foi utilizada uma população segregante oriunda de parentais contrastantes apenas para os tipos de crescimento (IND) e (SMD). Neste estudo apenas com os resultados da fenotipagem da geração F3 não possibilitou inferir a segregação

de 3:1 do alelo *Dt2*, mas com o auxílio da genotipagem utilizando os marcadores moleculares foi possível confirmar a dominância completa desse gene envolvido na expressão desta característica. Os três marcadores propostos tiveram elevada correspondência com a proporção de segregação esperada, ficando na média dos três marcadores de 2,97:1 para *Dt2* em relação a *dt2*.

Em um trabalho que utilizou linhagens isogênicas diferindo apenas para os três tipos de crescimento, o tipo semideterminado produziu plantas com características agronomicamente superiores ao tipo indeterminado como; maior massa de planta, diâmetro da haste, número de vagens, número de ramos e massa total de grãos e a altura de plantas foram superiores se comparadas às do tipo determinado (VICENTE et al., 2016).

3.6 CONCLUSÕES

Foram identificados polimorfismos no genoma da soja (tipo SNPs) e sintetizados marcadores moleculares ambos com elevado nível de acurácia ficando em torno de 86% para *Dt1* e *Dt2*.

Os resultados da PCR do painel associativo, possibilitou definir quais os marcadores propostos poderiam ser utilizados na população segregante, mas devido ao elevado número de materiais registrados erroneamente não foi possível definir qual seria o melhor marcador, justificando a necessidade do estudo complementar através de uma população segregante.

Neste trabalho foi proposto um método de avaliação fenotípica elaborado que utiliza a análise conjunta de dados morfológicos mensurados, (altura de plantas, distância entre os últimos internódios e número de vagens no rácemo terminal), que possibilitou distinguir as plantas para os tipos indeterminado e semideterminado viabilizando apurar a acurácia da predição dos marcadores para *Dt2*.

Os marcadores selecionados para *Dt2* chegaram muito próximos da proporção esperada para dominância completa de um gene, ambos foram eficientes para diferenciar as plantas para os dois tipos de crescimento na população segregante, validados pela análise conjunta de dados mensurados.

REFERÊNCIAS

- BERBEL, A.; FERRADIZ, C.; HECHT, V.; DALMAIS, M.; LUND, O.S.; SUSSIMILCH, F.C. *VEGETATIVE1* is a essential for development of the compound inflorescence in pea. **Nat Commum.** 3:797, 2012.
- BERNARD, R.L. Two genes affecting stem termination in soybeans. **Crop Sci.** 12: 235–239, 1972.
- BERNARD R.L.; C.R. CREMEENS C.R. Registration of 'Will' Soybean. **Crop Sci.** (Madison) 28(6):1028 Comment: CV-223, 1988.
- BRADLEY, D.; RATCLIFFE, O.; VINCENT, C.; CARPENTER, R.; COEN, E. Inflorescence commitment and architecture in Arabidopsis. **Science.**275: 80–83, 1997.
- CHANG, J.F; Green, D.E; Shibles R. Yield and agronomic performance of semi-determinate and indeterminate soybeans estem types. **Crop Sci.** 22(1):97-101, 1982.
- ELKIND, Y.; GURNICK, A.; KEDAR, N. Genetics of semideterminate growth habit in tomato. **HortScience**, 26: 1074–1075, 1991.
- FOUCHER, F.; MORIN, J.; COUTIADE, J.; CADIOUX, S.; ELLIS, N.; BANFIELD, M.J. *DETERMINATE* and *LATE FLOWERING* are two *TERMINAL FLOWER1/CENTRORADIALIS* homologs that control two distinct phases of flowering initiation and development in pea. **Plant Cell.**15(11):2742-54, 2003.
- FRANCO, J.; CROSSA, J.; RIBAUT, J.M.; BETRAN, J. A method for combining molecular markers and phenotypic attributes for classifying plant genotypes. **Theor. Appl. Genet.** 103: 944-952, 2001.
- FRIDAMAN, E.; LIU, Y.S.; CARMEL-GOREN, L.; GUR, A.; SHORESH, M.; PLEBAN, T.; ESHED, Y.; ZAMIR, D. . Two tightly linked QTLs modify tomato sugar content via different physiological pathways. **Mol. Genet. Genomics** 266: 821–82, 2002.
- GUSTAFSON-BROWN C.; SAVIDG B.; YANOFSKY M.F. Regulation of the Arabidopsis floral homeotic gene *APETALA1*. **Cell.** 76(1):131-43, 1994.
- HEATHELERLY, L.G; SMITH, J.R. Effect of soybean stem growth habit on height and node number after beginning bloom in the midsouthern USA. **Crop Sci.** 44(5):1855-8, 2004.

HEGDE, V.S. Morphology and genetics of a new found determinate genotype in chickpea. **Euphytica**. 182: 35–42, 2011.

HYMOWITZ, T. Speciation and cytogenetics. In: BOERMA, H.R.; SPECHT, J.E. (Ed.). **Soybeans: improvement, production and uses**. 4th ed. Madison: American Society of Agronomy, Crop Science Society of America and Soil Science Society of America, chap. 4, p. 97-136, 2004.

KEIM, P.; OLSON, T.C.; SHOEMAKER, R.C. A rapid protocol for isolating soybean DNA. **Soybean Genetics Newsletter**, v.15, p.150-152, 1988.

LILJEGREN, S.J.; GUSTAFSON-BROW, C.; PINYOPICH, A.; DITTA, G.S.; YANOFSKY, M.F. Interactions among *APETALA1*, *LEAFY*, and *TERMINAL FLOWER1* specify meristem fate. **Plant Cell**. 11(6):1007-18, 1999.

LIPKA, A.E.; et al. GAPIT: Genome association and prediction integrated tool. **Bioinformatics** 28(18):2397–2399, 2012.

LIU, B.; WATANABE, S.; UCHIYAMA, T.; KONG, F.; KANAZAWA, A.; XIA, Z.; NAGAMATSU, A.; ARAI, M.; YAMADA, T.; KITAMURA, K.; MASUTA, C.; HARADA, K.; ABE, J. The soybean stem growth habit gene *Dt1* is an ortholog of *Arabidopsis* *TERMINAL FLOWER1*. **Plant Physiol**. 153:198–210, 2010.

LIU Y.; ZHANG D.; PING J.; LI S.; CHEN Z.; MA J. Innovation of a Regulatory Mechanism Modulating Semideterminate Stem Growth through Artificial Selection in Soybean. **PLoS Genet**. 12(1), 2016.

MALDONADO S.V.J.; VALLIYODAN B.; JOSHI T.; KHAN S.M.; LIU Y.; WANG J.; ABDELNOOR R.V. Avaliação da variação genética entre cultivares brasileiras de soja por resequenciamento do genoma. **BMC Genomics**. 17:110, 2016.

PATIL G.; VALLIYODAN B.; DESHMUKHM R.; PRINCE S.; NICANDER B.; ZHAO M.; SONAH H.; SONG L.; LIN L.; CHAUDHARY J.; LIU Y.; JOSHI T.; XU D.; HENRY T. Soybean (*Glycine max*) SWEET gene family: insights through comparative genomics, transcriptome profiling and whole genome re-sequence analysis. **BMC Genomics**. 16:520, 2015.

PING, J.; LIU, Y.; SUN, L.; ZHAO, M.; LI, Y.; SHE, M.; SUI, Y.; LIN, F.; LIU, X.; TANG, Z.; NGUYEN, H.; TIAN, Z.; QIU, L.; NELSON, L. R.; CLEMENTE E. T.; SPECHT, E. J.;

MA, J. Dt2 is a gain-of-function MADS-domain factor gene that specifies semideterminacy in soybean. **Plant Cell** 26: 2831–2842, 2014.

PNUELI L.; CARMEL-GOREN L.; HAREVEN D.; GUTFINGER T.; ALVAREZ J.; GANAL M.; ZAMIR D.; LIFSCHITZ E. (1998). The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. **Development** 125: 1979–1989, 1998.

REPINSKI, S.L; KWAK, M.; GEPTS, P. The common bean growth habit gene *PvTFL1y* is a functional homolog of Arabidopsis *TFL1*. **Theor Appl Genet.** 124(8):1539-47, 2012.

SETIYONO, T.D.; WEISS, A.; SPECHT, J.; BASTIDAS, A.M.; CASSMAN, K.G.; DOBERMANN, A. Understanding and modeling the effect of temperature and daylength on soybean phenology under high-yield conditions. **Field Crops Res.** 100: 257–271, 2007.

SHANNON S.; MEEKS-WAGNER D.R.; A mutation in the Arabidopsis *TFL1* gene affects inflorescence meristem development. **Plant Cell.** 3(9):877-92, 1991.

TANAKA, Y. & SHIRAIWA, T., Stem growth habit affects leaf morphology and gas exchange traits in soybean. **Annals of Botany** 104: 1293–1299, 2009.

TIAN, Z.; WANG, X.; LEE, R.; LI, Y.; SPECHT, J.E.; NELSON, R.L. Artificial selection for determinate growth habit in soybean. **Proc Natl Acad Sci USA.** 107(19):8563-8, 2010.

VICENTE D.; SCHUSTER I.; LAZZARI F.; PARANZINI D.P.J.; OLIVEIRA R.A.M.; PRETE C.E.C.; Mapping and validation of molecular markers of genes Dt1 and Dt2 to determine the type of stem growth in soybean. **Acta Scientiarum.** 1679-9275, 2016.

WAGNER, D.; SABLONSKI, R.W.M.; MEYEROWITZ, E.M. Transcriptional activation of *APETALA1* by *LEAFY*. **Science.** 285(5427):582-4, 1999.

WEIGEL, D.; ALVAREZ, J.; SMYTH, D.R.; YANOFSKY, M.F.; MEYEROWITZ, E.M. *LEAFY* controls floral meristem identity in Arabidopsis. **Cell.** 69(5):843-59, 1992.

WINTER, C.M; AUSTIN, R.S.; BLAVILLAIN-BAUFUME, S.; REBACK, M.A.; MONNIAUX, M.; WU, M.F. *LEAFY* target genes reveal floral regulation logic, cis motifs, and a link to biotic stimulus response. **Dev. Cell.** 20(4):430-43, 2011.

WONG, C.E.; SINGH, M.B; BHALLA, P.L. Molecular processes underlying the floral transition in the soybean shoot apical meristem. **Plant Journal.** 57(5):832-45, 2009.