

ANA RAFAELA DA SILVA OLIVEIRA

Mapeamento citogenético comparativo em espécies de *Glycine* Willd.

Recife
2013

ANA RAFAELA DA SILVA OLIVEIRA

Mapeamento citogenético comparativo em espécies de *Glycine* Willd.

Dissertação apresentada ao Programa de Pós-Graduação em Agronomia “Melhoramento Genético de Plantas” da Universidade Federal Rural de Pernambuco, como parte dos requisitos para obtenção do grau de Mestre em Agronomia/Melhoramento Genético de Plantas.

COMITÊ DE ORIENTAÇÃO:

Professora Ana Christina Brasileiro-Vidal, Orientadora – UFPE

Professora Ana Maria Benko-Iseppon, Coorientadora – UFPE

Recife

2013

Ficha Catalográfica

O48m Oliveira, Ana Rafaela da Silva
Mapeamento citogenético comparativo em espécies de
Glycine Willd. / Ana Rafaela da Silva Oliveira. -- Recife, 2013.
86 f. : il.

Orientador (a): Ana Christina Brasileiro Vidal.
Dissertação (Mestrado em Melhoramento Genético de
Plantas) – Universidade Federal Rural de Pernambuco,
Departamento de Agronomia, Recife, 2012.
Inclui referências e apêndice.

1. Citogenética 2. BAC 3. FISH 4. *Glycine Max* 5. Mapa
cromossômico 6. Soja I. Vidal, Ana Christina Brasileiro,
Orientadora II. Título

CDD 581.15

Mapeamento citogenético comparativo em espécies de *Glycine* Willd.

ANA RAFAELA DA SILVA OLIVEIRA

Dissertação defendida e aprovada pela Banca Examinadora em 31/10/2012.

ORIENTADORA:

Profa. Dra. Ana Christina Brasileiro-Vidal – UFPE

EXAMINADORES:

Profa. Dra. Kyria Cilene de Andrade Bortoleti - UNIVASF

Profa. Dra. Andrea Pedrosa-Harand – UFPE

Prof. Dr. Reginaldo de Carvalho – UFRPE

**RECIFE – PE
Janeiro, 2013**

***A Deus e aos meus pais,
Dedico.***

Agradecimentos

Ao meu **Deus**, pela sua infinita bondade e misericórdia e pelo seu amor incondicional.

À minha mãe, **Maria da Paz da Silva Oliveira**, pelo amor e, sobretudo pelas lições de vida.

Ao meu pai, **Robson José de Oliveira**, pelo companheirismo, dedicação.

Aos meus irmãos, **Gercino Francisco da Silva Neto e Thaísa Gabrielle da Silva Oliveira**, pela amizade e apoio.

Meu sobrinho **Gercino Bisneto e companhia LTDA.**

À minha tia-mãe (Goga) **Gorette**, pela garra, dedicação e exemplo de mulher.

À minha tia **Fau**, por sua amizade.

À minha prima **Maria Júlia**, pela alegria contagiante que é te encontrar.

Aos **demais familiares**, pelo amor.

À Profa. Dra. **Ana Christina Brasileiro Vidal**, pela oportunidade, orientação, confiança e incentivo.

À Profa. Dra. **Ana Maria Benko Iseppon**, pelo apoio e atenção, e por sempre discutir e melhorar nossos trabalhos.

À Profa. Dra. **Andrea Pedrosa**, pela contribuição com a realização deste trabalho.

À Técnica do LGBV, **Vanessa Souza**, pela contribuição em manter o laboratório em sintonia e por saber lidar com os meus caprichos.

Aos amigos e colegas do LGBV **Ana Maria Félix, Anne Catherine, Artur Wanderley, Bruna Piereck, Diego Sotero, Emanuelle Vasconcelos, Flávia de Araújo, Geyner Cruz, Hayana Azevedo, Hévila Mendes, Isis Raphaella, João Pacífico, Lidiane Amorim, Luís Carlos Belarmino, Luiz Rodrigo Gazzaneo,**

Marília Gabriela, Marcos Antônio, Mitalle Matos, Pedro Marcos de Almeida, Pollyana Karla, Rômulo Dantas, Rosilda Cintra, Rodrigo Oliveira, Santelmo Vasconcelos, Silvany Araújo, Stephani Santos e Wilka Maria, pela felicidade que é conviver entre vocês.

Aos pesquisadores da Embrapa Soja, **Ricardo Vilela Abdelnoor e Marcelo Fernandes de Oliveira**, pela colaboração e fornecimento do material vegetal.

Aos professores do Programa de Pós-Graduação em Agronomia – UFRPE, pelos conhecimentos adquiridos.

Aos colegas de curso **Adriana, Lindomar, Marcina, Jayne, Kessiana, Renata, Allyson, José Rodolfo, João Filipi, Lucas, Guilherme, Hudson, Thiago Prates, Ramon, Ismael Gaião**, pelo prazer de conviver com vocês.

À secretária **Bernadete**, pela colaboração.

Às agências de fomento à pesquisa Conselho Nacional de Desenvolvimento Científico e Tecnológico (**CNPq**) e à Coordenação de Pessoal de Nível Superior (**CAPES**), pelo suporte financeiro imprescindível para a realização deste trabalho.

E a todos que contribuíram direta ou indiretamente na realização deste trabalho,

Muito obrigada!

Sumário

Lista de abreviaturas	i
Lista de figuras	ii
Lista de tabelas	iii
Resumo	iv
Abstract	v
1. Introdução	1
2. Revisão da literatura	4
2.1. O gênero <i>Glycine</i> Willd.	5
2.1.1 Filogenia do gênero <i>Glycine</i> Wild.	5
2.1.2 <i>Pools</i> gênicos	7
2.2 Importância econômica da soja	8
2.3 Origem e domesticação da soja	9
2.4 Construção de bibliotecas BACs e sequenciamento do tipo <i>shotgun</i>	10
2.4.1 Sequenciamento do genoma da soja	12
2.5 Sintonia entre soja e outras leguminosas	13
2.6 Estudos citogenéticos em <i>Glycine</i>	14
2.6.1 Análises convencionais	14
2.6.2 DNAr 5S e 45S	17
2.6.3 Hibridização <i>in situ</i> de outras sequências	19
2.6.4 Uso do BAC-FISH em análise genômico-comparativas	21
3. Referências bibliográficas	23
4. Capítulo	36
5. Conclusões	59
6. Apêndice	61

Lista de abreviaturas

8-HQ	<i>8-hydroxyquinoline</i> 8-hidroxiquinoleína
BAC	<i>Bacterial Artificial Chromosome</i> Cromossomo Artificial de Bactéria
BSA	<i>Bovine Serum Albumin</i> Albumina de Soro Bovino
cM	centiMorgans
CMA ₃	Chromomycin A3 Cromomicina A3
cv.	cultivar
Cy3	<i>Cyanine 3</i> Cianina 3
DAPI	<i>4',6-diamidino-2-phenylindole</i> 4',6-diamidino-2-fenilindol
DNA	<i>Desoxiribonucleic Acid</i> Ácido Desoxirribonucleico
DNA _r	<i>Ribosomal DNA</i> DNA ribossomal
EST	<i>Expressed Sequence Tags</i> Etiquetas de Sequências Expressas
FAO	<i>Food and Agriculture Organization</i> Organização para Agricultura e Alimentação
FISH	<i>Fluorescent In Situ Hybridization</i> Hibridização <i>In Situ</i> Fluorescente
FITC	<i>Fluorecein isotiocyanate</i> Fluoresceína isotiocianato
kb	Kilo bases
LG	<i>Linkage group</i> Grupo de Ligação
LGBV	Laboratório de Genética e Biotecnologia Vegetal
Mb	<i>Mega-base</i> Mega bases
Pb	<i>base pairs</i> Pares de bases
QTLs	<i>Quantitative Trait Loci</i> Locos de Caracteres Quantitativos
RFLP	<i>Restriction Fragment Length Polymorphism</i> Polimorfismo de Comprimento de Fragmentos de Restrição
SNPs	<i>Simple Nucleotide Polymorphisms</i> Polimorfismos de um Único Nucleotídeo
SSC	Citric acid–sodium citrate buffer Tampão Citrato de Sódio-Salino
SSR	<i>Sequence Simple Repeats</i> Sequências Simples Repetidas
UFPE	Universidade Federal de Pernambuco
UFRPE	Universidade Federal Rural de Pernambuco
UNIVASF	Universidade Federal do Vale do São Francisco

Lista de figuras

Revisão da literatura

- Figura 1.** Origem e distribuição geográfica da soja (*Glycine max*), incluindo áreas de introdução para o cultivo. 10
- Figura 2.** Metáfase convencional de *Glycine max* e *G. soja*. 15
- Figura 3.** Idiograma dos cromossomos de soja que correspondem aos cinco grupos baseado no padrão de condensação. 17
- Figura 4.** Metáfase de *G. max*, corada com CMA/DAPI (A) e hibridizada *in situ* com sondas de DNAr 45S (verde) e 5S (vermelho). 18

Capítulo

- Figura 1.** Hibridização *in situ* do BAC 173N15 do cromossomo *Gm18* de *Glycine max* em cromossomos mitóticos de *G. max* (A) e *G. soja* (B), contracolorados com DAPI (cinza). 45
- Figura 2.** Hibridização *in situ* de três clones BAC de *Glycine max* (*Gm16*): 145O12 (amarelo); 149P23 (azul escuro); 102N16 (laranja) e 88G15 (rosa), bem como de sondas de DNAr 5S (vermelho) e 45S (verde) em *G. tomentella*. Cromossomos mitóticos de *G. max*, $2n = 40$ (A); *G. soja*, $2n = 40$ (B); *G. syndetika*, $2n = 40$ (C, D); *G. tabacina*, $2n = 80$ (E) e *G. tomentella*, $2n = 78$ (F1 e F2). 46
- Figura 3.** Mapas genéticos (Satt) e de sequenciamento do cromossomo *Gm16* (J) de *G. max*. Representação esquemática dos cromossomos e cromossomos de homeologia de *G. max*, *G. soja*, *G. syndetika*, *G. tabacina* e *G. tomentella*. 47
- Figura 4.** Hibridização *in situ* de três clones BAC para o cromossomo 18 de *Glycine max*: 93L17 (vermelho); 138F23 (lilás), e 176B09 (azul claro). Cromossomos mitóticos de *G. max*, $2n = 40$ (A, B), *G. soja*, $2n = 40$ (C), *G. syndetika*, $2n = 40$ (D), *G. tomentella*, $2n = 78$ (E) e *G. tabacina*, $2n = 80$ (F). 48
- Figura 5.** Mapas genético (Satt ou Sat_) e de sequenciamento do cromossomo *Gm18* (G) de *G. max*. Representação esquemática dos cromossomos e cromossomos de homeologia de *G. max*, *G. soja*, *G. syndetika*, *G. tabacina* e *G. tomentella*. 50

Lista de tabelas

Revisão da literatura

Tabela 1. Espécies do gênero <i>Glycine</i> Willd. e seus respectivos números cromossômicos.	06
---	-----------

Capítulo

Tabela 1. Relação de espécies analisadas de <i>Glycine</i> , com seus respectivos números cromossômicos, genoma nuclear, grupo molecular, acesso e distribuição geográfica.	40
Tabela 2. Clone BAC, localização <i>in silico</i> no Soybase, região na pseudomolécula de <i>Glycine max</i> verificada <i>in silico</i> , tamanho em kb, região de sintenia em outros cromossomos de <i>G. max</i> e marcador(es) molecular(es) associado(s). Dados disponíveis no banco de dados SoyBase (http://soybase.org/gb2/gbrowse/gmax1.01/)	43
Tabela 3. Clones BAC hibridizados e resultados de suas respectivas localizações <i>in situ</i> em cinco espécies de <i>Glycine</i> .	50

Resumo

A soja [*G. max* (L.) Merr.] é considerada a principal oleaginosa cultivada, tendo sido a primeira leguminosa a ser sequenciada, gerando informações genômicas importantes. O presente trabalho teve por finalidade realizar um estudo genômico comparativo, mediante a localização *in situ* de oito BACs (*Bacterial Artificial Chromosomes*) pertencentes aos grupos de ligação G (pseudomolécula 18, quatro BACs) e J (pseudomolécula 16, quatro BACs) de *G. max* ($2n = 40$) em cromossomos de *G. soja* ($2n = 40$), *G. syndetika* ($2n = 40$), *G. tabacina* ($2n = 80$) e *G. tomentella* ($2n = 78$). As análises citogenéticas das sondas em *G. max*, *G. soja* e *G. syndetika* mostraram que sete BACs apresentaram marcação única em um ou dois pares cromossômicos, exceto para *G. soja* com o BAC 176B09 (três pares) e *G. syndetika* com o BAC93L17 (quatro pares). Para o BAC 173N15 em *G. max* e *G. soja* foram reveladas marcações de DNA repetitivo nas regiões centroméricas e pericentroméricas em todos os cromossomos de *G. max* e em 10 pares de *G. soja*, demonstrando diferenciação dessas regiões ao longo da evolução. Devido à presença do DNA repetitivo, esse BAC não foi hibridizado nas demais espécies. Em *G. tabacina*, dois BACs de cada grupo de ligação foram hibridizados em um par de cromossomos homeólogos equivalente, sendo visualizadas algumas marcas extras para os BACs do *Gm18*. Em relação a *G. tomentella*, apenas para os BACs de *Gm16* (102N16 e 145O12) foi possível identificar o par cromossômico de homeologia em relação à *G. max*, observando-se contudo a presença de sinais extras também referentes a possíveis alterações, como duplicação e translocação. Os BACs de *Gm18* (97L17 e 176B09) hibridizaram em diferentes cromossomos *G. tomentella* (dois pares cada), indicando prováveis rearranjos cromossômicos ao longo da evolução. Para os BACs 102N16 (*Gm16*) em *G. tomentella*, 93L17 (*Gm18*) em *G. syndetika* e *G. tomentella* e 176B09 (*Gm18*) em *G. soja*, um dos pares extras foi associado a um par de sítios de DNAr 45S decorrentes de possíveis divergências dos espaçadores intergênicos em sítios de DNAr. Foi observado também que para *Gm18*, a orientação do mapa de sequenciamento parece estar invertida em relação à morfologia cromossômica. De um modo geral, os resultados sugerem uma quebra de sintenia para *Gm16* em *G. syndetika*, *G. tomentella* e para *Gm18* em todas as espécies analisadas.

Palavras-chave: BAC, FISH, *Glycine max*, mapa cromossômico, soja

Abstract

Soybean [*G. max* (L) Merr.] is considered the main cultivated oil plant and also the first oilseed crop and legume sequenced, generating important genomic information. The present work aimed to perform a comparative genomic study, through *in situ* localization of eight BACs (Bacterial Artificial Chromosomes) sequences belonging to linkage groups G (pseudomolecule 18, four BACs) and J (pseudomolecule 16, four BACs) from *G. max* ($2n = 40$) in *G. soja* ($2n = 40$), *G. syndetika* ($2n = 40$), *G. tabacina* ($2n = 80$) and *G. tomentella* ($2n = 78$) chromosomes. Cytogenetic analyzes using the probes in *G. max*, *G. soja* and *G. syndetika* showed that seven BACs yielded single markers in only one or two chromosome pairs, except for *G. soja* with the BAC 176B09 (three pairs) and *G. syndetika* with BAC93L17 (four pairs). In *G. max* and *G. soja*, only 173N15 revealed a DNA repetitive mark when hybridized *in situ* in pericentromeric and centromeric regions in all chromosomes of *G. max* and 10 pairs of *G. soja*, showing differentiation of these regions throughout evolution. Due to the presence of repetitive DNA, the BAC has not been hybridized in the other species. In *G. tabacina* two BACs of each linkage group hybridized on a pair of homeologous chromosomes, equivalent being viewed some extra markers viewed for the *Gm18* BACs. In relation to *G. tomentella*, only for the BACs of *Gm16* (102N16 and 145O12) there was a conservation of synteny as compared with *G. max*, observing however the presence of some extra signals, such as duplication and translocation. The BACs of *Gm18* (97L17 and 176B09) hybridized on different chromosomes (two pairs each) in *G. tomentella*, what indicate probable chromosomal rearrangements during evolution. For BACs 102N16 (*Gm16*) in *G. tomentella*, 93L17 (*Gm18*) in *G. syndetika* and *G. tomentella* and 176B09 (*Gm18*) in *G. soja*, one of the extra pairs was associated with a pair of 45S rDNA sites related to possible differences in the intergenic spacers of rDNA sites. It was also observed that for *Gm18*, the map orientation sequencing appears to be inverted with respect to the chromosome morphology. The results suggest a break of synteny for *Gm16* in *G. syndetika* and *G. tomentella* and for *Gm18* in all analyzed species.

Keywords: BAC, FISH, *Glycine max*, chromosomal map, soybean

Introdução

1. Introdução

A soja [*Glycine max* (L) Merr.] é uma planta dicotiledônea, pertencente à subfamília Papilionoideae (Fabaceae), que tem grande importância econômica e ampla distribuição. Os grãos da soja são utilizados para muitos fins, porém, a maior importância da soja está relacionada à produção de óleo para a indústria e alimentação humana. Além disso, pode ser usada para a produção de farelo para ração animal e na recuperação de solos (fixação de nitrogênio atmosférico no solo), além de servir como matéria-prima para a produção de biocombustível (Costa Neto *et al.*, 2000; Graham e Vance, 2003; Chung e Singh, 2008).

Atualmente, o Brasil é o segundo maior produtor de soja, com vários programas de melhoramento visando à obtenção de cultivares com resistência a pragas, doenças e estresses ambientais, gerando plantas mais produtivas e/ou com características que atendam aos mercados industrial e alimentício (Embrapa, 2012). Recentemente, a conclusão do sequenciamento da soja permitiu o uso das informações moleculares para a construção de mapas genéticos e físicos, auxiliando no entendimento da estrutura e da organização genômica em alguns grupos de plantas, fornecendo informações úteis para a manipulação genética e consequente obtenção de genótipos superiores (Soybase, 2012). Dentre as técnicas citogenéticas que auxiliam no estudo do genoma, destaca-se a hibridização *in situ* fluorescente (FISH), a qual permite a localização física exata de sequências de DNA ao longo dos cromossomos e possibilita a construção de mapas citogenéticos e a sua integração com o mapa genético (Guerra, 2004; Fonsêca *et al.*, 2010).

Nesse sentido, cromossomos artificiais de bactérias (BACs), contendo marcas genéticas de interesse para a soja, podem ser selecionados *in silico* e obtidos através do site: <http://soybase.org/gb2/gbrowse/gmax1.01/> (Mudge *et al.*, 2005a) para uso como sondas na FISH. Análises BAC-FISH em soja, *G. soja* Sieb. e Zucc. e *G. gracilis* Skvortz, realizadas por Findley *et al.* (2010; 2011), por exemplo, têm servido como base para estudos de evolução cariotípica e de genômica comparativa, contribuindo para análise da macrossintenia no gênero e fornecendo informações importantes para programas de melhoramento. Dessa forma, o presente trabalho visou realizar um estudo genômico comparativo, mediante análise do padrão de distribuição de sequências BACs de soja em *G. max* ($2n = 40$), *G. soja* ($2n = 40$), *G. syndetika* B. E. Pfeil e Craven ($2n = 40$), *G. tabacina* Benth (Labill.) ($2n$

= 80), e *G. tomentella* Hayata ($2n = 78$), a fim de colaborar para o entendimento das relações de macrossintenia existentes nesse grupo de plantas.

Revisão da Literatura

2. Revisão da literatura

2.1 O gênero *Glycine* Willd.

Fabaceae é a terceira maior família das angiospermas, a qual corresponde às leguminosas. Possui uma ampla variedade de hábitos vegetativos, o que possibilita uma ocupação dos mais diversos tipos de habitats (Lewis *et al.*, 2005). Engloba cerca de 20.000 espécies distribuídas em 730 gêneros, entre os quais, destacam-se *Arachis* L., *Cicer* L., *Glycine* Willd., *Lotus* L., *Medicago* L., *Phaseolus* L., *Pisum* L. e *Vigna* Savi (Doyle e Luckow, 2003; Chung e Singh, 2008; APG III, 2010). Além da grande importância econômica, as leguminosas têm contribuído para o enriquecimento do solo, por meio da fixação de nitrogênio atmosférico mediante sua interação simbiótica com bactérias rizóbios, presentes no solo (Zhukov, 2009).

Glycine é um dos gêneros mais importantes entre as leguminosas. Está dividido dois subgêneros, *Glycine* Willd. (perene) e *Soja* (Moench) F.J. Herm (anual), com números cromossômicos somáticos variando entre 38 e 80 (Chung e Singh, 2008). O subgênero *Glycine* é composto por 26 espécies nativas da Austrália e de ilhas do oceano Pacífico. Dessas, 21 apresentam cariótipo com $2n = 40$ cromossomos (ver Tabela 1); duas espécies mostram $2n = 80$ (*G. dolichocarpa* Tateishi e Ohashi e *G. pescadrensis* Hayata); duas espécies possuem $2n = 40$ ou 80 (*G. hirticaulis* Tind. e Craven e *G. tabacina*), e uma (*G. tomentella*), que apresenta quatro citotipos diferenciados em: diploide, tetraploide e dois aneuploides, $2n = 40$, 80 , 38 e 78 , respectivamente. De um modo geral, as espécies do subgênero *Glycine* são extremamente diferentes em morfologia, citologia e genoma. Além de crescerem em climas muito diferentes, e terem uma distribuição geográfica bastante ampla, nunca foram domesticadas e permanecem como espécies perenes e selvagens (Chung e Singh, 2008).

O subgênero *Soja*, por sua vez, inclui apenas duas espécies: uma muito importante economicamente, a soja *G. max* com $2n = 40$ e a espécie selvagem, *G. soja*, com $2n = 40$, ambas tetraploides e anuais. Os genomas de *G. soja* e *G. max* são similares, sendo designados por G e G1, respectivamente (Singh *et al.*, 2007b).

Tabela 1. Espécies do gênero *Glycine* Willd. e seus respectivos números cromossômicos.

Espécies do subgênero <i>Glycine</i> Willd.	2n	Espécies do subgênero <i>Glycine</i> Willd.	2n
<i>G. albicans</i> Tindale and Craven	40	<i>G. microphylla</i> (Benth.) Tindale	40
<i>G. aphyonota</i> B. Pfeil	40	<i>G. montis-douglas</i> B.E.Pfeil and Craven	40
<i>G. arenaria</i> Tindale	40	<i>G. peratosa</i> B. E. Pfeil and Tindale	40
<i>G. argyrea</i> Tindale	40	<i>G. pindanica</i> Tindale and Craven	40
<i>G. canescens</i> F. J. Hermann	40	<i>G. pullenii</i> B. Pfeil, Tindale and Craven	40
<i>G. clandestina</i> Wendl.	40	<i>G. rubiginosa</i> Tindale and B. E. Pfeil	40
<i>G. curvata</i> Tindale	40	<i>G. stenophita</i> B. Pfeil and Tindale	40
<i>G. cyrotoloba</i> Tindale	40	<i>G. syndetika</i> B.E. Pfeil and Craven	40
<i>G. falcata</i> Benth.	40	<i>G. pescadrensis</i> Hayata	80
<i>G. gracei</i> B.E. Pfeil and Craven	40	<i>G. dolichocarpa</i> Tateishi and Ohashi	80
<i>G. latrobeana</i> (Meissn.) Benth.	40	<i>G. hirticaulis</i> Tindale and Craven	40 e 80
<i>G. lactovirens</i> Tindale and Craven	40	<i>G. tabacina</i> (Labill.) Benth.	40 e 80
<i>G. latifolia</i> (Benth.) Newell and Hymowitz	40	<i>G. tomentella</i> Hayata	38, 40, 78 e 80
Espécies do subgênero	2n		
<i>Soja</i> (Moench) F.J. Hermann			
<i>G. soja</i> Sieb. & Zucc.	40		
<i>G. max</i> (L.) Merr.	40		

2.1.1 Filogenia do gênero *Glycine* Wild.

Estudos com base nas análises de DNA nuclear, DNArn/ITS 5,8S, cloroplastidial e de histonas (H3-D) permitiram examinar a filogenia das espécies no gênero *Glycine* (Doyle *et al.*, 1990; 1999; 2002; Hsing *et al.*, 2001; Rauscher *et al.*, 2004; Pfeil *et al.*, 2006). De acordo com as análises de DNArn/ITS 5,8S feitas por Hsing *et al.* (2001), o subgênero *Glycine* apresenta-se organizado em cinco clados:

I. *G. cyrtoloba* e *G. curvata* (genoma C);

II. *G. albicans* e *G. lactovirens* (genoma I), *G. tomentella* 40, 78 e 80 (genoma variável), *G. arenaria* (genoma H), *G. hirticaulis* (genoma H), *G. pindanica*, *G. tomentella* (tom38);

III. *G. tabacina*, *G. dolichocarpa* (genoma B), *G. argyrea*, *G. canescens*, *G. clandestina* (genoma A), *G. latrobeana* (genoma A);

IV. *G. latifolia*, *G. tabacina*, *G. microphylla* (genoma B), e

V. *G. falcata* (genoma F).

Nesses estudos foram observados que os representantes de *G. tomentella* são polifiléticos (Brown *et al.*, 2002), de modo semelhante aos dados de histonas

H3-D, os quais também mostraram que existem vários grupos polifiléticos em *G. tabacina* e em *G. tomentella* (Doyle *et al.*, 2002, 2004). *Glycine tomentella*, por exemplo, consiste de seis 'raças' (T1-T6), cada um com uma combinação diferente de genomas diploides (Doyle *et al.*, 2004). Diferentemente, o subgênero *Soja* representado por *G. soja* e *G. max* (ambos com genoma G) forma um grupo monofilético.

As análises cloroplastídias no subgênero *Glycine*, por sua vez revelaram a presença de três grupos representados por três genomas cloroplastídias: Genoma A (*G. albicans*, *G. aphyonota*, *G. arenaria*, *G. argyrea*, *G. clandestina*, *G. falcata*, *G. hirticaulis*, *G. lactovirens*, *G. latrobeana*, *G. peratosa*, *G. pescadrensis*, *G. pindanica*, *G. pullenii*, *G. rubiginosa* e *G. tomentella*); Genoma B (*G. latifolia*, *G. microphylla*, *G. stenophita*, *G. dolichocarpa* e *G. tabacina*) e Genoma C (*G. cyrtoloba* e *G. curvata*) (Doyle *et al.*, 1990).

2.1.2 Pools gênicos

Baseado no sucesso de hibridação entre e/ dentre as espécies, ou seja, na facilidade de troca de genes entre elas, três *pools* gênicos podem ser considerados: primário (GP-1), secundário (GP-2) e terciário (GP-3) (Harlan e deWet 1971, citado por Chung e Singh, 2008, Borém *et al.*, 2005). O *pool* gênico primário consiste de espécies biológicas, que quando cruzadas, são capazes de produzir híbridos vigorosos, que exibem pareamento cromossômico meiótico normal e total fertilidade de sementes (Singh *et al.*, 2003). A segregação dos genes na F₁ é normal e a troca de genes é geralmente fácil. Para o gênero *Glycine*, o GP-1 é composto por todos os acessos de soja (*G. max*) e pela espécie selvagem *G. soja* (Singh *et al.*, 1998; Chung e Singh, 2008). O *pool* gênico secundário inclui espécies pertencentes ao mesmo gênero, as quais podem ser cruzadas com GP-1, porém com dificuldade e com fertilidade apenas parcial em F₁. Em soja, não há espécie conhecida para GP-2 (Chung e Singh, 2008).

Por sua vez, o GP-3 inclui cruzamentos entre espécies pouco aparentadas de outros gêneros ou espécies pouco relacionadas do mesmo gênero. Em geral, a hibridação é muito difícil e a esterilidade e/ou anomalia do híbrido são comuns, com necessidade de resgate de embriões e enxertia, entre outras medidas (Harlan e de Wet, 1971). Aparentemente, apenas *G. argyrea*, *G. canescens* e *G. tomentella* estão incluídas em GP-3. A produção de híbridos F₁ dessas espécies com soja é possível

mediante cultura de embriões, embora sejam estéreis (Singh *et al.*, 1998; Chung e Singh, 2008). Somente para cruzamentos entre *G. max* e *G. tomentella* ($2n = 78$) foi desenvolvida uma metodologia para a produção de linhagens férteis derivadas, permitindo a introgressão transversal de genes úteis a partir de GP-3 (Singh e Chung, 2007a).

O *pool* gênico quaternário (GP-4) é formado por espécies que apresentam barreiras pré e pós-hibridização que inibem o desenvolvimento do embrião e geram o aborto prematuro. Raramente a letalidade das plântulas híbridas e a inviabilidade de sementes híbridas e de plantas F_1 podem ser contornadas por cruzamentos-ponte dentro do gênero *Glycine* (Singh *et al.*, 2007a), observando-se que a maioria das 26 espécies do subgênero *Glycine* estão incluídas em GP-4. Entretanto, apenas algumas espécies de *Glycine* silvestres perenes foram hibridizadas com soja, como por exemplo, *G. clandestina*, *G. latifolia* e *G. tabacina* (Chung e Singh 2008).

2.2 Importância econômica da soja

Os grãos de soja são muito usados pela agroindústria, indústria química e de alimentos. Na alimentação humana, participam da composição de vários produtos como óleo refinado, temperos, preparo de bebidas, entre outros (Agnol *et al.*, 2010). Possuem elevados teores de proteína e energia para nutrição, apresentando cerca de 17 a 18% de óleo e 35 a 37% de proteína bruta de elevado valor biológico que oferece uma combinação excelente para uma dieta equilibrada dos seres humanos. Recentemente, a soja vem se destacando como fonte alternativa de biocombustível (Costa Neto *et al.*, 2000; Chung e Singh, 2008).

O cultivo da soja é anual e compreende as regiões subtropicais entre 50° de latitudes Norte e Sul (Li e Nelson, 2001), tendo como principais produtores os Estados Unidos, o Brasil, a Argentina e a China, que juntos, respondem por praticamente 87% da produção mundial, ou seja, 263,7 milhões de toneladas e um total de 103,5 milhões de hectares de área plantada (USDA, 2012, CONAB, 2012). Por ser uma espécie empregada como cultura de verão, é muito utilizada no cultivo de rotação com outras plantas, nas regiões Sul, Sudeste e Centro-Oeste, desde a década de 90 (Silva *et al.*, 2010). No comércio mundial, a soja é classificada como a primeira entre as principais oleaginosas como a colza, o amendoim, algodão, girassol, linhaça, gergelim e açafrão (FAO STAT, 2009). No Brasil, o comércio da

soja registrou uma redução de 8,95 milhões de toneladas, passando de 75,32 milhões de toneladas colhidas na safra 2010/11 para 66,37 milhões na safra atual, totalizando 24,2 milhões de hectares de área plantada e 3.106 Kg/ha de produtividade. Em 2011, as condições climáticas adversas causadas pelo fenômeno “La Niña” foram responsáveis pelo resultado negativo da safra, sobretudo nos estados da região Sul, parte do Sudeste e no Sudoeste de Mato Grosso do Sul, além da estiagem nos estados da região Nordeste, que causou perdas em todas as culturas (CONAB, 2012).

Atualmente, há vários programas de melhoramento trabalhando no desenvolvimento da cultura da soja no Brasil. No Paraná desde 1975 a Embrapa Soja (Empresa Brasileira de Pesquisa Agropecuária, Soja, Londrina) possui o segundo maior programa de melhoramento da soja do país, atrás apenas da empresa privada Tropical Melhoramento e Genética Ltda. (TMG) assistindo os produtores do Nordeste, Sudeste e do Centro-Oeste, lançando cultivares mais adaptadas ao cultivo sob as condições tropicais com baixas latitudes. No Sudeste brasileiro, a Companhia Monsanto é uma empresa privada, trabalhando desde 1963 na cadeia produtiva de soja no Brasil, desenvolvendo cultivares com várias características desejáveis para melhor atender às necessidades da região, como é o caso da soja transgênica com resistência ao herbicida glifosato, conhecida como soja RR (*Roundup Ready*®). Destacam-se ainda a Fundação de Apoio à Pesquisa Agropecuária de Mato Grosso (Fundação MT).

2.3 Origem e domesticação da soja

A soja cultivada é uma espécie paleotetraploide com $2n = 40$ cromossomos. Observa-se que duas rodadas de duplicações do genoma (autopoliploide) ou hibridações (alopoliploide) podem ter ocorrido. Estima-se que a primeira duplicação tenha ocorrido há 45 milhões de anos e a segunda há 14,5 milhões, e que os dois eventos de duplicação foram seguidos pela adição, perdas de genes, bem como numerosos rearranjos cromossômicos (Schlueter *et al.*, 2004; Schmutz *et al.*, 2010). Além dessas transformações, observa-se que a maioria dos alelos raros presentes na soja selvagem tenha sido perdida durante sua domesticação (Schlueter *et al.*, 2007; Hyten *et al.*, 2006).

Segundo Vavilov (1992) citado por (Guo *et al.*, 2010), o gênero *Glycine* tem como centro de origem, diversidade e domesticação o leste da Ásia, que

compreende a China e regiões adjacentes como o nordeste da Rússia, Península Coreana e Japão; onde foi domesticado há mais de 5.000 anos, a partir de seu progenitor selvagem anual, *G. soja* (Figura 1). Por essa razão a soja, é considerada uma das mais antigas plantas cultivadas. Posteriormente, as espécies de *Glycine* foram levadas do Japão e da China para o continente europeu em 1712, mas, somente em 1739 foi cultivada na França e depois introduzida na Inglaterra (1790) e Itália (1840) (Bonetti *et al.*, 1977 citado por Câmara, 2011).

A soja foi introduzida no continente americano em 1804, nos Estados Unidos da América, e chegou ao Brasil pelo estado da Bahia em meados da década de 1880 (Figura 1). Em 1908, imigrantes japoneses introduziram-na em São Paulo e em 1914 foi cultivada no Rio Grande do Sul. Somente na década 1970 se iniciou o cultivo em grande escala e a partir de 1990 o país tornou-se o segundo maior produtor mundial da cultura da soja, com 22,8% da produção mundial. Atualmente, é responsável por 27,1 % do plantio da soja (CONAB, 2012).

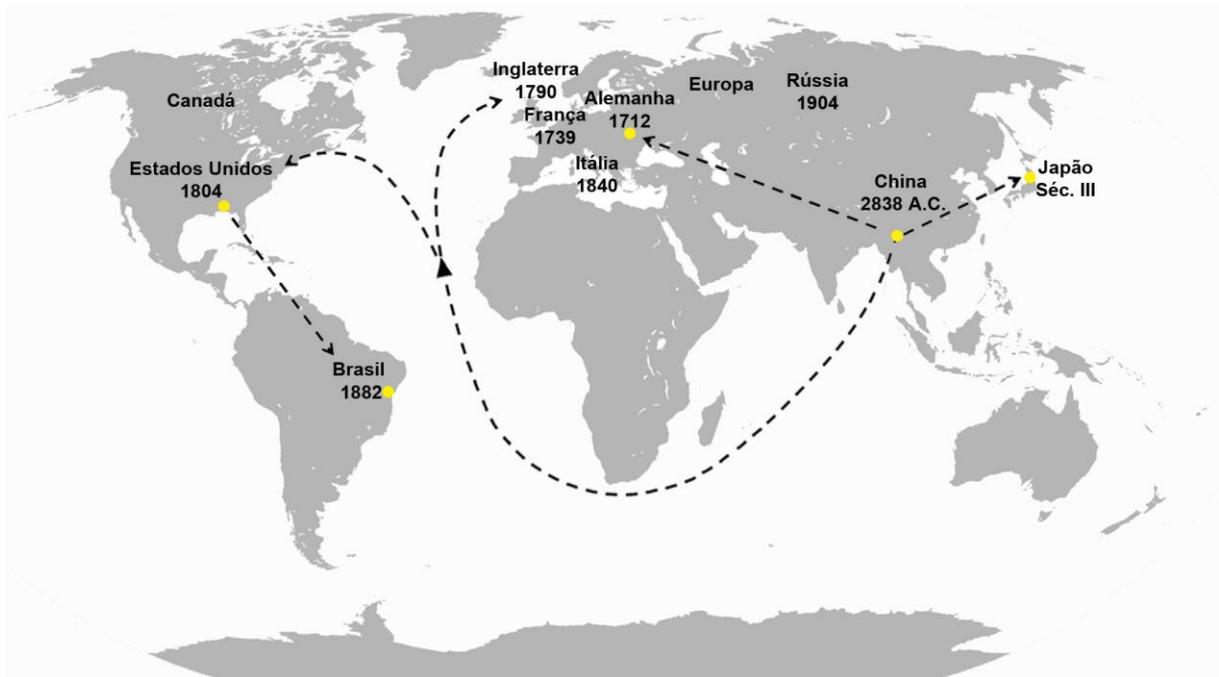


Figura 1. Origem e distribuição geográfica da soja (*Glycine max*), incluindo áreas de introdução para o cultivo. Fonte: Adaptado de Câmara (2011) cedido por Bonetti (1970).

2.4 Construção de bibliotecas BACs e sequenciamento tipo *shotgun*

Os cromossomos artificiais de bactérias (BACs) toleram inserções com cerca de 100 a 500 kb e são considerados vetores apropriados para clonagem de grandes

fragmentos de DNA na geração de uma biblioteca de DNA genômico (Tunster *et al.*, 2011). Para a construção de uma biblioteca BAC, faz-se necessário um grande número de clones BAC, o qual varia de acordo com o tamanho do genoma da espécie. Por exemplo, considerando uma média de inserções de BACs com 100 kb nos genomas de *Arabidopsis thaliana* (L.) Heynh. (125 Mb) e do arroz (450 Mb), foram necessários 7.000 e 23.000 clones BACs, respectivamente, para cobrir o genoma (Choi *et al.*, 1995; Zhang *et al.*, 1996). No entanto, espécies com grandes genomas, tais como a cevada (5000 Mb) e o trigo diploide (5600 Mb), requereram 250.000 e 280.000 clones BAC, respectivamente, para uma cobertura semelhante (Lijavetzky *et al.*, 1999; Yu *et al.*, 2000).

Em soja, a fonte de DNA para bibliotecas BACs foi a cultivar Williams 82, escolhida como o genótipo modelo pela comunidade científica para estudos genômicos (Stacey *et al.*, 2004). Três enzimas de restrição diferentes *HindIII*, *BstI*, e *EcoRI*, foram usadas para construir as três bibliotecas, GM_WBa, GM_WBb, e GM_WBc, respectivamente. O tamanho de inserção médio de GM_WBa, GM_WBb, GM_WBc, e GSS_Ba foram 150, 150, 131 e 150 kb e representam uma cobertura de 5,4; 12; 10,9 e 12 vezes o genoma, respectivamente (Ha *et al.*, 2012). Para o acesso PI468916 de *G. soja* foi construída uma biblioteca BAC, chamada GSS_Ba, com o uso da enzima *HindIII*.

O sequenciamento do tipo *shotgun* é formado a partir da fragmentação de sequências de DNA realizado mediante o uso de enzimas de restrição, ou das técnicas de sonicação ou nebulização, seguida da inserção dos fragmentos em vetores como BACs e organização de contígua (em *contigs*). Alguns desses *contigs* apresentam sequências semelhantes constituídas, em sua maioria, por DNA repetitivo, que se sobrepõem (*scaffolds*) e podem não ser sequenciados adequadamente e formar lacunas (*gaps*) nas regiões apropriadas. A presença dessas lacunas (*gaps*) representa um desafio substancial na montagem de fragmentos no sequenciamento *shotgun* impedindo o total entendimento da organização do genoma (Shoemaker *et al.*, 2008, Schmutz *et al.*, 2010).

As sequências obtidas também podem ser ancoradas em mapas genéticos de marcadores moleculares, como RFLP (*Restriction Fragment Length Polymorphism* - Polimorfismo em Comprimento de Fragmentos de Restrição), possibilitando a integração do mapa genético ao sequenciamento, permitindo assim, caracterizar e isolar os genes correspondentes às características desejáveis (Mudge *et al.*, 2004; 2005a).

2.4.1 Sequenciamento do genoma da soja

Glycine max foi a primeira leguminosa a ter seu genoma sequenciado e montado, permitindo o desenvolvimento de ferramentas essenciais para o melhoramento genético da cultura (Schmutz *et al.*, 2010). Para servir como uma fonte promissora de genes e de alelos que foram perdidos durante os gargalos da domesticação de *G. max*, a soja selvagem, *G. soja*, também teve seu genoma sequenciado (Kim *et al.*, 2010; Ha *et al.*, 2012).

O sequenciamento *shotgun* usado no mapeamento físico de *G. max* compreendeu cerca de 950 milhões de pares de bases (Mb), que ancorados representam 85% dos 1.115 Mb. A maior parte das sequências do genoma está montada em 20 pseudomoléculas e apresenta 397 sequências *scaffolds*, que se encontram ancoradas nas regiões do genoma de forma inequívoca (Arumuganathan e Earle, 1991; Schmutz *et al.*, 2010). Um conteúdo adicional de 17,7 Mb está presente em 1.148 sequências *scaffolds* não ancoradas, as quais são constituídas principalmente por DNA repetitivo e contém menos de 450 genes preditos (Schmutz *et al.*, 2010). Para montagem do mapa físico de *G. max* foram ancorados 4.628 marcadores genéticos compostos de 3.952 marcadores microssatélites do tipo SSRs (*Simple Sequence Repeats* - Repetições de Sequências Únicas) e 676 marcadores RFLP. Desses marcadores, 1.725 foram considerados como marcadores múltiplos (MHM), indicando que os marcadores foram ancorados em dois ou mais clones BAC (Ha *et al.*, 2012).

A ausência de sequências ancoradas em 15% do genoma é resultante de regiões repetitivas, onde há escassez de recombinação e de marcadores genéticos (Wu *et al.*, 2004; Schmutz *et al.*, 2010). O genoma da soja possui cerca de 57% de DNA repetitivo em regiões heterocromáticas pericentroméricas (regiões pobres em genes, onde são encontrados aproximadamente 22% de genes), com várias regiões duplicadas, apresentando uma média de 2,55 segmentos duplicados com até seis cópias por gene (Shoemaker *et al.*, 1996; Schmutz *et al.*, 2010), o que dificulta a montagem de *contigs* (Song *et al.*, 2004).

O genoma de *G. soja* foi sequenciado usando o analisador de genoma, *Genome Illumina*, resultando em 48,8 Gb de sequências, correspondendo a 52 vezes a cobertura do genoma. Compreende cerca de 915,4 Mb, ou seja, 97,65% de 937 Mb do genoma de *G. max*, com um total de 2,5 milhões de SNPs (*Single*

Nucleotide Polymorphism - Polimorfismos de um Único Nucleotídeo), previstos entre o genoma de *G. max* e *G. soja*. Cerca de 2.167.728 (86,5%) e 337.257 (13,5%) SNPs foram localizados em regiões não gênicas e gênicas, respectivamente (Kim *et al.*, 2010; Ha *et al.*, 2012).

O uso das informações de sequenciamento de culturas importantes, incluindo a soja, promete revolucionar a análise genômica, contribuindo especialmente para o melhoramento de plantas (Shultz *et al.*, 2006). Facilitará o monitoramento de características de interesse nos cruzamentos, a avaliação de alterações cromossômicas e o processo de escolha das melhores cultivares, além de permitir a identificação das afinidades cromossômicas de diferentes espécies (Wu *et al.*, 2008).

2.5 Sintenia entre soja e outras leguminosas

A sintenia e colinearidade podem ser definidas como conservação do conteúdo e da ordem gênica entre diferentes espécies (Mudge, 2005b). O mapeamento comparativo entre espécies foi estudado inicialmente em gramíneas, as quais têm apresentado uma conservação gênica relativamente alta, observando-se grandes blocos sintênicos entre diferentes espécies (Gale e Devos, 1998, Mudge *et al.*, 2005b).

Em leguminosas, análises genômico-comparativas também têm sido realizadas entre diferentes espécies (Choi *et al.*, 2004; Bertioli *et al.*, 2009; McClean *et al.*, 2010; Muchero *et al.*, 2009). A soja é considerada um excelente modelo para estudar a dinâmica evolutiva de um genoma paleopoliploide, especialmente por ter seu genoma sequenciado (Schmutz *et al.*, 2010), o qual pode ser comparado ao de outras leguminosas, mediante análise da sintenia, identificando regiões ortólogas (sequências gênicas homólogas diferenciadas por especiação e localizadas em genomas de diferentes organismos), como por exemplo, entre *G. max*, *Medicago* e/ou *Lotus* (Choi *et al.*, 2004).

Estudos entre *G. max*, *Pisum sativum* L., *Medicago truncatula* Gaertner, *Lotus japonicus* (Regel) K. Larsen e *Populus trichocarpa* (Torr. e Gray) (família Salicaceae), por exemplo, demonstraram que a sintenia é considerada elevada entre as espécies estreitamente relacionadas e diminui à medida que aumenta a distância filogenética (Bordat *et al.*, 2011). Outro trabalho relacionado ao estudo da sintenia entre *A. thaliana*, *G. max* e *M. truncatula* sugere que houve extensas adições e/ou supressões de genes após a divergência destas espécies (Schlueter *et al.*, 2008).

Entre *G. max* e *Phaseolus vulgaris* L., McClean *et al.* (2010) revelaram extensos blocos de loci compartilhados mediante estudos de sintenia com mapeamento usando RFLP e análise de EST (*Expressed Sequence Tags* - Etiquetas de Sequências Expressas). Estes autores observaram fortes evidências de que a soja teve uma duplicação há 11 milhões de anos, por meio de autoploidização, seguida por diversas quebras e rearranjos. Esses eventos foram observados na maioria dos cromossomos de soja por serem compostos de fragmentos de vários cromossomos do feijão comum (McClean *et al.*, 2010). Entre *Vigna unguiculata* (L.) Walp., *G. max* e *M. truncatula*, múltiplas regiões de sintenia também foram observadas. Os genomas de *G. max* e *M. truncatula* exibiram regiões gênicas homeólogas com cerca de 85 e 80% de SNPs mapeados em feijão-caupi, respectivamente, indicando que o genoma de feijão-caupi é mais semelhante ao de *G. max* do que *M. truncatula* (Lucas *et al.*, 2011). A integração de mapas de ligação e de sequenciamento com o mapa citogenético da *G. max* também tem sido usada para estudar a macrossintenia entre *G. max* e as espécies *G. soja* e *G. gracilis*, como será detalhado a seguir (Findley *et al.*, 2010; 2011).

2.6 Estudos citogenéticos em *Glycine* Willd.

2.6.1 Análises convencionais

Apesar da soja ser uma leguminosa muito valiosa, não é considerada uma planta modelo para estudos citogenéticos, por apresentar um grande número de cromossomos semelhantes ($2n = 40$), de pequeno tamanho (1,42-2,84 μm) (Sen e Vidyabhusan, 1960), ausência de marcas morfológicas distintivas e apenas um par de cromossomos portador de região organizadora de nucléolos (Yanagisawa *et al.*, 1991; Ohmido *et al.*, 2007; Figura 2A). Os cromossomos mitóticos de *G. soja* (Figura 2B), por sua vez, são cerca de 6-7% menores que os de *G. max* (Ahmad *et al.*, 1984) e apresentam também tamanho e morfologia semelhantes.

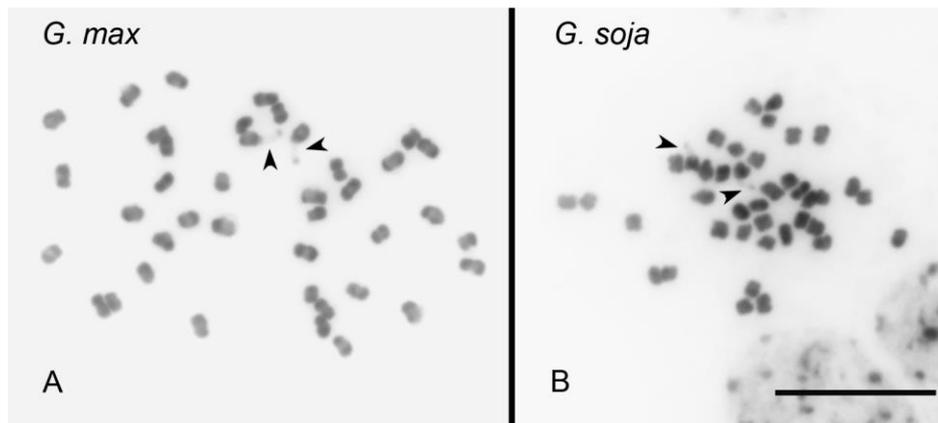


Figura 2. Metáfase convencional de *Glycine max* e *G. soja*. Cabeças desetas indicam a constrição secundária. A barra em B = 10 µm. Fonte: Ana Rafaela Oliveira

O primeiro mapa cromossômico da soja foi construído utilizando células meióticas em paquíteno, observando-se as diferenças de comprimento total dos cromossomos, relações entre os braços e proporções relativas de eucromatina e heterocromatina (Singh e Hymowitz, 1988). Os cromossomos 1 e 20 apresentaram 39,79 e 10,63 micrômetros (µm), respectivamente, e 35,8% do genoma de soja foi considerado heterocromático. Os braços curtos de seis dos 20 bivalentes (cromossomos 5, 7, 10, 18, 19 e 20) foram completamente heterocromáticos, e no cromossomo 13 foi relatada a presença do satélite associado ao nucléolo. Esta pesquisa pioneira serviu de base para identificar, desenvolver e utilizar indivíduos trissômicos ($2n = 2x + 1 = 41$) e monossômicos ($2n = 2x - 1 = 39$) de soja.

O desenvolvimento de trissômicos primários é uma ferramenta inestimável para auxiliar na localização de genes e associação de grupos de ligação molecular (*Molecular Linkage Groups* - MLG) a cromossomos específicos (Xu *et al.*, 2000; Zou *et al.*, 2003; Singh 2003). Na soja, as principais linhas de trissômicos primários ($2x + 1 = 41$; Xu *et al.*, 2000) foram usadas para associar 11 dos 20 MLGs com cromossomos citologicamente numerados (1 = D1a+q, 3 = N, 5 = A1, 8 = A2, 9 = K, 13 = F, 14 = C1, 17 = D2, 18 = G, 19 = L e 20 = I) (Cregan *et al.*, 2001; Zou *et al.*, 2003). Entretanto, aos nove MLGs não associados a cromossomos foram atribuídos números cromossômicos baseados no comprimento da distância de recombinação genética em centiMorgans cM, disponíveis no banco de dados SoyBase (<http://soybase.org/>), que não necessariamente correspondem às distâncias físicas em µm. Esta abordagem pode apresentar um resultado discrepante em relação à numeração crescente dos cromossomos de soja, por esta razão os nove

cromossomos não relacionados foram denominados pseudomoléculas (Findley *et al.*, 2011).

Uma das tentativas de montagem de um cariótipo de *G. max* através de mitose agruparam os cromossomos de células em prometáfase de acordo com os padrões de condensação (CP). Com base em CP, os cromossomos de *G. max* foram classificadas em cinco grupos. O grupo A é representado pelo par cromossômico com a razão de braço de 5,7 que apresenta uma região organizadora de nucléolo (RON) estreita localizada ao fim de seu braço curto (Figura 2) (Yanagisawa *et al.*, 1991) e apresenta uma pequena condensação entre as regiões proximais do braço curto e do braço longo. Tal par de cromossomos ocorre nas nove espécies do gênero *Glycine* estudadas por Ohmido *et al.* (2007) e é, portanto, considerado um cromossômico característico dentro do gênero *Glycine*, embora nem todos os satélites distais sejam visíveis. Cromossomos do grupo B exibem três regiões condensadas: uma no braço curto e duas no braço longo. O grupo C inclui 10 cromossomos com uma razão de braços maior do que 1,6 e uma região pericentromérica condensada. Esses cromossomos submetacêntricos são divididos em dois subgrupos: o primeiro com oito cromossomos apresentando uma região de maior condensação no braço longo enquanto que, o segundo com dois cromossomos exibindo uma extensa condensação no braço curto. O grupo E é composto por 12 cromossomos metacêntricos com razões de braços inferiores a 1,2 e uma região condensada pericentromérica: oito cromossomos possuem regiões condensadas de diferentes tamanhos e quatro de igual tamanho. O grupo D inclui 14 cromossomos com razões de braços entre a do grupo C e E, e com uma CP altamente variável. No entanto, as semelhanças observadas impediram uma completa identificação de todos os cromossomos, visto que apenas um par de cromossomos mostra uma razão elevada entre os braços de 5,7. Para os demais, as razões variavam entre 1,3 e 2,5 (Ohmido *et al.*, 2007).

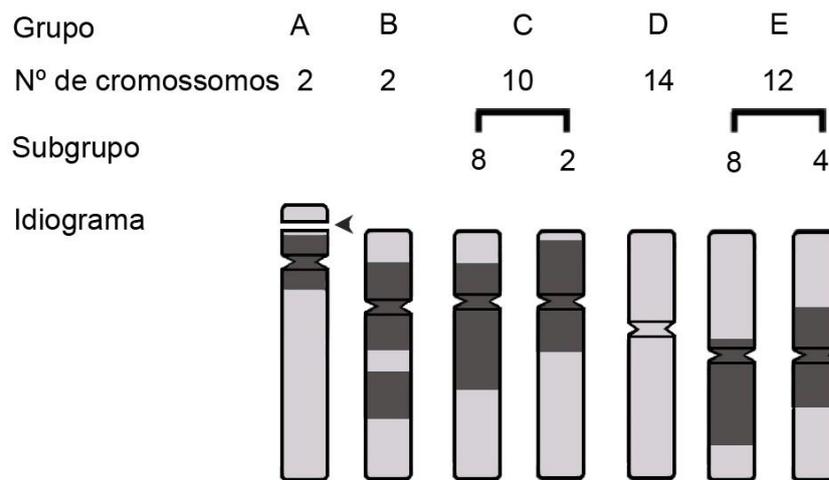


Figura 3. Idiograma dos cromossomos de soja que correspondem aos cinco grupos baseado no padrão de condensação. A cabeça de seta no grupo A indica a posição da região organizadora de nucléolo. Fonte: Adaptado de Ohmido *et al.* (2007).

Outra tentativa para auxiliar a diferenciação dos cromossomos, o emparelhamento dos homólogos e a montagem do cariógrama de *G. max* foi através de um método baseado na utilização de soluções para inibir a formação de microtúbulos, seguidos de maceração enzimática e uso da técnica de ventilação para fazer lâminas, a fim de mostrar constrições primárias e secundárias de cromossomos prometafásicos e metafásicos bem definidas. Os resultados mostraram que esta espécie possui sua natureza tetraploide com vinte pares de cromossomos, sendo seis metacêntricos (1, 2, 9, 10, 17 e 19) e 14 submetacêntricos (3, 4, 5, 6, 7, 8, 11, 12, 13, 14, 15, 16, 18 e 20), com apenas um par (13, submetacêntrico) portador da constrição secundária (Clarindo *et al.*, 2007).

2.6.2 Distribuição de DNAr 5S e 45S

Os primeiros estudos de FISH em soja usaram as sequências de DNAr 5S e 45S, por se apresentarem conservadas nas espécies, por serem de fácil hibridização e constituídas de múltiplas cópias em tandem, o que facilita a localização cromossômica das mesmas (Singh *et al.*, 2001). O gene ribossomal 5S (DNAr 5S) é uma região aparentemente menos conservada em eucariotos, quando comparada à região de DNAr 45S. Em *G. max*, a sequência de DNAr 5S localiza-se na região intersticial do braço curto do par cromossômico 19, *Gm19* (Figura 4B) (Gottlob-McHugh *et al.*, 1990; Shi *et al.*, 1996). A sonda de DNAr 45S, por sua vez, hibridiza apenas no braço curto do par cromossômico *Gm13* em posição terminal (Figura 4B),

correspondente ao cromossomo satelitado designado por Singh e Hymowitz (1988) e ao cromossomo que apresenta a região terminal marcada pela coloração com o fluorocromo base-específico cromomicina A₃ – CMA₃ (Figura 4A) (Skorupska *et al.*, 1989; Singh *et al.*, 2001, Bortoleti *et al.*, em preparação).

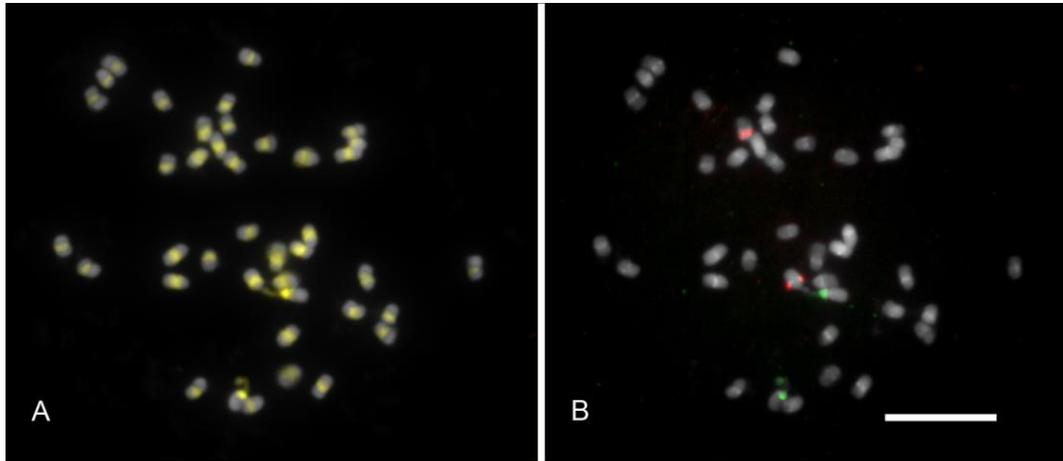


Figura 4. Metáfase de *G. max*, corada com CMA/DAPI (A) e hibridizada *in situ* com sondas de DNAr 45S (verde) e 5S (vermelho) Os cromossomos foram contracolorados com DAPI e pseudocoloridos em cinza. Barra em B = 10 μ m. Fonte: Bortoleti *et al.*, em preparação.

As espécies *G. soja* [(2n = 40), subgênero *Soja*], *G. canescens* [(2n = 40), subgênero *Glycine*] e *G. clandestina* [(2n = 40), subgênero *Glycine*] também apresentaram apenas um par de sítios de DNAr 5S, enquanto *G. tabacina* (2n = 80) e *G. tomentella* (2n = 80), ambas subgênero *Glycine*, possuíram dois pares de sítios de DNAr 5S. Para a sequência DNAr 18S-28S, foi detectado um par de sítios em *G. soja*, *G. canescens* e *G. clandestina*, bem como dois pares mais fortes e um par mais fraco em *G. tabacina* e *G. tomentella* (2n = 80) (Krishnan *et al.*, 2001). No mesmo período, Singh *et al.* (2001) examinaram a distribuição de loci de DNAr no gênero *Glycine* por FISH usando a região do espaçador interno transcrito (ITS1, 5,8S, ITS2; aproximadamente 700 nucleotídeos) como sonda. A hibridização da sonda de DNAr mostrou um par de sítios em 14 espécies com 2n = 40 cromossomos (*G. max*, *G. soja*, *G. albicans*, *G. arenaria*, *G. argyrea*, *G. canescens*, *G. clandestina*, *G. falcata*, *G. latifolia*, *G. latrobeana*, *G. microphylla*, *G. pindanica*, *G. tabacina* e *G. tomentella* [genoma D₃, genoma D (2n = 40)] e em dois acessos aneuploides [*G. tomentella* (2n = 38) e *G. tomentella* Hayata (2n = 78)]. Dois pares de sítios foram observados em: duas espécies com 2n = 40 (*G. cyrtoloba* e *G. curvata*); duas espécies com 2n = 80 (*G. tabacina*, genoma (BB₁ e AB); *G. tomentella*, genoma AD₃ e A?); e no anfiplóide (genoma AAD₃D₃) sintetizado (2n =

80) a partir do cruzamento entre *G. canescens* (genoma AA, $2n = 40$) e *G. tomentella* (genoma D₃D₃, $2n = 40$) (Singh *et al.*; 2001).

2.6.3 Hibridização *in situ* de outras sequências

O primeiro estudo de BAC-FISH (em cromossomos mitóticos) e de Fiber-FISH (em fibras estendidas de DNA) em soja foi realizado com dois clones BAC ancorados a regiões opostas do grupo de ligação E (*Gm15*) e duplicados nos grupos A2 e B2 (*Gm08* e *Gm14*, respectivamente), demonstrando a utilidade da FISH na integração dos mapas genético e citogenético em *G. max* (Pagel *et al.*, 2004).

Posteriormente, Lin *et al.* (2005), por meio do sequenciamento de DNA, FISH e Fiber-FISH, analisaram uma fração repetitiva do genoma da soja usando os BACs pericentroméricos 076J21 e 009M21 (grupo de ligação L). Ambos marcaram na região pericentromérica, porém não completamente coincidentes. Análise da sequência do BAC 076J21 mostrou que ele consiste principalmente de sequências repetitivas, incluindo fragmentos semelhantes a elementos transponíveis (Laten *et al.*, 1993, 2003; Wright e Voytas 2002). Algumas destas sequências são metiladas e não se encontram conservadas em *Vigna radiata* L. Wilczek e nas 14 espécies de soja perenes analisadas, com exceção de *G. latifolia* que hibridizou com uma das sequências repetitivas. Estes dados mostraram a presença de uma fração repetitiva altamente concentrada nas regiões pericentroméricas dos cromossomos de soja, detentora de uma rápida evolução das unidades de repetição em tandem intercaladas com retroelementos (Lin *et al.*, 2005). Posteriormente, Tek *et al.* (2010) mostraram que sequências relacionadas a retroelementos estão de fato presentes nos centrômeros funcionais de soja.

Estas duas sequências pericentroméricas (BACs 076J21 e 009M21), 19 clones BACs pertencentes ao grupo de ligação L (*Gm19*, portador do DNAr 5S) geneticamente ancorados a marcadores RFLP ou SSR (Marek *et al.*, 2001), além da sonda telomérica pAtT4 (com inserto isolado de *A. thaliana*) foram utilizados por Walling *et al.* (2006) como sondas para FISH em cromossomos paquitênicos com intuito de examinar a estrutura cromossômica, integrar mapas genético e citogenético e investigar a natureza poliploide da soja. Sete dos 19 BACs foram hibridizados como cópia única no braço longo de *Gm19*, mostrando que este braço é altamente eucromático. As marcas apresentaram homeologia com outro par cromossômico, embora com ausência de colinearidade para alguns *loci*, indicando uma natureza poliploide para a soja.

A origem aloploide da soja foi posteriormente estudada mediante estudos moleculares e citomoleculares com sondas centroméricas por Gill *et al.* (2009) e Tek *et al.* (2010). Três subfamílias de DNA satélite centromérico altamente repetitivo (CentGm-4 com 411-bp, CentGm-2 com 92bp e CentGm-1 com 91bp) marcaram mediante BAC-FISH dois subconjuntos distintos de centrômeros de soja e, em ao menos um caso, um par de homólogos. Sugere-se ainda que *G. max* tenha como um dos possíveis progenitores uma espécie extinta (Gill *et al.*, 2009; Findley *et al.*, 2010). Esses DNAs satélites centroméricos também estão presentes em *G. soja*, mas não puderam ser detectados nas 15 espécies perene de soja analisadas, nem em *Vigna radiata*, provavelmente indicando uma divergência rápida do DNA satélite centromérico dentro do gênero *Glycine* (Gill *et al.*, 2009 e Findley *et al.*, 2011).

A conclusão do sequenciamento em *G. max* tem fornecido recursos para a integração dos mapas genético, de sequenciamento e citogenético para a soja. Um coquetel de sondas tipo BAC foi hibridizado em soja e em *G. soja*, permitindo a identificação dos 20 cromossomos individuais de *G. max*, sua associação com cada grupo de ligação correspondente, assim como a identificação e caracterização de translocação cromossômica recíproca entre os cromossomos 11 e 13 em dois acessos de *G. soja* (Findley *et al.*, 2010). Esse sistema de cariotipagem permitiu posteriormente a identificação de cromossomos e de pontos de quebra aproximados presentes em outras cinco linhas de translocação de soja, resultantes de irradiação, de mutação espontânea ou de introgressões (Findley *et al.*, 2011).

Sondas de oligonucleotídeos com padrão SSR, formados por unidades de repetição de um a seis nucleotídeos, localizadas dispersas ou em tandem, também podem ser empregadas na identificação e caracterização cromossômica por FISH, associadas em geral a regiões heterocromáticas. Além disso, podem auxiliar em análises comparativas de cromossomos homeólogos de espécies distintas, permitindo indicar o quanto essas sequências estão conservadas entre espécies próximas e como elas têm evoluído (Cuadrado *et al.*, 2008). Em soja, FISH de sondas de oligonucleotídeos sintéticos (AG)₈, (AAC)₅, (AAG)₅, (ACC)₅, (CTC)₅ e (TGA)₆ revelou um padrão de distribuição disperso ao longo dos cromossomos da espécie, localizado preferencialmente nas regiões de heterocromatina pericentromérica. Essas marcas não puderam ser detectadas pela análise *in silico*, a partir de sequências disponibilizadas no banco de dados SoyBase [<http://www.soybase.org/gb2/gbrowse/gmax1.01/>], provavelmente por fazerem parte

das regiões de DNA repetitivo não sequenciadas adequadamente (Berlamino *et al.*, 2012; Bortoleti *et al.*, em preparação).

2.6.4 Uso de BAC-FISH em análises de genômica comparativa

A associação da citogenética ao uso de sequências de DNA dos genomas vegetais disponíveis tem auxiliado na identificação de cromossomos, na elucidação de possíveis rearranjos cromossômicos, bem como em análises comparativas entre táxons próximos. Tem permitido também a identificação de algumas alterações raras no genoma, demonstrando potencial para resolver relações de filogenia complexas e conflitantes (Lysak e Lexer, 2006).

Os primeiros estudos citogenéticos envolvendo comparações entre mapas genéticos e citogenéticos, mediante o uso de BAC-FISH, compreenderam as espécies de Brassicaceae (Xiong *et al.*, 2010), englobando diversas comparações cariotípicas intra e intergenéricas (Kuittinen *et al.*, 2004; Howell *et al.*, 2005; Park *et al.*, 2005; Lysak *et al.*, 2010). Uma análise de grande relevância para o grupo foi a comparação por pintura cromossômica de cinco espécies pertencentes a quatro gêneros diferentes, a fim de elucidar os mecanismos evolutivos que aconteceram na redução do número cromossômico de $n = 8$ (em *Arabidopsis lyrata* e seus ancestrais), para $n = 5$ (em *A. thaliana*). Uma conservação de grandes blocos sintênicos foi observada, entretanto, a presença de diversos rearranjos cromossômicos como fusões, inversões para- e pericêntricas e translocações recíprocas entre as espécies analisadas (Lysak *et al.*, 2006). A partir destas análises, foi proposto um cariótipo ancestral para as crucíferas (ACK) com oito cromossomos (AK1 a AK8) e 24 blocos sintênicos conservados (Lysak *et al.*, 2006; Schranz *et al.*, 2006). Posteriormente, a pintura cromossômica comparativa (CCP) foi utilizada para a reconstrução cariotípica de oito espécies com $x = 7$ cromossomos ($2n = 14, 28$), pertencentes a seis tribos de Brassicaceae, permitindo a reconstrução do cariótipo ancestral (Proto-Calepineae, PCK, $n = 7$) compartilhado por todas as espécies com $x = 7$ analisadas (Mandáková e Lysak, 2008).

Na família Fabaceae, BACs foram utilizados como sondas em FISH comparativa entre os mapas de *Phaseolus vulgaris* ($2n = 22$) com três espécies do clado Phaseoloids: *P. lunatus* L. (uma espécie filogeneticamente próxima, $2n = 22$), *P. microcarpus* Mart. (uma espécie selvagem e distante filogeneticamente, $2n = 22$) e *Vigna unguiculata* (L.) Walp. (espécie distante filogeneticamente, $2n = 22$).

Observou-se grande sintonia, com a ocorrência de raros eventos de translocação e inversão, cujo número de rearranjos aumentou com o aumento da distância genética (Fonsêca, 2010; Vasconcelos, 2011; Bonifácio *et al.*, 2012). O baixo número de rearranjos estruturais reflete uma estabilidade cariotípica na subtribo Phaseolinae, à qual pertencem os gêneros *Phaseolus*, *Vigna* e *Glycine*, confirmando sua aparente estabilidade cariotípica observada em estudos convencionais (Mercado-Ruaro e Delgado-Salinas, 1996; 1998; Fonsêca, 2010; Bonifácio *et al.*, 2012). Em soja, seis linhas de translocação cromossômica de *G. max*, *G. soja* ou *G. gracilis* foram caracterizadas citogeneticamente, permitindo a identificação dos cromossomos bem como dos pontos de quebra aproximados envolvidos nos rearranjos (Findley *et al.*, 2010; 2011). Como exemplo, pode-se citar a translocação em uma linhagem de *G. gracilis* envolvendo os cromossomos Gg01 e Gg08 (Findley *et al.*, 2011).

3. Referências bibliográficas

AGNOL, A. D.; LAZAROTTO, J. J.; HIRAKURI, M. H. **Desenvolvimento, mercado e rentabilidade da soja brasileira**. *Circular técnica da Embrapa* ISSN 2176-2864, 2010. Disponível em: http://www.cnpso.embrapa.br/download/CT74_eletronica.pdf. Acessado em: 22 de março, 2012.

AHMAD, Q. N.; BRITTEN, E. J.; BYTH, D. E. The karyotype of *Glycine soja* and its relationship to that of the soybean, *Glycine max*. **Cytologia**, v.49, p.645-658, 1984.

Angiosperm Phylogeny Group III (APG III) An update of the Angiosperm phylogeny group classification for the orders and families of flowering plants: APG III *Botanical Journal of the Linnean*, Paris, v.161, p.105–121 , 2010.

ARUMUGANATHAN, K.; EARLE, E. D. Nuclear DNA content of some important plant species. **Plant Molecular Biology. Reporter**, v.9, p.208-218, 1991.

BERTIOLI, D. J.; MORETZSOHN, M. C.; MADSEN, L. H.; SANDAL, N.; LEAL-BERTIOLI, S. C. M.; GUIMARÃES, P. M.; HOUGAARD. B. K.; FREDSLUND, J. L.; NIELSEN, A. M.; SATO, S.; TABATA, S.; CANNON, S. B.; STOUGAARD, J. An analysis of synteny of *Arachis* with *Lotus* and *Medicago* sheds new light on the structure, stability and evolution of legume genomes. **BioMed Central Genomics**, v.10, p.45, 2009.

BONETTI, L. P. Distribuição da soja no mundo. In: MIYASAKA, S.; MEDINA, J. C. **A soja no Brasil**. Campinas: Instituto de Tecnologia de Alimentos, 1977. p. 1–6.

BONIFÁCIO, E. M. Mapeamento comparativo entre o feijão comum e a fava. **(Tese de Doutorado)** – Universidade Federal Rural de Pernambuco. p.64, 2009.

BONIFÁCIO, E. M.; FONSÊCA, A.; ALMEIDA, C.; DOS SANTOS, K. G. B.; PEDROSA-HARAND, A. Comparative cytogenetic mapping between the lima bean (*Phaseolus lunatus* L.) and the common bean (*P. vulgaris* L.). **Theoretical and Applied Genetics**, v.124, p.1513–1520, 2012.

BORDAT, A.; SAVOIS, V.; NICOLAS, M.; SALSE, J.; CHAUVEAU, A.; et al: Translational genomics in legumes allowed placing *in silico* 5460 unigenes on the

pea functional map and identified candidate genes in *Pisum sativum* L. **G3, Genes, Genomes, Genetics**, v.2, p.93–103, 2011.

BORÉM, A.; VIEIRA, M. L. C. Glossário de Biotecnologia. 1. ed. Visconde do Rio Branco: **Editora Suprema**, v. 1, p.177, 2005.

BORTOLETI, K. C. A.; BENKO-ISEPPON, A. M.; SILVA, L. C. B.; OLIVEIRA, A.R.S.; ABDELNOOR, R.V.; RAMALHO, I.; MELO, N.F.; BRASILEIRO-VIDAL, A.C. Microsatellites: an important constituent or an enlargement factor of genome in legumes. *Em preparação*.

CÂMARA, G. M. S. **Introdução ao agronegócio soja**. Disponível em: <http://www.lpv.esalq.usp.br/ipv584/584%20Soja%2001%20%20-Apostila%20Texto%20%20Agronegocio%20Soja%202011.pdf>. Acessado em: 30 de julho de 2012.

CHOI, S.; CREELMAN, R. A.; MULLET, J. E.; WING, R. A. Construction and characterization of a bacterial artificial chromosome library of *Arabidopsis thaliana*. **Weeds World**, v.2, p.17–20, 1995.

CHOI, H. K.; MUN, J. H.; KIM, D. J.; ZHU, H. Y.; BAEK, J. M.; MUDGE, J.; ROE, B.; ELLIS, N.; DOYLE, J.; KISS, G. B.; YOUNG, N. D.; COOK, D. R. Estimating genome conservation between crop and model legume species. **Proceedings of the National Academy of Science of the USA**, v.101, p.15289–15294, 2004.

CHUNG, G.; SINGH, R. J. Broadening the Genetic Base of Soybean: A Multidisciplinary Approach. **Critical Reviews in Plant Sciences**, v.27, p.295–341, 2008.

CLARINDO, W. R.; CARVALHO, C. R.; ALVES, B. M. G. Mitotic evidence for the tetraploid nature of *Glycine max* provided by high quality karyograms. **Plant Systematic Evolution**, v.265, p.101-107, 2007.

COMPANHIA NACIONAL DE ABASTECIMENTO (CONAB). 2012. **Acompanhamento da safra brasileira: grãos, quarto levantamento, julho 2011 /2012** Companhia Nacional de Abastecimento. *Conab*, Brasília, p.29. Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/4d3fd0df22f81f0c48741a257174facd..pdf>. Acessado em: 22 de julho de 2012.

COMPANHIA MONSANTO. Disponível em: <http://www.monsanto.com.br/>. Acessado em: 30 de setembro de 2012.

COSTA NETO, P. R.; ROSSI, L. F. S.; ZAGONEL, G. F. E.; RAMOS, L. P. Produção de biocombustível alternativo ao óleo diesel através da transesterificação de óleo de soja usado em fritura. **Química Nova**, v.23, p.4, 2000.

CREGAN, P. B.; KOLLIPARA, K. P.; XU, S. J.; SINGH, R. J.; HYMOWITZ, T. Primary trisomics and SSR markers as tools to associate chromosomes with linkage groups in soybean. **Crop Science**, v.41, p.1262–1267, 2001.

CUADRADO, A.; CARDOSO, M.; JOUVE N. Physical organisation of simple sequence repeats (SSRs) in Triticeae: structural, functional and evolutionary implications. **Cytogenetic and Genome Research**, v.120, p.210–219, 2008.

DOYLE, J. J.; LUCKOW, M. A. The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. **Plant Physiology**, v.131, p.900–910, 2003.

FINDLEY, S. D.; CANNON, S.; VARALA, K.; DU, J.; MA, J.; HUDSON, M. E.; BIRCHLER, J.; STACEY, G. A FISH Karyotyping System for Soybean. **Genetics**, v.185, p.3727–3744, 2010.

FINDLEY, S. D.; PAPPAS, A. L.; CUI, Y.; BIRCHLER, J. A.; PALMER, R. G.; STACEY, G. Fluorescence *in situ* hybridization–based karyotyping of soybean translocation lines. **G3, Genes, Genomes, Genetics**, p.117–129, 2011.

FONSÊCA, A.; FERREIRA, J.; SANTOS, T. R. B.; MOSIOLEK, M.; BELLUCCI, E.; KAMI, J.; GEPTS, P.; GEFFROY, V.; SCHWEIZER, D.; SANTOS, K. G. B. *et al.* Cytogenetic map of common bean (*Phaseolus vulgaris* L.). **Chromosome Research**, v.18, p.487–502, 2010.

FONSÊCA, A. F. A. Evolução cariotípica no gênero *Phaseolus* L.: mapeamento comparativo entre *P. microcarpus* Mart. e o feijão comum (*P. vulgaris* L.). **(Dissertação de Mestrado)** – Universidade Federal de Pernambuco, p.84, 2010.

FUNDAÇÃO MATO GROSSO. Disponível em: <http://www.fundacaomt.com.br/> Acessado em: 30 de setembro de 2012.

GALE M.; DEVOS K. Comparative genetics in the grasses. **Proceedings of the National Academy of Science of the USA**, v.95, p.1971–1974, 1998.

GILL, N.; FINDLEY, S.; WALLING, J. G.; HANS, C.; MA, J.; DOYLE, J.; STACEY, G.; JACKSON, S. A. Molecular and chromosomal evidence for allopolyploidy in soybean, *Glycine max* (L.) Merr. **Plant Physiology**, v.151, p.1167–1174, 2009.

GOTTLOB-MCHUGH, S. G.; LEVESQUE, M.; MACKENZIE, K.; OLSON, M.; YAROSH, O.; JOHNSON, D. A. Organization of the 5S rRNA genes in the soybean *Glycine max* (L.) Merrill and conservation of the 5S rDNA repeat structure in higher plants. **Genome**, v.33, p.486–494, 1990.

GRAHAM, P. H.; VANCE C. P. Legumes: Importance and Constraints to Greater Use. **Plant Physiology**, v.131, p.872–877, 2003.

GUERRA, M. 2004. **Hibridização *in situ*: princípios básicos**, In: GUERRA, M. (Org.). *FISH – Conceitos e aplicações na citogenética*. Ribeirão Preto: Sociedade Brasileira de Genética, p.1–32.

GUO, J.; WANG, Y.; SONG, C.; ZHOU, J.; QIU, L.; HUANG, H.; WANG, Y. A single origin and moderate bottleneck during domestication of soybean (*Glycine max*): implications from microsatellites and nucleotide sequences. **Annals of Botany**, v.106, p.505–514, 2010.

HA, J.; ABERNATHY, B. NELSON, W.; GRANT, D.; WU, X.; NGUYEN, H. T.; STACEY, G.; YU, Y.; WING, R. A.; SHOEMAKER, R. C.; JACKSON S. A. Integration of the Draft Sequence and Physical Map as a Framework for Genomic Research in Soybean (*Glycine max* (L.) Merr.) and Wild Soybean (*Glycine soja* Sieb. and Zucc.). **G3, Genes, Genomes, Genetics**, p. 221–329, 2012.

HARLAN, J. R.; DE WET, J. M. J. Toward a rational classification of cultivated plants. **Taxon**, v.20, p.509–517, 1971.

HOWELL, E. C.; ARMSTRONG, S. J.; BARKER, G. C.; JONES, G. H.; KING, G. J.; RYDER, C. D.; KEARSEY, M. J. Physical organization of the major duplication on *Brassica oleracea* chromosome O6 revealed through fluorescence *in situ*

hybridization with *Arabidopsis* and *Brassica* BAC probes. **Genome**, v.48, p.1093–1103, 2005.

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 Science of the USA**, v.103, n.45, p.16666–16671, 2006.

KIM, M. Y.; LEE, S.; VAN, K.; KIM, T.-H.; JEONG, S.-C.; CHOI, I.-Y.; KIM, D.-S.; LEE, Y.-S.; PARK, D.; MA, J.; Kim, W.-Y.; Kim, B.-C.; Park, S.; Lee, K.-A.; Kim, D. H.; KIM, KIL H.; SHIN, JIN H.; JANG, Y. E.; KIM, K. D.; LIU, W. X.; CHAISAN, T.; KANG, Y. J.; LEE, Y.-H.; KIM, K.-H.; MOON, J.-K.; SCHMUTZ, J.; JACKSON, S. A.; BHAK, J.; LEE, S.-H. Whole-genome sequencing and intensive analysis of the undomesticated soybean (*Glycine soja* Sieb. and Zucc.) genome. **Proceedings of the National Academy of Science of the USA**, v.107, p.22032–22037, 2010.

KRISHNAN, P.; SAPRA, V. T.; SOLIMAN, K. M.; ZIPF, A. FISH mapping of the 5S and 18S-28S rDNA loci in different species of *Glycine*. **Journal of Heredity**, v.92, p.295-300, 2001.

KUITTINEN, H.; DE HAAN, A. A.; VOGEL, C.; OIKARINEN, S.; LEPPALA, J.; KOCH, M.; MITCHELL-OLDS, T.; LANGLEY, C. H.; SAVOLAINEN, O. Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *A. thaliana*. **Genetics**, v.168, p.1575–1584, 2004.

LATEN, H. M.; MORRIS, R. O. *SIRE-1*, a long interspersed repetitive DNA element from soybean with sequence similarity to retrotransposons: initial characterization and partial sequence. **Gene**, v.134, p.153–159, 1993.

LATEN, H. M.; HAVECKER, E. R.; FARMER, L. M.; VOYTAS, D. F. *SIRE1*, an endogenous retrovirus family from *Glycine max*, is highly homogeneous and evolutionarily young. **Molecular Biology Evolution**, v.20, p.1222–1230, 2003.

LEWIS, G.; SCHRIRE, B.; MACKINDER, B.; LOCK, M. Legumes of the world. **Royal Botanical Gardens**, Kew, Richmond, Surrey, UK, 2005.

LI, Z.; NELSON, R. L. Genetic Diversity among Soybean Accessions from Three Countries Measured by RAPDs. **Crop Science**, v.41, p.1337–1347, 2001.

LIJAVETZKY, D.; MUZZI, G.; WICKER, T.; KELLER, B.; WING, R.; DUBCOVSKY, J. Construction and characterization of a bacterial artificial chromosome (BAC) library for the A genome of wheat. **Genome**, v.42, p.1176–1182, 1999.

LIN, J. Y.; JACOBUS, B. H.; SANMIGUEL, P.; WALLING, J. G.; YUAN, Y.; SHOEMAKER, R. C.; YOUNG, N. D.; JACKSON, S. A. Pericentromeric regions of soybean (*Glycine max* L. Merr.) chromosomes consist of retroelements and tandemly repeated DNA and are structurally and evolutionarily labile. **Genetics**, v.170, p.1221–1230, 2005.

LYSAK, M. A.; LEXER, C. Towards the era of comparative evolutionary genomics in Brassicaceae. **Plant Systematic Evolution**, v.259, p.175–198, 2006.

LYSAK, M. A.; BERR, A.; PECINKA, A.; SCHMIDT, R. MCBREEN, K.; SCHUBERT I Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related Brassicaceae species. **Proceedings of the National Academy of Science of the USA**, v.103, n.13, p.5224–5229, 2006.

LYSAK, M. A.; MANDÁKOVÁ, T.; LACOMBE, E. Reciprocal and multi-species chromosome BAC painting in Crucifers (Brassicaceae). **Cytogenetic Genome Research**, v.129, p.184–189, 2010.

LUCAS, M. R.; DIOP, N.-N.; WANAMAKER, S.; EHLERS, J. D.; ROBERTS, P. A.; CLOSE, TIMOTHY J. Cowpea–soybean synteny clarified through an improved Genetic Map. **The Plant Genome**, v. 4, n. 3, p.218–225, 2011.

MANDÁKOVÁ, T.; LYSK, M. A. Chromosomal phylogeny and karyotype evolution in $x=7$ Crucifer species (Brassicaceae). **The Plant Cell**, v.20, p.2559–2570, 2008.

MAREK, L.F.; MUDGE, J.; DARNIELLE, L.; GRANT, D.; HANSON, N.; PAZ, M.; HUIHUANG, Y.; DENNY, R.; LARSON, K.; FOSTER-HARTNETT, D.; COOPER, A.; DANESH, D.; LARSEN, D.; SCHMIDT, T.; STAGGS, R.; CROW, J. A.; RETZEL, E.; YOUNG, N. D.; SHOEMAKER, R. C. Soybean genomic survey: BAC-end sequences near RFLP and SSR markers. **Genome**, v.44, p.572–581, 2001.

MCCLEAN, P. E.; MAMIDI, S.; MCCONNELL, M.; CHIKARA, S.; LEE, R. Synteny mapping between common bean and soybean reveals extensive blocks of shared loci. **BioMed Central Genomics**, v.11, p.184, 2010.

MERCADO-RUARO, P.; DELGADO-SALINAS, A. Karyological studies in several Mexican species of *Phaseolus* L. and *Vigna savi* (Phaseolinae, Fabaceae). In: Pickergill B, Lock JM (eds) Advances in legumes systematics 8: legumes of economic importance. Kew: **Royal Botanic Gardens**, p.83–87, 1996.

MERCADO-RUARO, P.; DELGADO-SALINAS, A. Karyotypic studies on species of *Phaseolus* (Fabaceae: Phaseolinae). **America Journal Botanic**, v.85, p.1–9, 1998.

MUCHERO, W.; DIOP, N. N.; BHAT, P. R.; FENTON, R. D.; WANAMAKERB, S.; POTTORFFB, M.; HEARNEC, S.; CISED, N.; FATOKUNC, C.; EHLERSB, J. D.; ROBERTS, P. A.; CLOSEB, T. J. A consensus genetic map of cowpea [*Vigna unguiculata* (L.) Walp.] and synteny based on EST-derived SNPs. **Proceedings of the National Academy of Science of the USA**, v.106, n.43, p.18159–18164, 2009.

MUDGE, J.; HUIHUANG, Y.; DENNY, R. L.; HOWE, D. K.; DANESH, D.; MAREK, L. F.; RETZEL, E.; SHOEMAKER, R. C.; YOUNG, N. D. Soybean bacterial artificial chromosome contigs anchored with RFLPs: insights into genome duplication and gene clustering. **Genome**, v.47, p.361–372, 2004.

MUDGE, J.; HUIHUANG, Y.; DENNY, R.L.; HOWE, D.K.; DANESH, D.; MAREK, L.F.; RETZEL, E.; SHOEMAKER, R.C.; YOUNG, N.D. Soybean bacterial artificial chromosome *contigs* anchored with RFLPs: insights into genome duplication and gene clustering. **BMC Plant Biology**, v.5, p.15, 2005a.

MUDGE, J.; CANNON, S.B.; KALO, P.; OLDROYD, G. E. D.; ROE, B. A.; TOWN, C. D.; YOUNG, N. D. Highly syntenic regions in the genomes of soybean, *Medicago truncatula*, and *Arabidopsis thaliana*. **BMC Plant Biology**, v.5, p.15, 2005b.

OHMIDO, N.; SATO, S.; TABATA, S.; FUKUI, K. Chromosome maps of legumes. **Chromosome Research**, v.15, p.97–103, 2007.

ORGANIZATION FOOD AND AGRICULTURE (FAO STAT). 2009. **Produção Mundial de Soja: Área Colhida, Produtividade, e Projeções de Longo Prazo.**

Disponível em: <http://faostat.fao.org/site/339/default.aspx>. Acessado em: 22 de maio de 2012.

PAGEL, J.; WALLING, J. G.; YOUNG, N. D.; SHOEMAKER, R. C.; JACKSON, S. A. Segmental duplications within the *Glycine max* genome revealed by fluorescence *in situ* hybridization of bacterial artificial chromosomes. **Genome**, v.47, p.764–768, 2004

PARK, J. Y.; KOO, D. H.; HONG, C. P.; LEE, S. J.; JEON, J. W.; LEE, S. H.; YUN, P. Y.; PARK, B. S.; KIM, H. R.; BANG, J. W.; PLAHA, P.; BANCROFT, I.; LIM, Y. P. Physical mapping and microsynteny of *Brassica rapa* ssp. *Pekinensis* genome corresponding to a 222 kbp gene-rich region of *Arabidopsis* chromosome 4 and partially duplicated on chromosome 5. **Molecular Genetics Genomics**, v.274, p.579–588, 2005.

SCHLUETER, J. A., P. DIXON, C. GRANGER, D. GRANT, L. CLARK et al. Mining EST databases to resolve evolutionary events in major crop species. **Genome**, v.47, p.868-876, 2004.

SCHLUETER, J. A.; LIN, J. Y.; SCHLUETER, S. D.; VASYLENKO-SANDERS, I. F.; DESHPANDE, S.; YI, J.; BLENESS, M. O.; ROE, B. A.; NELSON, R. T.; SCHEFFLER, B. E.; JACKSON, S. A.; SHOEMAKER, R. C. Gene duplication and paleopolyploidy in soybean and the implications for whole genome sequencing. **BioMed Central Genomics**, v.8, p.330, 2007.

SCHLUETER, J. A.; SCHEFFLER, B. E.; SCOTT JACKSON, S. A.; SHOEMAKER, R. C. Fractionation of Synteny in a Genomic Region Containing Tandemly Duplicated Genes across *Glycine max*, *Medicago truncatula* and *Arabidopsis thaliana*. **Journal of Heredity**, v.99, n.4, p.390–395, 2008.

SCHMUTZ, J.; CANNON, S. B.; SCHLUETER, 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.; SAKURAI, T.; UMEZAWA, T.; BHATTACHARYYA, M. K.; SANDHU, D.; VALLIYODAN, B.; LINDQUIST, E.; PETO, M.; GRANT, D.; SHU, S.; GOODSTEIN, D.; BARRY, K.; FUTRELL-GRIGGS, M.; ABERNATHY, B.; JIANCHANG DU, J.; TIAN, Z.; ZHU, L.; GILL, N.; JOSHI, T.; LIBAULT, M.; SETHURAMAN, A.; ZHANG, X.; SHINOZAKI, K.; NGUYEN, 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. **Theoretical and Applied Genetics**, v.114, p.1081–1090, 2010.

SCHRANZ, M. E.; LYSAK, M. A.; MITCHELL-OLDS, T. The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. **Trends Plant Science**, v.11, p.535–542, 2006.

SEN, N. K.; VIDYABHUSAN, R. V. Tetraploid soybeans. **Euphytica**, v.9, p.317–322, 1960.

SHI, L.; ZHU, T.; KEIM, P. Ribosomal RNA genes in soybean and common bean: chromosomal organization, expression, and evolution. **Theoretical and Applied Genetics**, v.93, p.136–141, 1996.

SHOEMAKER, R. C.; POLZIN, J.; LABATE, J.; SPECHT, J.; BRUMMER, E. C.; OLSON, T.; YOUNG, N.; CONCIBIDO, V.; WILCOX, J.; TAMULONIS, J. P.; KOCHERT, G.; BOERMAF, H. R. Genome duplication in soybean (*Glycine* subgenus *Soja*). **Genetics**, v.144, p.329–338, 1996.

SHOEMAKER, R. C.; GRANT, D.; OLSON, T.; WARREN, W. C.; WING, R.; YU, Y.; KIM, H.; CREGAN, P.; JOSEPH, B.; FUTRELL-GRIGGS, M.; NELSON, W.; DAVITO, J.; WALKER, J.; WALLIS, J.; KREMITSKI, C.; SCHEER, D.; CLIFTON, S. W.; GRAVES, T.; NGUYEN, WU, X.; LUO, M.; DVORAK, J.; NELSON, R.; CANNON, S.; TOMKINS, J.; SCHMUTZ, J.; STACEY, G.; JACKSON, S. Microsatellite discovery from BAC end sequences and genetic mapping to anchor the soybean physical and genetic maps. **Genome**, v.51, p.294–302, 2008.

SHULTZ, J. L.; KURUNAM, D. J.; SHOPINSKI, K. L.; IQBAL, M. J.; KAZI, S.; ZOBRIK, K.; BASHIR, R.; YAEGASHI, S.; LAVU, N.; AFZAL, A.; YESUDAS, C. R.; KASSEM, M. A.; WU, C.; ZHANG, H. B.; TOWN, C. D.; MEKSEM, K.; LIGHTFOOT, D. A. The soybean genome database (SoyGD): a browser for display of duplicated, polyploid regions and sequence tagged sites on the integrated physical and genetic maps of *Glycine max*. **Nucleic Acids Research**, v.34, p.758–765, 2006.

SILVA, J. B.; LAZARINI, E.; SÁ, M. E.; VIEIRA, R. D. Efeito da Irrigação sobre o Potencial Fisiológico de Sementes em Semeadura de Inverno. **Revista Brasileira de Sementes**, Londrina, v.32, n.2, p.73–82, 2010.

SINGH, R. J.; HYMOWITZ, T. The genomic relationship between *Glycine max* (L.) Merr. and *G. soja* Sieb. and Zucc. as revealed by pachytene chromosome analysis. **Theoretical and Applied Genetics**, v.76, p.705–711, 1988.

SINGH, R. J.; KOLLIPARA, K. P.; HYMOWITZ, T. The genomes of *Glycine canescens* F.J. Herm. and *G. tomentella* Hayata of Western Australia and their phylogenetic relationships in the genus *Glycine* Willd. **Genome**, v.41, p.669–679, 1998.

SINGH, R. J.; KIM, H. H.; HYMOWITZ, T. Distribution of rDNA loci in the genus *Glycine* Willd. **Theoretical and Applied Genetics**, v.103, p.212–218, 2001.

SINGH, R. J. **Plant Cytogenetics**, 2 ed. *CRC Press, Inc.*, Boca Raton, FL, 2003.

SINGH, R. J.; CHUNG, G. H.; NELSON, R. L. Landmark research in legumes. **Genome**, v.5, p.525–537, 2007a.

SINGH, R. J.; NELSON, R.; L.; CHUNG, G. H. Soybean (*Glycine max* (L.) Mer.). *In: Genetic Resources, Chromosome Engineering, and Crop Improvement Oilseed Crops.*, v.4, p.13–50, 2007b.

STACEY, G.; VODKIN, L.; PARROTT, W. A.; Shoemaker, R. C. National Science foundation-sponsored workshop report. Draft plan for soybean genomics. **Plant Physiology**, v.135, p.59–70, 2004.

SKORUPSKA, H.; ALBERTSEN, M. C.; LANGHOLZ, K. D.; ALMER, R. G. Detection of ribosomal RNA genes in soybean, *Glycine max* (L.) Merr.; by *in situ* hybridization. **Genome**, v.32, p.1091–1095, 1989.

SONG, Q. J.; MAREK, L. F.; SHOEMAKER, R. C.; LARK, K. G.; CONCIBIO, V. C.; DELANNAY, X.; SPECHT, J. E.; CREGAN, P. B.. A new integrated genetic linkage map of the soybean. **Theoretical and Applied Genetics**, v.109, p.122-128, 2004.

STACEY, G., VODKIN, L.; PARROTT, W. A.; SHOEMAKER, R. C. National Science Foundation-sponsored workshop report. Draft plan for soybean genomics. **Plant Physiology**, v.135, p. 59–70, 2004

TUNSTER, S. J.; PETTE, M. V. D.; JOHN, R. M. BACs as Tools for the Study of Genomic Imprinting. **Journal of Biomedicine and Biotechnology**, v.2011, p.10, 2011.

TEK, A. L.; KASHIHARA, K.; MURATA, M.; NAGAKI, K. Functional centromeres in soybean include two distinct tandem repeats and a retrotransposons. **Chromosome Research**, p.1–11, 2010.

TROPICAL MELHORAMENTO E GENÉTICA LTDA. (TMG) Disponível em: <http://www.tmg.agr.br/> Acessado em: 30 de setembro de 2012.

UNITED STATES DEPARTMENT OF AGRICULTURE (USDA) Disponível em www.usda.gov. Acessado em: 30 de setembro de 2012.

VASCONCELOS, E. V. Mapeamento Citogenético de *Vigna unguiculata* (L.) Walp. Mediante Hibridização *in situ* de Sequências de DNA de *Phaseolus vulgaris* L. **(Monografia de Graduação)** – Universidade Federal de Pernambuco, p.55, 2011.

VAVILOV, N. I. Origin and geography of cultivated plants. Cambridge: Cambridge University Press, 1992.

WALLING, J.; SHOEMAKER, G. R.; YOUNG, N.; MUDGE, J.; JACKSON, ; S. Chromosome-level homeology in paleopolyploid soybean (*Glycine max*) revealed through integration of genetic and chromosome maps. **Genetics**, v.172, p.1893–1900, 2006.

WRIGHT, D. A.; VOYTAS, D. F. *Athila 4* of *Arabidopsis* and *Calypso* of soybean define a lineage of endogenous plant retroviruses. **Genome Research**, v.12, p.122–131, 2002.

WU, C.; SUN, S.; NIMMAKAYALA, P.; SANTOS, F. A.; MEKSEM, K.; SPRINGMAN, R.; DING, K.; LIGHTFOOT, D. A.; ZHANG, H. B. A BAC- and BIBAC-based physical map of the soybean genome. **Genome Research**, v.14, p.319–326, 2004.

WU, X.; ZHONG, G.; FINDLEY, S. D.; CREGAN, P.; STACEY, G.; NGUYEN, H. T. Genetic marker anchoring by six-dimensional pools for development of a soybean physical map. **BioMed Central Genomics**, v.9, p.28, 2008.

XIONG, Z.; KIM, J. S.; PIRES, J. C. Integration of Genetic, Physical, and Cytogenetic Maps for *Brassica rapa* Chromosome A7. **Cytogenet Genome Research**, v.29, p.190–198, 2010.

XU, S. J.; SINGH, R. J.; HYMOWITZ, T. Monosomics in soybean: origin, identification, cytology, and breeding behavior. **Crop Science**, v.40, p.985–989, 2000.

YANAGISAWA, T.; TANO, S.; FUKUI, K.; HARADA, K. Marker chromosomes commonly observed in the genus *Glycine*. **Theoretical and Applied Genetics**, v.81, p.606-612, 1991.

YU, Y.; TOMKINS, J. P.; WAUGH, R.; FRISCH, D. A.; KUDRNA, D.; KLEINHOF, A.; BRUEGGEMAN, R. S.; MUEHLBAUER, G. J.; WISE, R. P.; WING, R. A. A bacterial artificial chromosome library for barley (*Hordeum vulgare* L.) and the identification of clones containing putative resistance genes. **Theoretical and Applied Genetics**, v.101, p.1093–1099, 2000.

ZHANG, H-B.; CHOI, S.; WOO, S-S.; LI, Z.; WING, R. A. Construction and characterization of two rice bacterial artificial chromosome libraries from the parents of a permanent recombinant inbred mapping population. **Molecular Breed**, v.2, p.11–24, 1996.

ZHUKOV, V. A.; SHTARK, O. Y.; BORISOV, A. Y.; TIKHONOVICH, I. A. Molecular Genetic Mechanisms Used by Legumes to Control Early Stages of Mutually Beneficial (Mutualistic) Symbiosis. **Russian Journal of Genetics**, v.45, n.11, p.1279–1288, 2009.

ZOU, J. J.; SINGH, R. J.; LEE, J.; XU, S. J.; CREGAN, P. B.; HYMOWITZ, T. Assignment of molecular linkage groups to soybean chromosomes by primary trisomics. **Theoretical and Applied Genetics**, v.107, p.745–750, 2003.

CAPÍTULO

Mapeamento citogenético comparativo em espécies de *Glycine* Willd.

Mapeamento citogenético comparativo em espécies de *Glycine* Willd.

Ana Rafaela da Silva, Oliveira^{a,b}, Emanuelle Varão Vasconcelos^b, Ana Maria Benko-Iseppon^b, Ricardo Vilela Abdelnoor^c, Marcelo Fernandes de Oliveira^c, Ana Christina Brasileiro-Vidal^{a,b}

^aDepartamento de Agronomia, Universidade Federal Rural de Pernambuco, Rua Dom Manoel de Medeiros s/nº, Dois Irmãos, CEP: 52171-900, Recife, PE, Brasil.

^bDepartamento de Genética, Universidade Federal de Pernambuco, Av. da Engenharia, s/nº, CEP 50740-600, Recife, PE, Brasil.

^cEmbrapa Soja, Caixa Postal 231, CEP 86001-970, Londrina, PR, Brasil.

^dAutor para correspondência: brasileirovidal.ac@gmail.com

Running title: Citogenética comparativa em *Glycine*

Resumo

Um estudo genômico comparativo foi realizado mediante a localização *in situ* de oito cromossomos artificiais de bactérias (BACs) pertencentes aos grupos de ligação G (pesudomolécula 18, quatro BACs) e J (pesudomolécula 16, quatro BACs) de *G. max* ($2n = 40$) em *G. soja* ($2n = 40$), *G. syndetika* ($2n = 40$), *G. tabacina* ($2n = 80$) e *G. tomentella* ($2n = 78$). As análises citogenéticas do BAC 173N15 em *G. max* e *G. soja* marcaram regiões de DNA repetitivo em todos os cromossomos de *G. max* e apenas 10 pares de cromossomos de *G. soja*, demonstrando que estas regiões têm se diferenciado ao longo da evolução. Em *G. soja*, *G. syndetika*, *G. tabacina* e *G. tomentella* outros sete BACs possibilitaram identificar os cromossomos homeólogos (exceto para *G. tomentella* com os BACs de *Gm18*) e algumas quebras da sintenia, decorrente de possíveis duplicações e translocações. Também foi observado que o *Gm18* apresenta orientação invertida ao mapa de sequenciamento. Para os BACs 176B09 (*Gm18*) em *G. soja*, 93L17 (*Gm18*) em *G. syndetika* e *G. tomentella* e 102N16 (*Gm16*) em *G. tomentella* um dos pares extras foi associado a um par de sítios de DNAr 45S decorrentes de possíveis divergências dos espaçadores intergênicos em sítios de DNAr. Os resultados sugerem uma quebra de sintenia para *Gm16* em *G. syndetika*, *G. tomentella* e para *Gm18* em todas as espécies analisadas.

Palavras-chave: BAC, FISH, *Glycine max*, mapa cromossômico, soja

Introdução

O gênero *Glycine* Willd. pertence à família Fabaceae (clado Phaseoloids), compreendendo aproximadamente 28 espécies, distribuídas em dois subgêneros: *Glycine* e *Soja*. O primeiro abrange 26 espécies (perenes e selvagens), enquanto o subgênero *Soja* inclui a soja [*Glycine max* (L) Merr.], cultivar de grande importância econômica, e a espécie selvagem *G. soja*, ambas de hábito anual [Chung e Singh, 2008].

Os grãos da soja são utilizados para muitos fins, porém, a maior importância da soja está relacionada à produção de óleo para a indústria e a alimentação humana. Além disso, pode ser usada para a produção de farelo para ração animal, na recuperação de solos (fixação de nitrogênio atmosférico no solo), além de servir como matéria-prima para a produção de biocombustível [Costa Neto et al., 2000; Graham e Vance, 2003; Chung e Singh, 2008].

Diversos grupos de pesquisa têm trabalhado para a obtenção de cultivares de soja mais adaptadas a estresses bióticos e abióticos e com características agrônômicas interessantes, especialmente associando-as a dados de genômica, transcriptômica e proteômica [Benko-Iseppon et al., 2012], bem como para a construção de mapas genéticos saturados [Cregan et al., 1999; Song et al., 2004; Choi et al., 2007]. O esforço crescente dos pesquisadores resultou no sequenciamento de seu DNA genômico, tratando-se da primeira leguminosa a ser sequenciada, com cerca de 85% do genoma linear sequenciado por *shotgun*, resultando em 20 pseudomoléculas cromossômicas, contendo 397 *scaffolds* com posições ordenadas [Schmutz et al., 2010]. Posteriormente, o genoma da soja selvagem (*G. soja*) também foi sequenciado utilizando duas plataformas de sequenciamento (MPS): Illumina Genome Analyzer (Illumina-GA) e Roche Genoma Sequencer FLX (GS-FLX) [Ha et al., 2012].

Estudos cariotípicos prévios em *Glycine* revelaram que o gênero é formado por espécies poliploides com números cromossômicos variando entre $2n = 38$ e 80, com morfologia semelhante (metacêntrica e submetacêntrica) e de pequeno tamanho [Chung e Singh, 2008]. Posteriormente, a hibridização *in situ* fluorescente (FISH) de sequências únicas e repetitivas (centroméricas, por exemplo) selecionadas *in silico* (<http://soybase.org/gb2/gbrowse/gmax1.01/>) permitiu a montagem do cariótipo da espécie; a associação dos cromossomos às pseudomoléculas sequenciadas; análises comparativas com cromossomos de

outras espécies, como *G. soja* Sieb. e Zucc. e *G. gracilis* Skvortz [Findley et al., 2010; 2011]; além de inferências sobre a distribuição de sequências repetitivas [Belarmino et al., 2012].

Em Fabaceae, análises comparativas intra- e intergenéricas têm sido realizadas com sucesso usando a técnica de BAC-FISH, especialmente entre espécies do clado Phaseoloids. Por exemplo, BACs de *Phaseolus vulgaris* L. hibridizados em cromossomos de *P. lunatus* L., *P. microcarpus* Mart. e *Vigna unguiculata* (L.) Walp. demonstraram uma forte sintenia para o grupo [Fonsêca, 2010; Bonifácio et al., 2012; Vasconcelos et al., em preparação], indicando que este tipo de inferência pode auxiliar no entendimento da evolução de espécies aparentadas nessa família.

Em vista do exposto, o presente estudo teve por finalidade contribuir para o entendimento da sintenia e dos eventos envolvidos na evolução cromossômica no gênero *Glycine*, mediante o uso de BACs de soja [*G. max* ($2n = 40$)] no mapeamento comparativo com *G. soja* ($2n = 40$) e espécies mais distantes filogeneticamente [*G. syndetika* B. E. Pfeil e Craven ($2n = 40$), *G. tabacina* (Labill.) Benth. ($2n = 80$) e *G. tomentella* Hayata ($2n = 78$)].

Materiais e métodos

Material vegetal

Sementes de *G. max*, *G. soja*, *G. syndetika*, *G. tabacina* e *G. tomentella* foram cedidas pelo banco de germoplasma da Embrapa Soja (Empresa Brasileira de Pesquisa Agropecuária, Soja, Londrina) (Tabela 1).

Tabela 1. Relação de espécies analisadas de *Glycine*, com seus respectivos números cromossômicos, acesso e distribuição geográfica.

Espécie	2n	Acesso	Distribuição
<i>G. max</i> (L.) Merr.	40	cv. Vitória da Conquista MG/BR-46	Amplamente cultivada
<i>G. soja</i> Sieb. & Zucc.	40	PI 597455	China, Japão, Rússia, Coreia, Taiwan
<i>G. syndetika</i> B.E. Pfeil & Craven	40	PI 441000, USDA, ou G1300, CSIRO	Austrália
<i>G. tabacina</i> (Labill.) Benth.	80	PI 339658, G1141	Austrália e ilhas do Oeste, central e sul do Pacífico
<i>G. tomentella</i> Hayata	78	PI 441001, G1133	Austrália

Coleta, pré-tratamento, fixação e digestão enzimática das raízes

Pontas de raízes coletadas a partir de sementes recém-germinadas foram pré-tratadas em 8-hidroxiquinoleína (8HQ) 2 mM por 4,5 h, a 18 °C, fixadas em metanol:ácido acético (3:1, v/v), por 2 a 24 h à temperatura ambiente e estocadas a -20 °C. Os meristemas foram lavados em água destilada e digeridos em uma solução de 2% (p/v) celulase 'Onozuka R-10' (Serva)/ 20% (v/v) pectinase (Sigma) por 3 h a 37 °C.

Preparação de lâminas

A metodologia utilizada seguiu o procedimento descrito por Carvalho e Saraiva (1993), com algumas modificações. O material foi lavado duas vezes em água destilada por 5 min e colocado próximo à extremidade de uma lâmina inclinada. Gotas de fixador (metanol: ácido acético, 3:1) gelado (-20 °C) foram pingadas com uma pipeta Pasteur sobre as raízes para total remoção da água. Em

seguida, foi procedida a maceração mecânica das raízes. Posteriormente, as lâminas foram secas ao ar com auxílio de um borrifador, mergulhadas em ácido acético 45% por 15 s e colocadas em uma superfície plana pré-aquecida (37 °C) por aproximadamente 5 min. As lâminas foram selecionadas em DAPI glicerol, descoradas em fixador (etanol: ácido acético, 3:1) por 30 min à temperatura ambiente e etanol absoluto por 1 h à temperatura ambiente ou 24 h entre 4 e 10 °C. As melhores lâminas foram selecionadas para a FISH.

Seleção dos BACs, extração e marcação de sondas

Na construção do mapa cromossômico, foram usadas sondas de clones BACs de soja contendo marcadores RFLPs mapeados nos grupos de ligação G (pseudomolécula 18) e J (pseudomolécula 16), que contêm genes e QTLs de resistência a doenças. Os BACs foram selecionados a partir do site <http://soybase.org/gb2/gbrowse/gmax1.01/> e são provenientes da University of Arizona, Department of Plant Sciences, Arizona Genomics Institute, adquiridas mediante o site: <http://www.genome.arizona.edu/cgi-in/orders/checkout.cg> (Tabela 2).

Também foram utilizadas as sequências D2, com um fragmento de 500 pb, contendo duas unidades de repetição de DNAr 5S oriunda de *Lotus japonicus* (Regel) K. Larsen [Pedrosa et al., 2002], e R2, contendo um fragmento de 6,5 Kb, com a unidade de repetição de DNAr 18S-5,8S-25S, proveniente de *Arabidopsis thaliana* (L.) Heynh. [Wanzenböck et al., 1997]. A extração do DNA dos BACs seguiu o protocolo do Kit de extração plasmidial da Qiagen (*Qiagen Plasmid Mini Kit*) com algumas adaptações. Os BACs foram marcados com Cy3-dUTP, o DNAr 5S com biotina-11-dUTP e o DNAr 45S com digoxigenina-11-dUTP por *nick translation* e usados como sonda.

Hibridização *in situ* fluorescente

A hibridização das lâminas seguiu o procedimento descrito por Pedrosa et al. (2001) com algumas alterações. A lavagem de estringência foi realizada em 0,1xSSC a 42 °C. As misturas de hibridização constituíram-se de: formamida 50% (v/v), dextran sulfato 10% (p/v), 20xSSC e 2-5 ng/μL de sonda, a qual foi desnaturada por 10 min a 70 °C. Para as sondas 102N16 e 176B09 que apresentaram marcações nas regiões centroméricas e pericentroméricas em *G. tomentella* foram utilizados DNA bloqueador (na proporção 50x ou 100x da fração

C₀t-100) na mistura de hibridização. As lâminas foram desnaturadas por 7 min a 70 °C com solução de formamida a 70% diluída em 2×SSC e hibridizadas de 18-48 h a 37 °C. Foram realizadas lavagens pós-hibridização e, em seguida, todas as preparações de sondas marcadas com Cy3 foram montadas com 8 µL de DAPI (1 µg/mL) em 50% de Vectashield (Vector). As sondas marcadas com biotina foram detectadas com avidina-rodamina (Vector Laboratories) em BSA 1% (p/v). Enquanto que as sondas marcadas com digoxigenina foram detectadas usando anti-digoxigenina de ovelha conjugado com fluoresceína isotiocianato FITC (Roche) e seguido de anti-IgG de ovelha conjugado com FITC (Sigma) também em BSA 1% (p/v). Todas as preparações foram montadas com 1 µg/mL de DAPI em 50% de Vectashield (Vector).

Tabela 2. Clone BAC, localização *in silico* no Soybase, região na pseudomolécula de *Glycine max* verificada *in silico*, tamanho em kb, região de sintenia em outros cromossomos de *G. max* e marcador(es) molecular(es) associado(s) a QTLs. Disponível no banco de dados SoyBase (<http://soybase.org/gb2/gbrowse/gmax1.01/>).

BAC GM_Wbc ¹	Localização <i>in silico</i>	Região ³	Tamanho (kb)	Sintenia	Satt, Sat ⁴	QTLs ⁵
0102N16	Gm16 ² :1231125..1382294	16S - T	151	Gm07	Satt405	Altura da planta Rendimento/ Maturação da vagem
0088G15	Gm16:27610047..27758800	16L - St	149	Gm02	Satt622	Número de vagens, produção
0149P23	Gm16:33265689..33383190	16L - St	117	Gm09	Satt244	Reação à <i>Phialophora gregata</i> / Reação à <i>Heterodera glycines</i> / Altura de planta
0145O12	Gm16:37192311..37309066	16L - T	117	Gm03 e 09	Satt712	-
0093L17	Gm18 ² :954362..1118734	18S - T	164	Gm11	Satt163	Altura de planta/ Reação à <i>H. glycines</i> / Teor de em ácido oleico nas sementes
0138F23	Gm18:4689845..4831719	18S - St	142	Gm11 e05	Satt235 e Satt 217	Conteúdo de nitrogênio das sementes/ Teor de cisteína nas sementes/ Teor de ácido linolênico nas sementes
0173N15	Gm18:10401526..10532467	18S - I	131	Gm08	Satt115	-
0176B09	Gm18:61071099..61214911	18L - T	144	Gm02	Sat_372	-

¹GM_Wbc, Biblioteca BAC de soja. ²Gm16 e Gm18, pseudomolécula 16 e 18 de *G. max*, respectivamente. ³St, Subterminal; T, Terminal; I, Intersticial; S, Short (braço curto); L, Long (braço longo). ⁴Satt ou Sat, Marcadores genéticos. ⁵QTL (*Quantitative Trait Loci* - Locos de Caracteres Quantitativos).

Análise dos dados

As melhores metáfases mitóticas foram capturadas através de uma câmera Leica DFC 340FX, utilizando o programa Leica CW 4000. As imagens foram otimizadas mediante ferramenta “brilho e contraste” com o Adobe Photoshop CS4 (Adobe Systems Incorporated). Após captura das imagens, todas as lâminas foram descoradas em etanol:ácido acético (3:1, v/v) por 30 min e etanol 100% por 1h e passaram por pelo menos uma re-hibridização para obtenção de novos resultados.

Para a sobreposição das imagens DAPI-BACs/ DNAr 5S ou 45S, a cor da imagem do DAPI foi alterada para tons de cinza e o fundo foi escurecido com a ferramenta *levels*. Os BACs foram pseudocoloridos em laranja, rosa, azul escuro, amarelo, azul claro, lilás ou vermelho e ambas as imagens foram sobrepostas, usando a ferramenta *lighten* do Adobe Photoshop CS4 (Adobe Systems Incorporated).

Para a construção do idiograma dos cromossomos homeólogos de *G. max*, *G. soja*, *G. syndetika*, *G. tabacina* e *G. tomentella* seguiu-se a metodologia estabelecida por Fonsêca et al. (2010), com adaptações. A identificação dos cromossomos *G. max* seguiu a nomenclatura determinada pelo site <http://www.soybase.org/>. Foi utilizado o software do programa MicroMeasure 3.3., disponível em <http://www.biology.colostate.edu/MicroMeasure> (Universidade do Estado do Colorado) para obter as medições cromossômicas. A posição de marcas foi medida a partir do telômero do braço longo (região inicial, correspondente ao 0) ao centrômero acrescida da distância do centrômero ao centro do sinal e do centro do sinal ao telômero (região final). O valor final consistiu de uma média das medições realizadas em pelo menos oito cromátides marcadas com mesmo padrão de condensação. Posteriormente, a montagem dos idiogramas foi feita mediante programa Adobe Flash CS4 Profissional da coleção Adobe Master Collection CS4.

Resultados

Localização das sequências BACs mediante FISH

Oito BACs de soja foram localizados *in situ* em cromossomos metafásicos de *G. max* (Gm , $2n = 40$), *G. soja* ($2n = 40$), *G. syndetika* ($2n = 40$), *G. tabacina* ($2n = 80$) e *G. tomentella* ($2n = 78$) (ver Tabelas 1 e 2). Em *G. max* e *G. soja*, todas as sondas evidenciaram marcações únicas em um ou dois pares cromossômicos, com exceção do BAC 173N15, que apresentou padrões de hibridização típicos de sequências de DNA repetitivo nas regiões centroméricas e pericentroméricas (Figura 1A, B), coincidentes com as regiões heterocromáticas reveladas pelo CMA₃⁺/DAPI em todo o complemento de *G. max* e em alguns cromossomos de *G. soja* (dados não mostrados).

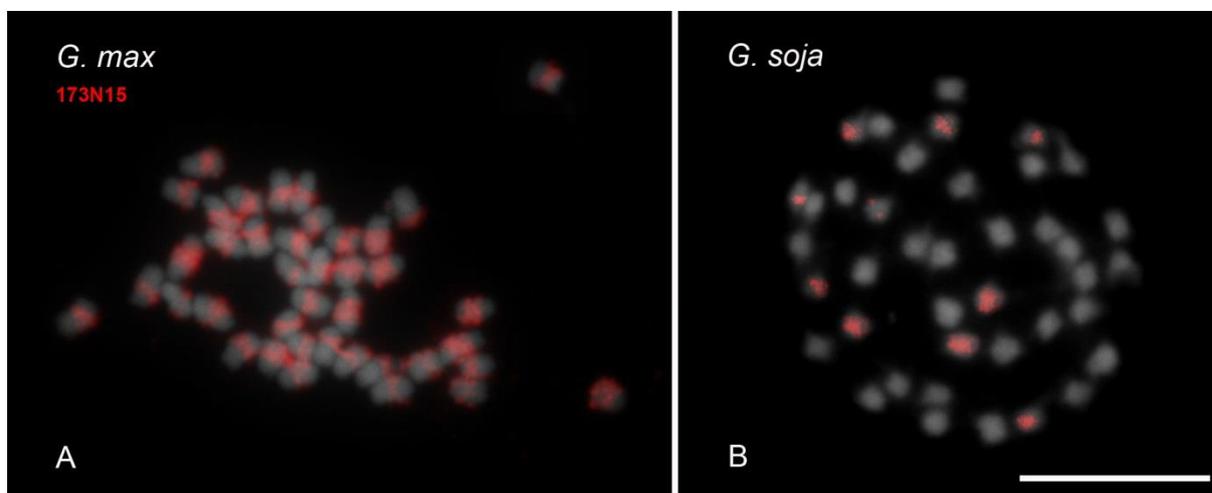


Figura 1. Hibridização *in situ* do BAC 173N15 do cromossomo *Gm18* de *Glycine max* em cromossomos mitóticos de *G. max* (A) e *G. soja* (B), contracolorados com DAPI (cinza). Barra em B = 10 μ m.

O BAC 102N16 mapeado geneticamente no grupo de ligação J (*Gm16*, posição terminal do braço curto) foi localizado *in situ* na região terminal do braço curto (S) de um par cromossômico em *G. max*, *G. soja* ($2n = 40$) e *G. tabacina* ($2n = 80$), permitindo a identificação dos cromossomos homeólogos portadores de *Gm16* nas duas últimas espécies (Figuras 2A, B, E; 3A, B, D). *Glycine syndetika* ($2n = 40$) exibiu dois pares de cromossomos com sítios terminais também no braço curto (Figuras 2C; 3C). *Glycine tomentella*, por sua vez, exibiu um par muito forte correspondente ao DNAr 45S e dois pares marcados na região terminal do braço curto, sendo um deles o provável homeólogo de *Gm16* (Figuras 2F1, F2; 3E). Vale ressaltar que nenhuma marcação proximal colocalizou com os dois pares cromossômicos portadores de DNAr 5S localizados na região proximal do braço curto (Figura 2F2).

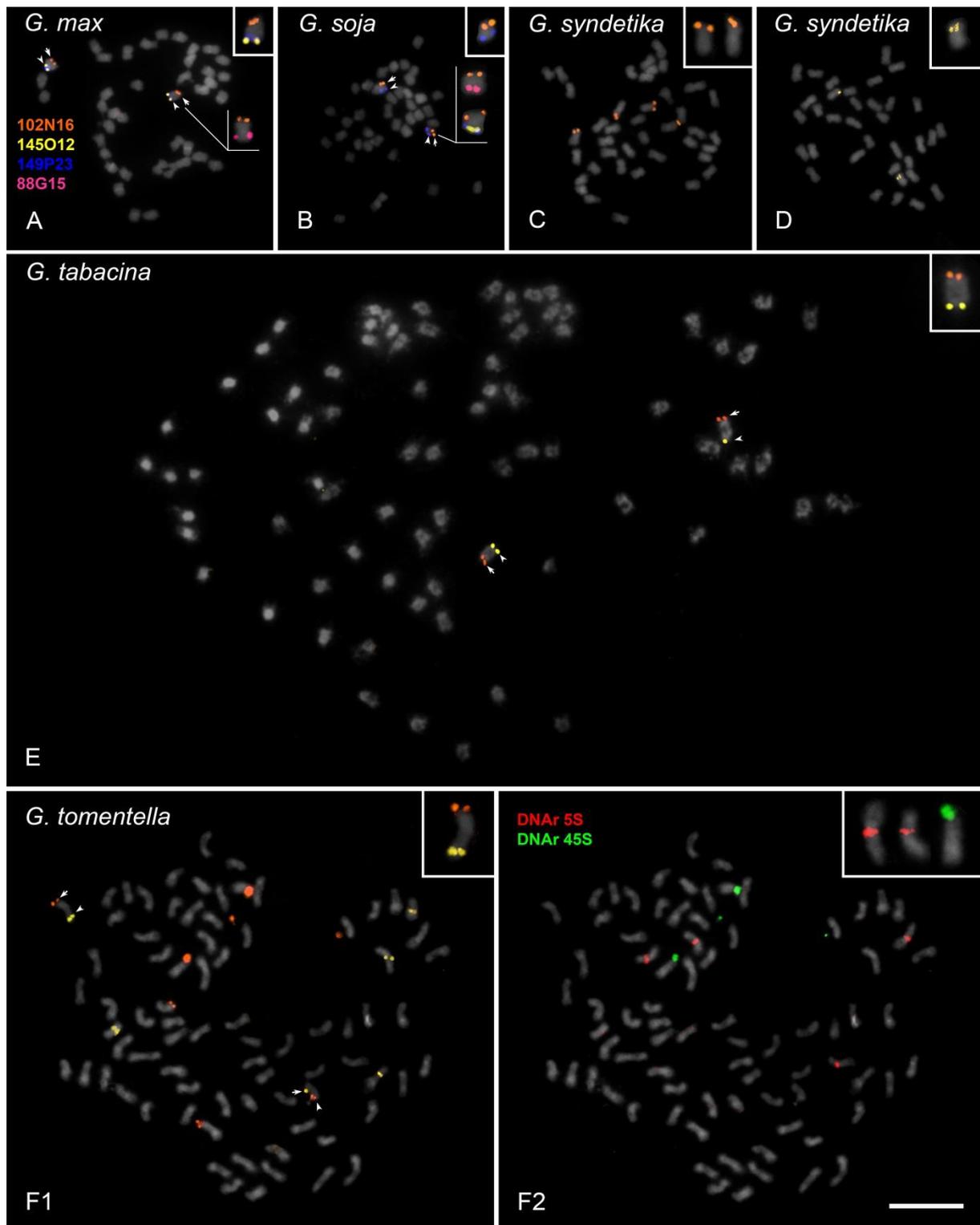


Figura 2. Hibridização *in situ* de quatro clones BAC de *Glycine max* (*Gm16*): 145O12 (amarelo); 149P23 (azul escuro); 102N16 (laranja) e 88G15 (rosa), bem como de sondas de DNAr 5S (vermelho) e 45S (verde) em *G. tomentella*. Cromossomos mitóticos de *G. max*, $2n = 40$ (A); *G. soja*, $2n = 40$ (B); *G. syndetika*, $2n = 40$ (C, D); *G. tabacina*, $2n = 80$ (E) e *G. tomentella*, $2n = 78$ (F1 e F2). Observa-se a colocalização de um par do BAC 102N16 com o par de sítios de DNAr 45S (F2). Os insertos mostram os cromossomos homeólogos ampliados, exceto para F2. Em A e B observa-se ainda inserto com marcação referente ao BAC 88G15 (*Gm16L*, rosa). Setas e cabeças de seta indicam as marcas que identificam os cromossomos homeólogos. Os cromossomos foram contracolorados com DAPI e pseudocoloridos em cinza. Barra em F2 = 10 μ m.

O clone BAC 145O12 [*Gm16*, região terminal do braço longo (L)] foi localizado *in situ* na região terminal do braço longo (L) de *Gm16* (Figura 3A). Um par com marcação mais fraca, às vezes não aparente, também foi visualizado em algumas células (sete de um total de 18 células), correspondente provavelmente a um dos cromossomos de sintenia recente (*Gm03* ou *Gm09*; Figura 2A; ver sintenia na Tabela 2). Para *G. soja* e *G. syndetika*, observou-se a presença de um par de sítios marcados fracamente na região terminal do braço longo e curto (dado não confirmado), respectivamente (Figuras 2B, D; 3B). Em *G. tabacina*, foram observadas marcações em um par de sítios na região terminal do braço longo (cromossomo metacêntrico, com razão de braços = 1,15), identificando o homeólogo (Figuras 2E; 3D). Para *G. tomentella*, este BAC evidenciou três pares de sinais fracos: dois proximais e um subterminal correspondente ao cromossomo de homeologia (Figura 2F1).

Adicionalmente, para os BACs 149P23 e 88G15 [*Gm16* braço longo (L), subterminal] que foram hibridizados apenas em cromossomos de *G. max* e *G. soja*, observou-se a presença de sinais na região subterminal (149P23) e intersticial (88G15) do braço longo de um par cromossômico (Figuras 2A, B; 3A, B).

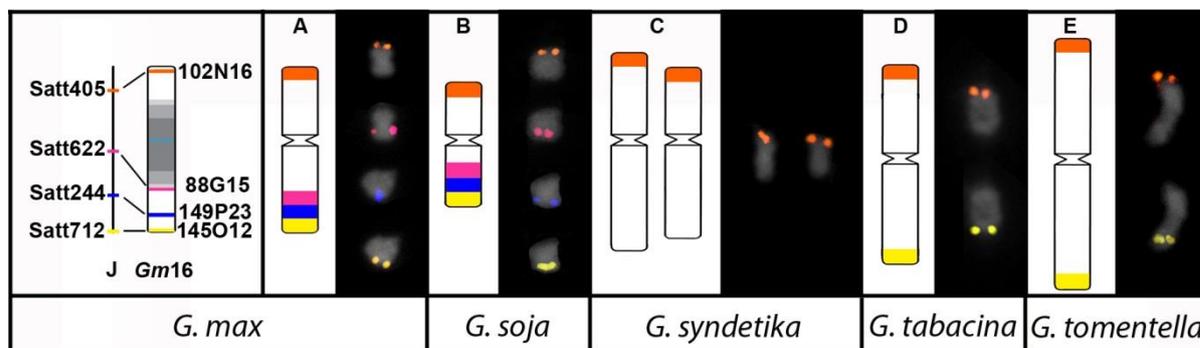


Figura 3. Mapas genéticos e de sequenciamento do cromossomo *Gm16* (J) de *G. max* com os clones BACs 102N16 (laranja), 88G15 (rosa), 149P23 (azul escuro) e 145O12 (amarelo). Representação esquemática dos cromossomos e cromossomos de homeologia de *G. max*, *G. soja*, *G. syndetika*, *G. tabacina* e *G. tomentella*.

Dentre os BACs mapeados geneticamente no grupo de ligação G (*Gm18*), 93L17 [terminal do braço curto (S)] foi localizado *in situ* na região terminal, do braço longo (L) em um par cromossômico de *G. max* e de *G. soja* ($2n = 40$), possibilitando a identificação do par homeólogo a *Gm18* em *G. soja* (Figura 5A, B), além do provável cromossomo de sintenia (*Gm11*) em *G. max*, detectado por uma marcação mais fraca na região terminal do braço curto (Figura 4A). *Glycine syndetika* ($2n = 40$) apresentou quatro pares: um deles colocalizando com o DNAr 45S, além de três terminais (um no braço longo e dois no braço curto), sendo que uma destas

marcações no braço curto é referente ao cromossomo homeólogo a *Gm18* (Figuras 4D; 5C). Em *G. tomentella* ($2n = 78$), este mesmo BAC hibridizou quatro pares cromossômicos: um deles colocalizando com o DNAr 45S terminal (inserto na Figuras 4E e 5E); outro proximal no braço curto, e os dois últimos em região terminal do braço curto (dados não mostrados). Em *G. tabacina* ($2n = 80$), foi visualizado um par de sítios marcado na região terminal do braço curto, em cromossomo correspondente ao de homeologia a *Gm18* (Figura 5D), além de três pares de marcações fracas (Figura 4F).

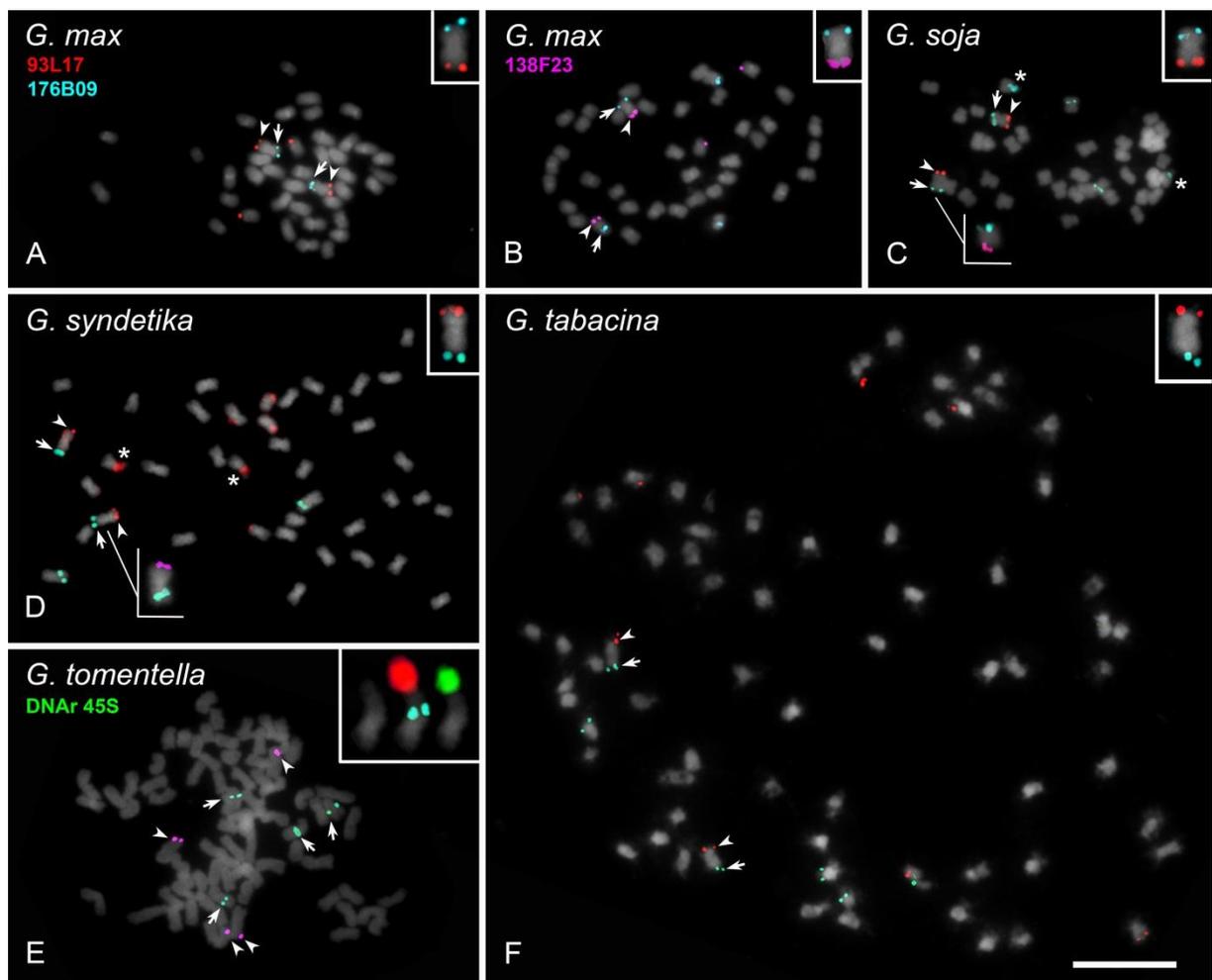


Figura 4. Hibridização *in situ* de três clones BAC para o cromossomo 18 de *Glycine max* (*Gm18*): 93L17 (vermelho); 138F23 (lilás), e 176B09 (azul claro). Cromossomos mitóticos de *G. max*, $2n = 40$ (A, B), *G. soja*, $2n = 40$ (C), *G. syndetika*, $2n = 40$ (D), *G. tomentella*, $2n = 78$ (E) e *G. tabacina*, $2n = 80$ (F). FISH de DNAr 45S (verde) em *G. tomentella*, colocalizado com BAC 93L17 (inserto em E). Setas e cabeças de seta indicam as marcas que identificam os cromossomos com homeologia. Asteriscos indicam a região organizadora de nucléolos. Os cromossomos foram contraincolorados com DAPI e pseudocolorados em cinza. Barra em F = 10 μ m.

O BAC 138F23 (grupo de ligação G, *Gm18*, braço curto terminal) hibridizou na região subterminal do braço longo de um par cromossômico de *G. max*

correspondente ao *Gm18* (Figura 5A) e na região terminal do braço curto do provável cromossomo sintênico *Gm11* com marcações mais fracas (Figura 4A). Em *G. soja* e *G. syndetika*, hibridizou em um par de cromossomos na região terminal do braço longo e curto, respectivamente, identificando os cromossomos de homeologia (inserto em Figuras 4C, D; 5B, C). Em *G. tomentella*, dois pares, um terminal no braço longo e um proximal no braço curto foram observados (Figuras 4E; 5E).

O clone BAC 176B09 (*Gm18L* terminal) foi localizado *in situ* em dois pares cromossômicos de *G. max*: *Gm18* em região subterminal no braço curto (Figura 5A) e o segundo, que mostra uma marcação fraca na região subterminal também do braço curto, não visualizada em todas as células, referente provavelmente à região de sintonia em *Gm02* (Figura 4A, B). *Glycine soja* apresentou uma marcação no par homeólogo (Figura 5B), além da presença de mais dois pares, um hibridizando uma provável região de DNAr 45S e outro par marcando uma região proximal do braço longo, sendo que esta última não foi visível na maioria das células (Figura 4C). Em *G. syndetika*, esta sonda hibridizou na região subterminal do braço longo do cromossomo homeólogo (Figura 5C) e em outro cromossomo em região subterminal do braço longo (Figura 4D). Em *G. tomentella*, estão presentes dois pares de marcações proximais no braço curto, um deles correspondente ao cromossomo portador de DNAr 45S (Figuras 4E; 5E). Em *G. tabacina*, a sonda possibilitou visualizar marcação na região terminal do braço longo do par de cromossomos correspondente ao homeólogo (Figura 5D), além de dois pares de marcas fracas subterminais e terminais (Figura 4F).

Vale ressaltar que para a FISH dos BACs 145O12, 93L17, 138F23 e 176B09 em cromossomos de *G. tomentella* foi utilizado DNA bloqueador (50x ou 100x da fração C0t-100) na mistura de hibridização, a fim de inibir a presença de marcações inespecíficas (Figuras 2F, 4E, 5).

As Figuras 3 e 5, bem como a Tabela 3 resumem os principais resultados obtidos no presente trabalho.

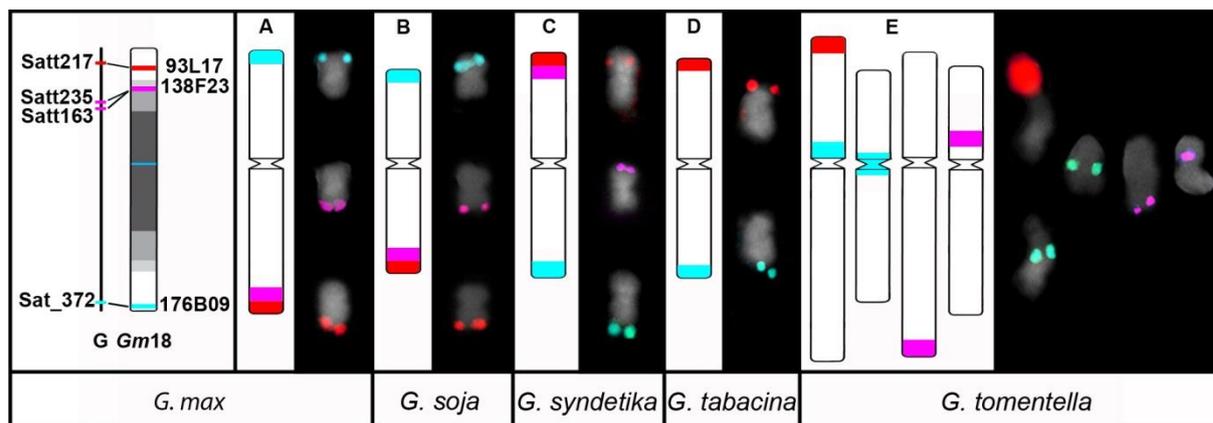


Figura 5. Mapas genético e de sequenciamento do cromossomo *Gm18* (G) de *G. max* com os clones BACs 93L17 (vermelho), 138F23 (lilás), 173B09 (azul claro). Representação esquemática dos cromossomos e cromossomos de homeologia de *G. max*, *G. soja*, *G. syndetika*, *G. tabacina* e *G. tomentella*.

Tabela 3. Clones BAC hibridizados e resultados de suas respectivas localizações *in situ* em cinco espécies de *Glycine*.

BAC GM_WBc ¹	Localização <i>in situ</i>				
	<i>G. max</i>	<i>G. soja</i>	<i>G. syndetika</i>	<i>G. tabacina</i>	<i>G. tomentella</i>
0102N16	16S ² -T ⁴	S-T	2 S-T	S-T	DNAr 45S/ 2 S-T
0088G15	16L ³ -I ⁶	L-I	-	-	-
0149P23	16L-St ⁵	L-St	-	-	-
0145O12	16L-T/ S-T	L-T	S-T 45S /	L-T S-T / várias	L-St / 2 L-P ⁷ 45S / S-P /
0093L17	18L-T/ S-T	L-T	2 S-T / L-T	marcações fracas	2 S-T
0138F23	18L-St/ S-T	L-T	S-T	-	S-P/L-T
0173N15	Pe ⁸	Pe	-	-	-
0176B09	18S-St/ S-St	DNAr 45S S-T/ L-P	2 L-St	L-T/ várias marcações fracas	S-P no cromossomo com DNAr 45S/S-P

¹GM_WBc, biblioteca BAC de *G. max*. ²L, (Long) braço longo; ³S, (Short) braço curto; ⁴T, terminal; ⁵St, subterminal; ⁶I, intersticial; ⁷P, proximal; ⁸Pe, pericentromérica..

Discussão

A técnica BAC-FISH, utilizando sequências de DNA de *G. max* como sondas, possibilitou o mapeamento citogenético comparativo de *Gm16* e *Gm18* entre os cromossomos de *G. max*, *G. soja*, *G. syndetika*, *G. tabacina* e *G. tomentella*, colaborando para os estudos de sintonia e para o entendimento da evolução cariotípica do grupo.

Importantes estudos de BAC-FISH vêm sendo realizados em análises genômico-comparativas inter- e intragenéricas de diversas espécies. Os dados indicam que espécies de um mesmo gênero, em geral, apresentam uma relativa estabilidade quanto à ocorrência de rearranjos cromossômicos, os quais são constituídos principalmente por fusões, inversões e translocações [Lysak et al., 2006; Mandáková e Lysak, 2008]. Em soja, por exemplo, estudos citogenéticos comparativos entre *G. max* e *G. soja* e entre *G. max* e *G. gracilis* foram realizados, permitindo a identificação de apenas um evento de translocação entre os cromossomos 11 e 13 de *G. soja* em relação a *G. max* [Findley et al., 2010] e entre os cromossomos 1 e 8 de *G. gracilis* [Findley et al., 2011].

No presente trabalho, sete BACs contendo sequências de cópia única foram hibridizados *in situ* em cromossomos de *G. max*, *G. soja*, *G. syndetika*, *G. tabacina* e *G. tomentella*, permitindo a localização citogenética de marcadores genéticos (Satt ou Sat) e a orientação dos grupos de ligação G (*Gm18*) e J (*Gm16*) em relação à estrutura dos cromossomos. A localização *in situ* dessas marcas possibilitou a análise da sintonia no genoma de cada espécie e entre as espécies analisadas (Figuras 3; 5). Para os dois cromossomos analisados, observou-se uma sintonia total das marcas entre as espécies, destacando quebra de sintonia para *Gm16* em *G. syndetika* e *G. tomentella* e para *Gm18* em todas as espécies analisadas.

Essa estabilidade cariotípica intragenérica e mesmo intergenérica também foi observada para outras espécies do clado Phaseoloids, no qual o gênero *Glycine* está inserido. BACs contendo sequências de *P. vulgaris*, por exemplo, foram usados em uma comparação genômica entre a referida espécie e três outras do clado Phaseoloids (*P. lunatus*, *P. microcarpus* e *V. unguiculata*). Observou-se neste estudo uma alta sintonia, com a ocorrência de alguns eventos de translocação e de inversão, cuja frequência aumentou com o aumento da distância filogenética [Bonifácio, 2009; Bonifácio et al., 2012; Vasconcelos *et al.*, em preparação]. Entre *P.*

vulgaris e *P. lunatus*, por exemplo, foram observadas algumas prováveis inversões pericêntricas nos cromossomos 2 e 9, bem como uma transposição ou inversão pericêntrica no cromossomo 10 [Bonifácio et al., 2012]. Entre *P. vulgaris* e *P. microcarpus*, foi ressaltada também uma alta sintenia, com apenas três quebras de colinearidade [Fonseca et al., 2010]. Já entre *P. vulgaris* e *V. unguiculata*, Vasconcelos et al. (em preparação) observaram uma sintenia parcial para FISH de BACs de quatro cromossomos de *P. vulgaris* em cinco de *V. unguiculata*, destacando a presença de eventos de duplicação, inversões e translocações.

Em Fabaceae, dados genômicos também apontam para a conservação de grandes blocos sintênicos entre as espécies [Choi et al., 2004], observando-se uma redução com o aumento da distância filogenética [Bordat et al., 2011]. No clado Phaseoloids, *G. max* e *P. vulgaris* mostraram extensos blocos de loci compartilhados mediante estudos de sintenia com mapeamento usando RFLP e análise de EST (*Expressed Sequence Tags* - Etiquetas de Sequências Expressas). Os autores observaram que a maioria dos cromossomos de soja é composta por fragmentos sintênicos com vários cromossomos do feijão comum [McClellan et al., 2010]. Também entre *V. unguiculata* e *G. max*, múltiplas regiões de sintenia foram observadas. O genoma de soja exibiu regiões gênicas homeólogas com cerca de 85% de SNPs mapeados em *V. unguiculata*, indicando uma grande semelhança entre os genomas [Lucas et al., 2011].

O aparecimento de mais de um par para algumas marcas para as espécies com $2n = 40$ e de marcações extras em *G. soja*, *G. syndetika*, *G. tabacina* e *G. tomentella* ($2n = 40, 40, 80$ e 78 , respectivamente) pode ser devido à natureza poliploide de seus genomas, com algumas regiões duplicadas, onde se espera que estas sequências apresentem regiões de homeologia com outros sítios do genoma [Gill et al., 2009; Schmutz et al., 2010]. Contudo, não pode ser descartada a possibilidade da presença de DNA repetitivo nessas regiões, mesmo usando fração C_{0t} como no caso de *G. tomentella* não bloqueado pela. Este último caso pode ser exemplificado pela marcação de alguns BACs colocalizando com as regiões de DNAr 45S, observada em *G. soja* (BAC 176B09, *Gm18*), *G. syndetika* (BAC 93L17, *Gm18*) e *G. tomentella* (BACs 0102N16, *Gm16*; 93L17, *Gm18*), não mostrada em *G. max*. A colocalização diferencial do DNA repetitivo com sítios de DNAr 45S das três referidas espécies pode ocorrer como consequência de uma divergência dos espaçadores intergênicos em sítios de DNAr [Garcia et al., 2010]. A associação de DNAr 45S a DNA repetitivo pode ser observada, por exemplo, em *Hordeum*

spontaneum (K. Koch) Thell [Belyayev et al., 2005], *Cestrum strigilatum* Ruiz e Pav. e *C. intermedium* Sendtn. [Fregonezi et al., 2007].

O único BAC com marcação repetitiva em *G. max* (173N15) hibridizou nas regiões pericentroméricas colocalizando ao menos em grande parte com a região de heterocromatina constitutiva de *G. max* [ver Bortoleti *et al.*, em preparação; Schmutz et al., 2010]. Contudo, em *G. soja*, este mesmo BAC marcou apenas cinco pares na região de heterocromatina pericentromérica [Silva et al., em preparação; Ha et al., 2012]. A presença de um número bem menor de marcações nas regiões pericentroméricas de *G. soja* demonstra que esta região tem se diferenciado de *G. max* ao longo da evolução [Du et al., 2012].

De um modo geral, os dados do presente trabalho sugerem a ocorrência de sintenia para o grupo, apesar da distância filogenética entre as espécies. Todos os BACs evidenciaram uma alta estabilidade cariotípica, observando apenas uma quebra de colinearidade em *G. tomentella* e outra possivelmente em *G. syndetika*, envolvendo possíveis inversões e/ou translocações.

Agradecimentos

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), no âmbito do Consórcio Brasileiro Genosoja da Empresa Brasileira de Pesquisa Agropecuária, Soja (Embrapa Soja) e à Coordenação de Pessoal de Nível Superior (CAPES) pelo apoio financeiro, bem como à Universidade do Arizona (Estados Unidos da América do Norte) pela disponibilização dos BACs.

Referências

- Belarmino, LC, Oliveira, ARS, Brasileiro-Vidal, AC, Bortoleti, KCA, Bezerra-Neto, JP, Abdelnoor, RV, Benko-Iseppon, AM Mining plant genome browsers as a mean for efficient connection of physical, genetic and cytogenetic mapping: an example using soybean. *Genet Mol Biol* 35(1): 335–347 (2012).
- Belyayev A, Raskina O, Nevo E: Variability of the chromosomal distribution of Ty3-gypsy retrotransposons in the populations of two wild Triticeae species. *Cytogenet Genome Res* 109: 43–49 (2005).
- Benko-Iseppon AM, Nepomuceno AL, Abdelnoor RV: GENOSOJA – The Brazilian Soybean Genome Consortium: High throughput omics and beyond. *Genet Mol Biol* 35: 2 (2012).
- Bonifácio EM Mapeamento comparativo entre o feijão comum e a fava. Tese (Doutorado em Ciências Biológicas) – Universidade Federal Rural de Pernambuco, Recife: 64 f (2009).
- Bonifácio EM, Fonsêca A, Almeida C, Santos KGB, Pedrosa-Harand A: Comparative cytogenetic mapping between the lima bean (*Phaseolus lunatus* L.) and the common bean (*P. vulgaris* L.). *Theor App Genet* 124: 1513–1520 (2012).
- Bordat A, Savoie V, Nicolas M, Salse J, Chauveau A, et al: Translational genomics in legumes allowed placing *in silico* 5460 unigenes on the pea functional map and identified candidate genes in *Pisum sativum* L. *G3* 2: 93–103 (2011). Bortoleti KCA, Benko-Iseppon AM, Silva LCB, Oliveira ARS, Abdelnoor RV, Ramalho I, MELO NF, Brasileiro-Vidal AC Microsatellites: an important constituent or an enlargement factor of genome in legumes. *Em preparação*.
- Carvalho CR e Saraiva LS: An air drying technique for maize chromosomes without enzymatic maceration. *Biotech Histochem* 68: 142–145 (1993).
- Choi IY, Hyten DL, Matukumalli LK, Song Q, Chaky JM, et al: A soybean transcript map: gene distribution, haplotype and single nucleotide polymorphism analysis. *Genetics* 176: 685–696 (2007).
- Choi HK, Mun JH, Kim DJ, Zhu HY, Baek JM, et al: Estimating genome conservation between crop and model legume species. *PNAS* 101: 15289–15294 (2004).

- Chung G, Singh RJ: Broadening the Genetic Base of Soybean: A Multidisciplinary Approach. *Crit Reviews in PI Sci* 27: 295–341 (2008)
- Costa Neto PR, Rossi LFS, Zagonel GF, Ramos LP: Produção de biocombustível alternativo ao óleo diesel através da transesterificação de óleo de soja usado em fritura. *Química Nova* 23: 4 (2000).
- Cregan PB, Jarvik T, Bush A.L, Shoemaker RC, Lark KG, et al: An integrated genetic linkage map of the soybean genome. *Crop Sci* 39: 1464–1490 (1999).
- Du J, Tian, Z, Sui, Y, Meixia Z, Song Q, et al: Pericentromeric Effects Shape the Patterns of Divergence, Retention, and Expression of Duplicated Genes in the Paleopolyploid Soybean. *Plant Cell* 24: 21–32 (2012).
- Findley SD, Cannon S, Varala K, Du J, Ma J, et al: A FISH Karyotyping System for Soybean. *Genetics* 185: 3727–744 (2010).
- Findley SD, Pappas AL, Cui Y, Birchler JA, Palmer RG, et al: Fluorescence *In Situ* Hybridization–Based Karyotyping of Soybean Translocation Lines. *G3* 1: 117–129 (2011).
- Fonsêca A, Ferreira J, Santos TRB, Mosiolek M, Bellucci E, et al: Cytogenetic map of common bean (*Phaseolus vulgaris* L.). *Chromosome Res* 18:487–502 (2010).
- Fregonezi JN, Vilas-Boas LA, Fungaro MHP, Gaeta ML, Vanzela ALL: Distribution of a Ty3/gypsy-like retroelement on the A and B-chromosomes of *Cestrum strigilatum* Ruiz & Pav. and *Cestrum intermedium* Sendtn. (Solanaceae). *Genet Mol Biol* 30: 599–604 (2007).
- Garcia S, Panero JL, Siroky J, Kovarik A: Repeated reunions and splits feature the highly dynamic evolution of 5S and 35S ribosomal RNA genes (rDNA) in the Asteraceae family. *BMC Plant Biol* 10: 176 (2010).
- Gill N, Findley S, Walling JG, Hans C, MA J, et al: Molecular and chromosomal evidence for allopolyploidy in soybean, *Glycine max* (L.) Merr. *Plant Physiol* 151: 1167–1174 (2009).
- Graham PH e Vance CP: Legumes importance and constraints to greater use. *Plant Physiol* 131: 872–877 (2003).

- Ha J, Abernathy B, Nelson W, Grant D, Wu X, et al: Integration of the Draft Sequence and Physical Map as a Framework for Genomic Research in Soybean (*Glycine max* (L.) Merr.) and Wild Soybean (*Glycine soja* Sieb. and Zucc.). *G3* 1: 221–329 (2012).
- Lucas MR, Diop N-N, Wanamaker S, Ehlers JD, Roberts PA, et al: Cowpea–Soybean Synteny Clarified through an Improved Genetic Map. *Plant Genome* 4: 218–225 (2011).
- Lysak MA, Berr A, Pecinka A, Schmidt R, McBreen K, Schubert I Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related Brassicaceae species *PNAS* 103: 5224–5229 (2006).
- Mandáková T, Lysak MA: Chromosomal Phylogeny and Karyotype Evolution in x=7 Crucifer Species (Brassicaceae). *Plant Cell* 20: 2559–2570 (2008).
- Mcclean PE, Mamidi S, Mcconnell M, Chikara S, Lee R: Synteny mapping between common bean and soybean reveals extensive blocks of shared loci. *BMC Genomics* 11: 184 (2010).
- Pedrosa A, Jantsch MF, Moscone EA, Ambros PF, Schweizer D: Characterisation of pericentromeric and sticky intercalary heterochromatin in *Ornithogalum longibracteatum* (Hyacinthaceae). *Chromosoma* 110: 203–213 (2001).
- Pedrosa A, Sandal N, Stougaard J, Schweizer D, Bachmair A: Chromosomal map of the model legume *Lotus japonicus*. *Genetics* 161: 1661–1672 (2002).
- Schmutz J, Cannon SB, Schlueter J, MA J, Mitros T, et al: Genome sequence of the palaeopolyploid soybean. *Theor Appl Genet* 463:178–183 (2010).
- Song QJ, Marek LF, Shoemaker RC, Lar KG, Concibio VC, et al: A new integrated genetic linkage map of the soybean. *Theor Appl Genet* 109: 122–128 (2004).
- SoyBase Disponível em: (<http://soybase.org/gb2/gbrowse/gmax1.01/>). Acessado em: 15 de setembro de 2012.
- Universidade do Arizona disponível em: <http://genomearizona.edu/cgi-in/orders/checkout.cg>. Acessado em: 30 de julho de 2010.

Vasconcelos EV: Mapeamento Citogenético de *Vigna unguiculata* (L.) Walp. Mediante Hibridização *in situ* de Sequências de DNA de *Phaseolus vulgaris* L. Monografia (Graduação em Ciências Biológicas) – Universidade Federal de Pernambuco, Recife: 55 f (2011).

Wanzenböck EM, Schöfer C, Schweizer D, Bachmair A: Ribosomal transcription units integrated via T-DNA transformation associate with the nucleolus and do not require upstream repeat sequences for activity in *Arabidopsis thaliana*. *Plant J* 11: 1007–1016 (1997).

Conclusões

Conclusões

1. Para *Gm18*, a orientação do mapa de sequenciamento está invertida em relação à morfologia cromossômica, provavelmente em função do padrão de condensação cromossômica.
2. Observa-se uma quebra de sintenia para *Gm16* em *G. syndetika* e *G. tomentella* e para *Gm18* em todas as espécies analisadas.
3. O aparecimento de mais de um par para algumas marcas para as espécies com $2n = 40$ e em *G. tabacina* e *G. tomentella* ($2n = 80$ e 78 , respectivamente) reflete a natureza poliploide de seus genomas.
4. Os sítios de DNAr 45S em *Glycine* parecem ter uma divergência de espaçadores intergênicos, evidenciada pela sua colocação diferencial com alguns BACs em *G. soja* (BAC 176B09, *Gm18*), *G. syndetika* (BAC 93L17, *Gm18*) e *G. tomentella* (BACs 0102N16, *Gm16*; 93L17, *Gm18*), não observada em *G. max*.
5. A região pericentromérica de *G. max* apresenta constituição diferenciada quando comparada à *G. soja*.

Apêndice