

MARIA EDUARDA FERRAZ VASCONCELOS

**CARACTERIZAÇÃO DE SEQUÊNCIAS REPETITIVAS EM *Phaseolus* L. GRUPO  
LEPTOSTACHYUS (Fabaceae), UM CLADO DE RÁPIDA EVOLUÇÃO  
CARIOTÍPICA**

RECIFE

2019

MARIA EDUARDA FERRAZ VASCONCELOS

**CARACTERIZAÇÃO DE SEQUÊNCIAS REPETITIVAS EM *Phaseolus* L. GRUPO  
LEPTOSTACHYUS (Fabaceae): UM CLADO DE RÁPIDA EVOLUÇÃO  
CARIOTÍPICA**

Dissertação apresentada ao Programa de Pós-Graduação em Biologia Vegetal da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de mestre em Biologia Vegetal.

Área de Concentração: Sistemática e Evolução.

Orientadora: Dra. Andrea Pedrosa-Harand

RECIFE

2019

Catálogo na Fonte:  
Bibliotecário Bruno Márcio Gouveia, CRB4/1788

Vasconcelos, Maria Eduarda Ferraz

Caracterização de sequências repetitivas em *Phaseolus* L. grupo *Leptostachyus* (Fabaceae) : um clado de rápida evolução cariotípica / Maria Eduarda Ferraz Vasconcelos. – 2019.

113 f. : il.

Orientadora: Dra. Andrea Pedrosa-Harand. .

Dissertação (mestrado) – Universidade Federal de Pernambuco. Centro de Biociências. Programa de Pós-graduação em Biologia Vegetal, Recife, 2019.

Inclui referências e anexos.

1. Genética vegetal. 2. Plantas - Evolução. 3. Cromossomos vegetais. I. Pedrosa-Harand, Andrea (orientadora). II. Título.

581.35

CDD (22.ed.)

UFPE/CB – 2022-129

MARIA EDUARDA FERRAZ VASCONCELOS

CARACTERIZAÇÃO DE SEQUÊNCIAS REPETITIVAS EM *Phaseolus* L. GRUPO  
LEPTOSTACHYUS (Fabaceae): UM CLADO DE RÁPIDA EVOLUÇÃO CARIOTÍPICA

APROVADA EM: 09 de Maio de 2019

BANCA EXAMINADORA:

---

Profª. Dra. Andrea Pedrosa-Harand (Orientadora) Departamento de Botânica – UFPE

---

Prof Dra. Ana Christina Brasileiro Vidal (Titular externa) Departamento de Genética – UFPE

---

Prof. Dra. Neide Santos (Titular externo) Departamento de Genética – UFPE

Recife - PE  
2019

Dedico este trabalho à minha família,  
PAME, por ser perfeita  
em meio às imperfeições.

## AGRADECIMENTOS

Como meu costume diário, agradeço primeiramente a Deus, por ter permitido que eu chegasse até aqui, colocando em meu caminho as pessoas certas que aqui serão citadas e tantas outras importantes à sua maneira;

Aos meus pais, Paulo de Tarso Vasconcelos e Mércia Ferraz Vasconcelos, por todo apoio, compreensão e esforço para me proporcionar o melhor que sempre puderam. E ao meu irmão Arthur, que mesmo de outro Estado, esteve do meu lado torcendo pelo meu sucesso. Sem vocês, eu nada seria;

À minha orientadora desde a graduação, Profa. Dra. Andrea Pedrosa, pela paciência na forma de lidar com as dúvidas e os problemas que surgem, mas principalmente, pela empatia e compreensão dos problemas que enfrentei durante esses últimos dois anos;

À Vânia, minha amiga e terapeuta, com quem aprendi a me conhecer, me cuidar, me respeitar e principalmente, me amar. Que esse exercício não cesse nunca e que todos possam ter uma Vânia em suas vidas;

À toda minha família, tios, tias, primos e primas (que são MUITOS), pela felicidade que é fazer parte disso;

À vovó Veneza (Maria Etelvina), que continua firme e forte conosco, mandando eu sempre priorizar os estudos, mas não esquecer de aproveitar a vida. Ao Dr. Simão Vasconcelos (tio Mano), por ser esse tio coruja cientista que todos gostariam de ter;

À todos os amigos da graduação na UFPE e seus agregados, aos recentes amigos da minha segunda graduação em andamento e a todos os amigos que a vida me trouxe. Todos vocês contribuíram e contribuem para o meu crescimento, cada um com sua particularidade;

Às grandes amigas que as faculdades me deram. Dressa e Djal, por todas as mensagens matinais de força e determinação, Djal com seus conselhos e Dressa com sua grande disponibilidade para ajudar. Paty, Nice e Haymée, por terem me incluído no grupo, mesmo eu não sendo da Ecologia, hahaha, vocês tornaram o mestrado mais divertido. A Gabi, por em menos de um ano

se tornar como uma irmã para mim, tão parecida e tão diferente, mas capaz de me entender até quando nem eu estou me entendendo. A todas, obrigada pelos risos, conversas profundas e sessões de terapia gratuitas. Estarei sempre com vocês;

À todos os colegas do laboratório de Citogenética e Evolução Vegetal pelo ambiente agradável e respeitoso para trabalho. Também aos professores do laboratório, Luiz Gustavo e Marcelo Guerra, pelos ensinamentos passados;

Um agradecimento especial a Mariela (Marie) por toda explicação para as minhas várias e incansáveis perguntas e a todos os meus “E agora, faço o que?” que foram “calmamente” respondidos com seu jeito Argentino de ser;

Agradeço também por aqueles que entraram mas já saíram da minha vida. As vezes não entendemos porque passamos por certas situações, mas, há um propósito maior nisso;

Por fim, agradeço à Universidade Federal de Pernambuco e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela infraestrutura e apoio financeiro.

Muito obrigada.

## RESUMO

Sequências repetitivas são responsáveis por grandes diferenças entre os genomas, pois variam mais rapidamente que sequências únicas. Além disso, eventos de recombinação entre sequências repetitivas similares podem levar a rearranjos cromossômicos. No gênero *Phaseolus*, o grupo *Leptostachyus*, com três espécies incluindo *P. leptostachyus* e *P. macvaughii*, apresenta cariótipo disploide devido a uma inserção cêntrica, além de várias translocações. O objetivo do trabalho foi investigar a fração de DNA repetitivo dessas duas espécies e analisar sua possível relação com a ocorrência de rearranjos. Para isso, suas sequências repetitivas foram comparadas com as de outras 11 espécies, três delas cariotipicamente estáveis. As sequências foram agrupadas por similaridade no RepeatExplorer e a ferramenta Tarean foi usada para identificação de DNAs satélites. As principais sequências foram localizadas *in situ*. Os elementos Ty3/gypsy, mais especificamente a linhagem Chromovirus, foram os mais abundantes em *Phaseolus* e tiveram uma distribuição cromossômica preferencialmente pericentromérica. A variação na abundância das principais linhagens de elementos transponíveis entre espécies de *Leptostachyus* foi similar à variação entre as mesmas e as demais espécies do gênero, sugerindo rápido *turnover* no grupo. Os DNAs satélites compartilhados revelaram padrões de distribuição cromossômica diferentes entre as espécies, confirmando sua rápida evolução no genoma. A distribuição atípica de um dos satélites se mostrou compatível com uma possível associação desse com os rearranjos cromossômicos frequentes no grupo.

Palavras-chaves: Chromovirus; DNA satélite; rearranjos cromossômicos; retrotransposons.

## ABSTRACT

Repetitive sequences are responsible for large differences between the genomes because they diverge faster than unique sequences. In addition, recombination events between similar repetitive sequences may lead to chromosomal rearrangements. In the *Phaseolus* genus, the Leptostachyus group, with three species including *P. leptostachyus* and *P. macvaughii*, presents dispoloid karyotype due to a centric insertion, in addition to many translocations. The aim of this work was to investigate the repetitive DNA fraction of these two species and to analyze the possible relation with the occurrence of rearrangements. For this, their repetitive sequences were compared with 11 other species, three of which were karyotypically conserved. The sequences were clustered by similarity in the RepeatExplorer pipeline and the TAREAN tool was used to identify satellite repeats. The main sequences were located *in situ*. The Ty3/gypsy elements, more specifically Chromovirus, were the most abundant in *Phaseolus* species and had preferentially a pericentromeric chromosomal distribution. The variation in abundance of the main transposable lineages between Leptostachyus species was similar to the variation between them and the other species of the genus, suggesting rapid turnover in this group. Shared satellite DNAs revealed different patterns of chromosome distribution among species, confirming their rapid evolution in the genome. The atypical distribution of one satellite family was compatible with a possible association of this and the frequent chromosomal rearrangements in the group.

Keywords: chromosomal rearrangements; chromovirus; retrotransposons; satellite DNA.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>10</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>12</b>
2.1	SEQUÊNCIAS REPETITIVAS DO GENOMA EUCARIOTO	12
2.1.1	Caracterização das sequências repetitivas do genoma	13
2.1.2	Evolução das sequências repetitivas e rearranjos cromossômicos	15
2.2	O GÊNERO PHASEOLUS	16
2.2.1	Taxonomia e Filogenia	16
2.2.2	Características morfológicas e importância econômica	16
2.2.3	Caracterização citogenética do gênero <i>Phaseolus</i>	17
2.2.3.1	O Clado <i>Leptostachyus</i>	18
2.2.4	Principais sequências repetitivas em <i>Phaseolus</i>	19
<b>3</b>	<b>ARTIGO 1: ANÁLISE COMPARATIVA DO DNA REPETITIVO E SUAS IMPLICAÇÕES EM UM GRUPO DE EVOLUÇÃO CARIOTÍPICA ACELERADA</b>	<b>22</b>
<b>4</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>55</b>
	<b>REFERÊNCIAS</b>	<b>56</b>
	<b>ANEXO A – NORMAS DA REVISTA</b>	<b>61</b>

## 1 INTRODUÇÃO

A especiação e a diversificação das plantas estão relacionadas a mudanças no genoma nuclear. O genoma é constituído por sequências de cópia única e por até 90% de sequências repetitivas, segmentos de DNA que estão presentes em múltiplas cópias. Essas sequências repetitivas são consideradas geradoras de grandes diferenças entre os genomas, uma vez que acumulam variações na sequência e número de cópias mais rapidamente que as sequências únicas durante a evolução, inclusive entre espécies próximas (Mehrotra e Goyal 2014).

As alterações no genoma também podem afetar desde um segmento cromossômico a um conjunto cromossômico inteiro, sendo tanto numéricas (poliploidia e disploidia) quanto estruturais (rearranjos cromossômicos) (Guerra 2008; Stormea e Masonb 2014). Geralmente, essas mudanças na configuração do cariótipo resultam principalmente de eventos de recombinação entre sequências repetitivas similares que podem levar a fusões, fissões, translocações e inversões cromossômicas (Schubert e Lysak 2011).

Mais recentemente, os avanços nas tecnologia de sequenciamento genômico de nova geração (NGS) têm permitido explorar genomas de um número cada vez maior de espécies, permitindo caracterizar inclusive a fração repetitiva do genoma (Novák *et al.* 2013; Veeckman *et al.* 2016). O DNA repetitivo pode estar distribuído de forma dispersa, como elementos transponíveis, ou restrito a locais específicos em uma configuração em tandem (Mehrotra e Goyal, 2014). Em muitas espécies os elementos transponíveis são o principal representante da fração repetitiva do genoma, chegando a representar cerca de 75% do genoma do milho, 40% do genoma humano e 20% do genoma de *Arabidopsis thaliana* (Slotkin e Martienssen, 2007).

Não se sabe ao certo porque alguns grupos apresentam taxas de evolução cariotípica maiores que outros. Em *Arabidopsis*, uma forte redução no número de cromossomos foi promovida por uma taxa acelerada de rearranjos, principalmente inversões e translocações (Mandáková *et al.* 2010). Como as espécies de Brassicaceae passaram por vários ciclos de poliplodização e diploidização do genoma, é possível que as altas taxas de rearranjos cromossômicos possam estar associadas à sua origem poliploide.

No gênero *Phaseolus*, pertencente à família Fabaceae, exclusivamente diploide e relativamente estável com  $2n = 22$  (Fonsêca e Pedrosa-Harand 2013; Bonifácio *et al.* 2012), um pequeno grupo monofilético de três espécies com cerca de 1,7-3,4 milhões de anos, o grupo *Leptostachyus*, apresenta o cariótipo disploide ( $2n = 20$ ) (Delgado-Salina *et al.* 2006). Inúmeros rearranjos estruturais, incluindo uma inserção cêntrica que levou a disploidia e diversas translocações, algumas compartilhadas e outras não, foram observados no grupo, sugerindo

uma taxa de evolução cariotípica aumentada recentemente (Ferraz *et al.* 2016; Fonsêca *et al.* 2016).

O gênero *Phaseolus* apresenta algumas espécies de grande importância socioeconômica, como o feijão comum (*P. vulgaris*) e a fava (*P. lunatus*) (Broughton *et al.* 2003). Devido a isso, várias ferramentas genômicas estão disponíveis, desde bibliotecas BACs à sequências completas de genomas. Além disso, apresenta desde espécies com poucos rearranjos cromossômicos até espécies, como no grupo *Leptostachyus*, que apresentam uma alta taxa de rearranjos cromossômicos. Uma vez que a poliploidia não pode ser a causa dessa maior taxa de rearranjos no grupo *Leptostachyus*, as perguntas que motivaram este trabalho foram: 1) Há diferenças na composição da fração repetitiva entre os genomas do grupo *Leptostachyus* e dos demais grupos, cariotipicamente mais estáveis? 2) Uma ou mais famílias de sequências repetitivas estão associadas aos rearranjos presentes no grupo? Nossa hipótese é que há associação entre a taxa de evolução de sequências repetitivas e de rearranjos estruturais no grupo, por isso nossa predição é que haveria uma maior divergência na abundância de algumas famílias de DNA repetitivo e uma possível associação espacial entre essas famílias mais divergentes e os rearranjos previamente mapeados.

Sendo assim, o objetivo do trabalho foi investigar, por meio de análises bioinformática e citogenética, o conjunto de sequências repetitivas de duas espécies do Clado *Leptostachyus*, *P. leptostachyus* e *P. macvaughii*, e da espécie do clado irmão *P. lunatus*, relativamente estável estruturalmente, e a possível associação dessas sequências com eventos de rearranjos cromossômicos no grupo. Adicionalmente, foi feita uma análise comparativa da fração repetitiva incluindo as três espécies sequenciadas nesse trabalho e mais 12 espécies do gênero, cujos genomas já se encontravam disponíveis (Rendón-Anaya *et al.* 2017).

## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 SEQUÊNCIAS REPETITIVAS DO GENOMA EUCARIOTO

As sequências repetitivas de DNA estão presentes em todas as plantas superiores e podem representar até 90% do tamanho do genoma em algumas espécies. Consistem em segmentos de DNA que estão presentes em múltiplas cópias no genoma e podem estar distribuídas de forma dispersa, como elementos transponíveis, ou restritos a locais específicos numa configuração em tandem, ou seja com as várias cópias adjacentes uma após a outra (Mehrotra e Goyal, 2014).

As sequências repetidas em tandem podem ser classificadas de acordo com o tamanho do monômero em três grupos: microssatélites, minissatélites e DNA satélites (Mehrotra e Goyal, 2014). As quantidades de repetições em tandem variam significativamente em diferentes espécies de plantas. Apesar de geralmente estar em menor proporção que os elementos transponíveis, no rabanete (*Raphanus sativus* L.), por exemplo, sequências satélites são os elementos mais abundantes, correspondendo a cerca de 13% do genoma (He *et al.* 2015).

O DNA satélite (DNAsat) é um componente principal das regiões heterocromáticas de genomas eucariotos, sendo encontrado principalmente em regiões centroméricas e subteloméricas dos cromossomos e geralmente apresenta uma elevada dinâmica evolutiva (Biscotti *et al.* 2015). Uma sequência particular de DNAsat pode ser específica de uma espécie ou estar presente em muitas espécies indicando que algumas sequências repetitivas evoluem rapidamente, enquanto outras podem ser conservadas. O DNA centromérico é geralmente a repetição em tandem mais abundante encontrada em plantas e animais (Melters *et al.* 2013). Em *Rhynchospora pubera*, espécie com cromossomos holocêntricos, foi identificado o DNA satélite *Tyba* presente ao longo de todo o holocentrômero e associado à CENH-3 (Marques *et al.* 2015). Já as famílias de sequências repetitivas subteloméricas geralmente são gênero ou espécie específicas, como é o caso da família TRI no cromossomo Y de *Silene latifolia* (Cermak *et al.* 2008).

Os elementos transponíveis são elementos móveis capazes de se inserir no genoma e se replicar de forma autônoma. Podem ser classificados de acordo com a composição e forma de replicação em: retrotransposons (Classe I), os quais se replicam através de intermediários de RNA e cDNA, ou transposons de DNA (Classe II), com excisão e reintegração do próprio elemento de DNA (Lambowitz *et al.* 2015). Esses elementos variam em abundância entre os

genomas de eucariotos, representando cerca de 75% do genoma do milho, 40% do genoma humano e 20% do genoma *Arabidopsis thaliana* (Slotkin e Martienssen, 2007).

Os elementos da classe I são divididos em retrotransposons LTR (*Long Terminal Repeats*; Longas Repetições Terminais) e retrotransposons não LTR, baseado na estrutura do elemento e no mecanismo de transposição. Os retrotransposons não LTR são frequentes nos genomas de mamíferos e correspondem aos elementos LINE (*Long Interspersed Nuclear Element*, Elementos Nucleares Interdispersos Longos), que são os elementos longos intercalados autônomos, e SINE (*Short Interspersed Nuclear Element*, Elementos Nucleares Interdispersos Longos) que são os elementos curtos intercalados não autônomos, uma vez que dependem dos LINEs para amplificação (Wessler, 2006).

Por sua vez, os retrotransposons LTR são classificados em Ty3-*gypsy*, Ty1-*copia*, retrovírus, BEL e DIRS1 (Goodwin e Poulter, 2002), caracterizados por apresentarem uma sequência terminal LTR e conter ao menos dois genes, *gag* e *pol*, responsáveis por codificar o capsídeo, a protease, a transcriptase reversa e a integrase (Wessler, 2006). Esses elementos são os principais responsáveis pela variação do tamanho do genoma em diversos grupos de plantas, como foi visto na tribo *Fabeae*, onde uma única linhagem de retrotransposons LTR Ty3-*gypsy*, os elementos Ogre, representa a maior parte dessa variação (Macas *et al.* 2015), e em *Panax*, onde a subfamília *PgDell*, também Ty3-*gypsy*, foi responsável por grande parte da variação do tamanho do genoma das espécies analisadas (Lee *et al.* 2017).

Evidências recentes sugerem que a fração repetitiva pode ter grandes efeitos na função e evolução dos genomas e cromossomos de plantas e animais e que diferenças nessa fração podem refletir as distâncias evolutivas entre as espécies (Mehrotra e Goyal 2014; Zhu *et al.* 2016; Ayres-Alves *et al.* 2017). Eventos de transposição também podem introduzir mutações e controlar a expressão gênica (Lambowitz *et al.* 2015), além de levar a rearranjos estruturais, como observado na espécie modelo *Arabidopsis thaliana* (Fransz *et al.* 2016). Tudo isso torna o estudo dessa fração do genoma um atrativo para pesquisadores interessados no assunto.

### 2.1.1 Caracterização das sequências repetitivas do genoma

As sequências repetitivas podem ser caracterizadas através de diferentes técnicas. Inicialmente, as análises se baseavam no princípio da cinética de reassociação do DNA (análise de  $C_0t$ ), uma vez que quanto mais repetida é uma sequência no genoma, mais rápida sua renaturação (Peterson *et al.* 2002). Utilizando-se da fração  $C_0t$  em *Arachis hypogaea* L. foi

possível identificar novos *repeats* que permitiram distinguir citogeneticamente os cromossomos correspondentes aos genomas A e B do amendoim tetraploide (Zhang *et al.* 2012).

Com o avanço da tecnologia, novas técnicas de sequenciamento além de novas metodologias de caracterização surgiram para otimizar esse processo, permitindo que mais sequências pudessem ser analisadas e de forma mais detalhada. No sequenciamento de nova geração (NGS) milhões de sequências são geradas simultaneamente e com grande qualidade e robustez, além de baixo ruído (Buermans e Dunnen 2014). Além disso, o preço do sequenciamento tem se apresentado mais acessível, o que torna a metodologia mais atrativa para o estudo dos genomas de um número cada vez maior de espécies.

O sequenciamento de genomas inteiro em baixas coberturas (por exemplo, 0,1× do genoma) permite a obtenção da fração repetitiva. Essas sequências podem ser analisadas utilizando ferramentas como a plataforma Galaxy/RepeatExplorer (<https://repeatexplorer-elixir.cerit-sc.cz/galaxy/>), que tem como princípio o agrupamento de sequências por similaridade (*clustering*) ao comparar par-a-par todos os *reads* gerados no sequenciamento. Com isso, permite caracterizar a composição e abundância dessas sequências no genoma (Novák *et al.* 2010, 2013). Outra ferramenta computacional mais específica para análise dos elementos dispostos em tandem é o TAREAN, que permite detecção de DNAs satélite de alta confiança (Novák *et al.* 2017). Os dados obtidos podem ser analisados individualmente ou ainda em abordagens genômicas comparativas. Na tribo Fabeae, por exemplo, o acúmulo diferencial de DNA repetitivo justificou 85% da variação de tamanho do genoma encontrada entre as 23 espécies analisadas, e a maior parte dessa diferença foi resultado de uma única linhagem de retrotransposons LTR (Macas *et al.* 2015).

As sequências repetitivas encontradas podem ser amplificadas por PCR com os *primers* específicos e visualizadas citogeneticamente através da técnica de hibridização *in situ* fluorescente (FISH), que consiste no pareamento de determinado segmento marcado de DNA ou RNA com uma sequência de nucleotídeos complementar situada dentro do núcleo ou citoplasma (Guerra, 2004). Em *Paphiopedilum*, o DNA satélite SatA foi identificado e caracterizado nas regiões centroméricas, pericentroméricas e subteloméricas dos cromossomos de todas as sete espécies do subgênero *Parvisepalum*, com padrão específico para cada espécie, o que revela a rápida evolução dessa sequência (Lee *et al.* 2018).

### 2.1.2 Evolução das sequências repetitivas e rearranjos cromossômicos

Sequências repetitivas são consideradas geradoras de grandes diferenças entre os genomas, uma vez que durante a evolução acumulam variações na sequência e no número de cópias mais rapidamente que as sequências únicas, inclusive entre espécies proximamente relacionadas (Mehrotra e Goyal, 2014). As mudanças no genoma também podem afetar desde um segmento cromossômico a um conjunto cromossômico inteiro, sendo tanto numéricas (poliploidia e disploidia) quanto estruturais (alterações cromossômicas) (Guerra, 2008; Stormea e Masonb, 2014).

Não se sabe ao certo porque alguns grupos apresentam taxas de evolução cariotípica maiores que outros. Em *Arabidopsis*, uma forte redução no número de cromossomos foi promovida por uma taxa acelerada de rearranjos, principalmente inversões e translocações (Mandáková *et al.* 2010). Geralmente, essas mudanças na configuração do cariótipo resultam principalmente de eventos de recombinação entre sequências repetitivas similares que podem levar a fusões, fissões, translocações e inversões cromossômicas (Schubert e Lysak, 2011). Na espécie modelo *A. thaliana*, a inversão paracêntrica encontrada no cromossomo 4 foi resultado da atividade do transposon Vandal que realocou um segmento de heterocromatina pericêntrica e mostrou forte associação a uma maior tolerância a seca (Fransz *et al.* 2016). Uma vez que a disploidia descendente (com redução no número cromossômico) é acompanhada pela perda de centrômeros e telômeros, a redução do número de cromossomos contribui para o encolhimento rápido e diploidização de um genoma poliploide (Schubert *et al.* 2016). Como as espécies de Brassicaceae passaram por vários ciclos de poliplodização e diploidização do genoma, é possível que as altas taxas de rearranjos cromossômicos possam estar associadas à sua origem poliploide (Dodsworth *et al.* 2016).

Mudanças nas sequências de DNA repetitivo e alterações cromossômicas são frequentemente relacionadas à diversificação de espécies e especiação. No entanto, embora a poliploidia seja frequentemente implicada na diversificação vegetal e relacionada à reestruturação genômica, não está claro qual o papel das sequências repetitivas para a evolução cromossômica ou quais os fatores mais relevantes em espécies diploides.

## 2.2 O GÊNERO *PHASEOLUS*

### 2.2.1 Taxonomia e Filogenia

O gênero *Phaseolus* L. pertence à família Fabaceae, terceira maior família entre as Angiospermas, está incluída na ordem Fabales, classe Magnoliopsida. Apresenta cerca de 100 espécies, popularmente conhecidas como feijões e distribuídas pelas Américas, que constitui seu centro de origem (Broughton *et al.* 2003).

Pela sua grande importância, tanto na nutrição humana quanto como forragem para gado, o gênero é foco de muitos estudos agrônômicos, sistemáticos e genéticos. Uma análise a partir de dados combinados de DNA ribossomal ITS/5,8S e de um loco plastidial *trnK* revelou seu monofiletismo e a existência de dois cladogramas principais (clados A e B), que se subdividem em oito cladogramas, além de quatro espécies independentes: *Phaseolus glabellus*, *P. macrolepis*, *P. microcarpus* e *P. oaxacanus* (Delgado-Salinas *et al.* 2006).

As espécies do clado A são distribuídas principalmente no México e não ocorrem na América do Sul nem em ilhas oceânicas. De forma geral, florescem apenas durante a estação chuvosa (exceto *P. microcarpus*), são sensíveis à perturbação do habitat e geralmente não toleram um longo período de geada. Por outro lado, as espécies do clado B são mais amplamente distribuídas pelas Américas e conseguem habitar ilhas. Florescem durante a estação seca ou chuvosa, não são sensíveis a perturbações e algumas toleram um longo período de geada (Delgado-Salinas *et al.* 2006).

Delgado-Salinas e colaboradores (2006) estimaram a idade aproximada do gênero *Phaseolus* entre 6-8 Ma. Entre os 8 cladogramas, o grupo *Vulgaris* é o mais antigo com cerca de 1,8-3,9 Ma, seguido de *Leptostachyus* com 1,7-3,4 Ma, enquanto que os mais novos incluem os grupos *Filiformis*, *Pedicellatus* e *Polystachios* com apenas 0,6-1,4 Ma.

### 2.2.2 Características morfológicas e importância econômica

De forma geral, o gênero é composto por plantas eudicotiledôneas com flores perfeitas que geralmente se autofecundam. São plantas herbáceas anuais, com um sistema radicular composto por uma raiz principal pivotante e sistema caulinar composto por uma haste principal ramificada ereta ou não. As folhas são compostas de três folíolos, sendo um terminal e dois laterais e opostos. As flores apresentam um cálice verde gamossépalo e a corola composta por cinco pétalas brancas, amareladas, rosadas ou violáceas, constituindo a maior o estandarte, as

médias, as duas asas e as duas últimas soldadas, a quilha. O gineceu é composto por um único pistilo com o ovário súpero e o androceu é composto por 10 estames. O fruto silvestre é deiscente do tipo legume, característico da família, onde as valvas se contorcem em espiral após deiscência e as sementes podem apresentar tamanho e coloração variada (Silva e Costa, 2003).

O gênero dos feijões inclui no clado B pelo menos cinco espécies cultivadas em todos os continentes do mundo, excetuando a Antártica, e utilizadas na alimentação humana: *P. vulgaris* L., *P. lunatus* L., *P. coccineus* L., *P. polyanthus* Greenman e *P. acutifolius* A. Gray. Entre elas, *P. vulgaris*, o feijão comum, compreende a fonte de proteína mais importante para diversas populações, principalmente na África e América Latina (Broughton *et al.* 2003), complementando outros alimentos ricos em carboidratos. Além de fonte de proteína, essa espécie também é fonte de vitaminas do complexo B, sais minerais, ferro, cálcio e fósforo, suprimindo aproximadamente 20% do requerimento diário desses nutrientes em um adulto. Outra espécie economicamente importante, *P. lunatus* (a fava), apesar de ser menos cultivada que o feijão comum, é uma importante alternativa de alimento e renda para populações locais em regiões do nordeste do Brasil (Oliveira *et al.* 2004).

### 2.2.3 Caracterização citogenética do gênero *Phaseolus*

Citogeneticamente, o gênero *Phaseolus* é caracterizado por ser exclusivamente diploide. A maioria das espécies apresenta  $2n = 22$  e número básico  $x = 11$  e um cariótipo simétrico composto predominantemente de cromossomos metacêntricos e submetacêntricos pequenos e morfologicamente similares (Mercado-Ruaro e Delgado-Salinas, 2000).

A partir de dados convencionais, os principais fatores relatados para explicar a evolução cariotípica em *Phaseolus* são rearranjos cromossômicos como inversões pericêntricas ou paracêntricas, translocações e perda ou ganho de cromatina (Sarbhoy, 1977, 1980; Sinha e Roy, 1979). Por meio da FISH com BACs e bacteriófagos, Pedrosa-Harand e colaboradores (2009) mapearam primeiramente três cromossomos da espécie. Posteriormente, utilizando além de BACs e bacteriófago, dois plasmídeos contendo as sequências do DNAr 5S e 35S, Fonsêca e colaboradores (2010) mapearam os oito pares restantes do feijão do genótipo mesoamericano 'BAT93'. Os resultados somados da comparação com os correspondentes mapas genéticos, tornaram disponíveis 43 pontos de ancoragem entre os mapas genético e físico citogenético.

O mapeamento comparativo realizado com a fava (*P. lunatus*) resultou no estabelecimento das homeologias cromossômicas e os mecanismos evolutivos que moldaram

seu cariótipo, quando comparado a *P. vulgaris*. Assim como o feijão comum, a fava está incluída no clado B, o que as tornam filogeneticamente próximas. Todos os 11 pares de cromossomos, com exceção do par 5 que foi identificado pela ausência de marcadores, foram identificados a partir da FISH utilizando 16 clones de cópia única. Os resultados revelam uma significativa conservação da sintenia, pois todos os BACs hibridizaram nos mesmos pares cromossômicos em ambas as espécies. Apesar disso, foram detectadas pelo menos três inversões pericêntricas (Bonifácio *et al.* 2012), com quebras de colinearidade das sequências.

A mesma conservação de sintenia foi observada para *P. microcarpus*, uma espécie selvagem do gênero pertencente ao clado A, e portanto mais distante filogeneticamente de *P. vulgaris*. Nessa espécie, 24 BACs apresentaram sinais únicos e foram utilizados para identificação dos 11 pares cromossômicos. No entanto, quatro quebras na colinearidade, provavelmente causadas por inversões cromossômicas, foram detectadas pela comparação dos dois genomas (Fonsêca e Pedrosa-Harand, 2013).

A FISH revelou ainda, para ambas as espécies analisadas, a presença de sítios de DNAr 5S e 35S nos cromossomos 10 e 6, respectivamente (Bonifácio *et al.* 2012; Fonsêca e Pedrosa-Harand, 2013), assim como observado em *P. vulgaris* (Fonsêca *et al.* 2010). Isso mostra que esses sítios se encontram conservados nesses cromossomos nas espécies do gênero, apesar de existirem alguns sítios adicionais não conservados em outros cromossomos em algumas espécies.

### 2.2.3.1 O clado *Leptostachyus*

Apesar da grande estabilidade numérica e relativa estabilidade estrutural do gênero (Bonifácio *et al.* 2012; Fonsêca e Pedrosa-Harand, 2013), um pequeno clado de três espécies, que constituem o grupo *Leptostachyus* (clado B), apresenta um cariótipo disploide com  $2n = 20$  e um par cromossômico relativamente maior que os demais (Mercado-Ruaro e Delgado-Salinas, 1998). O grupo inclui *P. leptostachyus*, *P. micranthus* e *P. macvaughii*, sendo a primeira de diversificação mais recente e a última mais basal na filogenia de Delgado-Salinas e colaboradores (2006).

Para o grupo *Leptostachyus*, Mercado-Ruaro e Delgado-Salinas (1998), a partir da análise do número e morfologia cromossômica, sugeriram que o evento de disploidia descendente também esteve envolvido na evolução do cariótipo dessas espécies, porém o mecanismo envolvido não havia sido esclarecido. A análise citogenética comparativa foi realizada até o momento em *P. macvaughii* e *P. leptostachyus* através da FISH com sequências

de *P. vulgaris* e revelou um evento único que deu origem a dispoloidia do grupo a partir de uma inserção cêntrica envolvendo os cromossomos 10 e 11. No entanto, rearranjos adicionais foram observados associados a esse evento, dentre eles várias translocações e inversões, algumas exclusivas e outras compartilhadas (Fonsêca *et al.* 2016; Ferraz *et al.* 2016).

Em *P. leptostachyus* foram observadas translocações envolvendo os cromossomos 1, 2, 3, 4, 6, 7, 8 e 9, além de duas inversões nos cromossomos 3 e 7, do tipo paracêntrica e pericêntrica respectivamente (Fonsêca *et al.* 2016). Por outro lado, *P. macvaughii* apresentou apenas uma translocação de parte do cromossomo 2 para o cromossomo 1 e inversões pericêntricas envolvendo os cromossomos 3 e 9. A presença de rearranjos exclusivos para as duas espécies após a inserção cêntrica, que levou a dispoloidia, sugere não um momento único de grande reestruturação genômica, mas uma taxa de evolução cariotípica elevada em um curto intervalo de tempo, com rearranjos sucessivos e independentes desde o surgimento do grupo há cerca de 1,7-3,4 milhões de anos (Ferraz *et al.* 2016).

#### 2.2.4 Principais sequências repetitivas em *Phaseolus*

Além de uma detalhada caracterização estrutural (Fonsêca e Pedrosa-Harand 2013; Bonifácio *et al.* 2012; Fonsêca *et al.* 2016; Ferraz *et al.* 2016), várias sequências repetitivas foram caracterizadas (David *et al.* 2009; Iwata *et al.* 2013; Richard *et al.* 2013; Gao *et al.* 2014; Ribeiro *et al.* 2017) e o genoma completo do feijão comum (*P. vulgaris*) e de outras espécies do gênero foram sequenciados (Schmutz *et al.* 2014; Rendón-Anaya *et al.* 2017) e estão disponíveis ([www.ncbi.nlm.nih.gov/genbank/](http://www.ncbi.nlm.nih.gov/genbank/)). Assim, o grupo se apresenta como bom modelo para estudar a evolução cariotípica em plantas diploides.

O genoma de *P. vulgaris* (feijão comum), espécie de maior importância econômica e mais bem estudada do gênero, é constituído por aproximadamente 45% de sequências repetitivas (Schmutz *et al.* 2014), com predomínio de blocos heterocromáticos pericentroméricos, subteloméricos e de DNAr (Fonsêca *et al.* 2010). Mais de 5.000 elementos transponíveis foram observados em *P. vulgaris* (Gao *et al.* 2014), como o pvCACTA1, aparentemente restrito ao gênero (Gao *et al.* 2014, 2016). A maioria dos elementos repetitivos do genoma são retrotransposons LTR, integrados nos últimos 2 milhões de anos, que cobrem cerca de 37% do genoma e estão classificados em 165 famílias, incluindo 65 Ty1-*copia*, 78 Ty3-*gypsy* e 22 famílias não classificadas. Cerca de 20% desses elementos foram inseridos no genoma do feijão comum a menos de 0,5 milhão de anos, sugerindo que os retrotransposons

LTR sofreram eventos de amplificação recentes nessa espécie, assim como em outras leguminosas (Schmutz *et al.* 2014).

A análise comparativa também revelou a predominância de retrotransposons LTR em outras espécies do gênero e em espécies do gênero relacionado *Vigna*, com exceção de *P. acutifolius* que apresentou sequências de DNA satélite como os elementos repetitivos mais abundantes do genoma. Dentre os retrotransposons, a superfamília Ty3-*gypsy*, em particular os elementos Ogre-Tat e Chromovirus, foi encontrada em maior proporção em todas as espécies estudadas, com exceção novamente de *P. acutifolius*, que apresentou predomínio da família Ty1-*copia* (Ribeiro *et al.* 2019).

Em eucariotos superiores, os centrômeros contêm tipicamente várias cópias de uma família de DNAsat. As sequências centroméricas CentPv1 e CentPv2, que evoluíram independentemente, estão localizadas em dois conjuntos distintos de cromossomos. O CentPv1 está organizado em estruturas de repetição de ordem superior (HOR) referidas como *Nazca* e são altamente abundantes nos centrômeros de *P. vulgaris* e encontradas em menor número de cópias em outras espécies do grupo *Vulgaris*, sugerindo uma rápida evolução das repetições centroméricas no gênero. Já o DNAsat CentPv2 é constituído de monômeros organizados de forma aleatória (Iwata *et al.* 2013; Iwata-Otsubo *et al.* 2016).

Uma sequência satélite originada do IGS do DNAr, denominada CC4, também foi isolada do feijão comum (Falquet *et al.* 1997). Essa sequência, exclusiva de *Phaseolus*, está presente na região centromérica ou pericentromérica das espécies do gênero, co-localizada ou não com o DNAr 35S (Almeida *et al.* 2012). Outras sequências centroméricas foram identificadas para *P. coccineus* e para *P. acutifolius* (Ribeiro *et al.* 2019).

Uma outra sequência repetitiva, denominada PvMeso, foi isolada e caracterizada sendo restrita à região terminal do cromossomo 7 nos acessos mesoamericanos de *P. vulgaris*. Os dados indicaram que uma repetição do cromossomo 11 foi amplificada no cromossomo 7 após o isolamento dos pools gênicos andinos e mesoamericanos (Ribeiro *et al.* 2011).

Uma proporção menor da heterocromatina está localizada em regiões subteloméricas contendo a sequência de satélites *kipu*, bem como o DNAr 35S. O DNA satélite *kipu*, ligado ao grupo de genes de resistência B4 do feijão comum (Geffroy *et al.*, 2009), é específico para o gênero e apresenta uma grande variação intragenérica no número de cópias (David *et al.* 2009; Richard *et al.* 2013; Ribeiro *et al.* 2017). A análise genômica dessa sequência revelou uma extensa recombinação entre cromossomos não homólogos nas regiões subteloméricas e também sugeriu a troca de sequências entre subtelômero e centrômero (Richard *et al.* 2013).

Inserido no espaçador não transcrito (NTS) do DNAr 5S no genoma de *P. vulgaris*, foi identificada a sequência *jumper*. Essa sequência está presente em todas as espécies de *Phaseolus*, mas é amplificada como DNAsat independente apenas em *P. microcarpus*, na qual se mostra associada ou não a *kipu* em subtelômeros e centrômeros (Ribeiro *et al.* 2017).

No geral, a distribuição das repetições de satélite centroméricos e subteloméricos é estável, ao passo que o número de cópias das repetições varia, indicando uma rápida amplificação/redução das repetições em regiões genômicas específicas (Iwata-Otsubo *et al.* 2016). Para o grupo *Leptostachyus*, apenas *kipu* foi investigado até o momento, revelando uma sequência abundante com sinais de hibridização em 18 dos 20 cromossomos de *P. leptostachyus* (Ribeiro *et al.* 2017) e ausente ou em baixa concentração em *P. macvaughii* (Ferraz *et al.* 2016).

### 3 ARTIGO 1 A SER SUBMETIDO À REVISTA CHROMOSOMA

## **Análise comparativa do DNA repetitivo e suas implicações em um grupo de evolução cariotípica acelerada**

Ferraz ME<sup>1</sup>, Ribeiro T<sup>2</sup>, Sader MA<sup>1</sup>, Pedrosa-Harand A<sup>1</sup>

<sup>1</sup>Laboratório de Citogenética e Evolução Vegetal, Departamento de Botânica, Universidade Federal de Pernambuco – Recife – PE/ Brasil

<sup>2</sup>Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso – Diamantino-MT/Brasil  
Autor para correspondência: Andrea Pedrosa-Harand andrea.harand@ufpe.br

### **Resumo**

A especiação e a diversificação das plantas estão relacionadas a mudanças no genoma nuclear. Mudanças estruturais no cariótipo resultam principalmente de eventos de recombinação entre sequências repetitivas. Em *Phaseolus*, cariotipicamente estável, o grupo *Leptostachyus* (*P. leptostachyus*, *P. micranthus* e *P. macvaughii*) apresenta alta taxa de rearranjos, incluindo uma inserção cêntrica compartilhada que levou a disploidia do clado. Este trabalho teve como objetivo analisar, de forma individual e comparada a outras espécies do gênero, a fração de DNA repetitivo de *P. leptostachyus* e *P. macvaughii* e sua possível relação com a ocorrência de rearranjos. Para isso, as sequências genômicas de baixa cobertura foram agrupadas por similaridade no RepeatExplorer e a ferramenta TAREAN foi usada para identificação de DNAs satélites. As principais sequências foram localizadas *in situ*. Retrotransposons LTR, mais precisamente Ty3/gypsy da linhagem Chromovirus, foram os elementos repetitivos mais abundantes. Diferenças marcantes em abundância das principais linhagens foram observadas entre as espécies de *Leptostachyus*. A maior divergência foi observada na fração de DNA satélite e três novas famílias foram caracterizadas. Os satélites revelaram padrão de hibridização

distintos nas espécies analisadas, confirmando a rápida evolução dessas sequências no genoma. A distribuição atípica de um dos satélites se mostrou compatível com uma possível associação dessa família e rearranjos cromossômicos frequentes no grupo.

Palavras-chaves: Chromovirus, DNA satélite, rearranjos cromossômicos, retrotransposons

## Introdução

Sequências repetitivas consistem em segmentos de DNA que estão presentes em múltiplas cópias no genoma de todas as espécies de plantas e podem chegar a até 90% do DNA total em algumas espécies (Mehrotra e Goyal 2014). Apesar de serem, por muito tempo, consideradas “DNA lixo”, evidências recentes sugerem que sequências repetitivas podem ter grandes efeitos na função e evolução dos genomas e cromossomos de plantas e animais (Mehrotra e Goyal 2014; Zhu *et al.* 2016; Ayres-Alves *et al.* 2017).

Essas sequências podem estar distribuídas de forma dispersa, como elementos transponíveis, ou em uma configuração em tandem, como os DNA satélites. As sequências repetidas em tandem são classificadas de acordo com o tamanho da unidade de repetição em microssatélites, minissatélites e DNA satélites (Mehrotra e Goyal 2014) e geralmente se apresentam em menor proporção no genoma que os elementos transponíveis. O DNA satélite (DNAsat) é um componente principal das regiões heterocromáticas, sendo encontrado principalmente em regiões centroméricas e subteloméricas dos cromossomos e geralmente apresenta uma rápida taxa de evolução, tanto em sequência como em número de cópias (Biscotti *et al.* 2015). Uma sequência particular de DNAsat pode ser específica de um cromossomo ou de uma espécie, ou estar presente em muitas espécies de um grupo, indicando que algumas sequências repetitivas evoluem mais rapidamente que outras (Robledillo *et al.* 2018; Cermak *et al.* 2008; Gong *et al.* 2012; Zhang *et al.* 2014).

Os elementos transponíveis são elementos móveis capazes de se inserir no genoma e se replicar e são classificados como retrotransposons (Classe I) ou transposons de DNA (Classe II) (Lambowitz *et al.* 2015). Os elementos da classe I são divididos em retrotransposons LTR (*Long Terminal Repeats*) e não LTR. Os retrotransposons não LTR são frequentes nos genomas de mamíferos e correspondem aos elementos LINE (*Long Interspersed Nuclear Element*) e SINE (*Short Interspersed Nuclear Element*) (Wessler 2006). Já os retrotransposons LTR, mais abundantes em vegetais, são classificados em Ty3-gypsy, Ty1-copia, Bel-Pao, retrovírus e retrovírus endógeno (ERV) (Wicker *et al.* 2007) caracterizados por apresentarem uma sequência terminal LTR e conter ao menos dois genes, *gag* e *pol* (Wessler 2006). Esses elementos são os principais responsáveis pela variação do tamanho do genoma em diversos grupos de plantas, como foi visto na tribo Fabaeae, onde uma única linhagem de retrotransposons LTR Ty3-gypsy, os elementos *Ogre*, representa a maior parte dessa variação (Macas *et al.* 2015).

O sequenciamento de nova geração (NGS) de genomas inteiros em baixa cobertura (*genome skimming*), associado a algoritmos especializados de bioinformática, tem permitido a ampla caracterização da fração repetitiva do genoma de um número cada vez maior de espécies (Novák *et al.* 2013). Diferenças nas proporções de elementos repetitivos podem refletir as distâncias evolutivas entre as espécies (Zhu *et al.* 2016; Ayres-Alves *et al.* 2017; Dodsworth *et al.* 2017). Além disso, eventos de transposição de elementos transponíveis, por exemplo, podem introduzir mutações e controlar a expressão gênica (Lambowitz *et al.* 2015), assim como levar a rearranjos estruturais (Raskina *et al.* 2004, Fransz *et al.* 2016, Siljak-Yakovlev *et al.* 2017).

Não se sabe ao certo porque alguns grupos apresentam taxas de evolução cariotípica maiores que outros. A poliploidia é frequentemente implicada na diversificação vegetal e relacionada à reestruturação genômica. Após eventos de poliploidia, os genomas passam, de forma mais ou menos rápida, por um fenômeno de diploidização, processo complexo que inclui muitas vezes vários rearranjos cromossômicos para redução numérica (Li *et al.* 2017). Em *Arabidopsis*, uma forte redução no número de cromossomos foi promovida por uma taxa acelerada de rearranjos, principalmente inversões e translocações (Mandáková *et al.* 2010). A passagem por vários ciclos de poliplodização e diploidização do genoma pode estar relacionada com as altas taxas de rearranjos cromossômicos nas espécies poliploides (Dodsworth *et al.* 2016). Contudo, como existe na família Brassicaceae poliploides estáveis (Mandáková *et al.* 2010b), outros fatores podem estar relacionados à ocorrência de uma maior taxa de rearranjos nesse grupo.

As mudanças estruturais do cariótipo resultam principalmente de eventos de recombinação entre sequências repetitivas similares que podem levar a fusões, fissões, translocações e inversões cromossômicas (Schubert e Lysak 2011). Em *A. thaliana*, a inversão paracêntrica encontrada no cromossomo 4 foi resultado da atividade do transposon Vandal, que realocou um segmento de heterocromatina pericêntrica e mostrou forte associação a uma maior tolerância à seca (Fransz *et al.* 2016). Não está claro, no entanto, qual o papel das sequências repetitivas para a evolução cromossômica de forma mais ampla ou quais os fatores mais relevantes em espécies diploides.

O gênero *Phaseolus*, que apresenta algumas espécies de grande importância socioeconômica, como o feijão comum (*P. vulgaris*) e a fava (*P. lunatus*) (Broughton *et al.* 2003), é caracterizado por ser exclusivamente diploide, com  $2n = 22$  na grande maioria das espécies (Fonsêca e Pedrosa-Harand 2017). Apesar de uma grande estabilidade numérica (Mercado-Ruaro e Delgado-Salinas 2000) e relativa estabilidade estrutural (Fonsêca e Pedrosa-

Harand 2013; Bonifácio *et al.* 2012), um pequeno grupo monofilético de três espécies com cerca de 1,7-3,4 milhões de anos, o grupo *Leptostachyus*, apresenta um cariótipo dispoloide ( $2n = 20$ ) (Mercado-Ruaro e Delgado-Salinas 1998). Inúmeros rearranjos estruturais, incluindo uma inserção cêntrica que levou à disploidia, e diversas translocações, algumas compartilhadas e outras não, foram observadas no grupo. Os rearranjos foram em menor número na espécie que divergiu primeiro, *P. macvaughii*, e mais numerosos na espécie de diversificação mais recente, *P. leptostachyus*, sugerindo uma taxa de evolução cariotípica aumentada recentemente (Ferraz *et al.* 2016; Fonsêca *et al.* 2016).

Além de uma detalhada caracterização estrutural (Fonsêca e Pedrosa-Harand 2017), várias sequências repetitivas foram identificadas no gênero (David *et al.* 2009; Iwata *et al.* 2013; Richard *et al.* 2013; Gao *et al.* 2014; Ribeiro *et al.* 2017). Um DNA satélite abundante no grupo *Vulgaris*, *kipu*, foi associado ao grande dinamismo da região subtelmérica e a rápida evolução de genes de resistência a doenças (Chen *et al.* 2018). O genoma completo do feijão comum (*P. vulgaris*) e de outras espécies do gênero foram sequenciados e estão disponíveis (Schmutz *et al.* 2014; Rendón-Anaya *et al.* 2017). Assim, o grupo se apresenta como bom modelo para estudar a evolução cariotípica em plantas diploides, particularmente investigando se há relação entre uma maior ocorrência de rearranjos cromossômicos com sequências repetitivas recém amplificadas.

Dessa forma, o objetivo do trabalho foi investigar, por meio de análises bioinformáticas e citogenéticas, o conjunto de sequências repetitivas de espécies do grupo *Leptostachyus* e a possível associação dessas sequências com eventos de rearranjos cromossômicos no grupo. Para isso, além da caracterização detalhada de três espécies, duas do grupo *Leptostachyus* e uma do grupo externo *Lunatus*, foi feita uma análise comparativa da fração repetitiva com mais 12 genomas do gênero *Phaseolus* (Rendón-Anaya *et al.* 2017; Ribeiro *et al.* 2019).

## Material e Métodos

### *Material vegetal*

Amostras de duas espécies do grupo *Leptostachyus*, *Phaseolus leptostachyus* (179671; CENARGEN – Brasília), *P. macvaughii* (G40656; CIAT-Colômbia), e uma do grupo *Lunatus*, *P. lunatus* ‘Vermelhinha’ (GL0135; EMBRAPA Arroz e Feijão - Góias), foram mantidas e multiplicadas no Jardim Experimental do Laboratório de Citogenética e Evolução Vegetal do Departamento de Botânica da Universidade Federal de Pernambuco (UFPE). *Phaseolus lunatus* foi usada como uma espécie mais próxima ao grupo *Leptostachyus* e cariotipicamente estável.

### *Preparação de lâminas*

Após germinação das sementes, as raízes foram coletadas, pré-tratadas em 8-hidroxiquinolina 2 mM por 20 h a 10 °C e fixadas em metanol:ácido acético (3:1, v/v). As raízes foram submetidas a uma digestão com 2% celulase (Onozuka) e 20% pectinase (Sigma) por 1 h e 30 min a 37°C. O preparo de lâminas foi realizado por secagem ao ar segundo o protocolo modificado de Carvalho e Saraiva (1993). Resumidamente, as raízes digeridas foram transferidas para lâminas inclinadas, lavadas cerca de 5 vezes com várias gotas de fixador gelado metanol:ácido acético (3:1) à medida que o material foi macerado com agulhas e secas com auxílio de uma bomba de ar. Por fim, as lâminas foram mergulhadas em ácido acético 45% por 5 min e incubadas para secar a 37 °C. As melhores lâminas foram, então, selecionadas após análise com coloração em 0,1 µg/µl DAPI em 50% glicerol, descoradas em Carnoy (etanol: ácido acético, 3:1) por 30 min seguido de etanol absoluto por 1 h e armazenadas a -20 °C.

### *Sequenciamento de nova geração e identificação de sequências repetitivas*

A extração de DNA genômico para as três espécies sequenciadas no presente trabalho foi realizada segundo protocolo de extração CTAB de Weising *et al.* (2005) utilizando material fresco. Os DNAs extraídos foram sequenciados na plataforma Illumina HiSeq2500 (BGI, Hong Kong) com a cobertura de 0,1x, considerando para *P. leptostachyus* 1C = 513 Mpb, *P. macvaughii* 1C = 440 Mpb e *P. lunatus* 1C = 685 Mpb (Roya Botanic Gardens, Kew, <http://data.kew.org/cvalues/>, Bennet e Leitch 2012). Os *reads paired-end* de 250 pb gerados foram analisados, por espécie, em relação às suas sequências repetitivas através da plataforma RepeatExplorer usando parâmetros padrão (Nóvak *et al.* 2010, 2013; <https://repeatexplorer-elixir.cerit-sc.cz>). Para auxiliar na anotação das sequências, um banco de dados personalizado, incluindo sequências previamente caracterizadas (CentPv1/Nazca e CentPv2 - Iwata *et al.* 2013; Iwata-Otsubo *et al.* 2016; CC4 - Almeida *et al.* 2012, *juniper* - Ribeiro *et al.* 2017, *kipu* - David *et al.* 2009 e satélites espécie-específicos - Ribeiro *et al.* 2019) foi incluído na análise.

As sequências obtidas foram filtradas por qualidade, os *reads* foram pareados e agrupados por similaridade. A ferramenta TAREAN foi utilizada para identificar possíveis DNA satélites de alta e baixa confiança (Novák *et al.* 2017) que foram confirmados com uma análise de Dotplot usando o programa DOTTER (Sonnhammer e Durbin 1995).

### *Análise comparativa da fração repetitiva em espécies de Phaseolus*

Além das três espécies sequenciadas nesse trabalho, outros 12 genomas de espécies de

*Phaseolus* que incluíram *P. vulgaris* (G5886; SRX1182008), *P. coccineus* (G35346; SRX1182014), *P. acutifolius* (G40001; SRX1182015) (também analisadas por Ribeiro *et al.* 2019), *P. maculatus* (PL-8841; SRS1938372), *P. polystachios* (G40782; SRS1938369), *P. hintonii* (SRS1938370), *P. leptostachyus* (PL-8829; SRS1938371), *P. microcarpus* (PL-8844; SRS1938373); *P. lunatus* (PL-8834; SRS1938374), *P. filiformis* (SRS1938376), *P. dumosus* (G36043; SRS1938378), *P. costaricensis* (G40811A; SRS1938379) (Rendón-Anaya *et al.* 2017) foram baixados do GenBank ([www.ncbi.nlm.nih.gov/genbank/](http://www.ncbi.nlm.nih.gov/genbank/)). Esses genomas foram incorporados em uma análise comparativa dos principais elementos repetitivos desse grupo utilizando a função *clustering* do RepeatExplorer 2 (Novák *et al.* 2013). As espécies previamente sequenciadas e disponíveis de *P. leptostachyus* e *P. lunatus* foram incluídas na análise por se tratarem de acessos diferentes, permitindo uma comparação intraespecífica da fração repetitiva.

Após serem filtradas por qualidade e pareadas, um *sampling* foi feito para amostrar aproximadamente 0,1× do genoma de cada espécie. Os *reads* de cada espécie foram concatenados e então agrupados por similaridade. As sequências dos *clusters* que não puderam ser anotados automaticamente foram comparadas com BLAST (Basic Local Alignment Search Tool) na base de dados do NCBI.

#### *Desenho de primers, amplificação e marcação de sondas*

Pares de *primers* foram desenhados (Tabela 1) para cada um dos satélites a partir da sequência consenso e voltados para fora, garantindo que a sequência só seria amplificada se disposta em tandem.

Os fragmentos foram amplificados por PCR a partir de reações de 50 µl contendo 10-20 ng de DNA genômico, 0,2 mM dNTP, 1× tampão de PCR, 0,4 µM de cada primer, 3 mM MgCl<sub>2</sub>, 1× TBT (150 mM trehalose, 0,2 mg/mL BSA, 0,2% Tween-20, 1,7 mM Tris-HCl pH 8.0) e 0,3 µl de TaqDNA polimerase caseira. As reações envolveram 30 ciclos de amplificação (1 min à 94° C, 1 min à temperatura de anelamento e 1 min à 72° C). Quando compartilhados entre espécies, os satélites foram amplificados a partir do DNA genômico de *P. leptostachyus*.

Os elementos Ogre-Tat e Chromovirus foram amplificados a partir do genoma de *P. vulgaris* usando *primers* previamente desenvolvidos (Ribeiro *et al.* 2019) e a sonda para o satélite *khipu* foi preparada a partir de uma mistura de unidades *khipu* de *P. vulgaris* (Pvkhpu mix; Richard *et al.* 2013). As sequências amplificadas foram marcadas por *Nick translation* com Cy3-dUTP (GE), com exceção do elemento Ogre-Tat que foi marcado com 488-5-dUTP (Alexa, Life Technologies).

### *Hibridização in situ fluorescente (FISH)*

As FISHs foram realizadas segundo Pedrosa *et al.* (2002). Resumidamente, as lâminas foram secas a 60°C por 30 min, pré-tratadas em RNase (100 µg/ml) a 37°C por 1 h, em 10 µg/ml de pepsina em 0,01 N HCl a 37°C por 20 min e fixadas em paraformaldeído 3,7% por 10 min. A mistura de hibridização foi composta por 50% de formamida, 10% de sulfato de dextrano, 2× SSC e 5-10 ng/µl de sonda. Após desnaturação da mistura por 10 min a 75°C, as lâminas foram desnaturadas com a mistura de hibridização por 5 min na mesma temperatura e então, incubadas em câmara úmida a 37°C por até três dias. Após hibridização, as lâminas foram lavadas em 2× SSC e 0,1× SSC a 42°C, sendo a estringência final 77%. Para a FISH de baixa estringência (40%), a mistura de hibridização foi composta de 10% de sulfato de dextrano, 6× SSC e 5-10 ng/µl de sonda e as lavagens pós hibridização foram realizadas em 6× SSC à temperatura ambiente. Por fim, as lâminas foram montadas em 2 µg/mL DAPI/Vectashield (Vector).

Para um dos satélites, exclusivo do grupo *Leptostachyus*, lâminas previamente hibridizadas com marcadores cromossomo-específicos clonados em BACs (cromossomos artificiais de bactérias) foram rehibridizadas para permitir a identificação dos cromossomos com sítios de DNA satélite (Ferraz *et al.* 2016; Fonsêca *et al.* 2016). As lâminas foram descoradas em 4× SSC/0,1% Tween 20 e as rehibridizações das lâminas foram realizadas de acordo com Heslop-Harrison *et al.* (1992).

As metáfases foram capturadas em microscópio de epifluorescência Leica DM5500B por meio do sistema de captura DFC345 FX (Leica). As melhores células fotografadas foram selecionadas para sobreposição e ajuste de brilho e contraste no Adobe Photoshop CS6. Os cromossomos foram identificados e numerados de acordo com a sentença com *P. vulgaris* (Fonsêca *et al.* 2010; Fonsêca *et al.* 2016).

## **Resultados**

### *Análise comparativa das sequências repetitivas*

A análise comparativa de 15 acessos de 13 espécies de *Phaseolus* (incluindo um acesso de *P. macvaughii*, dois de *P. leptostachyus* e dois de *P. lunatus*) gerou 149 *clusters* com abundância maior que 0,01% do genoma (Tabela 2). A proporção da fração repetitiva dos genomas variou de 24,13% em *P. macvaughii* até 46,62% em *P. polystachios*, estando *P.*

*macvaughii* e *P. leptostachyus* entre as espécies com menor proporção de sequências repetitivas e *P. lunatus* entre as espécies com maior proporção.

Em todas as espécies analisadas, os elementos repetitivos mais abundantes foram os retrotransposons LTR, que variaram de cerca de 5% em *P. microcarpus* até cerca de 27% em *P. lunatus* (Tabela 2). Com exceção de *P. filiformis*, que apresentou a mesma proporção de Ty3/gypsy e Ty1/copia, os elementos Ty3/gypsy foram predominantes em todas as espécies, variando de aproximadamente 3% em *P. microcarpus* até cerca de 21% em *P. lunatus* (Tabela 2; Fig. 1A, C e D).

Dentre os Ty3/gypsy, a linhagem Chromovirus, representada por CRM e Tekay, foi a mais abundante (Fig. 1C). O elemento CRM da linhagem Chromovirus variou de cerca de 2% em *P. microcarpus* até cerca de 9% em *P. hintonii* e foi o elemento mais abundante em *P. microcarpus*, *P. hintonii*, *P. filiformis*, *P. acutifolius*, *P. vulgaris*, *P. coccineus*, *P. costaricensis*, *P. dumosus* e *P. macvaughii*. Já a linhagem de Chromovirus Tekay variou de 0,26% do genoma de *P. microcarpus* a aproximadamente 12% em *P. lunatus* e foi o principal elemento dos genomas de *P. lunatus*, *P. polystachios* e *P. maculatus*. *P. macvaughii* apresentou uma proporção muito baixa de Tekay (apenas 0,52%) quando comparado aos dois acessos da espécie do mesmo grupo, *P. leptostachyus* (4,66% e 6,82%, respectivamente). Já *P. leptostachyus* apresentou valores similares de CRM e Tekay (4,78% e 4,66%, respectivamente, para *P. leptostachyus* e 5,08% e 6,82%, respectivamente, *P. leptostachyus* 2), revelando uma pequena variação entre os dois acessos dessa espécie. Entre os elementos não-Chromovirus, o elemento Tat foi predominante e variou de 0,37% em *P. filiformis* até cerca de 5% em *P. dumosus* (Tabela 2).

Elementos das linhagens Chromovirus e Ogre-Tat, previamente identificados pelo RepeatExplorer (Ribeiro *et al.* 2019), foram amplificados a partir do feijão comum (*P. vulgaris*) e localizados *in situ* em *P. leptostachyus*, *P. macvaughii* e *P. lunatus*. O elemento Chromovirus apresentou um padrão de hibridização pericentromérico semelhante entre as espécies (Fig. 2A, C e E). O Ogre-Tat apresentou um padrão mais disperso em todo o complemento, porém com enriquecimento na região pericentromérica (Fig. 2B, D e F). As regiões pericentroméricas com sinais mais fortes de hibridização geralmente apresentaram colocalização dessas duas sequências.

Os elementos Ty1/copia foram mais diversificados (seis linhagens), mas foram encontrados em menor número de cópias. O elemento SIRE foi o mais abundante em todas as espécies analisadas e variou de 1,2% do genoma de *P. microcarpus* até pouco mais de 6% do genoma de *P. vulgaris*. No entanto, pode-se observar que os acessos de *P. leptostachyus* 1 e 2

apresentaram uma menor proporção dessa sequência quando comparado a *P. macvaughii* (1,65%, 2,15% e 4,03%, respectivamente) e também quando comparados às demais espécies analisadas (Fig. 1D e Tabela 2).

Outros grupos de elementos móveis detectados incluíram retrotransposons não LTR, os LINEs, e também duas famílias de transposons de DNA, CACTA e Mutator. A abundância de CACTA variou de 0,2% no genoma de *P. microcarpus* até aproximadamente 1% em *P. polystachios*, sendo o principal representante dos transposons de DNA (Tabela 2).

Por fim, o DNA repetitivo em tandem consistiu do DNAr 35S e 5S e DNAs satélite. Na análise comparativa, a fração de satélites no genoma variou de 0,7% em *P. macvaughii* até mais de 11% em *P. microcarpus* e *P. acutifolius* (Tabela 2). Os satélites PacuSAT363 (Ribeiro *et al.* 2019), CC4 (Almeida *et al.* 2012), *juniper* (Ribeiro *et al.* 2017), *kipu* (David *et al.* 2009), PleSAT2514 e PluSAT991 (presente trabalho) foram identificados em representantes de todos os grupos, apesar de estarem em uma proporção significativamente baixa em algumas espécies, principalmente em *P. macvaughii* (Fig. 1B) Houve diferença na abundância de satélites entre os dois acessos de *P. lunatus* (2,14% em *P. lunatus* 1 e 0,97 em *P. lunatus* 2), sugerindo uma variação intraespecífica na proporção de satélites nessa espécie (Tabela 2).

#### *Análise in silico individual da fração repetitiva de P. leptostachyus, P. macvaughii e P. lunatus*

Para a análise mais detalhada da fração repetitiva de *P. leptostachyus*, *P. macvaughii* e *P. lunatus* foram utilizados 205.200, 176.000 e 274.000 dos *reads* gerados para cada espécie, respectivamente, correspondendo a 0,1× de cobertura do genoma. A análise individual gerou 203 *clusters* com abundância maior que 0,01% do genoma para *P. leptostachyus*, 164 para *P. macvaughii* e 194 para *P. lunatus*, indicando que a fração repetitiva corresponde a cerca de 37%, 32% e 45% do genoma dessas espécies, respectivamente.

A ferramenta TAREAN do Repeat Explorer foi utilizada para identificação de sequências de DNA satélites. A análise revelou a existência de dois satélites de alta confiança e um de baixa confiança para *P. lunatus*, três de alta e três de baixa confiança para *P. leptostachyus* e dois satélites de baixa confiança para *P. macvaughii* (Tabela 3). Para confirmar sua organização em tandem, foi feita uma análise de dotplot com todos os satélites encontrados pelo TAREAN, com exceção de dois de baixa confiança de *P. leptostachyus* (2514 pb e 5117 pb) e um de baixa confiança de *P. macvaughii* (28 pb), devido ao tamanho da unidade de repetição muito grande ou muito pequena (Fig. 3). O dotplot revelou alta similaridade entre as sequências PleSAT752 e PluSAT991, e entre PleSAT592 e PmaSAT648. O dotplot também

revelou alta similaridade entre as sequências PleSAT558, PleSAT549, PluSAT548 e *khipu*, confirmando sua presença nos genomas de *P. leptostachyus* e *P. lunatus*, mas não em *P. macvaughii*, muito embora tenha sido encontrado em baixa proporção na análise comparativa (apenas 0,04%). Após amplificação por PCR, foram confirmados três novos DNA satélites para o grupo de espécies analisadas (*PluSAT1149*, *PleSAT752/PluSAT991* e *PleSAT592/PmaSAT648*).

#### *Localização cromossômica dos principais DNA satélites*

A proporção extremamente baixa de *khipu* em *P. macvaughii* foi confirmada por FISH. Diferente de *P. leptostachyus*, que mostrou sinais subtelo méricos (Fig. 4A e B), a FISH em *P. macvaughii* não revelou nenhum sinal evidente (Fig. 4C e D).

Os três novos satélites identificados foram localizados por FISH. Em *P. lunatus*, a hibridização com seu satélite exclusivo *PluSAT1149* revelou um único par de cromossomos marcados fortemente na região do centrômero, colocalizado a uma das bandas heterocromáticas DAPI<sup>+</sup>, presente em dois pares cromossômicos (Fig. 5A). O *PleSAT752/PluSAT991* revelou também um único par de cromossomos marcados na região centromérica em *P. lunatus*, e colocalizado com a outra banda heterocromática DAPI<sup>+</sup> da espécie (Fig. 5A). Diferente do encontrado em *P. lunatus*, em *P. leptostachyus* o *PleSAT752/PluSAT991* se mostrou distribuído de forma dispersa em todos os cromossomos. Apesar de disperso, uma marcação mais forte foi detectada na região pericentromérica do par cromossômico maior (Fig. 5B). Apesar do TAREAN não ter detectado esse satélite entre os *clusters* mais abundantes de *P. macvaughii*, a análise comparada revelou essa sequência numa proporção de 0,04% em *P. macvaughii*. A hibridização cruzada nessa espécie mostrou um padrão semelhante ao encontrado em *P. leptostachyus*, porém com sinais de intensidade muito menor devido a sua baixa proporção (Fig. 5C e D). Com exceção do sinal pouco evidente do *PleSAT752/PluSAT991* em *P. macvaughii*, as demais hibridizações cruzadas com as sondas dos demais satélites não revelaram sinais de hibridização (dados não mostrados).

O *PleSAT592/PmaSAT648*, exclusivo para o grupo *Leptostachyus*, mostrou um padrão subtelo mérico semelhante em ambas as espécies do grupo, porém com número de sítios variados. Em *P. leptostachyus*, a hibridização revelou três pares de cromossomos marcados na região subtelo mérica, com sinais apenas em um dos braços cromossômicos (Fig. 5E). Devido à descondensação da RON no cromossomo *Ple6*, esse cromossomo foi identificado como portador de um dos três sítios desse satélite. O maior par cromossômico não apresentou sinal de hibridização para *PleSAT592/PmaSAT648*.

Em *P. macvaughii*, PleSAT592/PmaSAT648 apresentou sítios subteloméricos em ambos os braços de alguns cromossomos, outros com marcação em apenas um dos braços, e outros, ainda, sem nenhuma marcação (Fig. 5F). A identificação de cromossomos por BAC-FISH permitiu detectar que o maior par cromossômico, o Pma10/11, assim como os cromossomos Pma6 e Pma4 não apresentaram marcação com esse satélite (Fig. 6A e B). Os cromossomos Pma3 e Pma9 apresentaram sinal em apenas um dos braços cromossômicos, respectivamente os braços longo e curto (Fig. 6C e D). Já os cromossomos Pma1/2 e Pma8 apresentam sinais de hibridização em ambos os braços cromossômicos (Fig. 6E e F).

## Discussão

*Retrotransposons LTR Ty3/gypsy são os principais elementos repetitivos no genoma de Phaseolus*

A análise da fração repetitiva do genoma de duas das três espécies do grupo *Leptostachyus* revelou uma proporção menor de *repeats* (em média 29%) em relação ao total de espécies analisadas do gênero (37,46%), compatível com o menor tamanho do genoma dessas espécies (1C = 476,5 Mpb para o grupo *Leptostachyus* versus 1C = 598,54 para *Phaseolus*; Roy & Botanic Gardens, Kew, <http://data.kew.org/cvalues/>, Bennet e Leitch, 2012). Uma forte correlação positiva entre proporção de sequências repetitivas e tamanho do genoma foi encontrada para 23 leguminosas da tribo Fabeae (Macas *et al.* 2015) e para os genomas gigantes de *Fritillaria* (Liliaceae) (Kelly *et al.* 2015).

A análise comparativa com o RepeatExplorer 2 permitiu a caracterização das sublinhagens dos elementos Chromovirus, não obtida anteriormente. Além disso, a classificação recente separa Ogr-Tat em dois grupos distintos, permitindo com isso, uma melhor anotação da fração repetitiva. Apesar da nova classificação, os resultados obtidos para espécies de *Phaseolus* são condizentes com os obtidos anteriormente para *P. acutifolius*, *P. coccineus* e *P. vulgaris* (Ribeiro *et al.* 2019). Para *P. vulgaris*, houve uma discrepância na proporção de sequências repetitivas quando comparado ao genoma completo sequenciado (~ 35% versus 45%; Schmutz *et al.* 2014), indicando que tenha havido uma subestimativa através do método de *clustering* de uma subamostragem da fração repetitiva. No entanto, os principais elementos encontrados foram os mesmos.

Todas as espécies analisadas apresentaram retrotransposons LTR, mais precisamente Ty3/gypsy, como os elementos repetitivos mais abundante do genoma. *P. microcarpus*

apresentou uma proporção muito pequena desses elementos, talvez devido à maior fração de sequências (mais de 16%) que não puderam ser classificadas. Ty3/*gypsy* constituem o principal componente da fração repetida do genoma na maioria das espécies de plantas (Macas *et al.* 2015; Puterova *et al.* 2018; Gaiero *et al.* 2019; Ribeiro *et al.* 2019). Dentre eles, a linhagem Chromovirus, representada aqui pelas sublinhagens CRM e Tekay, foi a mais abundante para todas as espécies analisadas. A diferença do descrito previamente para espécies do gênero e gêneros relacionados, onde o elemento Ogre-Tat foi mais abundante (Ribeiro *et al.* 2019), deve-se, provavelmente, à anotação mais detalhada dos *repeats* no RepeatExplorer 2.

A hibridização com os elementos Chromovirus e Ogre-Tat revelou enriquecimento de ambos na região pericentromérica. Essa distribuição preferencial confirma dados prévios em *P. vulgaris* e *P. lunatus* por BAC-FISH que revelaram sinais pericentroméricos correspondentes a blocos de heterocromatina observados pela coloração CMA/DAPI (Fonsêca *et al.* 2010; Bonifácio *et al.* 2012). Assim, para as espécies de *Phaseolus*, a fração repetitiva está localizada preferencialmente na heterocromatina pericentromérica CMA<sup>+</sup>/DAPI<sup>-</sup>, onde os elementos Chromovirus e Ogre-Tat são mais abundantes. Outras espécies de genomas pequenos também mostraram o padrão pericentromérico para esses elementos, como em *Arabidopsis thaliana* (100 Mbp; Heslop-Harrison *et al.* 1997) e espécies de *Vigna* e *Cajanus*, com genomas de tamanho similar (Ribeiro *et al.* 2019). O mesmo foi observado para *Brachiara*, que apresenta um tamanho de genoma médio (~600 até 2000 Mbp; Santos *et al.* 2015). Por outro lado, em espécies de genoma grande esses elementos mostram uma distribuição mais dispersa, como foi visto para o elemento Ogre em *Lathyrus sativum* (8215 Mbp; Ceccarelli *et al.* 2012) e em *Vicia pannonica* (6601 Mbp; Neumann *et al.* 2006).

*DNAsat compartilhados revelaram padrões de distribuição cromossômica diferentes, confirmando sua rápida evolução no genoma*

A maior divergência entre as espécies de *Phaseolus* não esteve na composição de distribuição dos elementos transponíveis, mas sim na fração de DNAsat. No gênero *Phaseolus*, o DNAsat subtelomérico *hipu*, o mais conservado no gênero por estar presente em representantes de todos os grupos analisados, mostra significativa variação em número de cópias entre diferentes espécies (Ribeiro *et al.* 2017). Mesmo sob baixa estrigência (dados não mostrados), e diferente do observado para *P. leptostachyus*, a FISH não revelou nenhum sinal subtelomérico de *hipu* em *P. macvaughii*, compatível com sua baixa abundância. Pode-se concluir que o satélite *hipu* apresenta uma rápida evolução não apenas no gênero, mas particularmente no grupo *Leptostachyus*.

Os satélites PluSAT1449 e PleSAT752/PluSAT991 hibridizaram cada um em um dos dois pares cromossômicos portadores de uma banda centromérica heterocromática revelada pós-FISH em *P. lunatus*. Um desses pares provavelmente corresponde ao cromossomo *Plu3*, previamente mapeado (Almeida e Pedrosa-Harand 2013). No entanto, a localização centromérica dessas sequências só poderá ser confirmada após colocalização com a proteína centromérica CENH3. Apesar do DNAsat ser encontrado principalmente em regiões centroméricas e subteloméricas dos cromossomos (Biscotti *et al.* 2015), nenhum satélite centromérico foi encontrado pelo TAREAN para o grupo *Leptostachyus*. Os satélites centroméricos descritos para *P. vulgaris* CentPv1 e CentPv2 (Iwata *et al.* 2013; Iwata-Otsubo *et al.* 2016) não foram detectados em *Leptostachyus* e *P. lunatus* na análise comparativa, com exceção do CentPv2 em *P. macvaughii* que apresentou uma proporção muito baixa (0,02%). A ausência de repeats centroméricos também foi observada em espécies do gênero próximo *Vigna* (Iwata-Otsubo *et al.* 2016b; Ribeiro *et al.* 2019) e em espécies de *Solanum* (Zhang *et al.* 2014), reforçando a ideia da rápida evolução dessas sequências, aparentemente maior do que as sequências subteloméricas, como por exemplo *hipu*, aqui revelado conservado.

Diferente do observado para *P. lunatus*, o PleSAT752/PluSAT991 apresentou sinal fortemente disperso em *P. leptostachyus*, com uma banda mais forte na região proximal do maior par. *Phaseolus macvaughii* apresentou sinal de hibridização muito fraco, porém também disperso, revelando que essa sequência está presente numa proporção baixa do genoma, o que explica não ter sido detectada pelo Tarean. Esse maior par cromossômico foi resultado de uma disploidia descendente que levou a uma inserção cêntrica do cromossomo Pv10 no cromossomo Pv11 no grupo *Leptostachyus* (Ferraz *et al.* 2016; Fonsêca *et al.* 2016). O padrão disperso desse satélite revelado para as espécies do clado *Lepotstachyus* é diferente do único bloco centromérico observado em um par cromossômico de *P. lunatus*. Esse padrão de dispersão atípico no genoma do grupo *Leptostachyus* é principalmente evidente em *P. leptostachyus*, que apresentou uma amplificação recente deste DNAsat.

Apesar de apresentar um mesmo padrão subtelomérico em *Leptostachyus*, o PleSAT592/PmaSAT648 mostrou uma variação no número de sítios e intensidade dos sinais na FISH. Um maior número de sítios encontrados em *P. macvaughii* está de acordo com a abundância maior dessa sequência nessa espécie e permitiu caracterizar os cromossomos dessa espécie devido ao mapeamento prévio com BACs. Essa foi a única sequência típica do clado, exclusiva de seus representantes ou abundante o suficiente para permitir detecção apenas nestas espécies.

*A alta taxa de evolução cariotípica do Clado Leptostachyus pode estar relacionada ao rápido turnover dos elementos Tekay e SIRE ou a amplificação recente de DNAsat no genoma*

O grupo monofilético *Leptostachyus*, formado por *P. leptostachyus*, *P. macvaughii* e *P. micranthus*, apresenta alta taxa de evolução cariotípica, a qual é evidenciada por diversos rearranjos cromossômicos que foram previamente mapeados por BAC-FISH. Além da inserção cêntrica compartilhada que levou ao cariótipo disploide no grupo, inúmeros rearranjos ocorreram, em menor número na espécie de divergência mais antiga, *P. macvaughii* (Ferraz *et al.* 2016), e em maior número na espécie derivada mais recentemente, *P. leptostachyus* (Fonseca *et al.* 2016). Rearranjos cromossômicos geralmente ocorrem por recombinação de sequências repetitivas (Schubert e Lysak 2011). Acredita-se que sequências repetidas da heterocromatina podem afetar a evolução cariotípica facilitando rearranjos que tenham um impacto deletério mínimo sobre o genoma (Li *et al.* 2017). No entanto, o que causaria uma maior frequência de rearranjos em determinado grupo ou momento na evolução de um grupo e o mecanismo exato de como as sequências repetitivas estão envolvidas na evolução cromossômica permanecem desconhecidos.

De forma geral, a composição e abundância das diferentes classes de *repeats* não diferiu substancialmente entre o grupo *Leptostachyus* e os demais grupos analisados. Por serem do mesmo grupo, esperava-se que os principais elementos em *P. macvaughii* e *P. leptostachyus* se mostrassem mais semelhantes entre si. No entanto, os elementos mais abundantes das duas espécies (Tekay, CRM, Tat e SIRE) tiveram abundâncias tanto maiores quanto menores, revelando amplificação ou redução recente no grupo. Os acessos de *P. leptostachyus* analisados apresentaram proporção de SIRE muito baixas quando comparado a *P. macvaughii* e às demais espécies de outros grupos. Já *P. macvaughii* apresentou uma proporção muito baixa de Tekay quando comparado à *P. leptostachyus*. Essa variação foi maior que a variação observada entre espécies do grupo *Vulgaris*, com mais representantes amostrados e mais antigo (Delgado-Salinas *et al.* 2006). Isso pode ser explicado por um rápido *turnover* dessas sequências que evoluíram evidentemente mais rápido nas espécies desse clado do que nas espécies de outros grupos de *Phaseolus* e poderia estar associado com a maior taxa de rearranjos cromossômicos no grupo.

A única família de *repeats* aparentemente exclusiva do grupo *Leptostachyus* é o DNAsat PleSAT592/PmaSAT648. Sua menor abundância em *P. leptostachyus* que em *P. macvaughii*, embora esta espécie possua mais rearranjos cromossômicos (Ferraz *et al.* 2016; Fonseca *et al.* 2016) e sua distribuição subtelomérica, como outros DNAsat no gênero (por exemplo, *kipu* e *juniper*), não sugerem, no entanto, que sua amplificação possa ter propiciado um maior número de rearranjos. A hibridização deste DNAsat associado aos BACs para identificação

cromossômica, mostrou sua presença e ausência tanto em cromossomos que sofreram rearranjos quanto em cromossomos não rearranjados, indicando que não há relação direta entre a presença de sinais desse satélite e a ocorrência de rearranjos. O cromossomo resultante da inserção cêntrica, *Pma10/11*, por exemplo, não apresentou nenhuma marcação, assim como o *Pma4* e o *Pma6*, ambos não rearranjados. Por outro lado, os cromossomos *Pma1/2*, *Pma3* e *Pma9*, todos rearranjados, apresentaram sinais assim como o *Pma8*, não rearranjado. Em *P. leptostachyus*, foi possível ver que, embora o cromossomo maior *Ple10/11/6* (Fonsêca *et al.* 2016) não tenha apresentado marcação, o cromossomo *Ple6* mostrou sinal de hibridização. Além disso, não foi observada a presença de sítios em outras posições que não subteloméricas, embora várias translocações e inversões observadas possivelmente tenham envolvido regiões não terminais dos cromossomos. De todo modo, considerando que as regiões subteloméricas no feijão comum são bastante dinâmicas (Chen *et al.* 2018), rearranjos entre blocos subteloméricos podem explicar a grande diferença na distribuição de sítios entre *P. leptostachyus* e *P. macvaughii*.

Já o PleSAT752/PluSAT991, embora detectado na análise comparativa para todas as espécies, revelou uma amplificação no grupo *Leptostachyus*, além de uma distribuição atípica e dispersa. São poucos os DNAsat que apresentam tal padrão, mas isso foi observado para *Nicotiana* (Lim *et al.* 2004; Koukalova *et al.* 2010) e em animais, para *Locusta migratoria*, o gafanhoto-migratório (Ruiz-Ruano *et al.* 2016) e para o barbeiro (Pita *et al.* 2017). Essa organização dispersa pode ter propiciado rearranjos incomuns para o gênero, em especial em *P. leptostachyus*, cujo satélite esteve em maior proporção e que apresenta ainda mais rearranjos que *P. macvaughii*. DNAs satélites parecem estar associados a ocorrência de rearranjos em *Drosophila* (Khost *et al.* 2019) e em espécies de lagartos (Prakhongcheep *et al.* 2017). Em *Peromyscus*, um grupo de roedores, uma única sequência de DNA satélite conservada, apesar de variar no número de cópias e localização cromossômica entre as espécies, resulta em grandes blocos de cromatina homólogas, o que promove a ocorrência de rearranjos cromossômicos e variações polimórficas (Smalec *et al.* 2019). No grupo *Leptostachyus*, o mapeamento exato dos pontos de quebra e a anotação das sequências flangeadoras será necessário para confirmar o envolvimento de PleSAT752/PluSAT991 ou alguma das outras sequências repetitivas descritas neste trabalho nos rearranjos cromossômicos deste grupo de espécies.

## Agradecimentos

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq pelo suporte financeiro através de bolsas para Maria Eduarda Ferraz Vasconcelos e Andrea Pedrosa-Harand.

## Referências

- Almeida C, Fonsêca A, dos Santos KGB, Mosiolek M, Pedrosa-Harand A (2012) Contrasting evolution of a satellite DNA and its ancestral IGS rDNA in *Phaseolus* (Fabaceae). *Genome* 55:683–689. doi:10.1139/g2012-059
- Almeida C, Pedrosa-Harand A (2013) High macro-collinearity between lima bean (*Phaseolus lunatus* L.) and the common bean (*P. vulgaris* L.) as revealed by comparative cytogenetic mapping. *Theor Appl Genet* 126:1909-16. doi: 10.1007/s00122-013-2106-9
- Ayres-Alves T, Cardoso AL, Nagamachi CY, de Sousa LM, Pieczarka JC, Noronha RCR (2017) Karyotypic evolution and chromosomal organization of repetitive DNA sequences in species of *Panaque*, *Panaqolus*, and *Scobinancistrus* (Siluriformes and Loricariidae) from the Amazon Basin. *Zebrafish* 14, 251–260
- Biscotti MA, Olmo E, Heslop-Harrison JP (2015) Repetitive DNA in eukaryotic genomes. *Chromosome Res* 23:415–420
- Bennett MD, Leitch IJ (2012) Plant DNA C-values database (release 6.0, Dec. 2012) <http://www.kew.org/cvalues/>
- Bonifácio EM, Fonsêca A, Almeida C, dos Santos KGB, Pedrosa-Harand A (2012) Comparative cytogenetic mapping between the lima bean (*Phaseolus lunatus* L.) and the common bean (*P. vulgaris* L.). *Theor Appl Genet* 124:1513-20. doi: 10.1007/s00122-012-1806-x
- Broughton WJ, Hérnandez G, Blair MW, Beebe SE, Gepts P, Vanderleyden J (2003) Beans (*Phaseolus* spp.)—model food legumes. *Plant Soil* 252:55-128. doi: 10.1023/A:1024146710611
- Carvalho CR e Saraiva LS (1993) An air drying technique for maize chromosomes without enzymatic maceration. *Biotechnic e Histochemistry* 68:142-145. doi: 10.3109/10520299309104684

- Ceccarelli M, Sarri V, Polizzi E, Andreozzi G, Cionini PG (2013) Ogre retrotransposons in *Lathyrus* species. *Plant Biosystems* 147:99–104.
- Cermak T, Kubat Z, Hobza R, Koblizkova A, Widmer A, Macas J, *et al.* (2008) Survey of repetitive sequences in *Silene latifolia* with respect to their distribution on sex chromosomes. *Chromosome Res* 16:961–76
- Chen NWG, Thareau V, Ribeiro T, Magdelenat G, Ashfield T, Innes RW, Pedrosa-Harand A, Geffroy V (2018) Common Bean Subtelomeres Are Hot Spots of Recombination and Favor Resistance Gene Evolution. *Front Plant Sci.*, 14;9:1185
- David P, Chen NWG, Pedrosa-Harand A, Thareau V, Sévignac M, Cannon SB, Debouck D, Langin T, Geffroy V (2009) A nomadic subtelomeric disease resistance gene cluster in common bean. *Plant Physiol* 151:1048–1065. doi:10.1104/pp.109
- Delgado-Salinas A, Bibler R, Lavin M (2006) Phylogeny of the Genus *Phaseolus* (Leguminosae): A Recent Diversification in an Ancient Landscape. *Syst. Botany* 31(4)779-791. doi: 10.1600/036364406779695960
- Dodsworth S, Chase MW, Leitch AR (2016) Is post-polyploidization diploidization the key to the evolutionary success of angiosperms? *Bot. J. Linn. Soc.*, 180, 1–5.
- Dodsworth S, Jang TS, Struebig M, Chase MW, Schneeweiss HW, Leitch AR (2017) Genome-wide repeat dynamics reflect phylogenetic distance in closely related allotetraploid *Nicotiana* (Solanaceae). *Plant Syst Evol* 303:1013–1020
- Fonsêca A, Ferreira J, Santos TRB, Mosiolek M, Belluci E, Kami J, Gepts P, Geffroy V, Schweizer D, Santos KGB, Pedrosa-Harand A (2010) Cytogenetic map of common bean (*Phaseolus vulgaris* L.). *Chromosome Res* 18:487–502. doi: 10.1007/s10577-010-9129-8
- Ferraz ME, Fonsêca A, Pedrosa-Harand A. Mapeamento citogenético comparativo em *Phaseolus macvaughii* Delgado: mais do que uma simples dispoloidia. 2016. 66 f. Trabalho de Conclusão de Curso - Universidade Federal de Pernambuco, Recife
- Fonsêca A, Ferraz ME, Pedrosa-Harand A (2016) Speeding up chromosome evolution in *Phaseolus*: multiple rearrangements associated with a one-step descending dysploidy. *Chromosoma* 125(3)413-21. doi: 10.1007/s00412-015-0548-3
- Fonsêca A, Pedrosa-Harand A (2013) Karyotype stability in the genus *Phaseolus* evidenced by the comparative mapping of the wild species *Phaseolus microcarpus*. *Genome* 56:335-43. doi: 10.1139/gen-2013-0025

- Fonsêca A, Pedrosa-Harand A (2017) Cytogenetics and Comparative Analysis of Phaseolus Species. In: Marcelino Pérez de la Vega; Marta Santalla; Frédéric Marsolais.. (Org.). The Common Bean Genome. 1ed.Chan (ZG): Springer International Publishing AG, 2017, v. , p. 1-295.
- Fransz P, Linc G, Lee C-R, Aflitos SA, Lasky JR, Toomajian C, Ali H, Peters J, van Dam P, Ji X, Kuzak M, Gerats T, Schuber I, Schneeberger K, Colot V, Martienssen R, Koornneef M, Nordborg M, Juenger TE, de Jong H, Schranz ME (2016) Molecular, genetic and evolutionary analysis of a paracentric inversion in *Arabidopsis thaliana*. *Plant J*, 88:159-178
- Gaiero P, Vaio M, Peters SA, Schranz ME, Jong H, Speranza PR (2019) Comparative analysis of repetitive sequences among species from the potato and the tomato clades. *Annals of Botany* 123: 521–532, 2019 doi: 10.1093/aob/mcy186
- Gao D, Abernathy B, Rohksar D, Schmutz J, Jackson SA (2014) Annotation and sequence diversity of transposable elements in common bean (*Phaseolus vulgaris*). *Front Plant Sci*. 5:339. doi: 10.3389/fpls.2014.00339
- Gong Z, Wu Y, Koblizkova A, Torres GA, Wang K, Iovene M, Neumann P, Zhang W, Novak P, Buell CR, Macas J, Jiang J (2012) Repeatless and repeat-based centromeres in potato: implications for centromere evolution. *Plant Cell* 24:3559–3574. <https://doi.org/10.1105/tpc.112.100511>
- Heslop-Harrison JS, Harrison GE, Leitch IJ (1992) Reprobing of DNA: DNA in situ hybridization preparations. *Trends Genet* 8:372-373.
- Heslop-Harrison P, Brandes A, Taketa S et al. (1997) The chromosomal distributions of Tyl-copia group retrotransposable element in higher plants and their implications for genome evolution. *Genetica* 100:197-204.
- Iwata A, Tek AL, Richard MMS et al (2013) Identification and characterization of functional centromeres of the common bean. *Plant J* 76:47–60. doi: 10.1111/tpj.12269
- Iwata-Otsubo A, Radke B, Findley S et al. (2016) Fluorescence in situ hybridization (FISH)-based karyotyping reveals rapid evolution of centromeric and subtelomeric repeats in common bean (*Phaseolus vulgaris*) and relatives. *G3 (Bethesda)* 6:1013-1022. doi: 10.1534/g3.115.024984
- Iwata-Otsubo A, Lin JY, Gill N, Jackson SA (2016b) Highly distinct chromosomal structures in cowpea (*Vigna unguiculata*), as revealed by molecular cytogenetic analysis. *Chromosome Res* 24:197–216. doi: 10.1007/s10577-015-9515-3

- Kelly LJ, Renny-Byfield S, Pellicer J et al. (2015) Analysis of the giant genomes of *Fritillaria* (Liliaceae) indicates that a lack of DNA removal characterizes extreme expansions in genome size. *New Phytol* 208:596-607. doi: 10.1111/nph.13471
- Khost DE, Eickbush DG, Larracuenta AM (2017) Single-molecule sequencing resolves the detailed structure of complex satellite DNA loci in *Drosophila melanogaster*. *Genome Research*, 27(5), 709–721. doi:10.1101/gr.213512.116
- Koukalova B, Moraes AP, Renny-Byfield S, Matyasek R, Leitch AR, Kovarik A (2010) Fall and rise of satellite repeats in allopolyploids of *Nicotiana* over c. 5 million years. *New Phytol.* 186(1):148-60. doi: 10.1111/j.1469-8137.2009.03101
- Lambowitz AM, Gellert M, Chandler M, Craig NL, Sandmeyer SB, Rice PA (Eds) (2015) Mobile DNA III. American Society of Microbiology
- Li S, Su T, Cheng G, Wang B, Li X, Deng C, Gao W (2017) Chromosome Evolution in Connection with Repetitive Sequences and Epigenetics in Plants. *Genes (Basel)*, 8(10): 290
- Lim KY, Skalicka K, Koukalova B, Volkov RA, Matyasek R, Hemleben V, Leitch AR, Kovarik A (2004). *Dynamic Changes in the Distribution of a Satellite Homologous to Intergenic 26-18S rDNA Spacer in the Evolution of Nicotiana*. *Genetics*, 166(4), 1935–1946. doi:10.1534/genetics.166.4.1935
- Macas J, Novák P, Pellicer J et al. (2015) In depth characterization of repetitive DNA in 23 plant genomes reveals sources of genome size variation in the legume Tribe *Fabeae*. *PLoS ONE*10:e0143424. doi:10.1371/journal.pone.0143424
- Mandáková T, Heenan PB, Lysak MA (2010) Island species radiation and karyotypic stasis in *Pachycladon* allopolyploids. *Research article*, 10:367. doi: 10.1186/1471-2148-10-367
- Mandáková T, Joly S, Krzywinski M, Mummenhoff K, Lysak MA (2010) Fast Diploidization in Close Mesopolyploid Relatives of *Arabidopsis*. *The Plant Cell*, 22 (7): 2277–2290. doi: 10.1105/tpc.110.074526
- Mehrota S e Goyal V (2014) Repetitive Sequences in Plant Nuclear DNA: Types, Distribution, Evolution and Function. *Genomics, Proteomics and Bioinformatics* 12(4):164-171 doi: 10.1016/j.gpb.2014.07.003
- Mercado-Ruaro P, Delgado-Salinas A (1998) Karyotypic studies on species of *Phaseolus* (Fabaceae: Phaseolinae). *Am J Bot.* 85(1)1-9
- Mercado-Ruaro P, Delgado-Salinas A (2000) Cytogenetic studies in *Phaseolus* L. (Fabaceae). *Genetics and Molecular Biology* 23(4):985-987

- Neumann P, Koblížková A, Navrátilová A, Macas J (2006) Significant expansion of *Vicia pannonica* genome size mediated by amplification of a single type of giant retroelement. *Genetics* 173:1047-56.
- Novák P, Neumann P, Macas J (2010) Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. *BMC Bioinformatics* 11:378. doi:10.1186/1471-2105-11-378
- Novák P, Neumann P, Pech J, Steinhaisl J, Macas J (2013) Repeat Explorer: a galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* 29:792–793. doi: 10.1093/bioinformatics/btt054
- Novák P, Ávila LR, Koblížková A, Vrbová I, Neumann P, Macas J (2017) TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. *Nucleic Acids Res.* doi: 10.1093/nar/gkx257
- Pedrosa A, Sandal N, Stougaard J, Schweizer D, Bachmair A (2002) Chromosomal map of the model legume *Lotus japonicus*. *Genetics* 161(4): 1661–1672.
- Pita S, Panzera F, Mora P, Vela J, Cuadrado Á, Sánchez A, Palomeque T, Lorite P (2017). Comparative repeatome analysis on *Triatoma infestans* Andean and Non-Andean lineages, main vector of Chagas disease. *PLOS ONE*, 12(7), e0181635. doi:10.1371/journal.pone.0181635
- Prakhongcheep O, Thapana W, Suntronpong A, Singchat W, Pattanatanang K, Phatcharakullawarawat R, Muangmai N, Peyachoknagul S, Matsubara K, Ezaz T, Srikulnath K (2017) Lack of satellite DNA species-specific homogenization and relationship to chromosomal rearrangements in monitor lizards (Varanidae, Squamata). *BMC Evolutionary Biology*, 17(1) doi:10.1186/s12862-017-1044-6
- Puterova J, Kubat Z, Kejnovsky E, Jesionek W, Cizkova J, Vyskot B, Hobza R (2018) The slowdown of Y chromosome expansion in dioecious *Silene latifolia* due to DNA loss and male-specific silencing of retrotransposons. *BMC Genomics*. 19: 153
- Raskina O, Belyayev A, Nevo E (2004) Activity of the *En/Spm*-like transposons in meiosis as a base for chromosome repatterning in a small, isolated, peripheral population of *Aegilops speltoides* Tausch. *Chromosome Res.*, 12, 153–161.
- Rendón-Anaya M, Montero-Vargas JP, Saburido-Álvarez M, Vlasova A, Capella-Gutierrez S, Ordaz-Ortiz JJ, *et al* (2017) Genomic history of the origin and domestication of common bean unveils its closest sister species. *Genome Biology* 18:60

- Ribeiro T, dos Santos KG, Richard MM, Sévignac M, Thareau V, Geffroy V, Pedrosa-Harand A (2017) Evolutionary dynamics of satellite DNA repeats from *Phaseolus* beans. *Protoplasma* 254(2):791-801. doi: 10.1007/s00709-016-0993-8
- Ribeiro T, Vasconcelos E, dos Santos KG, Vaio M, Brasileiro-Vidal AC, Pedrosa-Harand A (2019) Diversity of repetitive sequences within compact genomes of *Phaseolus* L. beans and allied genera *Cajanus* L. And *Vigna* Savi. *Chromosome Res.* <https://doi.org/10.1007/s10577-019-09618-w>
- Richard MMS, Chen NWG, Thareau V *et al.* (2013) The subtelomeric *kipu* satellite repeat from *Phaseolus vulgaris*: lessons learned from the genome analysis of the andean genotype G19833. *Front Plant Sci* 4:109. doi:10.3389/fpls.2013.00109
- Robledillo LA, Koblížková A, Novák P, Böttinger K, Vrbová I, Neumann P, Schubert I, Macas J (2018) Satellite DNA in *Vicia faba* is characterized by remarkable diversity in its sequence composition, association with centromeres, and replication timing. *Scientific Reports*, 8: 5838
- Ruiz-Ruano FJ, López-León MD, Cabrero J, Camacho JPM (2016) High-throughput analysis of the satellitome illuminates satellite DNA evolution. *Sci. Rep.* 6: 28333
- Santos FC, Guyot R, do Valle CB, Chiari L, Techio VH, Heslop-Harrison P, Vanzela AL (2015) Chromosomal distribution and evolution of abundant retrotransposons in plants: gypsy elements in diploid and polyploid *Brachiaria* forage grasses. *Chromosome Res.* 23(3):571-82. doi: 10.1007/s10577-015-9492-6
- Schmutz J, McClean PE, Mamidi S *et al.* (2014) A reference genome for common bean and genome-wide analysis of dual domestications. *Nat Genet* 46:707–713. doi:10.1038/ng.3008
- Schubert I, Lysak Ma (2011) Interpretation of karyotype evolution should consider chromosome structural constraints. *Trends in Genetics*, v.27, n.6, p.207-16
- Siljak-Yakovlev S, Godelle B, Zoldos V, Vallès J, Garnatje T, Hidalgo O (2017) Evolutionary implications of heterochromatin and rDNA in chromosome number and genome size changes during dysploidy: A case study in *Reichardia* genus. *PLoS ONE*, 12, e0182318.
- Smalec BM, Heider TN, Flynn BL, O'Neill RJ (2019) A centromere satellite concomitant with extensive karyotypic diversity across the *Peromyscus* genus defies predictions of molecular drive. *Chromosome Research.* doi:10.1007/s10577-019-09605-1
- Sonnhammer EL, Durbin R (1995) A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene* 29;167(1-2):GC1-10

- Weising K, Nybom H, Wolff K, Kahl G (2005) CTAB protocol I. In: DNA fingerprinting in plants : principles, methods, and applications 2nd ed. Taylor & Francis Group. pp100-102
- Wessler SR (2006) Eukaryotic Transposable Elements: Teaching Old Genomes New Tricks. The Implicit Genome pp 138–165
- Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, et al. A unified classification system for eukaryotic transposable elements. Nat Rev Genet. 2007;8:973–82.
- Zhang H, Koblizkova A, Wang K, Gong Z, Oliveira L, Torres GA, Wu Y, Zhang W, Novak P, Buell CR, Macas J, Jiang J (2014) Boom-bust turnovers of megabase-sized centromeric DNA in Solanum species: rapid evolution of DNA sequences associated with centromeres. Plant Cell 26:1436–1447. <https://doi.org/10.1105/tpc.114.123877>
- Zhu Q, Cai Z, Tang Q, Jin W (2016) Repetitive sequence analysis and karyotyping reveal different genome evolution and speciation of diploid and tetraploid *Tripsacum dactyloides*. Crop J. 4, 247–255

Tabela 1- Sequências dos *primers* selecionados para os DNA satélites identificados em *P. leptostachyus*, *P. lunatus* e *P. macvaughii*, espécie utilizada para amplificação e temperatura de anelamento (Ta) das reações de PCR

Nome	Sequência dos primers	Espécie amplificada	Ta (°C)
PleSAT592/PmaSAT64 8	5'- CCTACACAATGAAATAGACCG T-3' 5'- TCCCGTGTACTAACATTTTCAG -3	<i>P. leptostachyus</i>	50
PleSAT752/PluSAT991	5'- TCGACAAAACACAGTAGATAC A-3' 5'- TTTCTTCCGATAGAAAAGTGG A-3	<i>P. leptostachyus</i>	50
PluSAT1149	5'- GACGGTGAACCCTCCATTC-3' 5'- AACTCAGACGTGTTGGCAGT-3	<i>P. lunatus</i>	55

Tabela 2. Proporção do genoma (%) de sequências repetitivas identificadas na análise de 15 genomas de *Phaseolus*

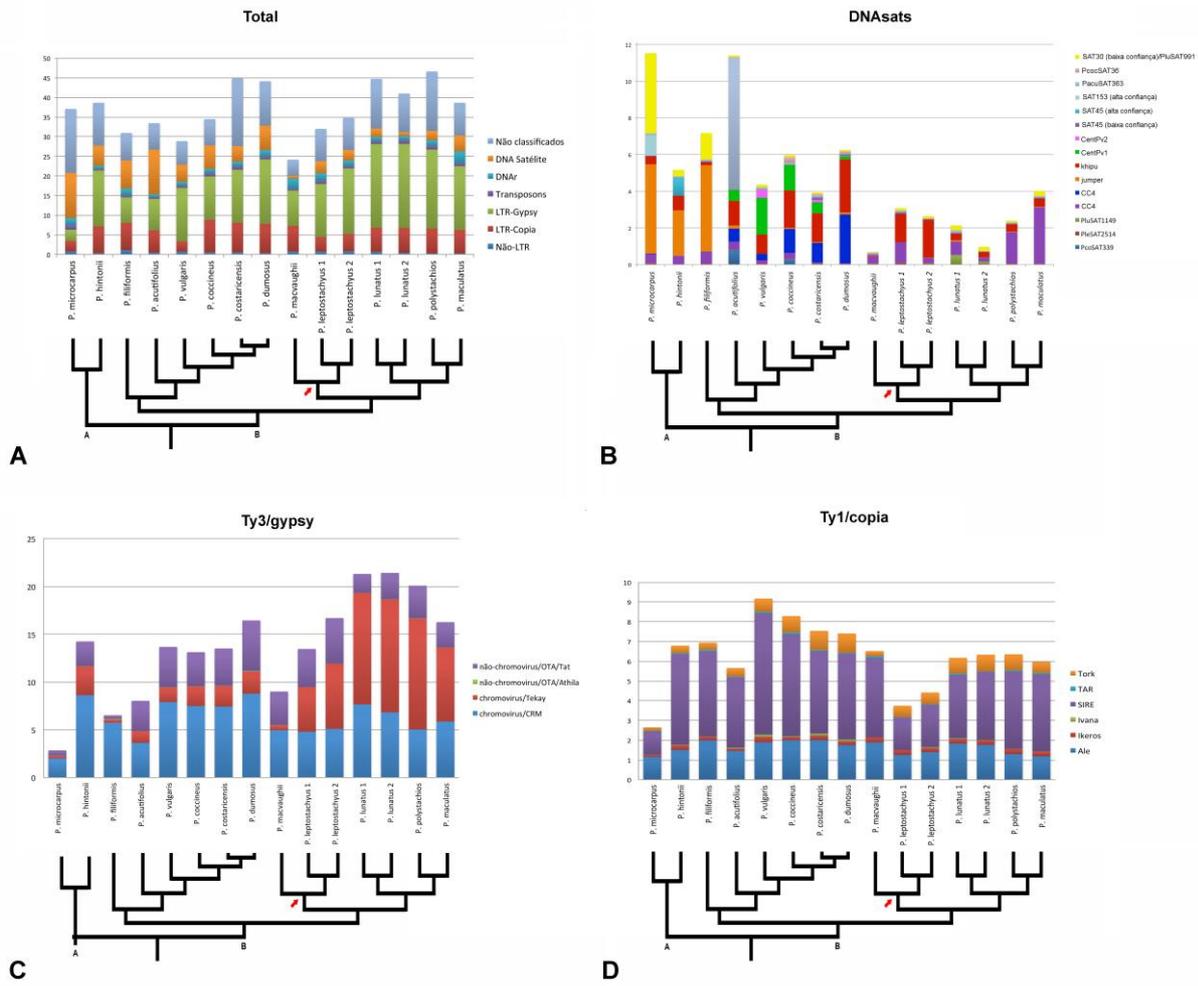
	<i>P. microcarpus</i>	<i>P. hintonii</i>	<i>P. filiformis</i>	<i>P. acutifolius</i>	<i>P. vulgaris</i>	<i>P. coccineus</i>	<i>P. costaricensis</i>	<i>P. dumosus</i>
<b>LTR</b>								
<i>Gypsy</i>								
CRM	1,99	8,61	5,74	3,62	7,90	7,48	7,42	8,78
Tekay	0,26	3,00	0,34	1,22	1,51	2,05	2,15	2,30
Athila	0,03	0,03	0,06	0,02	0,03	0,02	0,03	0,06
Tat	0,56	2,60	0,37	3,17	4,23	3,56	3,91	5,31
<b>Total Gypsy</b>	<b>2,84</b>	<b>14,25</b>	<b>6,52</b>	<b>8,03</b>	<b>13,67</b>	<b>11,05</b>	<b>13,51</b>	<b>16,45</b>
<i>Copia</i>								
Ale	1,15	1,51	1,99	1,46	1,89	2,00	2,00	1,75
Ikeros	0,10	0,23	0,19	0,12	0,27	0,17	0,23	0,19
Ivana	0,01	0,03	0,01	0,04	0,11	0,03	0,10	0,08
SIRE	1,19	4,60	4,30	3,56	6,15	5,16	4,18	4,35
Tar	0,03	0,05	0,08	0,04	0,07	0,09	0,06	0,05
Tork	0,17	0,38	0,36	0,44	0,67	0,83	0,97	0,98
<b>Total Copia</b>	<b>2,65</b>	<b>6,79</b>	<b>6,94</b>	<b>5,65</b>	<b>9,17</b>	<b>8,29</b>	<b>7,54</b>	<b>7,41</b>
<b>Total LTR</b>	<b>5,50</b>	<b>21,03</b>	<b>13,46</b>	<b>13,68</b>	<b>22,85</b>	<b>21,41</b>	<b>21,05</b>	<b>23,85</b>
<b>NÃO LTR</b>								
LINE	0,80	0,35	1,08	0,48	0,73	0,54	0,54	0,38
<b>DNA TRANSPOSON</b>								
CACTA	0,20	0,36	0,53	0,39	0,53	0,82	0,85	0,68
Mutator	1,01	0,01	0,01	0,01	0,03	0,01	0,01	0,01
<b>DNAr</b>								
35S	1,57	0,70	1,30	0,70	0,93	1,02	1,09	1,38
5S	0,19	0,12	0,40	0,02	0,06	0,09	0,10	0,25
<b>SATÉLITES</b>	<b>11,53</b>	<b>5,16</b>	<b>7,16</b>	<b>11,40</b>	<b>4,37</b>	<b>5,97</b>	<b>3,93</b>	<b>6,25</b>
<b>Não classificados</b>	<b>16,29</b>	<b>10,91</b>	<b>7,01</b>	<b>6,74</b>	<b>5,99</b>	<b>6,68</b>	<b>17,30</b>	<b>11,32</b>
<b>Total</b>	<b>37,08</b>	<b>38,64</b>	<b>30,95</b>	<b>33,43</b>	<b>35,48</b>	<b>36,54</b>	<b>44,88</b>	<b>44,12</b>

Tabela 2. Continuação

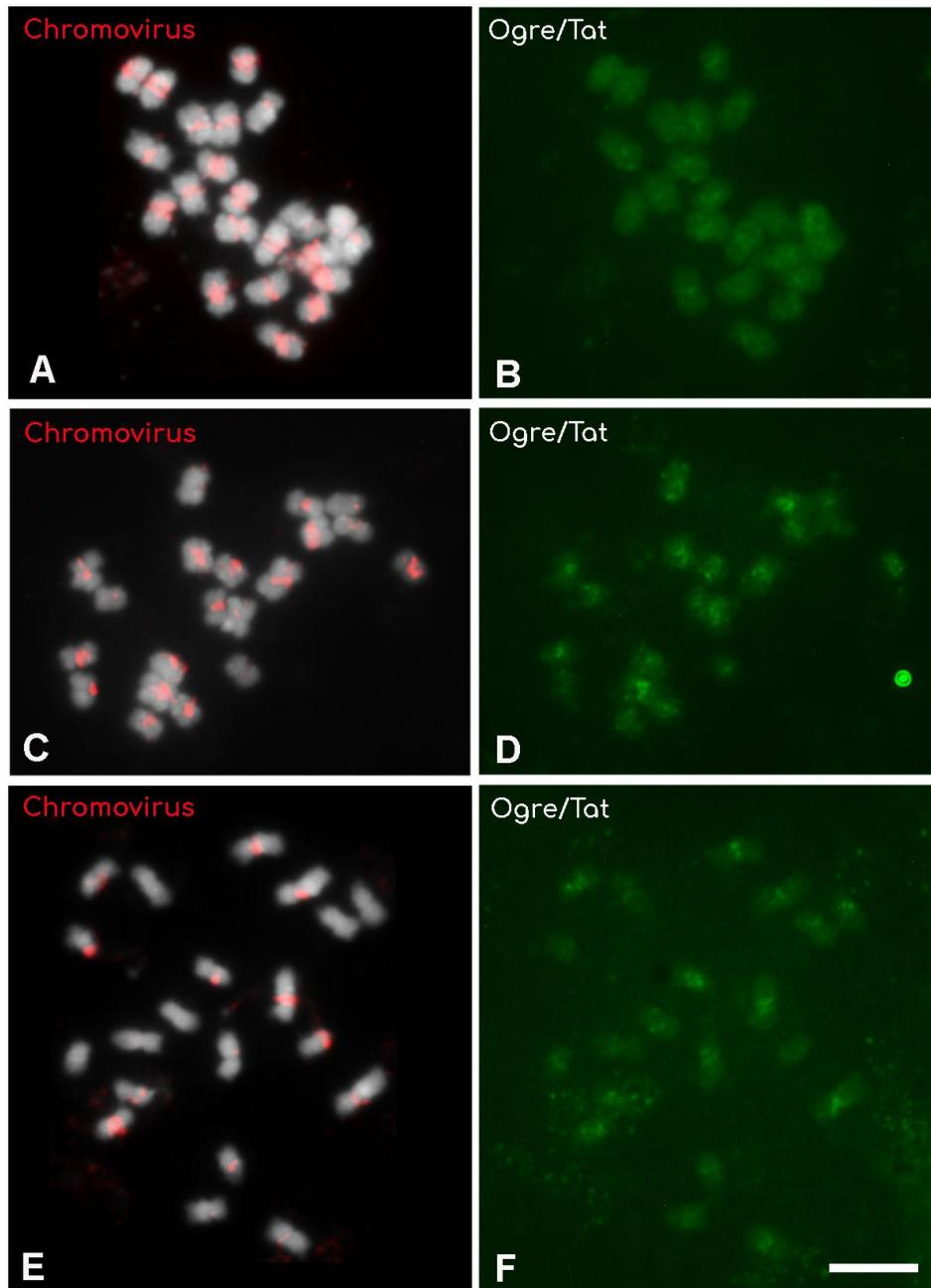
	P. macvaughii	P. leptostachyus 1	P. leptostachyus 2	P. lunatus 1	P. lunatus 2	P. polystachios	P. maculatus
<b>LTR</b>							
<i>Gypsy</i>							
CRM	4,95	4,78	5,08	7,64	6,82	5,03	5,86
Tekay	0,52	4,66	6,82	11,67	11,82	11,60	7,73
Athila	0,02	0,01	0,01	0,02	0,02	0,04	0,03
Tat	3,51	4,00	4,77	1,99	2,76	3,42	2,65
<b>Total Gypsy</b>	<b>9,00</b>	<b>13,46</b>	<b>16,69</b>	<b>21,32</b>	<b>21,42</b>	<b>20,09</b>	<b>16,26</b>
<i>Copia</i>							
Ale	1,89	1,25	1,39	1,82	1,76	1,29	1,18
Ikeros	0,26	0,25	0,22	0,23	0,23	0,25	0,24
Ivana	0,00	0,00	0,04	0,03	0,02	0,01	0,01
SIRE	4,03	1,65	2,15	3,21	3,46	3,92	3,91
Tar	0,08	0,02	0,03	0,07	0,02	0,06	0,06
Tork	0,26	0,57	0,60	0,80	0,85	0,81	0,56
<b>Total Copia</b>	<b>6,52</b>	<b>3,74</b>	<b>4,41</b>	<b>6,17</b>	<b>6,33</b>	<b>6,35</b>	<b>5,97</b>
<b>Total LTR</b>	<b>15,52</b>	<b>17,20</b>	<b>21,11</b>	<b>27,49</b>	<b>27,76</b>	<b>26,44</b>	<b>22,23</b>
<b>NÃO LTR</b>							
LINE	0,74	0,76	0,80	0,65	0,42	0,28	0,26
<b>DNA TRANSPOSON</b>							
CACTA	0,53	0,87	0,65	0,73	0,89	0,97	0,75
Mutator	0,00	0,01	0,01	0,01	0,02	0,00	0,01
<b>DNAr</b>							
35S	1,72	1,73	1,22	1,04	0,95	1,27	2,94
5S	0,80	0,11	0,17	0,02	0,14	0,12	0,08
<b>SATÉLITES</b>	<b>0,70</b>	<b>3,08</b>	<b>2,64</b>	<b>2,14</b>	<b>0,97</b>	<b>2,39</b>	<b>4,01</b>
<b>Não classificados</b>	<b>4,12</b>	<b>8,23</b>	<b>8,18</b>	<b>12,62</b>	<b>9,85</b>	<b>15,15</b>	<b>8,36</b>
<b>Total</b>	<b>24,13</b>	<b>31,99</b>	<b>34,78</b>	<b>44,70</b>	<b>40,98</b>	<b>46,62</b>	<b>38,65</b>

Tabela 3 – Resumo das famílias de DNA satélite identificadas no genoma de 3 espécies de *Phaseolus* usando o TAREAN

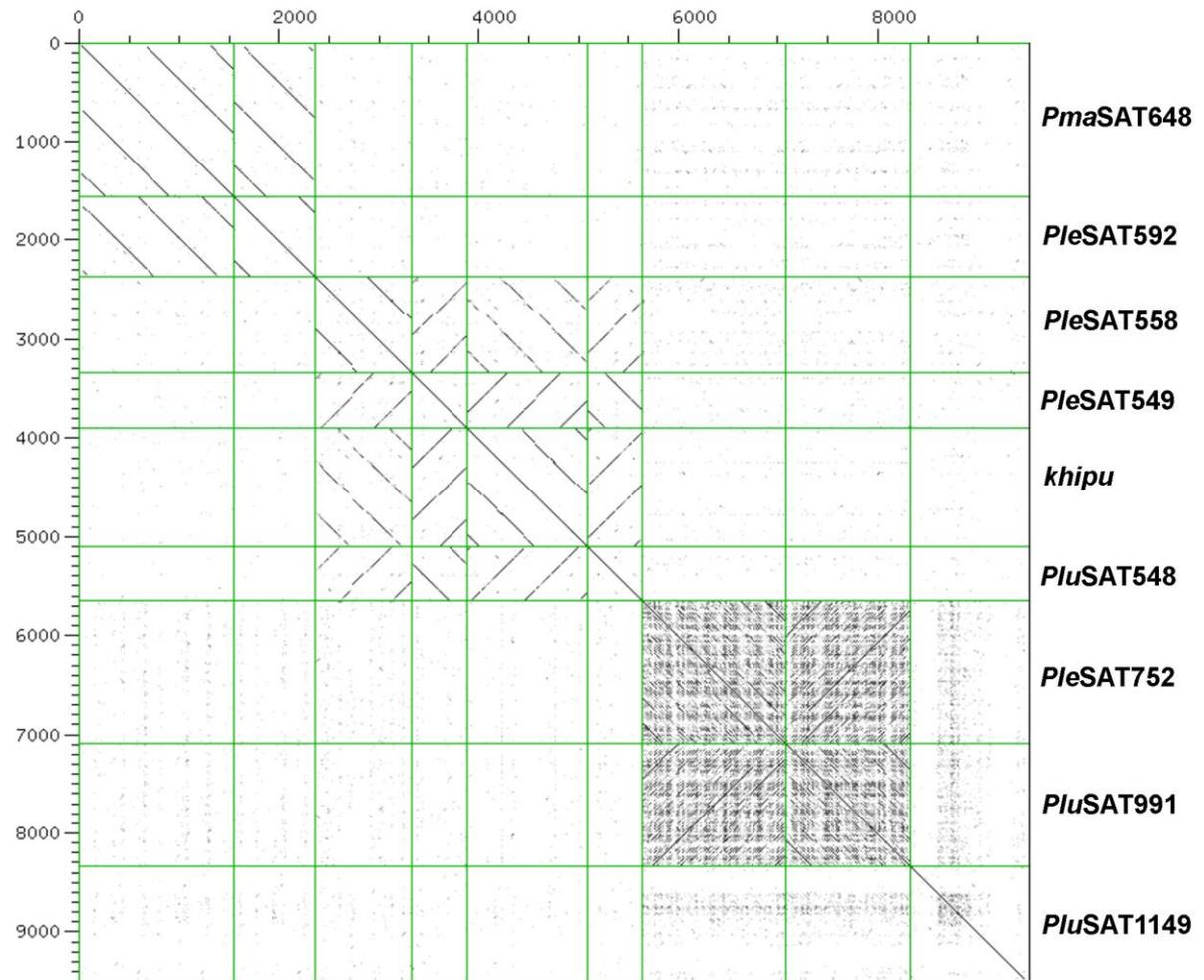
Espécie	Confiança	Tamanho (pb)	% genoma	Similaridade (%)
<i>P. leptostachyus (Ple)</i>	Alta	558	1,4	<i>khipu</i> (94,30)
		549	0,33	<i>khipu</i> (91,67)
		752	0,063	<i>PluSAT991</i> (98,01) <i>PmaSAT648</i>
	Baixa	592	0,016	(95,50)
		2514	0,056	
		5017	0,51	
<i>P. lunatus (Plu)</i>	Alta	991	0,12	<i>PleSAT752</i> (98,01)
		548	0,025	<i>khipu</i> (87,20)
	Baixa	1149	0,15	
<i>P. macvaughii (Pma)</i>	Baixa	28	0,38	
		648	0,028	<i>PleSAT592</i> (95,50)



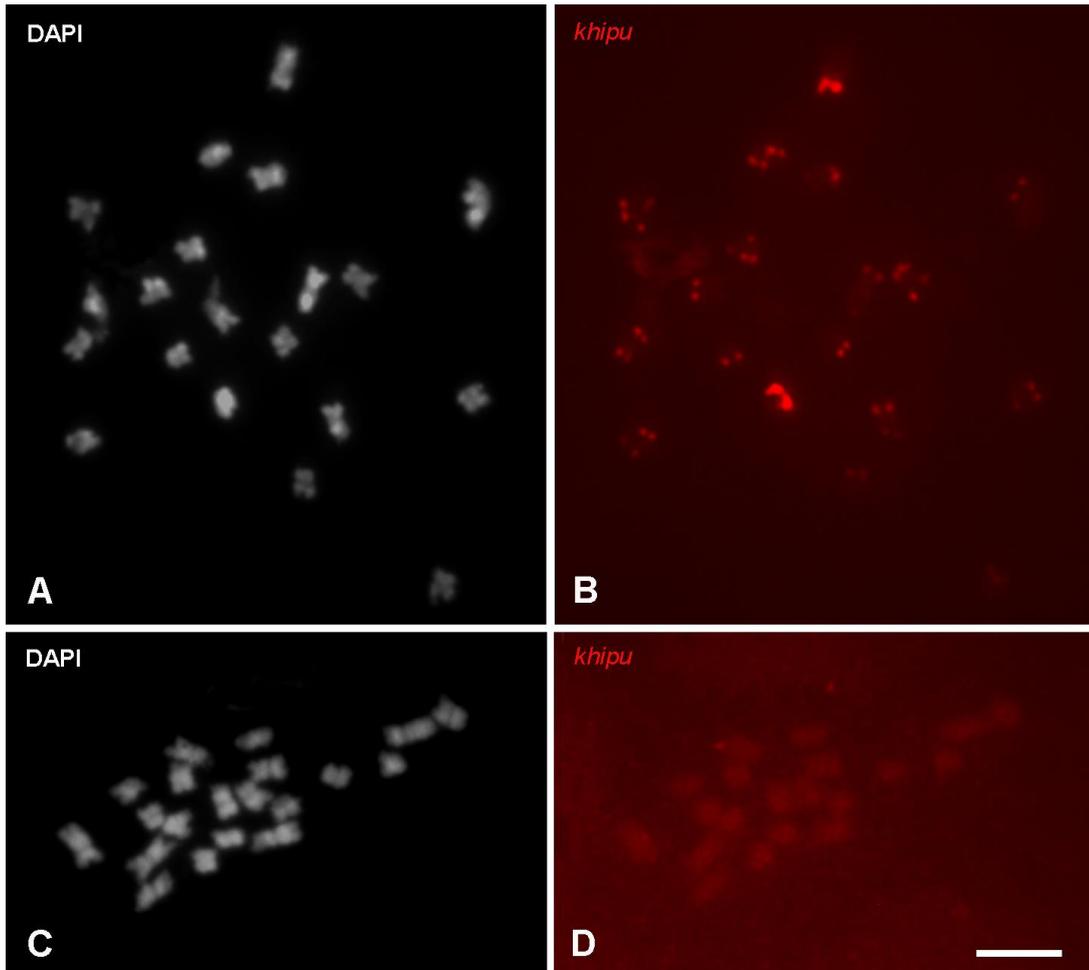
**Fig 1** Proporção dos principais elementos na fração repetitiva dos 15 genomas de *Phaseolus* analisados. Relações filogenéticas de acordo com Delgado-Salinas *et al.* (2006) separando o gênero em dois principais clados (A e B). Seta vermelha indica o grupo *Leptostachyus*



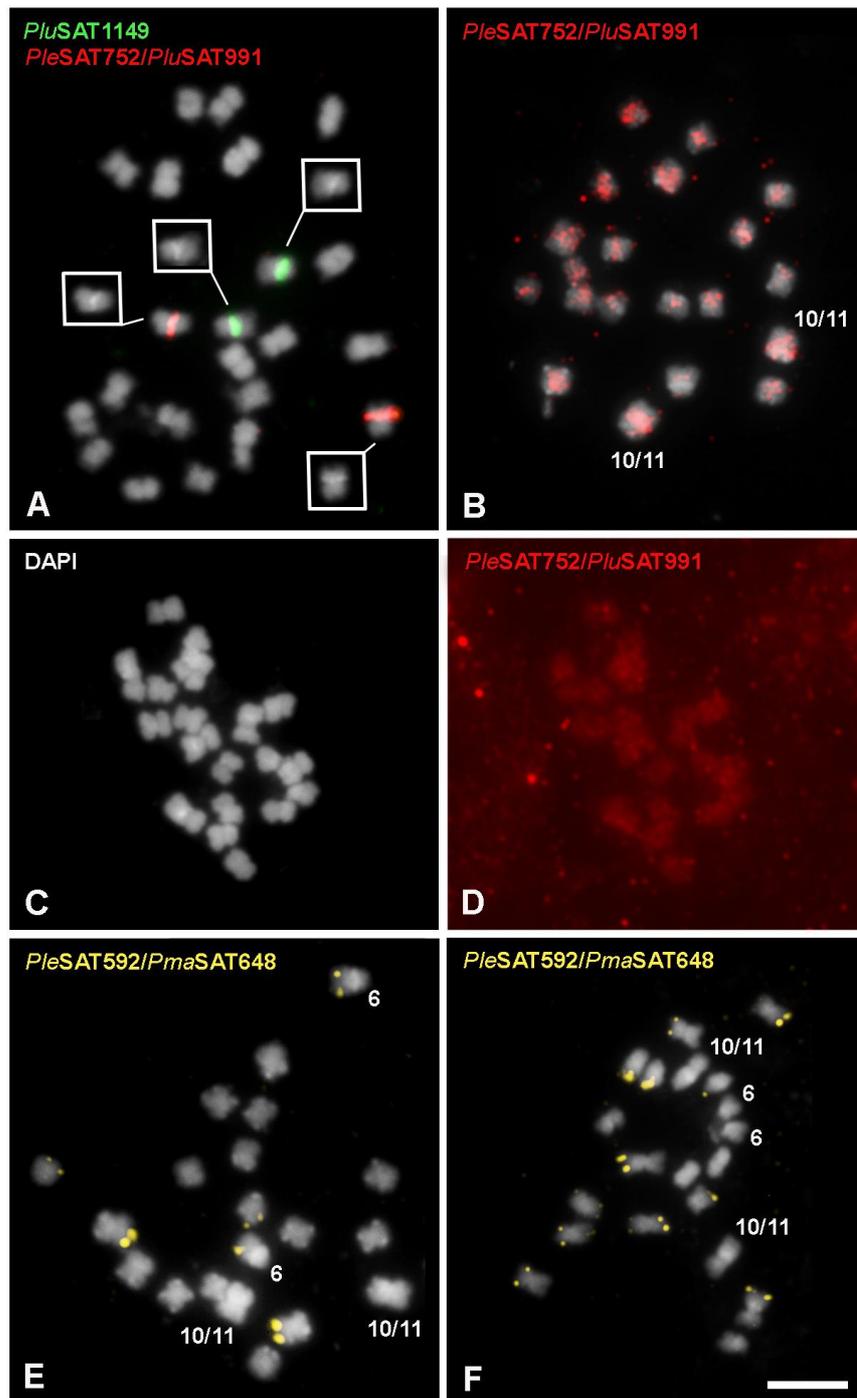
**Fig 2** Distribuição cromossômica das duas linhagens de Ty3/gypsy mais abundantes identificadas na análise *in silico* de *Phaseolus* hibridizadas em cromossômicos mitóticos de *P. lunatus* (A e B), *P. leptostachyus* (C e D) e *P. macvaughii* (E e F). Elemento Chromovirus (em vermelho, A, C e E) mapeado para a maioria das regiões pericentroméricas. Sinais para a sonda OGRE-Tat (em verde, B, D e F) também enriquecidos nessa região. Em A, C e E, os cromossomos foram contra-corados com DAPI e visualizados em cinza. Barra em F = 5  $\mu$ m



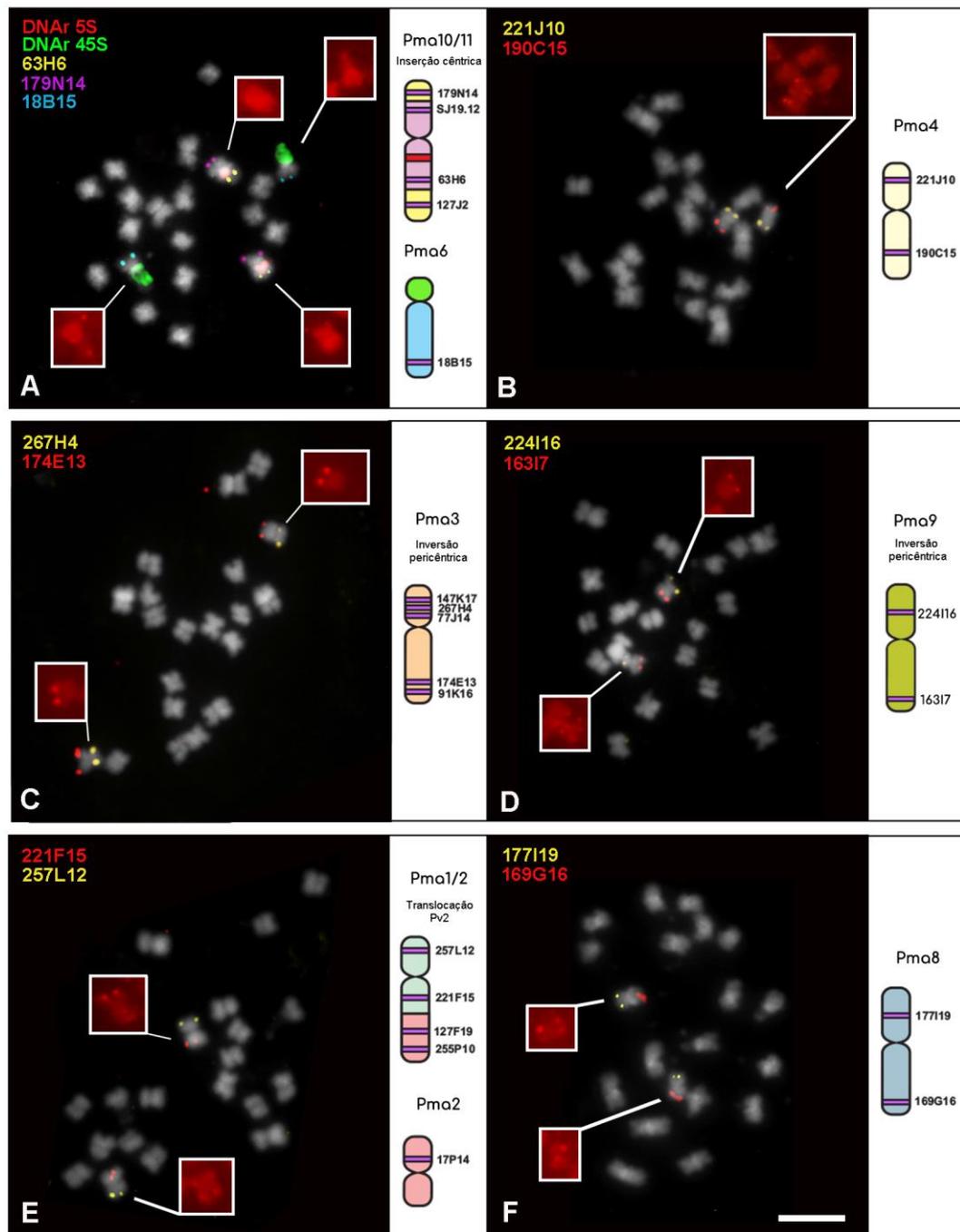
**Fig 3** Dotplot das repetições de satélite identificadas pelo TAREAN em *P. macvaughii*, *P. leptostachyus* e *P. lunatus* revelando similaridade entre as sequências PleSAT592 e PmaSAT648, PleSAT752 e PluSAT991, e entre PleSAT558, PleSAT549, PluSAT548 e *khipu* (David *et al.* 2009)



**Fig 4** Hibridização *in situ* fluorescente da sequência satélite *khipu* em cromossomos de *P. leptostachyus* (A e B) e *P. macvaughii* (C e D). Em B, sinais indicam a presença de *khipu* em *P. leptostachyus*. Em D, ausência de sinal de *khipu* em *P. macvaughii*. Barra em D = 5  $\mu$ m



**Fig 5** Localização cromossômica dos três DNA satélites revelados pelo TAREAN. Em **A**, *PluSAT1149* e *PleSAT752/PluSAT991* hibridizaram em *P. lunatus* como blocos nos pares de cromossomos portadores de uma banda heterocromática pós-FISH (inseto). Em *P. leptostachyus* (**B**) e *P. macvaughii* (**C** e **D**), o satélite *PleSAT752/PluSAT991* mostrou sinal disperso, com sinal mais forte no cromossomo maior (10/11) de *P. leptostachyus*. A sequência *PleSAT592/PmaSAT648* mostrou um padrão subtelo mérico com três pares marcados, incluindo o cromossomo 6, em *P. leptostachyus* (**E**), e seis pares em *P. macvaughii* (**F**). Os cromossomos foram contra-corados com DAPI e visualizados em cinza, com exceção de **D**. Barra em **F** = 5  $\mu$ m



**Fig 6** Identificação cromossômica por BAC-FISH em *P. macvaughii* (Ferraz *et al.* 2016) e rehibridização com *PleSAT592/PmaSAT648* (em vermelho, nos inserts). O *Pma10/11*, maior par cromossômico, e o *Pma6* (A) assim como o *Pma4* (B) não revelaram nenhum sinal. Os cromossomos *Pma3* (C) e *Pma9* (D) apresentaram sinal de hibridização em apenas um dos braços. Já os cromossomos *Pma1/2* (E) e *Pma8* (F) apresentam sítios em ambos os braços cromossômicos. A representação esquemática ao lado de cada célula mostra o mapeamento citogenético comparativo e seus rearranjos em relação a *P. vulgaris* (Fonsêca *et al.* 2010), indicando que não há relação entre a presença de sinais desse satélite e a ocorrência de rearranjos. Barra em F = 5  $\mu$ m

#### 4 CONSIDERAÇÕES FINAIS

- A análise comparativa da fração repetitiva de 15 genomas de *Phaseolus* revelou retrotransposons LTR, mais precisamente Ty3/*gypsy* da linhagem Chromovirus, como os elementos mais abundante do genoma.
- Os elementos Chromovirus e Ogre-Tat apresentaram enriquecimento na heterocromatina pericentromérica.
- Os elementos *Tekay* e SIRE apresentaram proporções muito diferentes no grupo *Leptostachyus* indicando um rápido turnover dessas sequências que poderiam estar associadas com a maior taxa de rearranjos cromossômicos.
- A maior divergência entre as espécies de *Phaseolus* esteve na fração de DNAsat. Três novos satélites foram revelados, sendo um exclusivo de *P. lunatus*, um compartilhado por *P. lunatus*, *P. leptostachyus* e em muito baixa proporção, *P. macvaughii*, e outro compartilhado apenas pelo grupo *Leptostachyus*.
- O satélite exclusivo de *Leptostachyus* apresentou distribuição subtelomérica. Rearranjos entre esses blocos poderiam explicar a diferença na distribuição de sítios entre *P. leptostachyus* e *P. macvaughii*, mas não parecem relacionados às inversões e translocações presentes no grupo.
- PleSAT752/PluSAT991, detectado na análise comparativa em todas as espécies, revelou uma distribuição dispersa atípica no grupo *Leptostachyus*, o que pode ter propiciado rearranjos incomuns, especialmente em *P. leptostachyus*, onde está presente em maior abundância.

## REFERÊNCIAS

- ALMEIDA C, FONSÊCA A, DOS SANTOS KGB, MOSIOLEK M, PEDROSA-HARAND A. Contrasting evolution of a satellite DNA and its ancestral IGS rDNA in *Phaseolus* (Fabaceae). *Genome* 55:683–689. 2012. doi:10.1139/g2012-059
- AYRES-ALVES T, CARDOSO AL, NAGAMACHI CY, DE SOUSA LM, PIECZARKA JC, NORONHA RCR. Karyotypic evolution and chromosomal organization of repetitive DNA sequences in species of *Panaque*, *Panaeolus*, and *Scobinancistrus* (Siluriformes and Loricariidae) from the Amazon Basin. *Zebrafish* 14, 251–260. 2017.
- BISCOTTI MA, OLMO E, HESLOP-HARRISON JP. Repetitive DNA in eukaryotic genomes. *Chromosome Res* 23:415–420. 2015.
- BONIFÁCIO EM, FONSÊCA A, ALMEIDA C, DOS SANTOS KGB, PEDROSA-HARAND A. Comparative cytogenetic mapping between the lima bean (*Phaseolus lunatus* L.) and the common bean (*P. vulgaris* L.). *Theor Appl Genet* 124:1513-1520. 2012. doi: 10.1007/s00122-012-1806-x
- BUERMANS, HPJ E DUNNEN JT. Next generation sequencing technology: Advances and applications. *Biochimica et Biophysica Acta (BBA) – Molecular Basis of Disease*. Vol 1842, issue 10 october 2014 1932- 1941. 2014.
- CERMAK T, KUBAT Z, HOBZA R, KOBLIZKOVA A, WIDMER A, MACAS J, et al. Survey of repetitive sequences in *Silene latifolia* with respect to their distribution on sex chromosomes. *Chromosome Res* 16:961–76. 2008.
- DAVID P, CHEN NWG, PEDROSA-HARAND A, THAREAU V, SÉVIGNAC M, CANNON SB, DEBOUCK D, LANGIN T, GEFFROY V. A nomadic subtelomeric disease resistance gene cluster in common bean. *Plant Physiol* 151:1048–1065. 2009. doi:10.1104/pp.109
- DELGADO-SALINAS A, BIBLER R, LAVIN M. Phylogeny of the Genus *Phaseolus* (Leguminosae): A Recent Diversification in an Ancient Landscape. *Syst. Botany* 31(4)779-791. 2006. doi: 10.1600/036364406779695960
- DODSWORTH S, CHASE MW, LEITCH AR. Is post-polyploidization diploidization the key to the evolutionary success of angiosperms? *Bot. J. Linn. Soc.* 180, 1–5. 2016.
- FALQUET J, CRESOUT F, DRON M. Molecular analysis of *Phaseolus vulgaris* rDNA unit and characterization of a satellite DNA homologous to IGS subrepeats. *Plant Physiol Biochem* 35:611–622. 1997.
- FERRAZ ME, FONSÊCA A, PEDROSA-HARAND A. Mapeamento citogenético comparativo em *Phaseolus macvaughii* Delgado: mais do que uma simples dispoloidia. 2016. 66 f. Trabalho de Conclusão de Curso - Universidade Federal de Pernambuco, Recife

- FONSÊCA A, FERRAZ ME, PEDROSA-HARAND A. Speeding up chromosome evolution in *Phaseolus*: multiple rearrangements associated with a one-step descending dysploidy. *Chromosoma* 125(3):413-21. 2016. doi: 10.1007/s00412-015-0548-3
- FONSÊCA A, FERREIRA J, DOS SANTOS TRB, MOSIOLEK M, BELLUCCI E, KAMI J, GEPTS P, GEFFROY V, SCHWEIZER D, DOS SANTOS, KGB, PEDROSA-HARAND A. Cytogenetic map of common bean (*Phaseolus vulgaris* L.). *Chromosome Res.* 18(4):487-502. 2010. doi: 10.1007/s10577-010-9129-8
- FONSÊCA A, PEDROSA-HARAND A. Karyotype stability in the genus *Phaseolus* evidenced by the comparative mapping of the wild species *Phaseolus microcarpus*. *Genome* 56(6):335-43. 2013. doi: 10.1139/gen-2013-0025
- FRANSZ P, LINC G, LEE C-R, AFLITOS SA, LASKY JR, TOOMAJIAN C, ALI H, PETERS J, VAN DAM P, JI X, KUZAK M, GERATS T, SCHUBERT I, SCHNEEBERGER K, COLOT V, MARTIENSSEN R, KOORNNEEF M, NORDBORG M, JUENGER TE, DE JONG H, SCHRANZ ME. Molecular, genetic and evolutionary analysis of a paracentric inversion in *Arabidopsis thaliana*. *Plant J*, 88:159-178. 2016.
- GAO D, ABERNATHY B, ROHKSAR D, SCHMUTZ J, JACKSON SA. Annotation and sequence diversity of transposable elements in common bean (*Phaseolus vulgaris*). *Front Plant Sci.* 5:339. 2014. doi: 10.3389/fpls.2014.00339
- GAO D, ZHAO D, ABERNATHY B, IWATA-OTSUBO A, HERRERA-ESTRELLA A, JIANG N, JACKSON SA. Dynamics of a Novel Highly Repetitive CACTA Family in Common Bean (*Phaseolus vulgaris*). *G3* 6(7):2091-101. 2016. doi: 10.1534/g3.116.028761
- GEFFROY V, MACADRÉ C, DAVID P, PEDROSA-HARAND A, SEVIGNAC M, DAUGA C, LANGIN T. Molecular analysis of a large subtelomeric nucleotide-binding-site-leucine-rich-repeat family in two representative genotypes of the major gene pools of *Phaseolus vulgaris*. *Genetics* 181: 405–419. 2009.
- GOODWIN TJ, POULTER RT. A group of deuterostome Ty3/gypsy-like retrotransposons with Ty1/copia-like pol-domain orders. *Mol. Genet. Genomics* 267: 481–491. 2002.
- GUERRA M. Hibridização in situ: princípios básicos. In: Conceitos e aplicações na citogenética. 1a edição, Ribeirão Preto, Sociedade Brasileira de Genética, p.1-32. 2004.
- GUERRA M. Chromosome numbers in plant cytotaxonomy: concepts and implications. *Cytogenetic and Genome Research* 120(3-4):339-350. 2008.
- HE Q, CAI Z, HU T, LIU H, BAO C, MAO W, et al. Repetitive sequence analysis and karyotyping reveals centromere-associated DNA sequences in radish (*Raphanus sativus* L.). *BMC Plant Biol.* 15, 105. 2015.
- IWATA A, TEK AL, RICHARD MMS, ABERNATHY B, FONSÊCA A, SCHMUTZ J, CHEN NW, THAREAU V, MAGDELENAT G, LI Y, MURATA M, PEDROSA-HARAND A, GEFFROY V, NAGAKI K, JACKSON SA. Identification and characterization of functional centromeres of the common bean. *Plant J* 76: 47–60. 2013. doi:10.1111/tpj.12269

- IWATA-OTSUBO A, RADKE B, FINDLEY S, ABERNATHY B, VALLEJOS CE, JACKSON SA. Fluorescence in situ hybridization (FISH)-based karyotyping reveals rapid evolution of centromeric and subtelomeric repeats in common bean (*Phaseolus vulgaris*) and relatives. *G3* 6(4):1013-22. 2016. doi: 10.1534/g3.115.024984
- LAMBOWITZ AM, GELLERT M, CHANDLER M, CRAIG NL, SANDMEYER SB, RICE PA (Eds). *Mobile DNA III*. American Society of Microbiology. 2015.
- LEE J et al. Rapid amplification of four retrotransposon families promoted speciation and genome size expansion in the genus *Panax*. *Scientific reports*, 7: 9045. LEWIS, G.P. *Legumes of Bahia*. Kew: Royal Botanic Gardens. XVI: 369. 1987. 2017.
- LEE YI, YAP JW, IZAN S, LEITCH IJ, FAY MF, LEE YC, et al. Satellite DNA in *Paphiopedilum* subgenus *Parvisepalum* as revealed by high-throughput sequencing and fluorescent in situ hybridization. *BMC Genomics* 19:578. 2018.
- MACAS J, NOVÁK P, PELLICER J et al. In depth characterization of repetitive DNA in 23 plant genomes reveals sources of genome size variation in the Legume tribe Fabeae. *PloS One* 10:e0143424. 2015.
- MANDÁKOVÁ T, JOLY S, KRZYWINSKI M, MUMMENHOFF K, LYSAK MA. Fast Diploidization in Close Mesopolyploid Relatives of *Arabidopsis*. *The Plant Cell*, 22 (7): 2277–2290. 2010. doi: 10.1105/tpc.110.074526
- MARQUES A, RIBEIRO T, NEUMANN P et al. Holocentromeres in *Rhynchospora* are associated with genome-wide centromere-specific repeat arrays interspersed amongst euchromatin. *Proceedings of the National Academy of Sciences of the USA* 112:13633–13638. 2015.
- MEHROTA S E GOYAL V. Repetitive Sequences in Plant Nuclear DNA: Types, Distribution, Evolution and Function. *Genomics, Proteomics and Bioinformatics* 12(4):164-171. 2014. doi: 10.1016/j.gpb.2014.07.003
- MELTERS DP, BRADNAM KR, YOUNG HA, TELIS N, MAY MR, RUBY JG, et al. Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biol* 14:R10. 2013.
- MERCADO-RUARO P, DELGADO-SALINAS A. Karyotypic studies on species of *Phaseolus* (Fabaceae: Phaseolinae). *Am J Bot*. 85(1)1-9. 1998.
- MERCADO-RUARO P, DELGADO-SALINAS A. Cytogenetic studies in *Phaseolus L.* (Fabaceae). *Genetics and Molecular Biology* 23(4):985-987. 2000.
- NOVÁK P, NEUMANN P, MACAS, J. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. *BMC bioinformatics* 11: 378. 2010.
- NOVÁK P, NEUMANN P, PECH J, STEINHAISSL J, MACAS, J. RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* 29: 792-793. 2013.

- NOVÁK P, ÁVILA ROBLLEDILLO L, KOBLÍŽKOV A, VRBOVÁ I, NEUMANN P, MACAS J. TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. *Nucleic Acids Res.* 2017. doi: 10.1093/nar/gkx257
- PEDROSA-HARAND A, KAMI J, GEPTS P, GEFFROY V, SCHWEIZER D. Cytogenetic mapping of common bean chromosomes reveals a less compartmentalized small-genome plant species. *Chromosome Research*, v.17, n.3, p.405–417. 2009.
- PETERSON AH, et al. Integration of Cot Analysis, DNA Cloning, and High- Throughput Sequencing Facilitates Genome Characterization and Gene Discovery. *Genome Research* 12: 795–807. 2002.
- RENDÓN-ANAYA M, MONTERO-VARGAS JP, SABURIDO-ÁLVAREZ M, VLASOVA A, CAPELLA-GUTIERREZ S, ORDAZ-ORTIZ JJ, et al. Genomic history of the origin and domestication of common bean unveils its closest sister species. *Genome Biology* 18:60; 2017.
- RIBEIRO T, DOS SANTOS KG, FONSÊCA A, PEDROSA-HARAND A. Isolation and characterization of a new repetitive DNA family recently amplified in the Mesoamer- ican gene pool of the common bean (*Phaseolus vulgaris* L., Fabaceae). *Genetica* 139:1135–1142. 2011.
- RIBEIRO T, DOS SANTOS KG, RICHARD MM, SÉVIGNAC M, THAREAU V, GEFFROY V, PEDROSA-HARAND A. Evolutionary dynamics of satellite DNA repeats from *Phaseolus* beans. *Protoplasma* 254:791-801. 2017. doi: 10.1007/s00709-016-0993-8
- RIBEIRO T, VASCONCELOS E, DOS SANTOS KG, VAIO M, BRASILEIRO-VIDAL AC, PEDROSA-HARAND A. Diversity of repetitive sequences within compact genomes of *Phaseolus* L. beans and allied genera *Cajanus* L. And *Vigna* Savi. *Chromosome Res.* 2019. <https://doi.org/10.1007/s10577-019-09618-w>
- RICHARD MMS, CHEN NWG, THAREAU V et al. The subtelomeric khipu satellite repeat from *Phaseolus vulgaris*: lessons learned from the genome analysis of the andean genotype G19833. *Front Plant Sci* 4:109. 2013. doi:10.3389/fpls.2013.00109
- SARBHOY R K. Cytogenetical studies in the genus *Phaseolus* Linn. III. Evolution in the genus *Phaseolus*. *Cytologia*, v.42, n.3-4, p.40-413. 1977.
- SARBHOY R K. Karyological studies in the genus *Phaseolus* Linn. *Cytologia*, v.45, n.3, p.363-373. 1980.
- SINHA SSN, ROY H. Cytological studies in the genus *Phaseolus*: Mitotic analysis of fourteen species. *Cytologia*, v.44, n.1, p.191-199. 1979.
- SCHMUTZ J, MCCLEAN PE, MAMIDI S, WU GA, CANNON SB, GRIMWOOD J, et al. A reference genome for common bean and genome-wide analysis of dual domestications. *Nat Genet*, 46(7):707-13. 2014. doi: 10.1038/ng.3008

SCHUBERT I E GIANG TH VU. Genome Stability and Evolution: Attempting a Holistic View. *Trends in Plant Science* 21 (9), 749-757. 2016. doi: 10.1016/j.tplants.2016.06.003

SCHUBERT I E LYSAK MA. Interpretation of karyotype evolution should consider chromosome structural constraints. *Trends in Genetics*, v.27, n.6, p.207-16. 2011.

SLOTKIN RK E MARTIENSSEN R. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet*, 8:272-285. 2007.

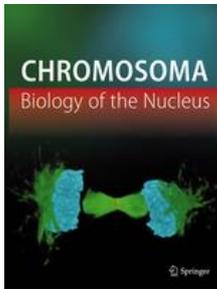
STORME A N E MASON B A. Plant speciation through chromosome instability and ploidy change: Cellular mechanisms, molecular factors and evolutionary relevance. *Current Plant Biology* 1:10–33. 2014.

WESSLER SR. Eukaryotic Transposable Elements: Teaching Old Genomes New Tricks. *The Implicit Genome* pp 138–165. 2006.

ZHANG L, XU C, YU W. Cloning and characterization of chromosomal markers from a Cot-1 library of peanut (*Arachis hypogaea* L.). *Cytogenet Genome Res* 137: 31–41. 2012.

ZHU Q, CAI Z, TANG Q, JIN W. Repetitive sequence analysis and karyotyping reveal different genome evolution and speciation of diploid and tetraploid *Tripsacum dactyloides*. *Crop J.* 4, 247–255. 2016.

## ANEXO A - NORMAS DA REVISTA



### Chromosoma - Instructions for Authors

#### AUTHORSHIP POLICY

Authorship should incorporate and should be restricted to those who have contributed substantially to the work in one or more of the following categories:

- Conceived of or designed study
- Performed research
- Analyzed data
- Contributed new methods or models
- Wrote the paper

#### TYPES OF PAPERS

- **Research Articles**  
For Research Articles there are no strict limits on the length of text, the number of Figures or the number of references. However, authors are asked to present manuscripts in a concise fashion and use electronic supplementary material where appropriate. Results and Discussion can be in separate sections or combined.
- **Reviews / Mini-Reviews**  
"Chromosoma" publishes both Mini-reviews and full Review articles. Mini-reviews are expected to comprise approximately 2,500 words and up to 4 display items (tables and figures). Full review articles may run up to 10,000 words (including references) and 10 display items. Reviews and mini-reviews should be preceded by an abstract (max. 300 words) that introduces the subject and summarizes the purpose and take-home message(s) of the article. The main text should be structured by using subtitles. The mini-review format is particularly appropriate for focusing on a recent highlight in a field, whereas the full review provides the opportunity for a more comprehensive survey of a research topic. Suggestions for review topics are welcome at all times, but review articles should be submitted only after prior approval by a member of the Editorial Board. All review articles will be subject to peer review.

#### MANUSCRIPT SUBMISSION

##### Manuscript Submission

Submission of a manuscript implies: that the work described has not been published before; that it is not under consideration for publication anywhere else; that its

Link: [https://www.springer.com/life+sciences/cell+biology/journal/412?print\\_view=true&detailsPage=pltc\\_1060507](https://www.springer.com/life+sciences/cell+biology/journal/412?print_view=true&detailsPage=pltc_1060507)