

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA VEGETAL

EVOLUÇÃO DO DNA SATÉLITE SUBTELOMÉRICO *hipu* NO GÊNERO
***Phaseolus* L. (FABACEAE)**

TIAGO RIBEIRO BARROS DOS SANTOS

RECIFE

2012

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA VEGETAL

EVOLUÇÃO DO DNA SATÉLITE SUBTELOMÉRICO *khipu* NO GÊNERO
Phaseolus L. (FABACEAE)

Dissertação de mestrado apresentada por Tiago Ribeiro Barros dos Santos ao programa de Pós-Graduação em Biologia Vegetal da Universidade Federal de Pernambuco, como um dos requisitos para a obtenção do título de Mestre em Biologia Vegetal, área de concentração Florística e Sistemática.

Orientadora: Profa. Dra. Andrea Pedrosa-Harand
(Dept. de Botânica, UFPE)

RECIFE
2012

Catálogo na fonte
Elaine Barroso
CRB 1728

Santos, Tiago Ribeiro Barros dos
Evolução do DNA satélite subtelométrico *hipu* no gênero *Phaseolus*
L. (Fabaceae)/ Tiago Ribeiro Barros dos Santos– Recife: O Autor, 2012.

77 folhas: Il., fig., tab.

Orientadora: Andrea Pedrosa-Harand

Dissertação (mestrado) – Universidade Federal de
Pernambuco. Centro de Ciências Biológicas. Biologia
Vegetal, 2012

Inclui bibliografia e anexo

- 1. Genética vegetal 2. Feijão 3. Leguminosa I. Pedrosa- Harand,**
Andrea (orientadora) II. Título

581.35

CDD (22.ed.)

UFPE/CCB-2015-195

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA VEGETAL

EVOLUÇÃO DO DNA SATÉLITE SUBTELOMÉRICO *khipu* NO GÊNERO
Phaseolus L. (FABACEAE)

Aprovado em 27/02/2012

COMISSÃO EXAMINADORA:

Profa. Dra. Andrea Pedrosa-Harand (Orientadora), Depto. de Botânica, CCB, UFPE

Profa. Dra. Ana Christina Brasileiro-Vidal, Dept. de Genética, CCB, UFPE

Profa. Dra. Ana Emília Barros e Silva, Depto. de Biologia, CCA, UFPB

Agradecimentos

Aos meus familiares, em especial aos meus pais, Antônio e Fátima (em memória), por todo amor, carinho e incentivo.

A todos que contribuíram, direta ou indiretamente, para a realização desse trabalho, meus sinceros agradecimentos:

À Andrea pela paciência, valiosas discussões e excelente orientação.

Aos colaboradores Valérie Geffroy e Vicent Thareau (Université Paris Sud) pela troca de informações e discussões.

A Marcos de Moraes e André Ribas (Universidade Federal de Pernambuco) pelo auxílio no procedimento de clonagem.

Aos amigos do laboratório: Liliane, Artur, Sandra, Silvokleio, Gilvanice, Gustavo, Mariana, André, T. Exposito, Magdalena e Lidiane.

Ao CNPq e ao PPGBV-UFPE pelo suporte financeiro.

Obrigado a todos!

Resumo

No gênero *Phaseolus* (Fabaceae) os estudos abordando a fração repetitiva de DNA do genoma ainda são poucos. A caracterização do DNA satélite *hipu*, inicialmente realizada no feijão comum (*P. vulgaris*), demonstrou que o mesmo poderia estar presente em outras espécies do gênero, provavelmente com bastante variação entre elas. Nesse trabalho, uma análise mais detalhada deste DNA satélite foi realizada, envolvendo o isolamento da sequência presente em diferentes espécies (*P. leptostachyus*, *P. lunatus* e *P. microcarpus*), mapeamento por hibridização *in situ* fluorescente (FISH) e *Southern blot*. Um coquetel de clones de *hipu* do feijão comum (*hipu* mix) hibridizados *in situ* sob 40% de estringência demonstrou a presença de sequências relacionadas a *hipu* na região subterminal dos cromossomos das outras espécies, com variação na quantidade e intensidade de terminais hibridizados. Fragmentos da sequência foram amplificados dos DNAs genômicos das três espécies utilizando *primers* específicos de regiões mais conservadas e ligados a vetores de clonagem. Quando mapeados por FISH, os clones obtidos demonstraram diferenças na marcação dependendo da espécie utilizada. O clone de *P. lunatus* (*Plukhipu1*), por exemplo, foi localizado em um par cromossômico em *P. lunatus* e *P. leptostachyus*, enquanto que em *P. vulgaris* o mesmo hibridizou em diversos pares. Análise baseada em máxima verossimilhança incluindo 68 clones de *P. vulgaris*, 30 de *P. leptostachyus*, nove de *P. lunatus* e 21 de *P. microcarpus* revelou um grande clado contendo a maioria dos clones de *P. vulgaris* e todos de *P. lunatus*. Enquanto as unidades de *P. microcarpus* e *P. leptostachyus* foram mais divergentes e ficaram separadas em diversos outros ramos. A hibridização *Southern blot* revelou uma unidade de repetição de 500 pb, com múltiplos presentes em todas espécies. Em conjunto, os dados confirmam *hipu* como uma família diversa de DNA satélite subtelomérico presente em diversas espécies de *Phaseolus*. As diferenças observadas nos padrões de hibridização e tamanhos das unidades de repetição sugerem a presença de diferenças significativas desta sequência entre diferentes espécies do gênero.

Palavras-chave: DNA satélite, *hipu*, subtelômeros, *Phaseolus*

Abstract

In the genus *Phaseolus* (Fabaceae), studies about the repetitive fraction of its genome are still scarce. The characterization of the satellite DNA *khipu*, performed for the common bean (*P. vulgaris*), showed that it might be present in other species of the genus, probably with large variation among them. In the present work, a more detailed analysis of this satellite DNA was performed, including the isolation of this sequence from different species (*P. leptostachyus*, *P. lunatus* and *P. microcarpus*), localization by fluorescent *in situ* hybridization (FISH) and Southern blot hybridization. A mix of *khipu* clones from common bean hybridized *in situ* at low stringency (40%) showed the presence of *khipu*-like sequences on the subterminal region of the chromosomes from the three species, with variation in intensity and number of signals. Fragments of the sequence were amplified and cloned from *P. leptostachyus*, *P. lunatus* and *P. microcarpus* genomic DNAs using specific primers for more conserved regions. The obtained clones demonstrated differences in respect to their FISH patterns depending on the species. The clone from *P. lunatus* (*khipuPlu1*), for example, was only located on a chromosome pair in *P. lunatus* and *P. leptostachyus*, while in *P. vulgaris* the same clone hybridized to several chromosome pairs. Maximum likelihood unrooted dendogram including 68 clones from *P. vulgaris*, 30 from *P. leptostachyus*, nine from *P. lunatus* and 21 from *P. microcarpus* showed most of the clones from *P. vulgaris* and all from *P. lunatus* grouping in one larger clade. While those unit from *P. microcarpus* and *P. leptostachyus* were more divergent and split in several other branches. Southern blot pattern showed a *khipu* unit of ~ 500 bp, with presence of multimers in all species. Taken together, the present data confirm *khipu* as a diverse subtelomeric satellite DNA family present in several *Phaseolus* species. Differences in hybridization patterns and repeat unit sizes suggest the presence of relevant differences in this sequence among different species of the genus.

Keywords: satellite DNA, *khipu*, subtelomeres, *Phaseolus*

SUMÁRIO

Agradecimentos.....	4
Resumo.....	5
Abstract	6
1. Apresentação	11
2. Revisão de Literatura	14
2.1. Organização genômica nos eucariotos: eu- e heterocromatina	15
2.2. DNA repetitivo e organização cromossômica longitudinal.....	17
2.2.1. Centrômero e sequências repetitivas centroméricas.....	17
2.2.2. Telômero, telomerase e sequência repetitiva telomérica.....	19
2.2.3. Subtelômero e DNA repetitivo subtelomérico	20
2.3. Legumes	21
2.4. O Gênero <i>Phaseolus</i>	22
2.4.1. Posicionamento taxonômico	22
2.4.2. Importância sócio-econômica	23
2.4.3. O feijão comum (<i>P. vulgaris</i>).....	23
2.4.4. A fava (<i>P. lunatus</i>)	24
2.4.5. <i>Phaseolus microcarpus</i> e <i>P. leptostachyus</i> Benth.....	24
2.4.6. Citogenética em <i>Phaseolus</i>	25
2.4.7. Sequências repetitivas em <i>Phaseolus</i>	26
3. Referências bibliográficas	28
4. Manuscrito submetido ao periódico <i>Annals of Botany</i>	36
5. Conclusões	74
6. Anexo	79

LISTA DE TABELAS E FIGURAS

- Tabela 1.** Distribuição cromossômica de uma mistura de *khipu* de *P. vulgaris* (*Pvkhipu mix*) em espécies de *Phaseolus* sob diferentes estringências.....59
- Figura 1.** Distribuição cromossômica de um conjunto de *khipu* de *P. vulgaris* (*Pvkhipu mix*) hibridizado *in situ* sob baixa e alta estringência em diferentes espécies de *Phaseolus*: *P. vulgaris* (a e e); *P. leptostachyus* (b e f); *P. lunatus* (c e g) e *P. microcarpus* (d e h). Em a-d, as imagens são mostradas sobrepostas ao DAPI (cinza). Asteriscos em b, c e d indicam os cromossomos com os sinais de hibridização mais fortes e as setas em b, c, d, f, g e h apontam os cromossomos com regiões organizadoras do nucléolo (RON). Em f, g e h cabeças-de-seta indicam todos os terminais marcados e em a apenas os mais fracos. Barra em h = 5 µm.....60
- Figura 2.** Tipos cromossômicos identificados em *P. vulgaris* e outras espécies de *Phaseolus* (*P. leptostachyus*, *P. lunatus* e *P. microcarpus*) baseado no padrão de hibridização com um conjunto de *khipu* de *P. vulgaris*. Tipo I, marcação em ambos os braços; tipo II, marcação em um dos braços e tipo III, nenhuma marcação.....61
- Figura 3.** Árvore filogenética de sequências de *khipu* de diferentes espécies de *Phaseolus* baseada em máxima verossimilhança. Cores azul, laranja e verde identificam as sequências de *P. leptostachyus*, *P. microcarpus* e *P. lunatus*, respectivamente. Sequências em preto foram extraídas do genoma de *P. vulgaris*. Setas indicam os clones utilizados como sonda em FISH na Figura 5.....62
- Figure 4.** Organização molecular de *khipu* em *Phaseolus* visualizada por hibridização *Southern blot*. DNAs foram hibridizados sob alta (a) e baixa (b) estringência com uma mistura de *khipu* de *P. vulgaris*. Relações filogenéticas mostradas no topo da figura foram baseadas em Delgado-Salinas et al. (2006).....63
- Figura 5.** FISH sob alta estringência de clones *khipu* isolados de diferentes espécies de *Phaseolus*. Distribuição do clones *Plukhipu7* em *P. lunatus* (a) e *P. leptostachyus* (b); (c-d) localização de *Pmkhipu1* em *P. microcarpus* (c) e *Plekhipu12* em *P. leptostachyus* (d); (e-f) hibridização de *Pmkhipu2* em *P. lunatus* (e) e *P. leptostachyus* (f). Barra em f = 5 µm.....64
- Figura 6.** Organização genômica de *jumper* em *P. vulgaris*. (a) Distribuição de *jumper* em dois pseudocromossomos (cromossomos 5 e 6) mostrando a sequência como parte de um fragmento maior (crom. 6) ou arranjada em tandem (crom. 5); (b) Alinhamento

de interjumper com o DNAr 5S de diferentes espécies mostrando sua similaridade com a região codificante (CDS); (c) Estrutura da unidade de repetição do DNAr 5S de *P. vulgaris*.....65

Figura 7. FISH de *jumper* (vermelho) e *khipu* (amarelo) de *P. microcarpus* e do DNAr 5S (azul) em cromossomos mitóticos de *P. microcarpus* (cinza). Cabeças-de-seta apontam sítios de colocação de *khipu-jumper* e asteriscos os cromossomos nos quais as sequências aparecem em terminais diferentes. P: pericentrômeros; S: subtelômeros Barra = 5 µm.....66

Figura 8. Southern blot de DNAs genômicos de *Phaseolus* (Po: *P. oligospermus*; Pm: *P. microcarpus*; Plu: *P. lunatus*; Ple: *P. leptostachyus* e Pv: *P. vulgaris*) e outros legumes (Cc: *Cajanus cajan* e Vr: *Vigna radiata*) digeridos com *HindIII* e hibridizados sob baixa estrigência com uma sonda de *jumper* de *P. microcarpus*.....67

Tabela S1. BACs subteloméricos analisados por southern blot. Estes clones mostraram a mesma marcação subtelomérica quando hibridizados *in situ* nos cromossomos do feijão comum (Pedrosa-Harand et al., 2009; Fonsêca et al., 2010).....68

Tabela S2. Primers usados para amplificação de sequências relacionadas a *khipu* em espécies de *Phaseolus*.....69

Tabela S3. Número de unidades de *jumper* no genoma de *P. vulgaris*.....70

Fig. S1. Southern blot de BACs subteloméricos hibridizados com *khipu* sob alta estrigência. Amostras foram coradas com SYBR gold (a e c) e então hibridizadas com *khipu* (b e d). Enzima de restrição usadas: H= *HindIII*, R= *RsaII*, E= *EcoRI*. BAC 147K17 é um BAC que apresenta sinais únicos após FISH e foi usado como controle.71

Fig. S2. Hibridização Southern blot de clones *khipu* de diferentes espécies de *Phaseolus* sob alta estrigência. (a) Clone *khipuPle2* de *P. leptoscahyus* e (b) clone *khipuPm1* de *P. microcarpus*. Cc: *Cajanus cajan*; Vu: *Vigna radiata*; Pm: *P. microcarpus*; Po: *P. oligospermus*; Pv: *P. vulgaris*; Ple: *P. leptoschyus* e Plu: *P. lunatus*.....72

Fig. S3. (a) Representação esquemática do clone de *P. microcarpus* (*Pmkhipu2*) contendo *jumper*. Região em cinza claro mostra o fragmento de *khipu* (450 pb), enquanto a cinza escuro representa as duas unidades de *jumper*; (b) alinhamento dos dois monômeros de *jumper* identificados no presente clone. Nucleotídeos marcados em negrito representam polimorfismos ao longo da sequência; (c) Amplificação de *jumper* por PCR em diferentes espécies de *Phaseolus* (Po: *P. oligospermus*; Pm: *P. microcarpus*; Pv: *P. vulgaris* ‘BAT93’ e ‘JaloEPP’; Ple: *P. leptostachyus*; Plu: *P. lunatus*) e quatro espécies relacionadas (Ma: *Macroptilium atropurpureum*; Vr: *Vigna*

radiata; *Vu*: *V. unguiculata* e *Cc*: *Cajanus cajan*). Barras na parte esquerda da figura representam os tamanhos dos fragmentos.....73

Fig. S4. Alinhamento múltiplo de um conjunto de NTS extraídos do DNAr 5S de dados genômicos de *P. vulgaris*. *Pv06*: *Pv* cromossomo 6 e *Pv10*: *Pv* cromossomo 10. Linha horizontais na lateral esquerda marcam os NTS sem *jumper*.....74

Fig. S5. Análise de máxima verossimilhança do satélite *jumper* de diferentes espécies de *Phaseolus*. *P. vulgaris* (*Pv*; preto), *P. oligospermus* (*Po*; vermelho) and *P. microcarpus* (*Pm*; azul).....76

1. Apresentação

1. Apresentação

O genoma dos eucariotos está constituído basicamente por dois tipos de sequências de DNA: sequências não repetitivas, dentre estas a maioria dos genes, e sequências repetitivas, as quais são geralmente subdivididas em dois tipos com base na organização e padrão de distribuição no genoma. O primeiro tipo corresponde às sequências dispersas, intercaladas com sequências únicas e distribuídas ao longo do genoma. O outro tipo corresponde às sequências organizadas em *tandem*, abrangendo o DNA satélite, minissatélite e microsatélite, as quais formam blocos, em todos ou em alguns cromossomos, com até milhares de unidades de repetição organizadas na mesma direção (*head-to-tail*). O DNA satélite vem sendo extensivamente estudado, mostrando que esse tipo de sequência está presente em diferentes regiões dos cromossomos (pericentrômero, subtelômero e regiões intercalares) e associado a diferentes processos celulares, tais como a manutenção da coesão entre cromátides e a correta segregação cromossômica (Sharma e Raina, 2005; Hemleben *et al.*, 2007; Gong *et al.*, 2009).

O gênero *Phaseolus* L. (Fabaceae) é monofilético e constituído por cerca de 70 espécies organizadas filogeneticamente em dois grandes clados (A e B) formados por espécies selvagens e cultivadas (Delgado-Salinas *et al.*, 2006). No clado B, cinco espécies vêm sendo domesticadas (*P. vulgaris* L., *P. lunatus* L., *P. coccineus* L., *P. polyanthus* Greenman e *P. acutifolius* A. Gray) (Broughton *et al.*, 2003) e duas delas apresentam grande importância econômica a nível mundial e regional. O feijão comum (*P. vulgaris*) é a principal leguminosa utilizada para consumo humano direto, sendo uma importante fonte protéica para populações da África e América Latina. No Brasil, *P. lunatus*, o feijão fava, é extensamente consumido pelas populações nordestinas.

Os estudos citogenéticos no grupo se iniciaram em 1925, quando da descrição do número cromossômico em *P. acutifolius*, *P. coccineus*, *P. lunatus* e *P. vulgaris* (revisado por Mercado-Ruaro e Delgado-Salinas *et al.*, 2000), permanecendo nas décadas seguintes restritos a novas contagens cromossômicas e descrições cariotípicas simples (Mok e Mok, 1976; Zheng *et al.*, 1991). Isso foi devido principalmente ao diminuto tamanho apresentado por seus cromossomos e à inexistência de técnicas capazes de distingui-los. Com o desenvolvimento da hibridização *in situ* fluorescente (FISH) e sua utilização como técnica rotineira nos estudos citogenéticos em *Phaseolus* foi possível pela primeira vez estabelecer, em conjunto com dados de tamanho cromossômico e padrão de heterocromatina, uma nomenclatura numérica para os

cromossomos de espécies do gênero (Moscone *et al.*, 1999). Mais recentemente, utilizando BACs (*Bacterial Artificial Chromosomes*) selecionados de uma biblioteca desenvolvida para o feijão comum (Kami *et al.*, 2006), foi estabelecido um mapa citogenético detalhado para a espécie (Pedrosa-Harand *et al.*, 2009; Fonsêca *et al.*, 2010) e diversos estudos comparativos vêm sendo realizados a partir de então em *P. lunatus* (Almeida, 2006; Bonifácio, 2009), *P. microcarpus* (Fonsêca, 2010) e *P. leptostachyus* (Fonsêca A, comunicação pessoal).

A fração repetitiva de DNA no gênero vem sendo, entretanto, ainda pouco estudada. No feijão comum aproximadamente 50% do genoma é constituído por sequências repetitivas (Schlueter *et al.*, 2009), embora apenas cinco famílias tenham sido caracterizada em detalhes até o momento (Falquet *et al.*, 1997; Métais *et al.*, 1998; Garber *et al.*, 1999; David *et al.*, 2009; Ribeiro *et al.*, 2011). O DNA satélite *khipu* foi identificado durante a caracterização molecular do loco B4 de resistência à antracnose do feijão comum, presente na região subterminal do braço curto do cromossomo 4 (David *et al.*, 2009). A sequência apresenta unidade consenso de 528 pb com conteúdo de GC de 38% e está distribuída ao longo da região subtelomérica da maioria dos cromossomos do feijão comum. Por meio de análises em membrana foi demonstrado que *khipu* apresenta uma grande variação dentro do gênero, com espécies mais relacionadas e outras mais dissimilares à *P. vulgaris* apresentando diferentes quantidades da sequência, podendo essas variações refletirem diferenças quantitativas ou ao nível de sequência. Embora *khipu* tenha sido detectado nesse trabalho em *P. lunatus* e *P. microcarpus*, clones de *P. vulgaris* que contem *khipu* não detectaram essa sequência ao serem hibridizados *in situ* nessas duas espécies (Almeida, 2006; Bonifácio, 2009; Fonsêca, 2010).

Desta maneira, o presente trabalho, teve por objetivo caracterizar o DNA satélite *khipu* em diferentes espécies de *Phaseolus*. Por meio de hibridização *in situ* fluorescente, *Southern blot* e análises de sequências, as unidades de repetição, assim como os padrões de distribuição de *khipu*, foram caracterizados em *P. leptostachyus*, *P. lunatus* e *P. microcarpus*.

2. Revisão de Literatura

2. Revisão de Literatura

2.1. Organização genômica nos eucariotos: eu- e heterocromatina

Os genomas eucarióticos consistem essencialmente de DNA associado a proteínas compartimentalizados como heterocromatina ou eucromatina nos cromossomos (Grewal & Elgin, 2002; Heslop-Harrison & Schwarzacher, 2011). Classicamente, a heterocromatina tem sido distinguida da eucromatina com base no seu padrão de condensação na intérfase, se apresentando como regiões altamente condensadas e com replicação tardia na fase S. Essas regiões consistem predominantemente de DNA repetitivo, incluindo DNAs satélites, elementos transponíveis e retrovírus, e tipicamente apresentam baixa densidade gênica, sendo referidas como heterocromatina constitutiva. Entretanto, em alguns casos, a heterocromatina também pode ser encontrada em locos regulados pelo desenvolvimento, onde o estado conformacional da cromatina pode ser alterado em resposta a sinais celulares e ativação gênica. O mecanismo de compensação de dose observado para o cromossomo X de mamíferos é um exemplo clássico da formação de heterocromatina facultativa. Ao contrário, as regiões eucromáticas se apresentam mais descompactadas, com um ciclo de condensação-descondensação paralelo ao ciclo celular, sendo ricas em genes, embora estes estejam irregularmente distribuídos, e mais acessíveis aos fatores de transcrição (Rennig, 1999; Chow e Brown, 2003; Grewal e Jia, 2007).

Embora esteja distribuída ao longo do cromossomo, a eucromatina tem sido dividida em diferentes categorias com base na presença de uma vasta gama de propriedades e num padrão de diferenciação longitudinal visualizado citologicamente. Dentre as técnicas que vem sendo mais extensamente utilizadas para a caracterização dessa fração do genoma, especialmente em mamíferos, estão os bandeamentos G e R. O bandejamento G é o principal método de análise citogenética clínica, bem como para outros estudos citogenéticos em, e em combinação com o bandejamento R, o qual apresenta um padrão de bandejamento complementar (revisado por Sumner, 2002), permite obter um panorama geral a respeito da organização da eucromatina de um genoma. As bandas G tipicamente são caracterizadas por apresentar cromatina rica em AT e com replicação tardia, sendo uma região com baixa densidade gênica; ao passo

que as bandas R tendem a ser ricas em GC, replicarem cedo na fase S e serem as regiões ricas em genes da eucromatina.

A heterocromatina constitutiva mostra pouca ou nenhuma atividade gênica, havendo uma tendência de se referir a toda cromatina geneticamente inativa e condensada como heterocromatina. De forma geral, todos os cromossomos apresentam heterocromatina, embora em algumas espécies a mesma não seja facilmente detectada pelas técnicas rotineiramente utilizadas, como o bandeamento C ou os fluorocromos base-específicos (Norris *et al.*, 2004; Sumner, 2002). Blocos heterocromáticos estão presentes na região pericentromérica na grande maioria dos cromossomos, embora também possam ser visualizados em posições intersticiais e terminais (Hansen, 1973; King *et al.*, 1986; Guerra, 2000).

O estabelecimento e a manutenção dessas regiões são realizados por um vasto grupo de proteínas responsáveis pela metilação de DNA, metilação de histonas e replicação da própria heterocromatina, nesse último caso em conjunto com a própria DNA polimerase. Além disso, tem sido relatada a associação de RNAs de interferência (RNAi) como um mecanismo de disparo, sendo a sua presença, juntamente com repressores transcricionais, um fator essencial para o estabelecimento da heterocromatina em eucariotos, como *Drosophila melanogaster* (Craig, 2004; Grewal e Jia, 2007). Essa fração da cromatina apresenta uma característica-chave, a habilidade de propagação, influenciando a expressão de uma determinada região de forma sequência-independente, num processo conhecido como silenciamento. A onda de heterocromatinização tende a se espalhar pelo genoma até ser suprimida pela presença de elementos limitadores ou *insulators*, sequências de DNA que atuam como barreira neutra impedindo a propagação do silenciamento para regiões adjacentes. Além disso, por meio do efeito posicional, PEV (*Positional Effect Variegation*), genes eucromáticos podem ser deslocados para heterocromatina ou para próximo dela e também serem bloqueados (Grewal e Elgin, 2002; Sumner, 2003; Grewal e Jia, 2007).

Embora durante muito tempo a heterocromatina tenha sido relatada como uma fração do genoma sem funcionalidade (revisado por Rennig, 1999), trabalhos recentes tem demonstrado justamente o contrário. A heterocromatina é responsável pelo grande aumento e variação no conteúdo de DNA total em muitos eucariotas, além de estar envolvida em diversos processos celulares, tais como pareamento cromossômico, recombinação e manutenção da coesão entre cromátides-irmãs e segregação cromossômica (Brutlag, 1980; Bernard *et al.*, 2001; Bernard e Allshire, 2002).

2.2. DNA repetitivo e organização cromossômica longitudinal

As sequências repetitivas são componentes universais dos genomas eucarióticos, podendo estar distribuídas de forma mais compacta em uma região específica dos cromossomos ou estarem espalhadas ao longo do complemento. Ao nível de sequência, as mesmas são subdivididas em duas grandes categorias: sequências repetidas em *tandem* e sequências repetidas dispersas. Nas sequências repetitivas em tandem (DNAs satélites, minissatélites e microsatélites), a unidade de repetição básica, denominada *repeat*, se encontra repetida dezenas ou milhares de vezes, distribuindo-se adjacientemente umas às outras, numa organização denominada *head-to-tail* e geralmente formando blocos nos cromossomos. As sequências repetidas dispersas, as quais se encontram mais comumente espalhadas no genoma, interespaçadas com as sequências gênicas, são compostas predominantemente por elementos transponíveis (revisado por Sharma e Raina, 2005).

Em plantas, as diferentes classes de sequências repetitivas parecem apresentar um padrão de distribuição genômica bem característico, o que contribuiu para o estabelecimento de um modelo de distribuição das mesmas ao longo dos cromossomos (Schmidt e Heslop-Harrison, 1998). Nesse modelo os genes estão organizados como blocos discretos entre os vários motivos de sequências repetitivas. Os elementos dispersos são localizados ao longo do genoma, podendo formar agrupamentos em algumas regiões. Por outro lado, as regiões centroméricas/pericentroméricas, intercalares, subterminais e teloméricas dos cromossomos tendem a apresentar um massivo acúmulo de sequências repetidas em *tandem*, do tipo DNA satélite, sendo muitos desses sítios correspondentes a blocos de heterocromatina constitutiva visualizados pelo bandejamento C (Schmidt e Heslop-Harrison, 1998).

2.2.1. Centrômero e sequências repetitivas centroméricas

O centrômero é uma região especializada do cromossomo responsável pela manutenção da coesão entre as cromátides-irmãs até o início da anáfase, além de ser também o local de montagem do cinetócoro e da associação dos microtúbulos. Em organismos mais simples, como a levedura *Saccharomyces cerevisiae*, essa região é

composta por apenas três elementos de DNA centromérico (CDEI, CDEII e CDEIII), sem a presença de sequências repetitivas. *Schizosaccharomyces pombe*, por sua vez, mostra uma expansão da região centromérica em alguns kilobases, com *repeats* organizados de forma invertida e que parecem ser necessários para o total funcionamento dos centrômeros (Sumner, 2002). Em outros eucariotos, entretanto, o centrômero (pericentrômero) apresenta uma estrutura mais complexa, sendo formado por uma grande quantidade de sequências repetidas, como DNAs satélites, retrotransposons, SINEs e LINEs (Smith, 2002; Panchenko e Black, 2009).

Diferentes famílias do DNA satélite alfa (*α -satellite*) compõem a região centromérica dos cromossomos de primatas e, em humanos, essas sequências parecem participar da função centromérica (revisado por Alexandrov *et al.*, 2001). Da mesma forma, em plantas, DNAs satélites, juntamente com retroelementos, têm sido identificados como os principais componentes da região centromérica e pericentromérica, indicando que as mesmas sejam tipicamente formadas por sequências repetitivas (Jiang *et al.*, 2003; Neumann *et al.*, 2011). No arroz (*Oryza sativa*), por exemplo, Dong *et al.* (1998) observaram que a região centromérica consiste de um complexo de elementos repetitivos formado por pelo menos sete distintas famílias, sendo uma delas um DNA satélite (família RCS2) com unidade de repetição de 168 pb e presente em mais de 6.000 cópias no genoma. Utilizando um clone de DNA repetitivo, pSau3A9, isolado do sorgo (*Shorgum bicolor*), Jiang *et al.* (1996) realizaram um estudo citogenético comparativo em diferentes cereais (arroz, aveia, centeio, cevada, milho e trigo) e observaram a conservação desse elemento repetitivo como componente centromérico em todos os cromossomos.

Nas solanáceas, diferentes *repeats* satélites têm sido isolados e caracterizados no genoma da batata, *Solanum tuberosum*, e espécies relacionadas, como *Solanum bulbocastanum*. Embora mapeados predominantemente na região pericentromérica, eles mostram bastante variação quanto à sua presença dentro do gênero e em relação ao número de *repeats*. Alguns têm sido identificados em um grande número de espécies, como o elemento 2D8 (Stupar *et al.*, 2002). Outros, entretanto, são extremamente específicos. O *repeat* Sobo, amplificado recentemente dentro do gênero, está presente apenas na subespécie *bulbocastanum* de *S. bulbocastanum* (Tek *et al.*, 2005). Em alguns casos, como em *Arabidopsis thaliana*, a identificação de uma grande quantidade dessas sequências na região pericentromérica foi utilizada como indício de que as mesmas sejam elementos funcionais no complexo centrômero/cinetócoro (Nagaki *et al.*, 2003).

2.2.2. Telômero, telomerase e sequência repetitiva telomérica

Os telômeros são estruturas nucleoprotéicas presentes nos terminais cromossômicos eucarióticos que protegem os mesmos contra fusões e degradação, sendo assim essenciais para a manutenção da integridade do genoma. Além disso, tem sido relatada sua associação com a organização em *bouquet* na meiose e na organização do núcleo interfásico mitótico (configuração de Rabl) (McKnight *et al.*, 1997; McEachern *et al.*, 2000).

A estrutura telomérica da maioria das plantas é bastante similar à observada nos vertebrados e, ao nível de sequência, os telômeros de grande parte dos eucariotas são formados por arranjos em *tandem* de uma sequência de DNA curta, o DNA telomérico. Essa sequência é em geral replicada pela telomerase e está tipicamente conservada em uma vasta quantidade de organismos (Sumner, 2002). As telomerases são nucleoproteínas com atividade de transcriptase reversa e que apresentam um RNA molde em seu interior. Esse RNA serve como iniciador para a síntese de novas moléculas de DNA telomérico em terminais cromossômicos quebrados, bem como nos telômeros existentes, possibilitando dessa forma a manutenção dos terminais cromossômicos (McKnight *et al.*, 1997; McEachern *et al.*, 2000).

O hexanucleotídeo TTAGGG está presente como sequência telomérica em diferentes quantidades nos telômeros da maioria dos vertebrados e alguns protozoários, embora outras sequências repetidas também tenham sido encontradas como componentes teloméricos em outras espécies (ver Sumner, 2002). No musaranho ibérico, *Sorex granarius*, por exemplo, *repeats* de DNAr 45S ocorrem de forma inter espaçada com os *repeats* teloméricos nos cromossomos acrocêntricos (Zhdanova *et al.*, 2007). Nas drosófilas, o DNA satélite típico foi substituído por retroelementos tipo LINEs (Mason e Biessmann, 1995) ou, no caso de *Drosophila virilis* e outras espécies do grupo *virilis*, por um outro DNA satélite com unidade de repetição de 370 pb (Biessmann *et al.*, 2000).

O DNA telomérico vegetal, o heptanucleotídeo TTTAGGG, foi originalmente caracterizado em *Arabidopsis thaliana* e é considerado o arquétipo da sequência telomérica em plantas, estando extensamente distribuído desde algas unicelulares a angiospermas (Zellinger e Riha, 2006). Entretanto, algumas notáveis exceções são observadas no grupo. A ordem Asparagales (monocotiledônea) compreende três grupos

filogeneticamente distintos quanto à composição dos telômeros (Sykorová *et al.*, 2006). Diversos gêneros como *Aloe*, *Hyacinthella*, *Othocallis* e *Allium*, demonstram a perda do DNA telomérico do tipo de *Arabidopsis* (Adams *et al.*, 2001) e em pelo menos três deles o DNA telomérico de vertebrados tem sido identificado como componente dos terminais cromossômicos (Weiss e Scherthan, 2002; Puizina *et al.*, 2003; Weiss-Schneeweiss *et al.*, 2004). Nas eudicotiledôneas, a perda do heptanucleotídeo TTTAGGG tem sido relatada apenas para *Cestrum* e outros dois gêneros próximos da família Solanaceae, embora a sequência que passe a exercer essa função não tenha sido determinada (Sykorová *et al.*, 2003a; Sykorová *et al.*, 2003b).

Em alguns casos, esse DNA telomérico pode ocorrer fora dos telômeros formando os sítios ectópicos ou sítios intersticiais de DNA telomérico (ITS - *Interstitial Telomere Sites*), os quais são encontrados principalmente nas regiões centromérica e pericentromérica dos cromossomos. Esses sítios têm sido observados em diferentes grupos de animais e plantas e podem ser originados por inversões paracêntricas com ponto de quebra envolvendo os *repeats* teloméricos, fusões cromossômicas em *tandem* ou por meio de translocações robertsonianas, levando nesses dois últimos casos a alterações no número cromossômico (ver, por exemplo, Fagundes *et al.*, 1997; Lee *et al.*, 1993). Em *Pinus densiflora*, por exemplo, diversos sítios intersticiais formados por variantes do *repeat* telomérico de *Arabidopsis* foram identificados (Shibata *et al.*, 2005).

2.2.3. Subtelômero e DNA repetitivo subtelomérico

Todos os organismos possuem imediatamente abaixo dos *repeats* teloméricos, numa posição mais proximal, uma região denominada subtelômero ou região subtelomérica. Essa região tem se mostrado estruturalmente similar quanto à composição de sequências em diferentes grupos, embora a sua extensão e a organização dos elementos repetitivos geralmente presentes na mesma variem bastante. Essas regiões funcionam como uma zona de transição entre sequências cromossomo-específicas e os arranjos de *repeats* teloméricos e nelas têm sido encontrados desde sequências teloméricas normais ou degeneradas, retrotransposons e genes, até arranjos de DNAs satélites bastante variáveis quanto à sua complexidade (McEachern *et al.*, 2000).

Nos vegetais, as regiões subteloméricas têm sido caracterizadas pela presença de uma organização nucleossomal diferente da apresentada pelos telômeros, sendo em alguns casos, como em *Nicotiana tabacum*, considerados domínios estruturalmente distintos (Sykorová *et al.*, 2001). Essas regiões são constituídas por diferentes famílias de elementos repetitivos, como DNAs satélites, DNAr 45S ou retroelementos, que podem apresentar-se altamente conservados dentro de um grupo, serem genoma ou mesmo espécie-específicos, além de mostrarem uma ampla variação quantitativa e qualitativa ao nível cromossômico (revisado por Sharma e Raina, 2005). Kazama *et al.* (2006) caracterizaram citomolecularmente os terminais cromossômicos de *Silene latifolia*, identificando que os mesmos são compostos por quatro diferentes subfamílias de DNA satélite diferencialmente distribuídas nessa região e indicando que cada uma delas tem evoluído independentemente no genoma da espécie, fazendo com que cada subtelômero apresente uma composição de DNA satélite característica. Em espécies do gênero *Beta*, a região subtelomérica é formada por uma família de DNA satélite com unidade de repetição entre 344-362 pb, contendo dois *subrepeats* internos, e que está presente como um marcador cromossomo-específico em *B. nana* e na maioria dos terminais de *B. vulgaris* e *B. corolliflora* (Dechyeva e Schmidt, 2006).

2.3. Legumes

Leguminosae ou Fabaceae é a terceira maior família de plantas com flores e a segunda mais importante economicamente, atrás apenas das gramíneas (Poaceae). Tradicionalmente, a família é subdividida em três subfamílias (Caesalpinioideae, Mimosoideae e Papilionoideae), as quais são formadas por cerca de 18.000 espécies distribuídas em mais de 600 gêneros, incluindo plantas utilizadas para fins ornamentais, medicinais e agrofloretais, além de forrageiras e outras de grande importância alimentícia. Diferentes espécies de feijão (*Cajanus*, *Phaseolus*, *Vicia* e *Vigna*) representam valiosas fontes protéicas na dieta humana e animal, enquanto outras espécies, como o amendoim (*Arachis hypogea*) e a soja (*Glycine max*), são as principais matérias-primas na produção de óleos vegetais (Graham e Vance, 2003; Singh *et al.*, 2007; Judd *et al.*, 2009). Outra marcante característica das leguminosas é seu papel central na ciclagem do nitrogênio. A habilidade de fixação de N₂ por meio de associação simbiótica com bactérias do gênero *Rhizobium* propicia que, por ano, cerca

de 40 a 60 milhões de toneladas desse importante macronutriente esteja disponível no solo para as atividades agrícolas (Graham e Vance, 2003).

O grande valor econômico atrelado a muitas espécies de leguminosas fez com que várias delas se consolidassem como organismos modelos nos campos da citologia, genética e biologia molecular (Sing *et al.*, 2007; Rispaud *et al.*, 2010). Diversos estudos baseados em interações planta-patógeno, padrões metabólicos, fisiológicos e de desenvolvimento, bem como análises de transcriptoma e proteoma, vêm sendo desenvolvidos em diferentes legumes, como *Glycine. max* (a soja) e *Pisum sativum* (a ervilha), além das espécies modelo *Lotus japonicus* e *Medicago truncatula*, fornecendo maiores detalhes sobre a base molecular por trás de diversos processos biológicos (Cook, 1999; Gepts *et al.*, 2005; Rispaud *et al.*, 2010). De forma semelhante, a comparação de regiões sintênicas entre legumes, e também entre estes e espécies não relacionadas, tem se mostrado uma excelente abordagem para identificação dos padrões evolutivos de conservação e divergência genômicas e na caracterização gênica funcional (ver por exemplo Li *et al.*, 2011).

2.4. O Gênero *Phaseolus*

2.4.1. Posicionamento taxonômico

O gênero *Phaseolus* L. pertence à família Fabaceae, subfamília *Papilionoideae*, tribo *Phaseoleae*. Essa tribo é uma das mais relevantes economicamente por englobar também os gêneros *Glycine* (soja) e *Vigna* (feijões asiáticos e africanos) (Lackey, 1985). *Phaseolus* é formado por cerca de 70 espécies e constitui um grupo monofilético sustentado por diferentes tipos de marcadores (ver, por exemplo, Delgado-Salinas *et al.*, 1999). Na filogenia mais recente para o gênero (Delgado-Salinas *et al.*, 2006), as espécies são subdivididas em oito grupos (*Tuerckheimii*, *Pauciflorus*, *Pedicellatus*, *Filiformis*, *Vulgaris*, *Leptostachyus*, *Lunatus* e *Polystachius*) e estes formam dois grandes clados, denominados A e B. No clado A estão a maioria das espécies consideradas selvagens, entre elas *P. microcarpus* Mart., e no clado B todas as cinco que vêm sendo domesticadas: *P. vulgaris* L., *P. lunatus* L., *P. coccineus* L., *P. polyanthus* Greenman e *P. acutifolius* L. Gray, entre outras (Broughton *et al.*, 2003).

2.4.2. Importância sócio-econômica

Os feijões, de um modo geral, são as leguminosas mais importantes para consumo humano direto, sendo uma das principais fontes de proteínas para as populações da África e América Latina. A *faseolina*, principal proteína de reserva da semente, pode representar até 50% das proteínas do feijão. Além disso, o alto conteúdo de minerais, especialmente ferro (Fe) e zinco (Zn) é uma característica que torna o seu consumo extremamente importante, sobretudo em regiões onde há uma alta prevalência de deficientes em micronutrientes (Pereira e Souza, 1992; Broughton *et al.*, 2003).

Consumido de diferentes maneiras, como grãos secos ou verdes ou na forma de vagens (legumes), o feijão apresenta produção total anual de mais de 23 milhões de toneladas, sendo sete milhões produzidas apenas na América Latina e África. Analisado *per capita*, o consumo varia bastante entre diferentes regiões, com destaque para o feijão comum (*P. vulgaris*) e a fava (*P. lunatus*). Na Bolívia o consumo por pessoa é de 24 kg ano⁻¹; em regiões rurais do Quênia esse valor chega a 60 kg ano⁻¹. Em outros países da África, como Ruanda e Burundi, a média é de cerca de 40 kg ano⁻¹. No Brasil, as duas regiões com maior consumo são o Nordeste e o Sudeste, com 20,8 e 18,2 kg ano⁻¹, respectivamente (Broughton *et al.*, 2003).

2.4.3. O feijão comum (*P. vulgaris*)

O feijão comum é considerado uma das mais importantes espécies de *Phaseolus* pelo grande valor agrônomo e sócio-econômico agregado. A espécie é originária das Américas, em uma região entre o Equador e o norte do Peru (Kami *et al.*, 1995), de onde se dispersou formando dois principais conjuntos gênicos, um Andino e um Mesoamericano, o primeiro compreendendo a região central e o sul da América do Sul e o segundo se estendendo do México ao norte da América do Sul.

A domesticação tem ocorrido de forma independente nessas duas regiões, onde a espécie vem sendo cultivada a mais de 4.000 anos. Os conjuntos gênicos têm sido organizados em raças de origens bem controversas, sendo quatro Mesoamericanas (Durango, Guatemala, Jalisco e Mesoamerica) e três Andinas (Chile, Nueva Granada e Peru). Análises baseadas em caracteres morfo-anatômicos, adaptações ecológicas e

diferentes marcadores moleculares dão suporte tanto para um processo de domesticação único, como para a ocorrência de múltiplos eventos de domesticação em cada conjunto gênico (revisado por Chácon *et al.*, 2005).

2.4.4. A fava (*P. lunatus*)

A fava (*P. lunatus*), também conhecida como feijão-lima ou feijão-fava, é uma das cinco espécies do gênero *Phaseolus* explorada comercialmente, ficando atrás apenas do feijão comum em importância. A espécie, que está posicionada na filogenia do gênero (Delgado-Salinas *et al.*, 2006) no grupo *Lunatus*, apresenta-se geneticamente estruturada em dois conjuntos tradicionalmente reconhecidos, Andino e Mesoamericano (Salgado *et al.*, 1995), assim como *P. vulgaris*. O Andino está restrito ao Equador e norte do Peru, enquanto que o Mesoamericano compreende acessos extensamente distribuídos ao longo do continente americano, englobando áreas na América Central, Cuba, Colômbia, sul do Peru e na Argentina. Recentemente, Serrano-Serrano *et al.* (2010) propuseram a existência de um conjunto gênico adicional para a espécie. Por meio da análise de dados do espaçador interno transcrito (ITS) do DNA ribossomal e do DNA plastidial em uma extensa quantidade de acessos de *P. lunatus* e espécies relacionadas, os autores identificaram a presença de um novo conjunto Mesoamericano, ocorrente na região do Istmo de Tehuantepec, sudeste do México. Além disso, devido ao estreito relacionamento encontrado com outras espécies andinas do gênero, foi possível estabelecer essa região como centro de origem para *P. lunatus* (Serrano-Serrano *et al.*, 2010).

2.4.5. *Phaseolus microcarpus* e *P. leptostachyus* Benth

Embora não sejam espécies com potencial econômico associado, *P. microcarpus* e *P. leptostachyus* têm sido utilizadas em diversos estudos taxonômicos, filogenéticos e citogenéticos em *Phaseolus*. A primeira espécie, originária das montanhas do México (Silva e Costa, 2003), está posicionada filogeneticamente num clado distinto (clado A) daquelas que vêm sendo domesticadas, sendo considerada uma das espécies mais basais e, portanto, de divergência precoce, na evolução do gênero. *Phaseolus leptostachyus*, espécie do clado B, é considerada grupo-irmão de *P. lunatus* e, juntamente com os

outros componentes do grupo *Leptostachyus*, apresentam uma divergência cariotípica mais acentuada em relação ao resto do gênero (Delgado-Salinas *et al.*, 1998; Delgado-Salinas *et al.*, 2006).

2.4.6. Citogenética em *Phaseolus*

Os primeiros registros de análises citogenéticas para o gênero *Phaseolus* datam da década de 1920, quando Karpetschenko descreveu os números cromossômicos de *P. acutifolius*, *P. coccineus*, *P. lunatus* e *P. vulgaris* como $2n = 22$ (revisado por Mercado-Ruaro e Delgado-Salinas, 2000). Desde então, uma grande quantidade de estudos citogenéticos têm sido realizados nas espécies do grupo, variando desde contagens cromossômicas e descrições cariotípicas simples (ver, por exemplo, Mercado-Ruaro e Delgado-Salinas, 1996) à construção de mapas cromossômicos detalhados (ver, por exemplo, Pedrosa-Harand *et al.*, 2009).

O gênero apresenta uma grande estabilidade cariotípica, com a grande maioria das espécies descritas até o momento apresentando cromossomos com morfologia variando de metacêntrica à submetacêntrica (Mercado-Ruaro e Delgado-Salinas, 1996; 1998). Após uma extensiva análise citogenética numérica no gênero, esses mesmos autores propuseram $x = 11$ como número básico para *Phaseolus*. A exceção à regra está no grupo composto por *P. leptostachyus*, *P. micranthus* e *P. macvaughii* (grupo *Leptostachyus* na filogenia do gênero; Delgado-Salinas *et al.*, 2006), no qual o número cromossômico descrito foi $2n = 20$ (Mercado-Ruaro e Delgado-Salinas *et al.*, 1998).

Devido ao reduzido tamanho dos cromossomos nas espécies do gênero, o que dificulta bastante a distinção entre pares cromossômicos (Mercado-Ruaro e Delgado-Salinas, 2000), os estudos iniciais ficaram restritos a descrições numéricas, morfológicas e de padrões de bandas C como as realizadas em *P. vulgaris* e *P. coccineus* (Mok e Mok, 1976; Zheng *et al.*, 1991). Com o avanço das técnicas em citogenética molecular e o advento da hibridização *in situ* fluorescente (FISH), as análises em *Phaseolus* se ampliaram. Zheng *et al.* (1994) descreveram pela primeira vez, por meio da hibridização *in situ*, o posicionamento e quantidade de sítios de DNA ribossomal 45S (DNAr 45S) em diferentes espécies de *Phaseolus* (*P. angularis*, *P. calcaratus*, *P. coccineus* e *P. vulgaris*), algumas hoje consideradas *Vigna*. Posteriormente, Moscone *et al.* (1999), combinando a localização dos sítios de DNA

ribossomal (5S e 18-25S), o padrão de distribuição da heterocromatina e a morfologia cromossômica, identificaram a maioria dos cromossomos de *P. acutifolius* var. *latifolius*, *P. coccineus*, *P. lunatus* e *P. vulgaris* ‘Saxa’ e ‘Wax’, propondo uma nomenclatura cromossômica numérica para as espécies do gênero. Mais recentemente, como um dos esforços para o melhoramento genético da espécie, uma biblioteca de BACs (*Bacterial Artificial Chromosome*) foi desenvolvida para o feijão comum (Kami *et al.*, 2006). Selecionados a partir de marcadores moleculares (Bngs e Legs) mapeados geneticamente (Vallejos *et al.*, 1992; Hougaard *et al.*, 2008), alguns desses clones BACs foram utilizados no estabelecimento de mapas cromossômicos para duas variedades do feijão comum, BAT93 (Mesoamericana, Pedrosa-Harand *et al.*, 2009; Fonsêca *et al.*, 2010) e G19833 (Andina, Altrock *et al.*, 2011). Esses clones passaram a constituir marcadores cromossômicos e vêm sendo utilizados nos estudos de evolução cromossômica em outras espécies do gênero, como *P. lunatus* (Almeida, 2006; Bonifácio, 2009) e *P. microcarpus* (Fonsêca, 2010). De forma geral, as análises têm revelado uma grande conservação de sintenia entre os cromossomos homeólogos. A presença de apenas pequenos rearranjos (inversões paracêntricas e pericêntricas), identificados pela mudança posicional do centrômero ou de BACs localizados citologicamente, corroboram a já anteriormente sugerida estabilidade cariotípica do grupo.

2.4.7. Sequências repetitivas em *Phaseolus*

De forma geral, os dados referentes à fração repetitiva do genoma de *Phaseolus* não têm sido explorados de forma intensa. Até o momento, nenhuma espécie do grupo apresenta seu genoma totalmente sequenciado, embora esse processo já tenha sido iniciado em dois acessos do feijão comum (G19833 e BAT93, Scott Jackson, comunicação pessoal). Utilizando-se da análise de sequências de terminais de BACs (*BAC-end sequence*) oriundos de uma biblioteca construída para *P. vulgaris*, Schlueter *et al.* (2009) estimaram a composição de sequências do genoma da espécie. Essa abordagem permitiu identificar diferentes classes de sequências repetitivas nos terminais de BACs analisados, como elementos transponíveis (transposons e retrotransposon), além de sequências similares a DNA repetitivo centromérico, DNA telomérico, entre outros, as quais correspondem a cerca de 50 % do genoma.

Além disso, alguns trabalhos de caracterização citomolecular de famílias de DNA repetitivo, já foram realizados no grupo. Um DNA minissatélite (OPG9; Métais *et al.*, 1998), duas famílias de DNA satélite, CC4 (Falquet *et al.*, 1997) e *khipu* (David *et al.*, 2009), um retrotransposon, Tpv2 (Garber *et al.*, 1999) e PvMeso (Ribeiro *et al.*, 2011) tiveram sua organização e distribuição cromossômica descritas. O CC4 foi caracterizado como um DNA satélite relacionado ao espaçador intergênico do DNA ribossomal (IGS-like) do feijão comum e está presente na região proximal de dois ou três pares de cromossômicos de diferentes acessos de *P. vulgaris* (Falquet *et al.*, 1998; Almeida CCS; dados não publicados). O minissatélite OPG9 apresenta unidade de repetição de 15 pb (Métais *et al.*, 1998), enquanto que o retrotransposon Tpv2, a primeira família de retroelementos isolada do feijão comum, apresenta tamanho em torno de 4 kb e está distribuída dispersamente ao longo do genoma, mas também como blocos em algumas regiões cromossômicas (Garber *et al.*, 1999; Pedrosa-Harand, A dados não publicados). PvMeso é uma sequência repetitiva subtelomérica altamente amplificada no braço longo do cromossomo 7 em acessos mesoamericanos de *P. vulgaris* e pode ser útil na genotipagem desses acessos (Ribeiro *et al.*, 2011).

O feijão comum apresenta uma série de *clusters* gênicos de resistência contra patógenos em seu genoma, e foi durante a caracterização molecular de um deles, o cluster B4 de resistência contra o fungo patogênico *Colletotrichum lindemuthianum*, que uma nova sequência repetitiva foi inicialmente identificada (Geffroy *et al.*, 2009). O B4, um dos maiores *clusters* da espécie, está localizado na região terminal do braço curto do cromossomo 4 e com a análise mais detalhada do mesmo o DNA satélite *khipu* foi descrito (David *et al.*, 2009).

Khipu apresenta unidade consenso de 528 pb e 38,2 % de conteúdo de GC, embora tenham sido descritas unidades (92 no total) com tamanho entre 311 e 564 pb e identidade variando entre 40,2 e 99,8 %. As análises moleculares, por meio de hibridização em membrana, revelaram a especificidade da sequência a *Phaseolus* e a presença de grande variabilidade na intensidade dos sinais entre as espécies analisadas. Alguns delas, como *P. lunatus* e *P. leptostachyus*, pertencentes ao clado B e, portanto, mais relacionadas à *P. vulgaris*, mostraram intensidade de hibridização similar à *P. microcarpus*, espécie selvagem e do clado A na filogenia do grupo (Delgado-Salinas *et al.*, 2009). Esses dados podem refletir apenas diferenças quantitativas, mas também podem representar divergência de sequência entre as espécies. Ao nível cromossômico, *khipu* foi caracterizado apenas em um acesso do feijão comum (BAT93), sendo

localizado na região subterminal da maioria dos cromossomos da espécie (David *et al.*, 2009).

3. Referências bibliográficas

- Acosta MJ, Marchal JA, Fernández-Espartero C, Romero-Fernández I, Rovatsos MT, Giagia-Athanasopoulou G, Gornung E, Castiglia R, Sánchez A (2010). Characterization of the satellite DNA Msat-160 from species of *Terricola* (*Microtus*) and *Arvicola* (Rodentia, Arvicolinae). *Genetica* 138: 1085-1098.
- Adams SP, Hartman TPV, Lim KY, Chase MW, Bennett MD, Leitch IJ, Leitch AR (2001). Loss and recovery of Arabdopsis-type telomere repeat sequence 5'-(TTTAGGG)_n-3' in the evolution of a major radiation of flowering plants. *Plant Mol Biol* 60: 633-646.
- Alexandrov I, Kazakov A, Tumeneva I, Shepelev V, Yurov V (2001). Alpha-satellite DNA of primates: old and new families. *Chromosoma* 110: 253-266.
- Almeida CCS (2006). Mapeamento físico e análise evolutiva em *Phaseolus vulgaris* L. e *Phaseolus lunatus* L., utilizando hibridização fluorescente *in situ* (FISH). Tese de Doutorado. Recife, 104p.
- Altrock S, Fonsêca A, Pedrosa-Harand A (2011). Chromosome identification in the Andean common bean accession G19833 (*Phaseolus vulgaris* L., Fabaceae). *Genet Mol Biol* 34: 459-463.
- Barros e Silva AM, Marques A, Santos KGB, Guerra M (2010). The evolution of CMA bands in *Citrus* and related genera. *Chromosome Res* 18: 503-514.
- Bernard P, Maure JF, Partridge JF, Genier S, Javerzat JP, Allshire RC (2001). Requirement of heterochromatin for cohesion at centromeres. *Science* 294: 2539-2542.
- Bernard P, Allshire RC (2002). Centromeres become unstuck without heterochromatin. *Trends Cell Biol* 12: 419-424.
- Biessmann H, Zurovcova M, Yao JG, Lozovskaya E, Walter MF (2000). A telomeric satellite in *Drosophila virilis* and its sibling species. *Chromosoma* 109: 372-380.
- Bonifácio EM (2009). Mapeamento comparativo entre o feijão comum e a fava. Trabalho de conclusão de Curso. Recife, 2009. 64p.

- Broughton WJ, Hernández G, Blair M, Beebe S, Gepts P, Vanderleyden J (2003). Beans (*Phaseolus* spp.) - model food legumes. *Plant Soil* 252: 55-128.
- Brutlag DL (1980) Molecular arrangement and evolution of heterochromatin DNA. *Ann Rev Genet* 14: 121-144.
- Chacón MI, Pickersgill SB, Debouck DG (2005). Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races. *Theor Appl Genet* 110:432-444.
- Charlesworth B, Sniegowski P, Stephan W (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371: 215-220.
- Chow JC, Brown CJ (2003). Forming facultative heterochromatin: silencing of an X chromosome in mammalian females. *Cell Mol Life Sci* 60 (2003): 2586–2603
- Cohen S, Segal D (2009). Extrachromosomal circular DNA in eukaryotes: possible involvement in the plasticity of tandem repeats. *Cytogenet Genome Res* 124: 327-338.
- Cohen S, Agmon N, Sobol O, Segal D (2010). Extrachromosomal circles of satellite repeats and 5S ribosomal DNA in human cells. *Mobile DNA* 1:11.
- Cook DR (1999). *Medicago truncatula* — a model in the making!. *Curr Opin Plant Biol* 2: 301–304.
- Craig JM (2004). Heterochromatin—many flavours, common themes. *Bioessays* 27: 17-28.
- David P, Chen NWG, Pedrosa-Harand A, Thareau V, Sévignav 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.
- Dechyeva D, Schmidt T (2006). Molecular organization of terminal repetitive DNA in *Beta* species. *Chromosome Res* 14: 881-897.
- Delgado-Salinas A, Turley T, Richman A, Lavin M (1999). Phylogenetic analysis of the cultivated and wild species of *Phaseolus* (Fabaceae). *Syst Bot* 24: 438-460.
- Delgado-Salinas A, Bibler R, Lavin M (2006). Phylogeny of the genus *Phaseolus* (Leguminosae): a recent diversification in an ancient landscape. *Syst Bot* 31: 779-791.
- Dong F, Miller JT, Jackson SA, Wang GL, Ronald PC, Jiang J (1998). Rice (*Oryza sativa*) centromeric regions consist of complex DNA. *Proc Natl Acad Sci USA* 95: 8135-8140.

- Falquet J, Creusot F, Dron M (1997). Molecular analysis of *Phaseolus vulgaris* rDNA unit and characterization of a satellite DNA homologous to IGS subrepeats. *Plant Physiol Biochem* 35: 611-622.
- Fonsêca A (2010). Evolução cariotípica no gênero *Phaseolus* L.: mapeamento comparativo entre *P. microcarpus* Mart. e o feijão comum (*P. vulgaris* L.). Dissertação de Mestrado. Recife, 85p.
- Fonsêca A, Ferreira J, Santos TRB, Mosiolek M, Bellucci 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.
- Fry K, Salser W (1977). Nucleotide sequences of HS- α satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. *Cell* 12: 1069-1084.
- Garber K, Bilic I, Pusch O, Tohme J, Bachmair A, Schweizer D, Jantsch V (1999). The Tpv2 family of retrotransposons of *Phaseolus vulgaris*: structure, integration characteristics, and use for genotype classification. *Plant Mol Bio* 39: 797-807.
- Geffroy V, Macadré C, David P, Pedrosa-Harand A, Sévignac M, Dauga C, Langin T (2009). 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.
- Gepts P, Beavis DB, Brummer EC, Shoemaker RC, Stalker HT, Weeden NF, Young ND (2005). Legumes as a model plant family. Genomics for food and feed report of the cross-legume advances through genomics conference. *Plant Physiol* 137: 1228-1235.
- Graham PH, Vance CP (2003). Legumes: Importance and Constraints to Greater Use. *Plant Physiol* 131: 872-877.
- Grewal SIS, Elgin SCR (2002). Heterochromatin: new possibilities for the inheritance of structure. *Curr Opin Gen Dev*, 12:178–187
- Grewal SIS, Jia S (2007). Heterochromatin revisited. *Nat Rev* 8: 35-46.
- Guerra M (2000). Patterns of heterochromatin distribution in plant chromosomes. *Genet Mol Biol* 23: 1029-1041.
- Hansen KM (1973). Heterochromatin (C bands) in bovine chromosomes. *Hereditas* 73: 65-70.
- Heslop-Harrison JS, Schwarzacher T (2011). Organisation of the plant genome in chromosomes. *Plant J* 66: 18-33.

- Hougaard BK, Madsen LH, Sandal N, Moretzsohn MC, Fredslund J, Schauser L, Nielsen AM, Rohde T, Sato S, Tabata S, Bertoli DJ, Stougaard J (2008). Legume anchor markers link syntenic regions between *Phaseolus vulgaris*, *Lotus japonicus*, *Medicago truncatula* and *Arachis*. *Genetics* 179: 2299-2312.
- Jiang J, Nasuda S, Dong F, Scherrer CW, Woo SS, Wing RA, Gill BS, Ward DC (1996). A conserved repetitive DNA element located in the centromeres of cereal chromosomes. *Proc Natl Acad Sci USA* 93: 14210-14213.
- Jiang J, Birchler JA, Parrott WA, Dawe RK (2003). A molecular view of plant centromeres. *Trends Plant Sci* 8: 570-575.
- Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ (2009). *Sistemática Vegetal: Um Enfoque Filogenético* 3a ed. Porto Alegre: Artmed, pp 612.
- Kami J, Velásquez VB, Debouck DG, Gepts P (1995). Identification of presumed ancestral DNA sequences of phaseolin in *Phaseolus vulgaris*. *Proc Natl Acad Sci USA* 92: 1101-1104.
- Kami J, Poncet V, Geffroy V, Gepts P (2006). Development of four phylogenetically-arrayed BAC libraries and sequence of the APA locus in *Phaseolus vulgaris*. *Theor Appl Genet* 112: 987-998.
- Kazama Y, Sugiyama R, Suto Y, Uchida W, Kawano S (2006). The clustering of four subfamilies of satellite DNA at individual chromosome ends in *Silene latifolia*. *Genome* 49: 520-530.
- King M, Honeycutt R, Contreras N (1986). Chromosomal repatterning in crocodiles: C, G and N-banding and the *in situ* hybridization of 18S and 26S rRNA cistrons. *Genetica* 70: 191-201.
- King K, Jobst J, Hemleben V (1995). Differential homogenization and amplification of two satellite DNAs in the genus *Cucurbita* (Cucurbitaceae). *J Mol Evol* 41: 996-1005.
- Lackey JA (1985). Tribe 10. Phaseoleae DC. In: Rolhill, R.M.; Raven, P.H. (Eds.) *Advances in legume systematic. Part 1*. Royal Botanic Gardens, London, p.301-327.
- Lee C, Sasi R, Lin CC (1993). Interstitial localization of telomeric DNA sequences in the Indian muntjac chromosomes: further evidence for tandem chromosome fusions in the karyotypic evolution of the Asian muntjacs. *Cytogenet Cell Genet* 63: 156-159.
- Li J, Dai X, Liu T, Zhao PX (2011) LegumeIP: an integrative database for comparative genomics and transcriptomics of model legumes. *Nucleic Acids Res* 1-9.

- Mason JM, Biessmann H (1995). The unusual telomeres of *Drosophila*. TIG 11: 58-62.
- McEachern MJ, Krauskopf A, Blackburn EH (2000). Telomerase and their control. Annu Rev Genet 34: 331-58.
- McKnight TD, Fitzgerald MS, Shippen DE (1997). Plant Telomeres and Telomerases. A Review. Biochemistry 62: 1224-1231.
- Mercado-Ruaro P, Delgado-Salinas A (1996). Karyological studies in several Mexican species of *Phaseolus* L. and *Vigna* Savi (Phaseolinae, Fabaceae). In: B Pickersgill and JM Lock (eds.) Advance in Legume Systematics 8: Legumes of Economic Importance, pp. 83-87.
- Mercado-Ruaro P e Delgado-Salinas A (1998). Karyotypic studies on species of *Phaseolus* (Fabaceae: Phaseolinae). Am J Bot 85: 1-9.
- Mercado-Ruaro P, Delgado-Salinas A (2000). Cytogenetic studies in *Phaseolus* L. (Fabaceae). Genet Mol Biol 23: 985-987.
- Meštrović N, Plohl M, Mravinac B, Ugarković D (1998). Evolution of satellite DNAs from the genus *Palorus* – experimental evidence for the “Library” Hypothesis”. Mol Biol Evol 15: 1062-1068.
- Métais I, Aubry C, Hamon B, Peltier D, Jalouzot R (1998). Cloning, quantification and characterization of a minisatellite DNA sequence from common bean *Phaseolus vulgaris* L.. Theor Appl Genet 97: 232-237.
- Mok DWS e Mok MC (1976). A modified Giemsa technique for identifying bean chromosomes. J Hered 67: 187-188.
- Moscone EA, Klein F, Lambrou M, Fuchs J, Schweizer D (1999). Quantitative karyotyping and dual-color FISH mapping of 5S and 18-25S rDNA probes in the cultivated *Phaseolus* species (Leguminosae). Genome 42: 1224-1233.
- Nagaki K, Talbert PB, Zhong CX, Dawe RK, Henikoff S, Jiang J (2003). Chromatin immunoprecipitation reveals that the 180-bp satellite repeat is the key functional DNA element of *Arabidopsis thaliana* centromeres. Genetics 163: 1221-1225.
- Nei M, Rooney AP (2005). Concerted and Birth-and-Death evolution of multigene families. Annu Rev Genet 39: 121-152.
- Neumann P, Navrátilová A, Koblízková A, Kejnovský E, Hřibová E, Hobza, R, Widmer A, Doležel J, Macas J (2011). Plant centromeric retrotransposons: a structural and cytogenetic perspective. Mobile DNA 2: 4.

- Norris TB, Rickards GK, Daugherty CH (2004). Chromosomes of tuatara, *Sphenodon*, a chromosome heteromorphism and an archaic reptilian karyotype. *Cytogenet Genome Res* 105: 93-99.
- Panchenko T, Black BE (2009). The Epigenetic Basis for Centromere Identity in Centromere Structure and Evolution. In: Đurđica Ugarković (ed.). *Centromere structure and evolution*. Springer, Berlin.
- Pedrosa-Harand A, Kami J, Gepts P, Geffroy V, Schweizer D (2009). Cytogenetic mapping of common bean chromosomes reveals a less compartmentalized small-genome plant species. *Chromosome Res* 17: 405-417.
- Pereira PAA, De Souza CRB (1992). Tipos de Faseolina em raças ‘crioulas’ de feijão no Brasil. *Pesq Agropec Bras* 28: 1219-1221.
- Pons J, Juan C, Petitpierre E (2002). Higher-order organization and compartmentalization of satellite DNA PIM357 in species of the coleopteran genus *Pimelia*. *Chromosome Res* 10: 597-606.
- Puizina J, Weiss-Schneeweiss H, Pedrosa-Harand A, Kamenjarin J, Trinajstić I, Riha K, Schweizer D (2003). Karyotype analysis in *Hyacinthella dalmática* (Hyacinthaceae) reveals vertebrate-type telomere repeats at the chromosome ends. *Genome* 46: 1070-1076.
- Rennig W (1996). Heterochromatin. *Chromosoma* 108: 1-9.
- Ribeiro T, Santos KGB, Fonsêca A, Pedrosa-Harand A (2011). Isolation and characterization of a new repetitive DNA family recently amplified in the Mesoamerican gene pool of the common bean (*Phaseolus vulgaris* L., Fabaceae). *Genetica* 139: 1135-1142.
- Rispaud N, Kaló P, Kiss GB, Ellis THN, Gallardo K, Thompson RD, Prates E, Larrainzar E, Ladrera R, González EM, Arrese-Igor Cesar, Ferguson BJ, Gresshoff PM, Rubiales D (2010). Model legumes contribute to faba bean breeding. *Field Crops Res* 115: 253-269.
- Salgado AG, Gepts P, Debouck DG (1995). Evidence for two gene pools of the Lima bean, *Phaseolus lunatus* L., in the Americas. *Genetic ResourCrop Evol* 42: 15-28.
- Schlueter JA, Goicoechea JL, Collura K, Gill N, Lin J-Y, Yu Y, Kudrna D, Zuccolo A, Vallejos CE, Muñoz-Torres M, Blair MW, Tohme J, Tomkins J, McClean P, Wing RA, Jackson AS (2009). BAC-end sequence analysis and draft physical map of the common bean (*Phaseolus vulgaris* L.) genome. *Trop Plant Biol* 1: 40-48.

- Schmidt T, Heslop-Harrison JS (1998). Genomes, genes and junk: the large-scale organization of plant chromosomes. *Trends Plant Sci* 3: 195-199.
- Serrano-Serrano ML, Hernández-Torres J, Castillo-Villamizar G, Debouck DG, Sánchez MIC (2010). Gene pools in wild lima bean (*Phaseolus lunatus* L.) from the Americas: evidences for an Andean origin and past migrations. *Mol Phylogenet Evol* 54: 76–87.
- Sharma S, Raina SN (2005). Organization and evolution of highly repeated satellite DNA sequences in plant chromosomes. *Cytogenet Genome Res* 109: 15-26.
- Shibata F, Matsusaki Y, Hizume M (2005). AT-rich sequences containing *Arabidopsis*-type telomere sequence and their chromosomal distribution in *Pinus densiflora*. *Theor Appl Genet* 110: 1253-1258.
- Singh RJ, Chung GH, Nelson RL (2007). Landmark research in legumes. *Genome* 50: 525-537.
- Silva HT, Costa AO (2003). Caracterização botânica de espécies silvestres do gênero *Phaseolus* L. (Leguminosae). Documentos/Embrapa Arroz e Feijão, Santo Antônio de Goiás, 40 p.
- Smith MM (2002). Centromeres and variant histones: what, where, when and why?. *Curr Opin Cell Biol* 14: 279–285.
- Sykorová E, Fajkus J, Ito M, Fukui K (2001). Transition between two forms of heterochromatin at plant subtelomeres. *Chromosome Res* 9: 309-323.
- Sykorová E, Lim KY, Chase MW, Knapp S, Leitch IJ, Leitch AR, Fajkus J (2003). The absence of *Arabidopsis*-type telomeres in *Cestrum* and closely related genera *Vestia* and *Sessea* (Solanaceae): first evidence from eudicots. *Plant J* 34: 283-291.
- Sykorová E, Lim KY, Fajkus J, Leitch AR (2003b). The signature of the *Cestrum* genome suggests an evolutionary response to the loss of (TTTAGGG)_n telomeres. *Chromosoma* 112: 164-172.
- Sykorová E, Leitch AR, Fajkus J (2006). Asparagales telomerases which synthesize the human type of telomeres. *Plant Mol Biol* 60: 633–646.
- Stupar RM, Song J, Tek AL, Cheng Z, Dong F, Jiang J (2002). Highly condensed potato pericentromeric heterochromatin contains rDNA-related tandem repeats. *Genetics* 162: 1435-1444.
- Sumner AT (2002). Chromosomes organization and function 1a ed. Berlin: Blackwell Publishing, pp 294.

- Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ (2009). *Sistemática Vegetal: Um Enfoque Filogenético* 3a ed. Porto Alegre: Artmed, pp 612.
- Tek AL, Song J, Macas J, Jiang J (2005). Sobo, a Recently Amplified Satellite Repeat of Potato, and Its Implications for the Origin of Tandemly Repeated Sequences. *Genetics* 170: 1231-1238.
- Ugarković D, Plohl M (2002). Variation in satellite DNA profiles - causes and effects. *The EMBO Journal* 21: 5955-5959.
- Viguera E, Canceill, D, Ehrlich SD (2001). Replication slippage involves DNA polymerase pausing and dissociation. *EMBO J* 20: 2587-2595.
- Weiss H, Scherthan H (2002) *Aloe* spp. - plants with vertebrate-like telomeric sequences. *Chromosome Res* 10: 155-164.
- Weiss-Schneeweiss H, Riha K, Jang CG, Puizina J, Scherthan H, Schweizer D (2004). Chromosome termini of the monocot plant *Othacallis siberica* are maintained by telomerase, which specifically synthesizes vertebrate-type telomere sequences. *Plant J* 37: 484-493.
- Zellinger B, Riha K (2006). Composition of plant telomeres. *Biochim Biophys Acta* 1769: 399-409.
- Zhdanova NS, Minina JM, Karamisheva TV, Draskovic I, Rubtsov NB, Londoño-Vallejo JA (2007). The very long telomeres in *Sorex granarius* (Soricidae, Eulipothyphla) contain ribosomal DNA. *Chromosome Res* 15: 881- 890.
- Zheng J, Nakata M, Uchiyama H, Morikawa H, Tanaka R (1991). Giemsa C-banding patterns in several species of *Phaseolus* L. and *Vigna* Savi, Fabaceae. *Cytologia* 56: 459-466.
- Zheng J, Irifune K, Hirai K, Nakata M, Tanaka R, Morikawa H (1995). In situ hybridization to metaphase chromosomes in six speceis of *Phaseolus* and *Vigna* using ribosomal DNA as the probe. *J Plant Res* 107: 365-369.

4. Manuscrito submetido ao periódico *Annals of Botany*

Artigo Original

Caracterização citogenética e dinâmica evolutiva de DNAs satélites em feijões do gênero *Phaseolus*

Tiago Ribeiro¹, Karla G.B. dos Santos¹, Manon M. S. Richard², Mireille Sévignac², Vincent Thareau², Valérie Geffroy^{2,3} and Andrea Pedrosa-Harand^{1,*}

¹Laboratório de Citogenética e Evolução Vegetal, Departamento de Botânica, Universidade Federal de Pernambuco, Recife, PE, Brasil, ²Institut of Plant Sciences Paris-Saclay (IPS2), UMR 9213/UMR1403, CNRS, INRA, Université Paris-Sud, Université d'Evry, Université Paris-Diderot, Sorbonne Paris-Cité, Bâtiment 630, 91405 Orsay, France ³Unité Mixte de Recherche de Génétique Végétale, Institut National de la Recherche Agronomique, Gif-sur-Yvette, France

Título resumido: DNAs satélites em *Phaseolus*

*Para correspondência. E-mail andrea.pedrosaharand@pesquisador.cnpq.br

Resumo

Introdução e objetivos As regiões subteloméricas dos cromossomos do feijão comum (*Phaseolus vulgaris*) são sítios de acúmulo para *khipu*, o principal DNA satélite identificado até o momento nesse genoma. No presente trabalho, a organização genômica de *khipu* no feijão comum e em outras espécies de *Phaseolus*: *P. leptostachyus*, *P. lunatus* and *P. microcarpus* foi investigada. Adicionalmente, uma nova sequência satélite (*jumper*), relacionada ao DNAr 5S, foi identificada.

Métodos Clonagem, hibridização *Southern blot* e hibridização *in situ* fluorescente (FISH) foram combinadas para analisar a organização e distribuição desses satélites em *Phaseolus*.

Principais resultados Uma mistura de clones de *khipu* oriunda de *P. vulgaris* foi hibridizada *in situ* e confirmou a presença de sequências silimares nas regiões subterminais dos cromossomos em todas as espécies analisadas, com diferenças no número e intensidade dos sinais. *Khipu* apresentou monômero de ~500 pb e a análise de sequência de fragmentos clonados relevou distintas relações entre as unidades das diferentes espécies. Além disso, um novo DNA satélite de 170 pb foi identificado. Denominada de *jumper*, a sequência mostrou similaridade com o espaçador não transcrito (NTS) do DNA ribossomal 5S em *P. vulgaris*, mas foi detectada em alta abundância apenas nos cromossomos de *P. microcarpus*, estando presente nos subtelômeros e pericentrômeros.

Conclusões Os dados demonstram *khipu* como um importante DNA satélite no gênero *Phaseolus* e a complexidade na composição de repeats no genoma de *P. microcarpus*. Adicionalmente, o panorama de distribuição diferencial e evolução dessas sequências reiteram a presença de uma biblioteca de DNAs satélites em *Phaseolus*.

Palavras-chave: heterocromatina, Amplificação de DNA satélite, subtelômeros, pericentrômeros, FISH, *Southern blot*, DNAr 5S, *Vigna*, *Macroptilium*, *Cajanus*, *Phaseolus*.

INTRODUÇÃO

DNAs satélites compreendem longos arranjos de uma sequência básica, monômero, numa organização em tandem denominada *head-to-tail*. Os monômeros são variáveis em tamanho, geralmente entre 160-180 pb ou 320-370 pb (Macas *et al.*, 2002; Hemleben *et al.*, 2007). Estas sequências são especialmente abundantes na heterocromatina localizada nos centrômeros, pericentrômeros e subtelômeros. Entretanto, também podem ocorrer em regiões intersticiais ao longo do cromossomo (Sharma and Raina, 2005; Hemleben *et al.*, 2007). Originalmente tratados como DNA lixo (*junk DNA*), os DNAs satélites têm sido identificados como atuantes em diferentes processos celulares, como manutenção da coesão entre cromátides irmãs e na segregação cromossômica (Nagaki *et al.*, 2003; Gong *et al.*, 2009).

O processo de evolução dessas sequências ocorre de forma conjunta (evolução em concerto), embora a homogeneização possa ser incompleta ou ocorre lentamente (Nei and Rooney, 2005; Hemleben *et al.*, 2007). Como resultado disso, e dependendo da idade da sequência, as famílias de DNA satellite podem apresentar perfis cromossomo- ou espécie-específicos, como o JNK de centeio (*Secale cereale*, Nagaki *et al.*, 1999), enquanto outros podem estar presentes de forma mais abrangente em um grupo, como as famílias CsSat em *Citrus* (Barros e Silva *et al.*, 2010) ou a E180 em *Medicago* e gêneros relacionados (Rosato *et al.*, 2012).

Phaseolus é um gênero pertencente à Fabaceae e compreende cerca de 70 espécies, algumas com largo potencial socio-econômico, como o feijão comum (*P. vulgaris*) e o feijão fava (*P. lunatus*) (Broughton *et al.*, 2003; Delgado-Salinas *et al.*, 2006). O grupo é filogeneticamente resolvido em dois grandes clados, denominados A e B (Delgado-Salinas *et al.*, 2006). O A compreende a maioria das espécies nativas, incluindo *P. microcarpus*, enquanto o clade B engloba espécies nativas e cultivadas, como *P. leptostachyus*, *P. lunatus* e *P. vulgaris*. A maioria das espécies de *Phaseolus* possui $2n = 22$ cromossomos, exceto o grupo formado por *P. leptostachyus*, *P. micranthus* e *P. macvaughii* (o grupo *Leptostachyus*), onde uma redução para $2n = 20$ foi detectada (Mercado-Ruaro and Delgado-Salinas, 1998; Delgado-Salinas *et al.*, 2006).

A heterocromatina no gênero, principalmente investigada em espécies de interesse econômico, está distribuída na região pericentromérica e subtelomérica dos cromossomos (Moscone *et al.*, 1999; Fonsêca *et al.*, 2010). Em *P. vulgaris*, ela é visualizada como bloco ricos em bases CG nos pericentrômeros da maioria dos

cromossomos (Fonsêca *et al.*, 2010), mas também é visualizada nos terminais cromossômicos em estágio paquitênico (David *et al.*, 2009; Geffroy *et al.*, 2009). A caracterização de sequências repetitivas no gênero está limitada a *P. vulgaris*, onde diferentes DNAs satélites, CC4 (Falquet *et al.*, 1997; Almeida *et al.*, 2012), OPG9 (Métais *et al.*, 1998) e *khipu* (David *et al.*, 2009), bem como outros elementos repetitivos como *PvMeso* (Ribeiro *et al.*, 2011) e *Tpv2* (Garber *et al.*, 1999), foram descritos. Dados genômicos e cromossômicos, entretanto, sugerem que cerca de 50% do genoma da espécie seja composto de elementos repetitivos, sugerindo que apenas uma pequena fração desses elementos foi investigada até então (Schlueter *et al.*, 2009; Fonsêca *et al.*, 2010).

O DNA satélite *khipu* foi inicialmente descrito durante a caracterização molecular do locus de genes de resistência B4 do feijão comum (Geffroy *et al.*, 2009) e caracterizado em detalhes por David *et al.* (2009) e Richard *et al.* (2013). A sequência apresenta um monômero de 528 pb, conteúdo de GC de 38,2 % e está localizada nos subtelômeros da maioria dos cromossomos de *P. vulgaris*. Hibridização em membrana com clones *khipu* de *P. vulgaris* mostrou a especificidade da sequência para o gênero e uma grande variação intraespecífica (David *et al.*, 2009). O que pode representar apenas diferença no número de cópias entre espécies analisadas ou significar mudanças significativas ao nível de sequência. Dados de mapeamento por BAC-FISH de clones contendo *khipu* oriundos de *P. vulgaris* em *P. lunatus* e *P. microcarpus*, entretanto, não mostraram qualquer sinal subtelomérico, sugerindo uma baixa frequência do satélite no genoma ou a divergência de sequência entre estas espécies (Bonifácio *et al.*, 2012; Fonsêca and Pedrosa-Harand, 2013). No presente trabalho, a presença, distribuição cromossômica e organização de *khipu* foi analisada em diferentes espécies de *Phaseolus* (*P. leptostachyus*, *P. lunatus*, *P. microcarpus* e *P. vulgaris*, representando linhagens divergentes no gênero) a fim de entender a sua evolução. Além disso, um novo satélite, localizado adjacente a *khipu* e relacionado ao DNA ribossomal 5S, foi descrito e analisado em diferentes espécies do gênero.

MATERIAL E MÉTODOS

Material vegetal e extração de DNA genômico

Sementes de *P. leptostachyus* ‘179671’ (*Ple*), *P. lunatus* ‘G26256’ (*Plu*), *P. oligospermus* ‘40691’ (*Po*) e *P. microcarpus* ‘G40535’ e ‘179698’ (*Pm*), bem como de *P. vulgaris* (‘BAT93’ e ‘JaloEEP558’, *Pv*), *Vigna radiata* ‘China1529’, *Vigna*

unguiculata ‘BR14-Mulato’, *Macroptilium atropurpureum* ‘4413’ e *Cajanus cajan* ‘Petrolina’ foram obtidas da Embrapa Arroz e Feijão (Goiás, Brasil), Embrapa Semi-Árido (Pernambuco, Brasil) e do Centro Internacional para Agricultura Tropical - CIAT (Cali, Colombia). DNA genômico foi isolado de folhas jovens utilizando o kit *DNeasy Plant Mini Kit* (Qiagen).

Seleção de clones BACs

Sete clones BACs da biblioteca de *P. vulgaris* ‘BAT93’ (Kami *et al.*, 2006), selecionados com marcadores genéticos cópia-única, mas mostrando sinais de hibridização subteloméricos nos experimentos de FISH em *P. vulgaris* (Pedrosa-Harand *et al.*, 2009; Fonsêca and Pedrosa, 2013), foram analisados para confirmar a presença de *kipu* ou outras sequências subteloméricas (Tabela dado suplementar S1).

Isolamento, clonagem e sequenciamento de fragmentos kipu-like em Phaseolus

Duas estratégias foram usadas para amplificação de fragmentos de *kipu* em outras espécies de *Phaseolus*. Baseado em um alinhamento prévio de unidades de *kipu* de *P. vulgaris* (David *et al.* 2009), um par de primers degenerados foi desenhado utilizando o programa *Primo profile 3.4* (trial version, ChangeBioscience). Adicionalmente, primers não degenerados desenhados para alinhar em diferentes unidades de *kipu* foram feitos com base no mesmo alinhamento (Tabela dado suplementar S2). Diferentes combinações foram usadas em PCR (*kipudegF* x *kipudegR*; *kipuF1* x *kipuR2*; *kipuF1* x *kipuR3*; *kipuF2* x *kipuR1*; *kipuF2* x *kipuR3*; *kipuF3* x *kipuR2* and *kipuF4* x *kipuR2*) e os fragmentos obtidos foram clonados em vetor pCR2.1-TOPO (Invitrogen) ou pGEM-T/pGEM-T easy (Promega). O sequenciamento foi realizado para ambas as fitas de DNA nos serviços da Macrogen (Coréia do Sul), da plataforma de sequenciamento LABCEN/CCB/UFPE (Brasil) ou no GATC (Alemanha).

Identificação e clonagem de jumper em Phaseolinae

Um dos clones da biblioteca de plasmídeos de *P. microcarpus* continha um outro DNA satélite adjacente a *kipu*. Esta nova sequência foi nomeada de *jumper*. Um par de primers (JumpF: TATGAAAGTTGCTCGGAACG and JumpR: AACTACTGACAAATCGGCAC) foi desenhado e usado para amplificar esta sequência do genoma de diferentes espécies de *Phaseolus* e outras *Phaseolinae*. Os

amplicons de *P. microcarpus*, *P. oligospermus*, *V. unguiculata* e *C. cajan* foram clonados como descrito anteriormente.

Validação e análises filogenéticas

Análises genômicas foram conduzidas nas versões disponíveis dos genomas de feijão comum "*Phaseolus vulgaris* v1.0" (Schmutz *et al.*, 2014), soja, "*Glycine max* v1.1" (Schmutz *et al.*, 2010), e grão-de-bico, "*Cicer arietinum* GCA_000331145.1" (Jain *et al.*, 2013).

Para análises envolvendo *khipu*, cada um dos fragmentos clonados das diferentes espécies de *Phaseolus* foi checado quanto a sua similaridade com *khipu* com o *hmmsearch* (<http://hmmer.janelia.org>; Eddy, 1998) usando um HMM *profile* baseado num alinhamento múltiplo de 105 unidades de *khipu* (David *et al.*, 2009). As sequências positivas de *P. leptostachyus*, *P. lunatus* and *P. microcarpus* foram alinhadas com 68 *khipu* de *P. vulgaris*, representando a diversidade dessa sequência no genoma da espécie (Richard *et al.*, 2013), usando o programa *Muscle* (Edgar, 2004a, b). A árvore de máxima verossimilhança foi gerada de acordo com Richard *et al.* (2013).

Sequências similares a *Jumper* foram pesquisadas nos genomas do feijão comum, soja e grão-de-bico com HMMer usando um modelo HMM baseado em 74 *jumper* como arquivo de entrada, incluindo 30 *jumper* de *P. microcarpus* e 44 *jumper* de *P. oligospermus*. Três unidades de *jumper* foram isoladas de cada agrupamento da sequência presente no genoma de *P. vulgaris*, resultando num total de 71 *jumper* oriundo de *P. vulgaris*. As mesma foram alinhadas com os clones previamente identificados em *P. oligospermus* (44 sequências) e *P. microcarpus* (30 sequências) e a árvore construída por máxima verossimilhança como anteriormente descrito. A edição final foi realizada usando o pacote R.

Hibridização Southern blot

Para análise do BACs, o DNA foi extraído usando o Plasmid Mini Kit (Qiagen) e digerido com *EcoRI*, *HindIII* ou *RsaI* (Fermentas). Posteriormente, foi separado por eletroforese em gel de agarose e blotado em membrana de náilon Hybond N+ (Roche Diagnostics). Para análise de DNA genômicos, os DNAs de *Phaseolus* species, *V. radiata* e *C. cajan* foram digeridos apenas com *HindIII*. As sondas usadas foram: quatro clones de diferentes espécies de *Phaseolus* selecionados aqui (*Pmkhipu1*, *Pmkhipu2*, *Plekhipu2* e *Plukhipu7*); uma mistura de unidades de *khipu* de *P. vulgaris* (*Pvkhipu mix*;

David *et al.*, 2009); um produto de PCR de *jumper* amplificado do genoma de *P. microcarpus*; e uma unidade específica de *khipu* (um fragmento de 500 pb isolado do clone 1H04; David *et al.*, 2009, para a hibridização de BAC apenas). A marcação foi realizada por random primer com digoxigenina-11-dUTP (Roche Diagnostics). A hibridização foi realizada a 37°C como descrito em Ribeiro *et al.* (2011) e a lavagem a 45°C e/ou 68°C para baixa e alta estringência, respectivamente.

Preparação de lâminas e hibridização fluorescente in situ (FISH)

As lâminas e a FISH com 76% de estringência foram realizadas segundo Fonsêca *et al.* (2010). Para a hibridização em baixa estringência (40%), o mix foi composto por 10% (w/v) dextrano sulfato, 6x SSC and 50-100 ng de sonda e as lavagens foram realizadas em temperatura ambiente. As sondas foram as mesmas descritas acima, exceto para *P. leptostachyus* onde outro clones foi utilizado (*Plekhipu12*), e foram marcadas por *nick translation* com Cy3-dUTP (Amersham). Quando necessário, a rehibridização das lâminas foi realizada de acordo com Heslop-Harrison *et al.* (1992). As fotos foram feitas usando um microscópio de epifluorescência Leica DMLB equipado com uma câmera de vídeo Cohu.

RESULTADOS

Com a intenção de descobrir se *khipu* era o DNA satélite mais abundante dos subtelômeros dos cromossomos de *P. vulgaris*, sete clones BAC (de quatro terminais cromossômicos distintos) com padrão de distribuição subtelomérico foram selecionados e analisados quanto à sua composição de sequências. Hibridização em membrana utilizando uma sonda de *khipu* e DNA dos BACs digeridos com *HindIII*, *EcoRI* e *RsaI* revelou diversas bandas marcadas em todos os BACs (Figura dado suplementar Fig.1). Estes fragmentos foram os mesmos visualizados quando a fração repetitiva *C₀t-1* foi usada como sonda (não mostrado), sugerindo que *khipu* é, se não a única sequência, a família mais abundante presente nos insertos desses BACS e nos subtelômeros do feijão comum.

A presença e distribuição de sequências relacionadas à *khipu* em outras três espécies de *Phaseolus* (*P. leptostachyus*, *P. lunatus* and *P. microcarpus*) foi avaliada por meio da hibridização *in situ* de uma mistura de unidades de *khipu* do feijão comum sob alta e baixa estringência. No primeiro caso, sinais subteloméricos foram

visualizados em todas as espécies, mas em intensidade menor a encontrada em *P. vulgaris* sob as mesmas condições (forte marcação nos terminais em ambos os braços na maioria dos cromossomos; não mostrado). Sinais adicionais foram vistos sob baixa estringência. Em *P. leptostachyus*, a marcação estava presente em 18 dos 20 cromossomos, embora mais fraca do que a visualizada em *P. vulgaris* (Figs. 1a, b, e e f e Tabela 1). *P. lunatus* mostrou oito terminais cromossômicos marcados, com dois sinais mais fortes que os outros (Figs. 1c e g), enquanto *P. microcarpus* exibiu de 12-14 cromossomos hibridizados em um dos braços (Figs. 1d e 1h). De acordo com o padrão de FISH, três tipos cromossômicos, variando em quantidade a depender da espécie, foram definidos com base na presença de *khipu* em dois (tipo I), um (tipo II) ou nenhum (tipo III) dos braços cromossômicos (Fig. 2; Tabela 1).

Sequências *khipu-like* foram obtidas de outras espécies de *Phaseolus* por meio de uma estratégia baseada em PCR. Nove primers foram desenhados a partir de um alinhamento múltiplo de sequências de *khipu* de *P. vulgaris* (David *et al.*, 2009) e usados para amplificação. Vinte e um clones de *P. microcarpus* (*Pmkhipu1-Pmkhipu21*), nove de *P. lunatus* (*Plukhipu1-Plukhipu9*) e 30 de *P. leptostachyus* (*Plekhipu1-Plekhipu30*) foram obtidos e usados em uma reconstrução filogenética, em combinação com um conjunto de 68 *khipu* de *P. vulgaris*. Embora alguns clados espécie-específicos tenham sido encontrados, sequências das diferentes espécies apareceram misturados de forma geral. Clones de *P. leptostachyus* foram mais relacionados com os de *P. microcarpus* do que com os *P. lunatus* (Fig. 3), contrastando com as relações filogenéticas pré-estabelecidas.

A organização molecular de *khipu* foi investigado por hibridização em membrana usando DNAs genômicos de diferentes espécies digeridos com *HindIII*. A mistura de *khipu* de *P. vulgaris* (*Pvkhipu mix*) mostrou hibridização em todas as espécies de *Phaseolus* (Fig. 4a, b), resultando num padrão de bandas em escada (*ladder-like*) variando entre 0,5 e 2 kb. O clone de *P. lunatus* (*Plukhipu7*) apresentou forte marcação apenas no genoma de *P. vulgaris* (não mostrado), enquanto o oriundo de *P. leptostachyus* (*Plekhipu2*) hibridizou apenas na espécie de origem (dados suplementar Fig. S2a). Uma banda de alto peso molecular foi vista no genoma de *P. microcarpus* e *P. leptostachyus*, enquanto um fragmento de 0,5 kb foi observado em *P. oligospermus*, quando o clone *Pmkhipu1* (*P. microcarpus*) foi utilizado (dados suplementar Fig. S2b). *V. radiata* and *C. cajan* não mostraram sinais de hibridização com todas as sondas testadas.

A distribuição cromossômica das sequências de *khipu* obtidas de outras espécies foi avaliada por meio de FISH. Os clones foram mapeados nos subtelômeros em todos os casos, mas com clara variação interespecífica no número e intensidade dos sinais (Fig. 5). *Plukhipu7* mostrou sinal em um único par cromossômico de *P. lunatus* (Fig. 5a) e *P. leptostachyus* (Fig. 5b), enquanto em *P. vulgaris* a marcação foi detectada em 12-14 cromossomos (não mostrado). O clone de *P. leptostachyus* (*Plekhipu12*) mostrou marcação apenas na própria espécie, com diferenças na intensidade entre os cromossomos (Fig. 5d). Adicionalmente, dois clones de *P. microcarpus* foram localizados em um dos braços cromossômicos em seis pares da própria espécie (*Pmkhipu1*; Fig. 5c) ou em um par de *P. leptostachyus* e *P. lunatus* (*Pmkhipu2*; Fig. 5e-f). Quando hibridizado nos cromossomos da espécie de origem, *Pmkhipu2* mostrou marcação nos subtelômeros e pericentrômeros de vários cromossomos. Uma análise mais detalhada desse clone revelou a presença de outro DNA satélite, adjacente a *khipu*, o qual foi responsável pelo padrão de marcação mais disperso ao longo dos cromossomos.

A nova sequência, nomeada *jumper*, foi identificada na forma de duas unidades de ~ 170 pb no clone *Pmkhipu2*. Uma triagem por PCR revelou a ocorrência desse novo satélite no genoma de diferentes espécies de *Phaseolus*. Amplicons também foram observados em outros gêneros (*Cajanus*, *Macroptilium* and *Vigna*), embora o padrão sugira uma organização distinta da sequência (dados suplementar Fig. S3).

Uma análise mais detalhada da organização molecular de *jumper* foi realizada pela checagem de sua distribuição no genoma de diferentes legumes (feijão comum, grão-de-bico e soja). No feijão comum, 1138 unidades de *jumper* foram identificadas em 10 dos 11 pseudomoléculas (exceto no pseudocromossomo 4) (dados suplementar Tabela S3). Estas unidades estavam organizadas em arranjos em tandem *head-to-tail*, eventualmente adjacente à *khipu*, ou como duas unidades de separadas por outra sequência, aqui chamada *interjumper*, em um modo contínuo (Fig. 6a). Nos genomas de soja e grão-de-bico não foram detectadas sequências similares a *jumper*.

Interessantemente, parte da sequência de *interjumper* encontrada em *P. vulgaris* mostrou similaridade com o DNA ribossomal 5S (Fig. 6b) e a organização *jumper-interjumper-jumper* foi encontrada apenas nos pseudocromossomos 6 e 10, aqueles que possuem sítios de DNAr 5S no feijão comum. A relação entre *jumper* e o DNAr 5S foi então extensamente analisada a partir do isolamento e comparação de várias sequências de DNAr 5S dos pseudocromossomos 6 and 10. *Jumper* foi encontrado como parte do

espaçador não transcrito NTS (Non-Transcribed Spacer) de todas as sequências de DNAr 5S rDNA do pseudocromossomo 6 (Fig. 6c) e várias do pseudocromossomo 10 (dado suplementar Fig. S4). Embora o NTS, e portanto, *interjumper* varie em tamanho, *jumper* foi encontrado sempre na mesma posição ao longo de todos os NTS analisados (dado suplementar Fig. S4). As unidades de *jumper* isoladas do genoma do feijão comum se mostraram bastante divergentes daquelas obtidas de *P. microcarpus* e *P. oligospermus*, o que pode representar um artefato do procedimento de clonagem ou ser devido a recente amplificação de *jumper* nas últimas duas espécies. Excluindo umas poucas unidades, todas as sequências de *P. oligospermus* and *P. microcarpus* agruparam em clados espécie-específicos (dado suplementar Fig. S5).

A distribuição citológica de *jumper* foi analisada por FISH usando um produto de PCR de *P. microcarpus* e revelou, mesmo sob baixa estrigência, sinais de hibridização apenas em *P. microcarpus*, corroborando a divergência de sequências. Todos os cromossomos apareceram marcados com diferentes intensidades nos pericentrômeros, subtelômeros ou ambos os domínios. A colocalização com o DNA satélite *kipu* foi visualizada em 12 subtelômeros, mas não com os loci de DNAr 5S (Fig. 7). Por hibridização em membrana, a presença de um padrão em escada foi detectado apenas nos genomas de *P. microcarpus* and *P. oligospermus* (sob baixa estrigência) quando a mesma sonda foi utilizada (Fig. 8).

DISCUSSÃO

Estudos prévios de mapeamento comparativo em *P. lunatus* (Bonifácio *et al.*, 2012) e *P. microcarpus* (Fonsêca and Pedrosa-Harand, 2013) usando BACs contendo *kipu* de *P. vulgaris* (Fonsêca *et al.*, 2010) sugeriram a ausência desta sequência repetitiva nos subtelômeros dessas espécies. Entretanto, os dados citomoleculares do presente trabalho, combinados com as análises prévias realizadas por David *et al.*, (2009), demonstram que *kipu* é de fato uma sequência satélite presente nos subtelômeros de todas as espécies de *Phaseolus* analisadas. Assim, a variação observada em FISH com a mistura de *kipu* de *P. vulgaris* *kipu* sob diferentes estrigências nas variadas espécies dão suporte para divergências ao nível de sequência. Adicionalmente, os dados também indicam variações no número e tamanho dos blocos

de *khipu* entre as espécies, com a sequência sendo mais abundante em *P. vulgaris* and *P. leptostachyus* e menos abundante em *P. lunatus*.

A estratégia baseada em PCR para amplificação e clonagem de sequências relacionadas a *khipu* foi eficiente. As análises de sequência mostraram similaridade significativa entre os fragmentos obtidos aqui e as unidades de *khipu* do *P. vulgaris* (David *et al.*, 2009; Richard *et al.*, 2013). Entretanto, mesmo usando primers desenhados para amplificar regiões mais conservadas das unidades, sequências divergentes também foram obtidas. O consenso apresentou alto percentual de AT, como já conhecido para *P. vulgaris* (David *et al.*, 2009) e diversas outras famílias de DNA satélite (Macas *et al.*, 2002).

A análise de máxima verossimilhança revelou que *khipu* oriundo das diferentes espécies agruparam de forma mista. Apesar disso, alguns clones de *P. lunatus* foram mais relacionados às unidades de *P. vulgaris* (Richard *et al.*, 2013), enquanto as de *P. leptostachyus* foram mais próximas às de *P. microcarpus*, uma espécie basal do gênero posicionada no clado A (Delgado-Salinas *et al.*, 2006). Diversos estudos têm demonstrado que a evolução em concerto da unidade de repetição leva, em geral, a uma menor variação intraespecífica do que entre espécies relacionadas (ver, por exemplo, Fernández *et al.*, 2001; Acosta *et al.*, 2010). Algumas vezes, entretanto, esta homogeneização pode ocorrer de forma lenta ou incompleta, levando ao surgimento de subfamílias de uma mesma sequência em um genoma. Em *Silene latifolia*, por exemplo, quatro diferentes subfamílias de um DNA satélite subtelomérico foram identificadas, com cada subtelômero apresentando uma composição diferentes de cada uma delas (Kazama *et al.*, 2006). Os agrupamentos de sequências encontradas nas análises realizadas aqui sugerem que *khipu* não está sendo completamente homogeneizado.

A análise por FISH usando diferentes clones *khipu* demonstraram a especificidade da sequência aos subtelômeros, como previamente relatado para *P. vulgaris* (David *et al.*, 2009). Estas regiões são consideradas uma zona de transição entre sequências cromossomo-específicas e o DNA telomérico (McEachern *et al.*, 2000), sendo um local comum para formação de blocos heterocromáticos e acumulação de DNAs satélites (Sharma and Raina, 2002). Em *Citrus* (Rutaceae) e grupos relacionados (*Fortunella* and *Poncirus*), por exemplo, Barros e Silva *et al.*, (2010) identificaram um DNA satélite, nomeado CsSat, como a principal sequência presente nos terminais cromossômicos. Em *P. vulgaris*, estes terminais são particularmente enriquecidos em *khipu* e genes de resistência, como os loci Co-2 e B4 (Geffroy *et al.*,

1998; Geffroy *et al.*, 2009), e susceptíveis a frequentes recombinações ectópicas (David *et al.*, 2009).

As variações observadas no número e intensidade dos sinais de FISH dentro e entre espécies de *Phaseolus* podem resultar de diferenças no número de cópias por sítio e/ou devido à divergência entre as sequências presentes em diferentes cromossomos, já que DNAs satélite são conhecidos como elementos de rápida evolução (Hemleben *et al.*, 2007). Por exemplo, o clone *Plukhipu7* foi localizado em uma parte dos cromossomos de *P. vulgaris*, em um número maior que na própria espécie de origem. Nenhum dos clones usados como sonda neste estudo detectou vários blocos de *khipu* em *P. lunatus*, sugerindo que esta sequência esteja subrepresentada neste genoma. Por outro lado, este mesmo clone foi mapeado em um par cromossômico de *P. leptostachyus*, enquanto um clone (*Plekhipu12*) da própria espécie marcou outro par, confirmando a presença de variantes de *khipu* em diferentes terminais cromossômicos. Na batata, os subtelômeros são locais de acúmulo de pelo menos duas diferentes famílias de DNA satélite (CL14 e CL34), diferentemente distribuídas a depender do subtelômero (Torres *et al.*, 2011). As quatro famílias identificadas em *Silene latifolia* também mostraram um padrão de distribuição peculiar, variando o suficiente para produzir padrões específicos em cada terminal cromossômico (Kazama *et al.*, 2006).

Interessantemente, os subtelômeros de *P. microcarpus* possuem não apenas *khipu*, mas duas distintas famílias de DNA satélite. *Khipu* and *jumper* estão intercalados ou adjacentes em pelo menos 10 subtelômeros, suportando a presença de uma biblioteca de DNAs satélites no gênero. Esta biblioteca representaria uma fonte permanente de sequências que podem independentemente ser amplificadas em cada espécie (Plohl *et al.* 2012). De fato, enquanto *khipu* foi citologicamente detectado em todas as espécies, *jumper* sofreu uma amplificação pontual apenas em *P. microcarpus* e espécies relacionadas. Esses satélites variaram no tamanho da unidade de repetição (~528 pb x ~170 pb para *khipu* e *jumper*, respectivamente), número de cópias (2,460 x 353 cópias em *P. vulgaris*, Richard *et al.* 2013 e presente trabalho), distribuição (subtelômerica x subtelômerica and pericentromérica), tempo de amplificação (mais antigo para *khipu*) e grau de homogeneização intraespecífica (maior para *jumper*), confirmando a evolução independente.

Análises genômicas em *P. vulgaris*, *G. max* and *C. arietinum* revelaram que *jumper* compartilha similaridade com o NTS do DNAr 5S no feijão comum. Este DNA satélite apresentou conservação de sequência e foi encontrado sempre na mesma

posição ao longo do NTS do pseudocromossomo 6 em alguns do pseudocromossomo 10, com regiões flanquadores conservadas. No momento, não é possível elucidar a origem desse DNA satélite. A alta similaridade entre as sequências presentes no NTS e aquelas na forma de arranjos em tandem nos outros cromossomos sugerem uma origem recente de *jumper* de parte do NTS. Ao passo que a ausência da sequência em parte dos NTS do cromossomo 10 mostraria uma origem independente do DNA satélite, seguida de inserção no NTS e homogeneização incompleta neste cromossomo. Diferentes DNAs satélites originados do DNA ribossomal têm sido identificados (Jo *et al.* 2011, Martins *et al.* 2006; Vittorazzi *et al.* 2011), principalmente do DNAr 45S. Na maioria dos casos, a sequência recém surgida tende a divergir. Usualmente, ela não é mapeada no local de origem (sítios de DNAr), mas se espalha por outros domínios (ver, por exemplo, Jo *et al.* 2009; Almeida *et al.* 2012). Os sinais de *jumper* não foram identificados nos loci de DNAr 5S, mas sim nos subtelômeros e/ou pericentrômeros de todos os cromossomos de *P. microcarpus*. Frequentemente, os arranjos de DNAs satélites são encontrados nesses locais, mas não em ambos ao mesmo tempo. Um padrão similar foi identificado para o satélite em *Medicago* e grupos relacionados (Rosato *et al.* 2011), sugerindo a ocorrência eventual de recombinação ectópica entre a heterocromatina pericentromérica e subtelomérica. Este fenômeno parece ser frequente em *P. microcarpus* para explicar a extensa distribuição de *jumper* no genoma da espécie. Alternativamente, as duas sequências teriam atingido o padrão de distribuição observado independentemente, após poucos eventos de recombinação.

DNAs satélites apresentam geralmente unidades de repetição entre 160-180 ou 320-370 pb (Hemleben *et al.*, 2007), entretanto diferentes famílias com tamanhos incomuns têm sido descritas, como Cent4 em milho, ~740 pb (Page *et al.*, 2001), e *kipu* em *P. vulgaris* (David *et al.*, 2009). Aqui, o tamanho incomum foi confirmado para *kipu* em outras espécies de *Phaseolus*, enquanto *jumper* (170 pb) se encaixou na média geral. Em muitos casos, esta variação na unidade de repetição aparece devido a mudanças no reconhecimento dos sítios de restrições ao longo do arranjo, levando ao surgimento de unidades mais complexas denominadas *higher order repeats* (HOR). Em *P. vulgaris*, um DNA satélite centromérico, Nazca, está organizado desta forma, com uma unidade complexa de 528 pb composta por três diferentes subrepeats (Awata *et al.*, 2013). Para *kipu*, que apresenta uma unidade de ~ 500 pb, este tipo de organização não foi observado nas espécies analisadas.

Em *P. vulgaris*, *hipu* foi identificado em ~1% das sequências de terminais de BACs analisadas e está relacionado a genes de resistência, indicando que o mesmo exerça um papel evolutivo significativo (David *et al.*, 2009). Os dados moleculares e citológicos obtidos aqui demonstram a importância de estratégias complementares para o estudo da fração repetitiva do genoma. *Khipu* representa um importante componente da heterocromatina subtelomérica in *Phaseolus*, embora divergente dentro e entre espécies. Além disso, a identificação de dois DNAs satélites no genoma de *P. microcarpus*, como ambos ocupando os mesmos ou diferentes subtelômeros, sugere que uma biblioteca de DNAs satélites seja o mecanismo mais comum de evolução dessas sequências no gênero, como já mostrado para diferentes grupos de plantas e animais.

AGRADECIMENTOS

Oa autores agradecem a Marcos Morais e André Ribas (Departamento de Genética, UFPE) pelo auxílio com os experimentos de clonagem, a plataforma de sequenciamento - LABCEN/CCB na UFPE pelo uso de suas facilidades, Labex SPS (INRA, France), 3P project and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brasil, pelo suporte financeiro.

DADOS SUPLEMENTARES

Figura S1: Hibridização *Southern blot* de clones BAC com uma sonda de *hipu*. Figura S2: Clones *hipu* de diferentes espécies de *Phaseolus* hibridizados em DNA genômico. Figura S3: Caracterização do clone de *P. microcarpus* contendo *jumper* e amplificação por PCR em diferentes espécies. Figura S4: Alinhamento de diferentes NTS mostrando a posição conservada de *jumper*. Figura S5: Árvore de *jumper* isolados de diferentes espécies de *Phaseolus*. Tabela S1: Clones BAC subteloméricos usados nas análises de hibridização *Southern blot*. Tabela S2: Lista de primers usados para amplificação de *hipu*. Tabela S3: Unidades de *jumper* presente no genoma de *Phaseolus vulgaris*.

LITERATURA CITADA

Acosta MJ, Marchal JA, Fernández-Espartero C, et al. 2010. Characterization of the satellite DNA Msat-160 from species of *Terricola* (*Microtus*) and *Arvicola* (Rodentia, Arvicolinae). *Genetica* **138**: 1035-1085

- 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.
- Awata A, Tek AL, Richard MMS, et al. 2013.** Identification and characterization of functional centromeres of the common bean. *The Plant Journal* **76**: 47-60.
- Barros e Silva AE, Marques A, Santos KGB, Guerra M. 2010.** evolution of CMA bands in *Citrus* and related genera. *Chromosome Research* **18**: 503-514.
- 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.). *Theoretical Applied Genetics* **124**: 1513-20.
- Broughton WJ, Hernández G, Blair M, Beebe S, Gepts P, Vanderleyden J. 2003.** Beans (*Phaseolus* spp.) - model food legumes. *Plant and Soil* **252**: 55-128.
- David P, Cheng NWG, Pedrosa-Harand A, et al. 2009.** A nomadic subtelomeric disease resistance gene cluster in common bean. *Plant Physiology* **151**: 1048-1065.
- Delgado-Salinas A, Bibler R, Lavin M. 2006.** Phylogeny of the genus *Phaseolus* (Leguminosae): a recent diversification in an ancient landscape. *Systematic Botany* **31**: 779-791.
- Eddy SR. 1998.** Profile Hidden Markov Models. *Bioinformatics* **14**: 755-763.
- Edgar RC. 2004.** MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 113.
- Edgar RC. 2004.** MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**: 1792-1797.
- Falquet J, Cresout F, Dron M. 1997.** Molecular analysis of *Phaseolus vulgaris* rDNA unit and characterization of a satellite DNA homologous to IGS subrepeats. *Plant Physiology and Biochemistry* **35**: 611-622.
- Felsenstein J. 1989.** PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics* **5**: 164-166.
- Fernández R, Barragán MJL, Bullejos M, et al. 2001.** Molecular and cytogenetic characterization of highly repeated DNA sequences in the vole *Microtus cabreræ*. *Heredity* **87**: 637-646.
- Fonsêca A, Ferreira J, Santos TRB, et al. 2010.** Cytogenetic map of common bean (*Phaseolus vulgaris* L.). *Chromosome Res* **18**: 487-502.

- 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**: 335-343.
- Garber K, Bilic I, Pusch O, Tohme J, et al. 1999.** The Tpv2 family of retrotransposons of *Phaseolus vulgaris*: structure, integration characteristics, and use for genotype classification. *Plant Molecular Biology* **39**: 797-807.
- Geffroy V, Cresout F, Falquet J, et al. 1998.** A family of LRR sequences in the vicinity of the Co-2 locus for anthracnose resistance in *Phaseolus vulgaris* and its potential use in marker-assisted selection. *Theoretical and Applied Genetics* **96**: 494-502.
- Geffroy V, Macadré C, David P, et al. 2009.** 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.
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010.** New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59**: 307-321.
- Gong Z, Yu H, Huang J, Yi C, Gu M. 2009.** Unstable transmission of rice chromosomes without functional centromeric repeats in asexual propagation. *Chromosome Research* **17**: 863-872.
- Gouy M, Guindon S, Gascuel O. 2010.** SeaView version 4: A Multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* **27**: 221-224.
- Hembelen V, Kovarik A, Torrez-Ruiz RA, Volkov RA, Beridze T. 2007.** Plant highly repeated satellite DNA: molecular evolution, distribution and use for identification of hybrids. *Systematics and Biodiversity* **5**: 277-289.
- Heslop-Harrison JS, Harrison GE, Leitch IJ. 1992.** Reprobing of DNA:DNA in situ hybridization preparations. *TIGG* **8**:372–373.
- Jain M, Misra G, Patel RK, et al. 2013.** A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). *The Plant Journal* **74**: 715-729.
- Jo SH, Koo DH, Kim JF, et al. 2009.** Evolution of ribosomal DNA-derived satellite repeat in tomato genome. *BMC Plant Biology* **9**: 42.

- Jo SH, Park HM, Kim SM, Kim HH, Hur CG, Choi D. 2011.** Unraveling the sequence dynamics of the formation of genus-specific satellite DNAs in the family solanaceae. *Heredity* **106**: 876-885.
- Kami J, Poncet V, Geffroy V, Gepts P. 2006.** Development of four phylogenetically-arrayed BAC libraries and sequence of the APA locus in *Phaseolus vulgaris*. *Theoretical and Applied Genetics* **112**: 987-998.
- Kazama Y, Sugiyama R, Suto Y, Uchida W, Kawano S. 2006.** The clustering of four subfamilies of satellite DNA at individual chromosome ends in *Silene latifolia*. *Genome* **49**:520-530.
- Keane TM, Creevey CJ, Pentony MM, Naughton TJ, McInerney JO. 2006.** Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evolutionary Biology* **6**: 29.
- Macas J, Mészáros T, Nouzová M. 2002.** PlantSat: a specialized database for plant satellite repeats. *Bioinformatics* **18**: 28-35.
- Martins C, Ferreira IA, Oliveira C, Foresti F, Galleti Jr PM. 2006.** A tandemly repetitive centromeric DNA sequence of the fish *Hoplias malabaricus* (Characiformes: Erythrinidae) is derived from 5S rDNA. *Genetica* **127**: 133-141.
- McEachern MJ, Krauskopf A, Blackburn EH. 2000.** Telomeres and their control. *Annual Review of Genetics* **34**: 331-358.
- Mercado-Ruaro P, Delgado-Salinas A. 1998.** Karyotypic studies on species of *Phaseolus* (Fabaceae: Phaseolinae). *American Journal of Botany* **85**: 1-9.
- Métais I, Aubry C, Hamon B, Peltier D, Jalouzot R. 1998.** Cloning, quantification and characterization of a minisatellite DNA sequence from common bean *Phaseolus vulgaris* L.. *Theoretical and Applied Genetics* **97**: 232-237.
- Moscone EA, Klein F, Lambrou M, Fuchs J, schweizer D. 1999.** Quantitative karyotyping and dual-color FISH mapping of 5S and 18-25S rDNA probes in the cultivated *Phaseolus* species (Leguminosae). *Genome* **42**: 1224-1233.
- Nagaki K, Tsujimoto H, Sasakuma T. 1999.** A novel repetitive sequence, termed the JNK repeat family, located on an extra heterochromatic region of chromosome 2R of Japanese rye. *Chromosome Research* **6**: 95-101.

- Nagaki K, Talbert PB, Zhong CX, Dawe RK, Henikoff S, Jiang J. 2003.** Chromatin immunoprecipitation reveals that the 180-bp satellite repeat is the key functional DNA element of *Arabidopsis thaliana* centromeres. *Genetics* **163**: 1221-1225.
- Nei M, Rooney AP. 2005.** Concerted and birth-and-death evolution of multigene families. *Annual Review in Genetics* **39**: 121-152.
- Page BT, Wanous MK, Birchler JA. 2001.** Characterization of a maize chromosome 4 centromeric sequence: evidence for an evolutionary relationship with the B chromosome centromere. *Genetics* **159**: 291-302.
- Pedrosa-Harand A, Kami J, Gepts P, Geffroy V, Schweizer D. 2009.** Cytogenetic mapping of common bean chromosomes reveals a less compartmentalized small-genome plant species. *Chromosome Research* **17**: 405-417.
- Plohl M, Meštrović N, Mravinac C. 2012.** Satellite DNA Evolution. In: Garrido-Ramos MA, ed. *Repetitive DNA*. Basel: Karger, 126–152.
- Ribeiro T, Santos KGB, Fonsêca A, Pedrosa-Harand A. 2001.** Isolation and characterization of a new repetitive DNA family recently amplified in the Mesoamerican gene pool of the common bean (*Phaseolus vulgaris* L., Fabaceae). *Genetica* **139**: 1135-1142.
- 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. *Frontiers in Plant Science* **4**: 109.
- Rosato M, Galián JA, Rosselló JA. 2012.** Amplification, contraction and genomic spread of a satellite DNA family (E180) in *Medicago* (Fabaceae) and allied genera. *Annals of Botany* **109**: 773-82.
- Schlueter JA, Goicoechea JL, Collura K, et al. 2008.** BAC-end sequence analysis and a draft physical map of the common bean (*Phaseolus vulgaris* L.) genome. *Tropical Plant Biology* **1**: 40-48.
- Schmutz J, Cannon SB, Schlueter J, et al. 2010.** Genome sequence of the palaeopolyploid soybean. *Nature* **463**: 178-183.
- Schmutz J, McClean P, Mamidi S, et al. 2014.** A reference genome for common bean and genome-wide analysis of dual domestications. *Nature Genetics* **46**: 707-713.
- Sharma S, Raina SN. 2002.** Organization and evolution of highly repeated satellite DNA sequences in plant chromosomes. *Cytogenetics and Genome Research* **109**: 15-26.

Torres GA, Gong Z, Iovene M, et al. 2011. Organization and evolution of subtelomeric satellite repeats in the potato genome. *G3* **1**: 85-92.

Vittorazzi SE, Lourenço LB, Del-Grande ML, Recco-Pimentel SM. 2011. Satellite DNA Derived from 5S rDNA in *Physalaemus cuvieri* (Anura, Leiuperidae). *Cytogenetic and Genome Research* **134**: 101-107.

Tabela 1. Distribuição cromossômica de uma mistura de *kipu* de *P. vulgaris* (*Pvkipu* mix) em espécies de *Phaseolus* sob diferentes estringências

Espécie	Tipos cromossômicos ^a	
	Alta estringência (76%)	Baixa estringência (40%)
<i>P. vulgaris</i>	14 I + 4 II + 4 III	16 I + 4 II + 2 III
<i>P. leptostachyus</i>	2 II + alguns terminais fracamente marcados	12 I + 6 II + 2 III
<i>P. lunatus</i>	Alguns terminais fracamente marcados	2 I + 6 II + 14 III
<i>P. microcarpus</i>	Alguns terminais fracamente marcados	12-14 II + 8-10 III

^a I: Sinais de *kipu* em ambos os braços cromossômicos; II: Sinais em apenas um dos braços; III: cromossomos sem sinais de *kipu*

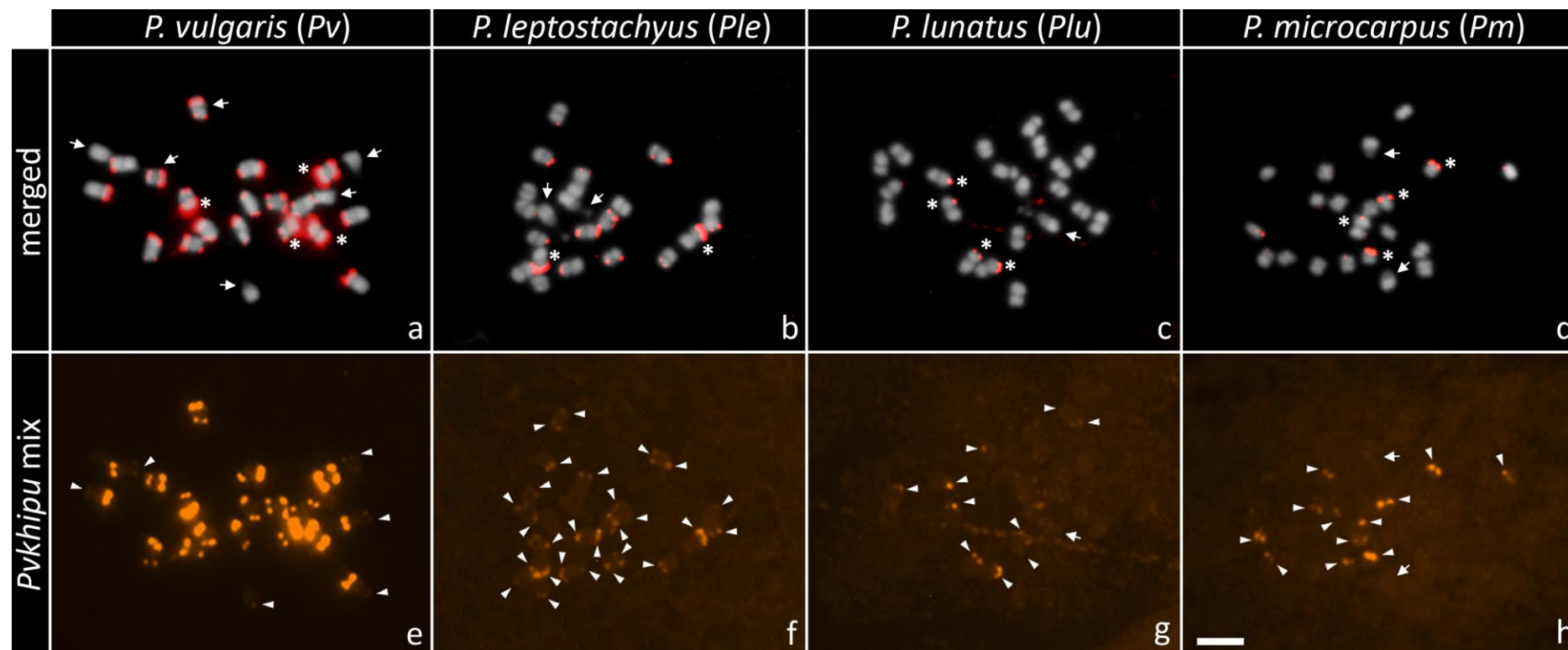


Figura 1. Distribuição cromossômica de um conjunto de *khipu* de *P. vulgaris* (*Pvkhipu mix*) hibridizado *in situ* sob baixa e alta estringência em diferentes espécies de *Phaseolus*: *P. vulgaris* (a e e); *P. leptostachyus* (b e f); *P. lunatus* (c e g) e *P. microcarpus* (d e h). Em a-d, as imagens são mostradas sobrepostas ao DAPI (cinza). Asteriscos em b, c e d indicam os cromossomos com os sinais de hibridização mais fortes e as setas em b, c, d, f, g e h apontam os cromossomos com regiões organizadoras do nucléolo (RON). Em f, g e h cabeças-de-seta indicam todos os terminais marcados e em a apenas os mais fracos. Barra em h = 5 μ m.

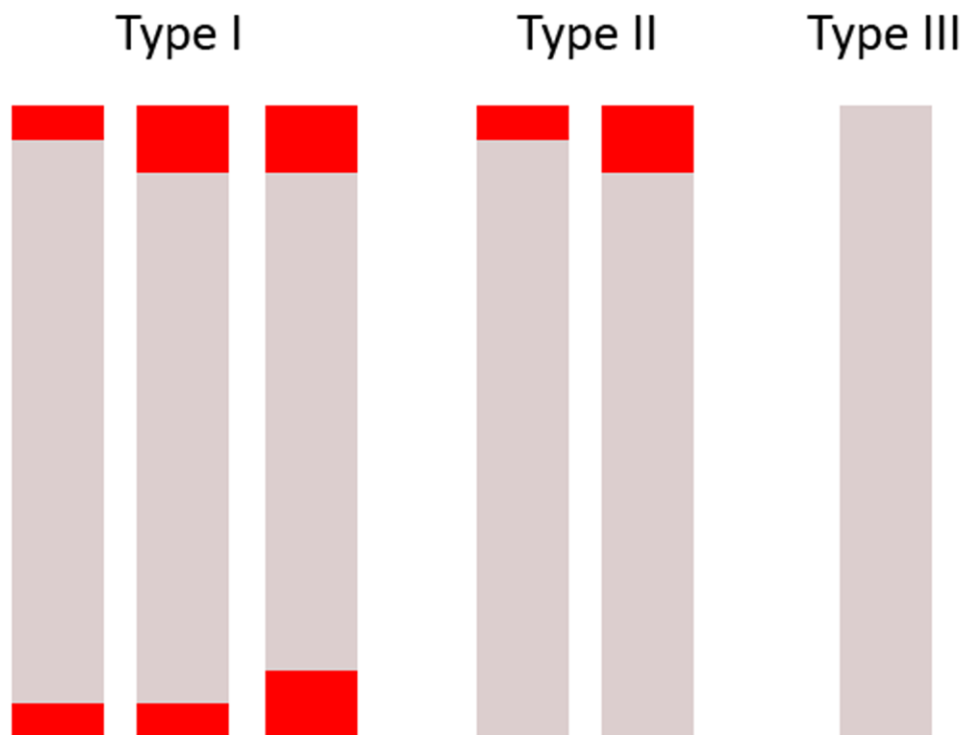


Figura 2. Tipos cromossômicos identificados em *P. vulgaris* e outras espécies de *Phaseolus* (*P. leptostachyus*, *P. lunatus* e *P. microcarpus*) baseado no padrão de hibridização com um conjunto de *khipu* de *P. vulgaris*. Tipo I, marcação em ambos os braços; tipo II, marcação em um dos braços e tipo III, nenhuma marcação.

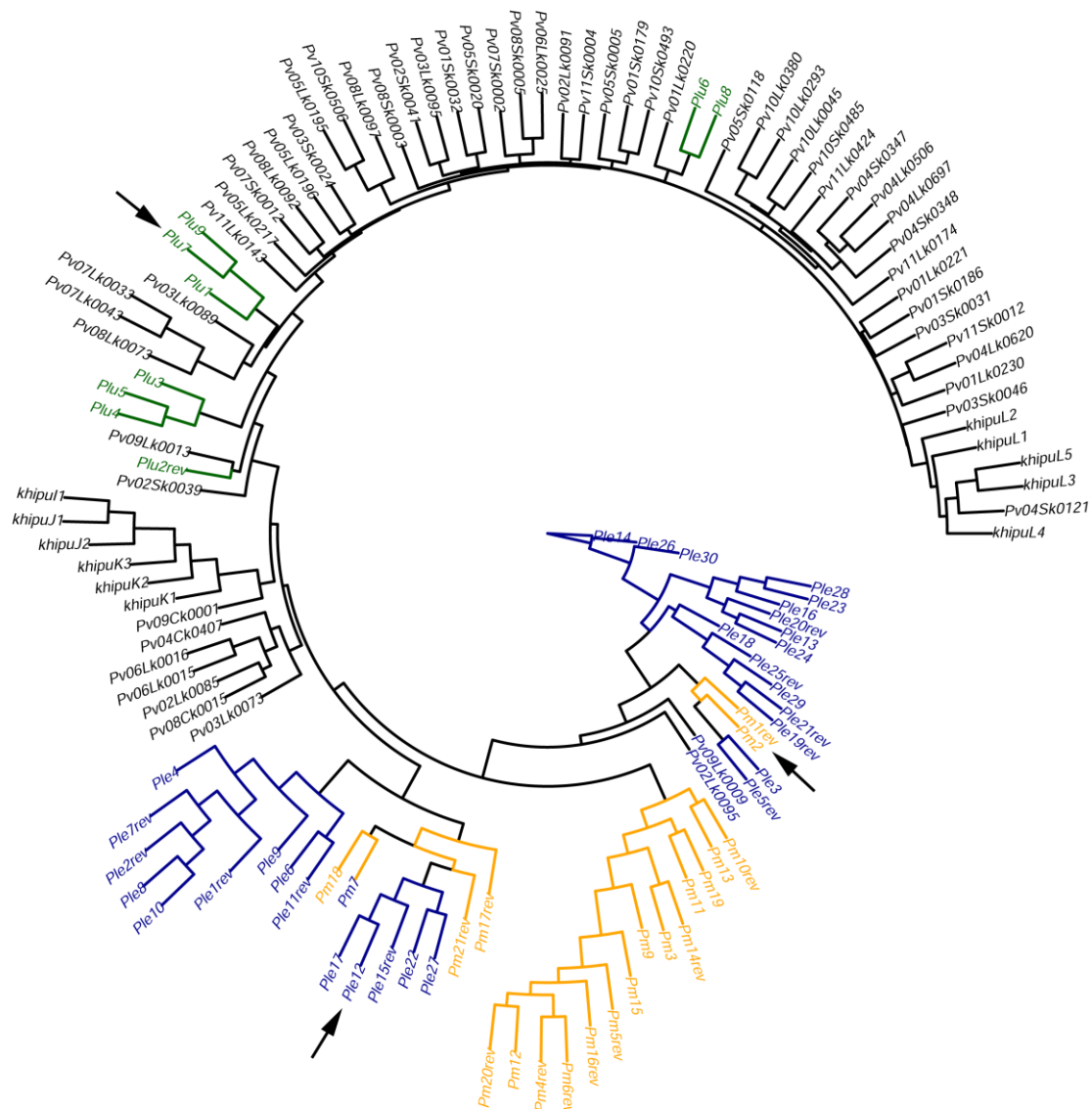


Figura 3. Árvore filogenética de seqüências de *khipu* de diferentes espécies de *Phaseolus* baseada em máxima verossimilhança. Cores azul, laranja e verde identificam as seqüências de *P. leptostachyus*, *P. microcarpus* e *P. lunatus*, respectivamente. Seqüências em preto foram extraídas do genoma de *P. vulgaris*. Setas indicam os clones utilizados como sonda em FISH na Figura

5.

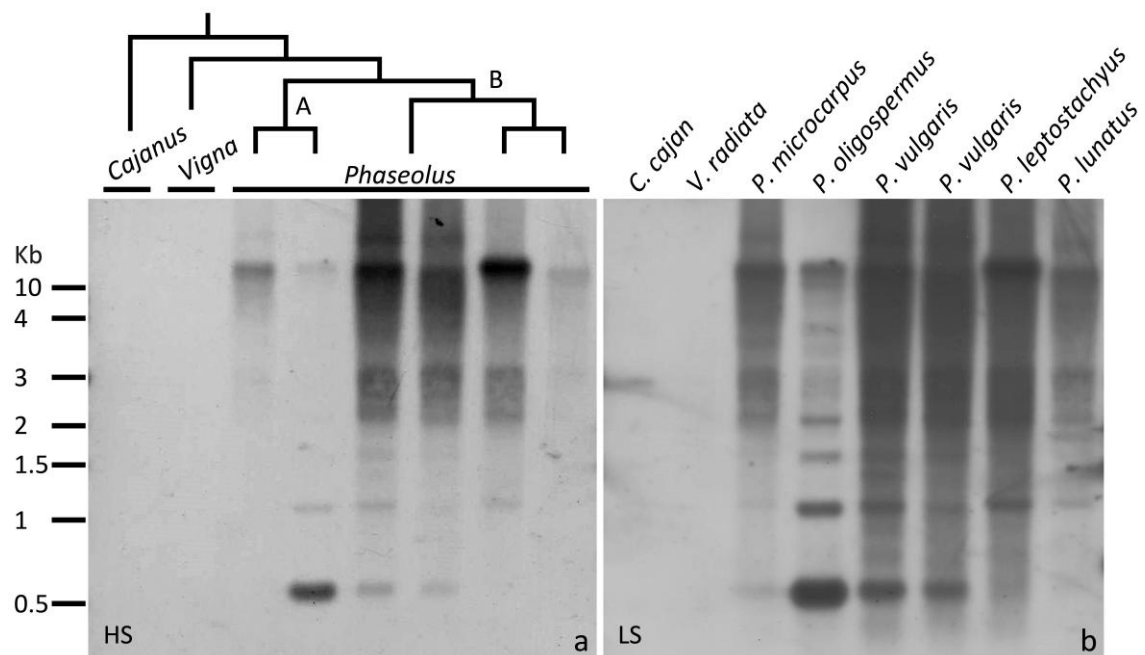


Figure 4. Organização molecular de *kipu* em *Phaseolus* visualizada por hibridização *Southern blot*. DNAs foram hibridizados sob alta (a) e baixa (b) estringência com uma mistura de *kipu* de *P. vulgaris*. Relações filogenéticas mostradas no topo da figura foram baseadas em Delgado-Salinas et al. (2006).

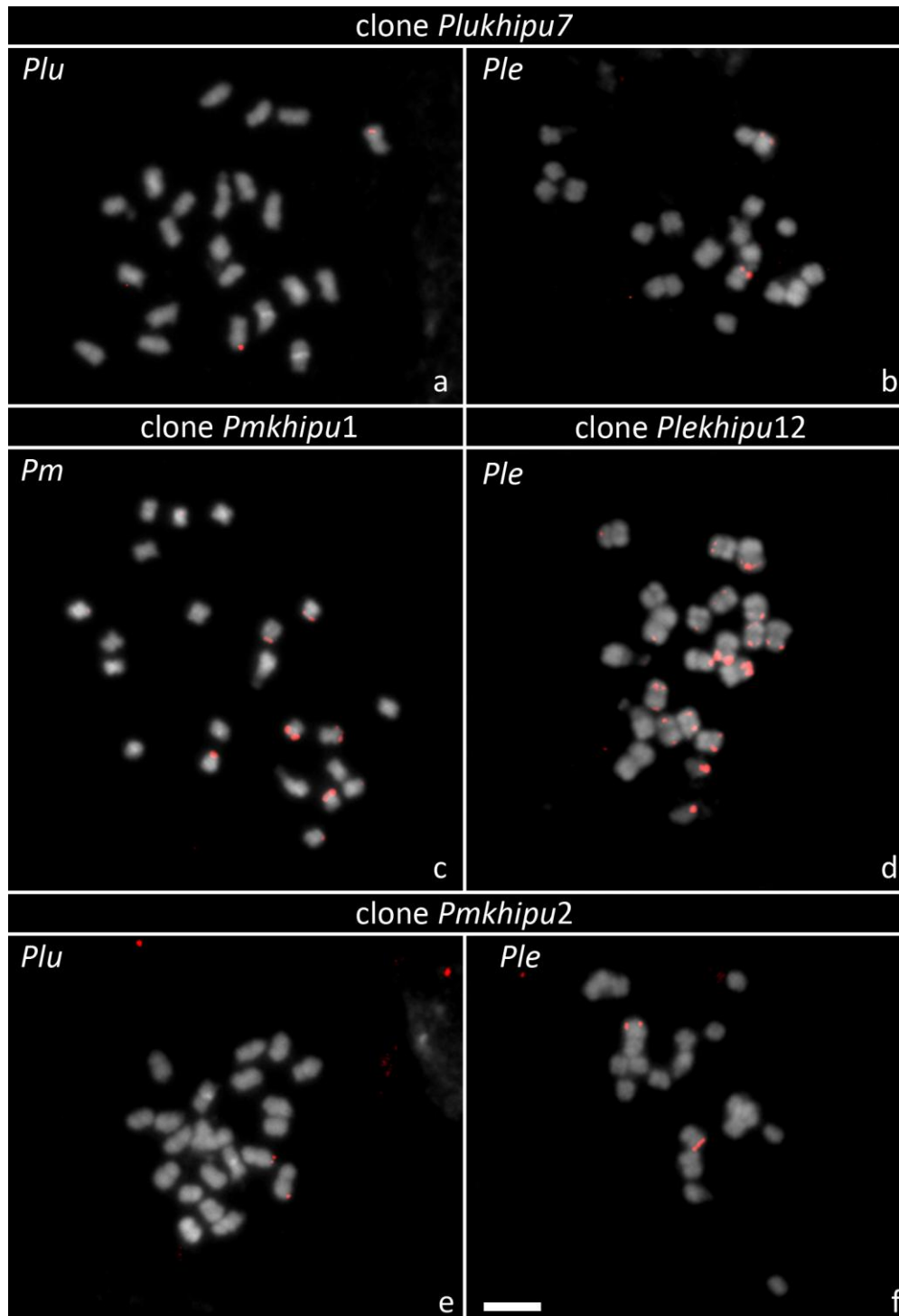


Figura 5. FISH sob alta estringência de clones *hipu* isolados de diferentes espécies de *Phaseolus*. Distribuição do clones *Plukhipu7* em *P. lunatus* (a) e *P. leptostachyus* (b); (c-d) localização de *Pmkhipu1* em *P. microcarpus* (c) e *Plekhipu12* em *P. leptostachyus* (d); (e-f) hibridização de *Pmkhipu2* em *P. lunatus* (e) e *P. leptostachyus* (f). Barra em f = 5 μ m.

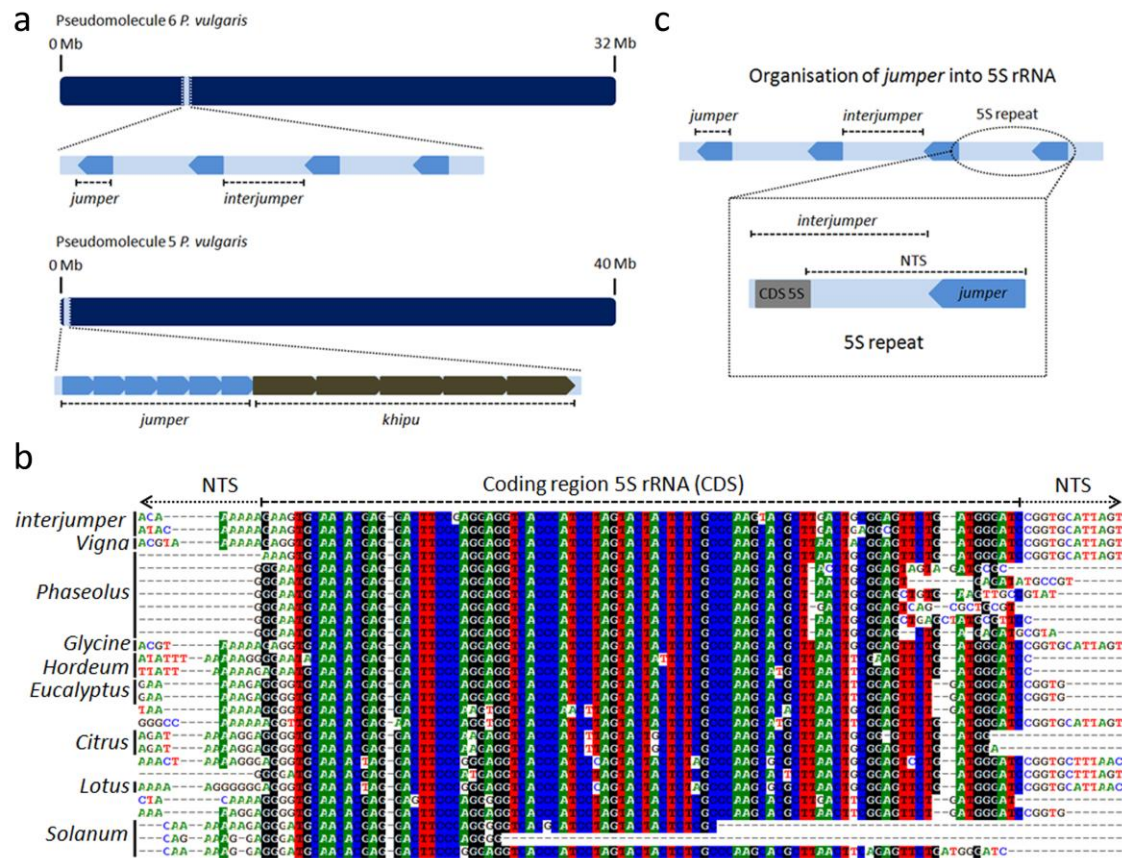


Figura 6. Organização genômica de *jumper* em *P. vulgaris*. (a) Distribuição de *jumper* em dois pseudocromossomos (cromossomos 5 e 6) mostrando a sequência como parte de um fragmento maior (crom. 6) ou arranjada em tandem (crom. 5); (b) Alinhamento de interjumper com o DNAr 5S de diferentes espécies mostrando sua similaridade com a região codificante (CDS); (c) Estrutura da unidade de repetição do DNAr 5S de *P. vulgaris*.

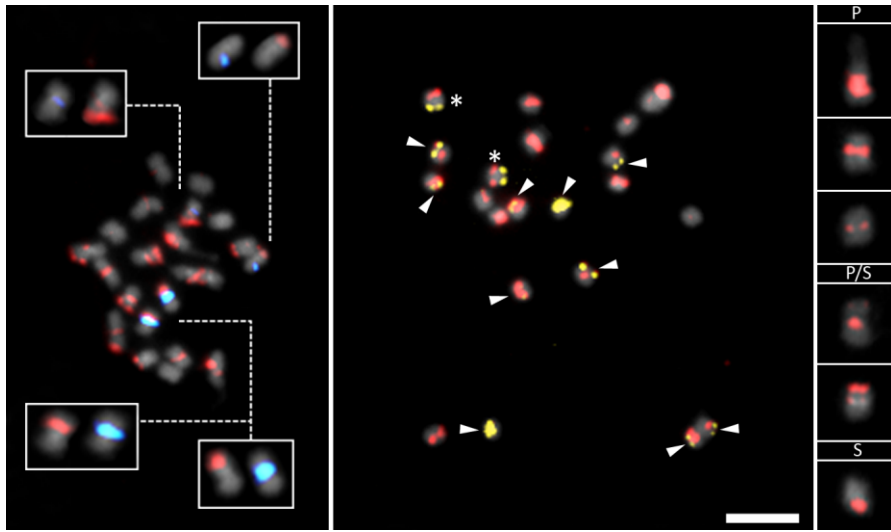


Figura 7. FISH de *jumper* (vermelho) e *khipu* (amarelo) de *P. microcarpus* e do DNAr 5S (azul) em cromossomos mitóticos de *P. microcarpus* (cinza). Cabeças-de-seta apontam sítios de colocação de *khipu-jumper* e asteriscos os cromossomos nos quais as sequências aparecem em terminais diferentes. P: pericentrômeros; S: subteloômeros Barra = 5 μ m.

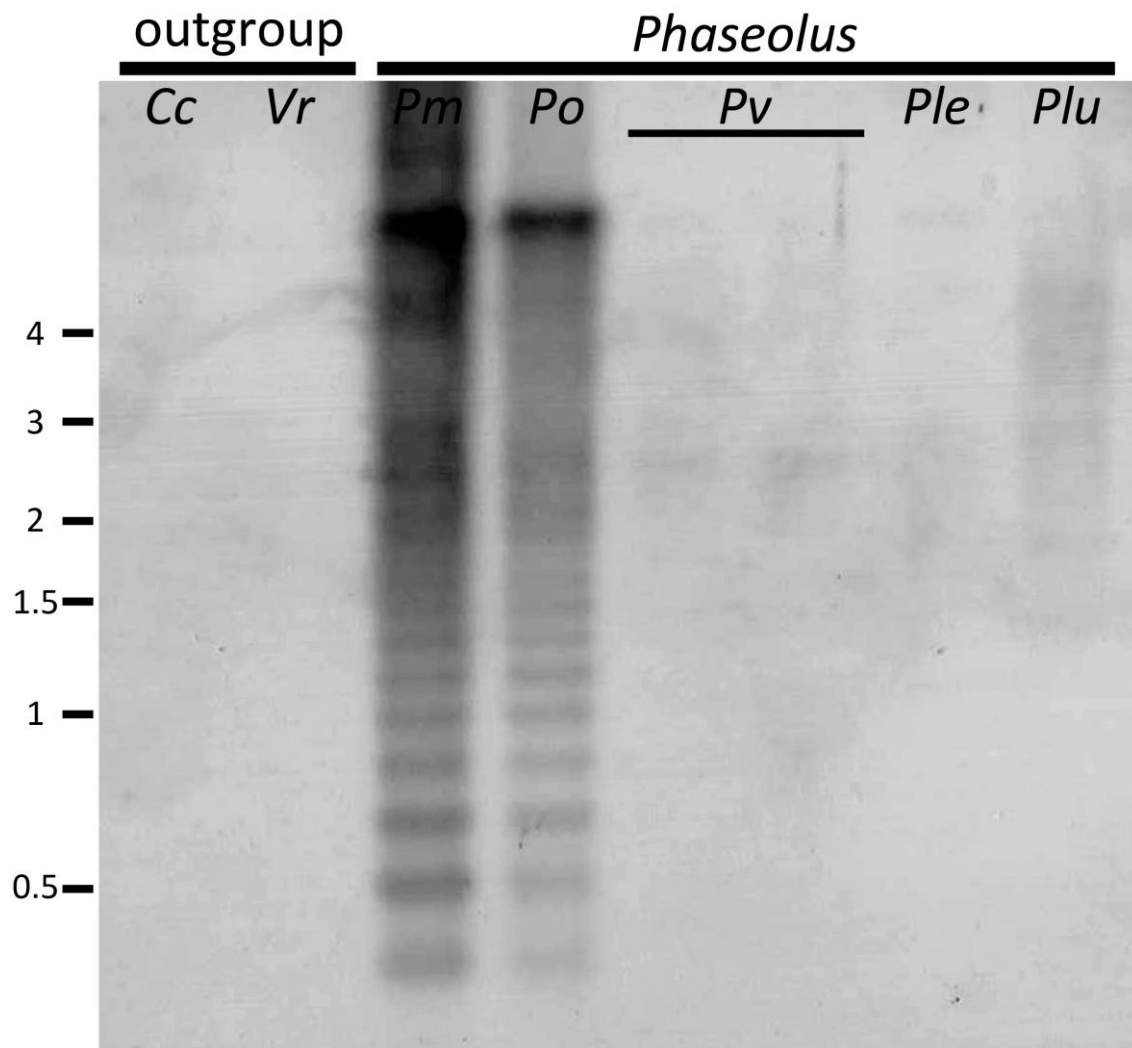


Figura 8. Southern blot de DNAs genômicos de *Phaseolus* (*Po*: *P. oligospermus*; *Pm*: *P. microcarpus*; *Plu*: *P. lunatus*; *Ple*: *P. leptostachyus* e *Pv*: *P. vulgaris*) e outros legumes (*Cc*: *Cajanus cajan* e *Vr*: *Vigna radiata*) digeridos com *Hind*III e hibridizados sob baixa estrigência com uma sonda de *jumper* de *P. microcarpus*.

FIGURAS SUPLEMENTARES

Tabela S1. BACs subteloméricos analisados por *southern blot*. Estes clones mostraram a mesma marcação subtelomérica quando hibridizados *in situ* nos cromossomos do feijão comum (Pedrosa-Harand et al., 2009; Fonsêca et al., 2010)

Cromossomo	Clone BAC	Marcador genético correspondente
2	14F02	Bng 57
4	78L24	Bng 160
4	190C15	Bng 184
4	221J10	Bng 151
8	234P20	Bng 96
10	63H06	Bng 218
10	119E19	L177

Tabela S2. Primers usados para amplificação de sequências relacionadas a *khipu* em espécies de *Phaseolus*.

Primer	Sequência (5' - 3')
<i>khipu</i> Fdeg	CACGTAAGAATCNS
<i>khipu</i> Rdeg	TTGTGCAAACCANNG
<i>khipu</i> F1	TTTCAGGTCAAACGGATGAGT
<i>khipu</i> R1	TTTGCACCCAGGTAAGGAG
<i>khipu</i> F2	CACACACAAACGAACCTTCC
<i>khipu</i> R2	CCCAGGTAAGGAGGCACGTCCA
<i>khipu</i> F3	AGACACAGTGACGAATGTCTGG
<i>khipu</i> R3	GGACTTACTACGCCTTACTCCTTG
<i>khipu</i> F4	ACCCAAGGAGCAAGCCGTAGT

Tabela S3. Número de unidades de *jumper* no genoma de *P. vulgaris*.

Cromossomo	Número de cópias
Crom01	124
Crom02	30
Crom 03	36
Crom04	0
Crom05	144
Crom06	254
Crom07	55
Crom08	105
Crom09	179
Crom10	94
Crom11	117
Total	1,138

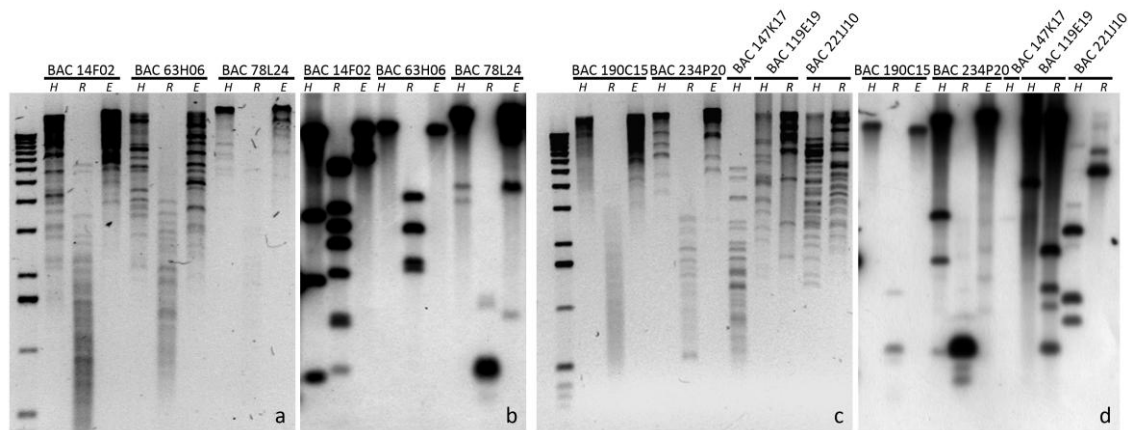


Fig. S1. Southern blot de BACs subteloméricos hibridizados com *khipu* sob alta estringência. Amostras foram coradas com SYBR gold (a e c) e então hibridizadas com *khipu* (b e d). Enzima de restrição usadas: H= *Hind*III, R= *Rsa*II, E= *Eco*RI. BAC 147K17 é um BAC que apresenta sinais únicos após FISH e foi usado como controle.

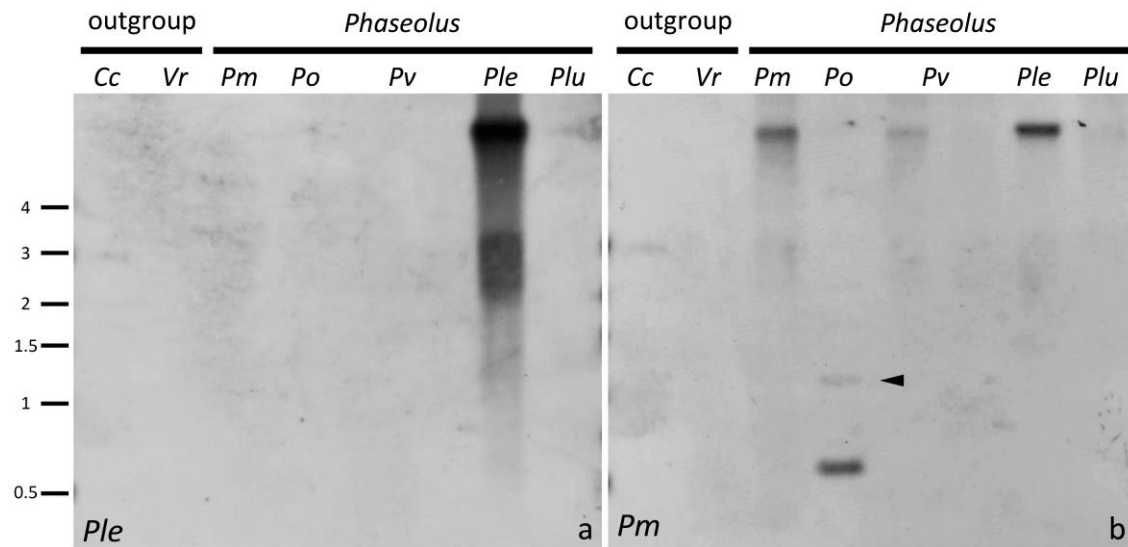


Fig. S2. Hibridização *Southern blot* de clones *hipu* de diferentes espécies de *Phaseolus* sob alta estringência. (a) Clone *khpuPle2* de *P. leptoschahyus* e (b) clone *khpuPm1* de *P. microcarpus*. Cc: *Cajanus cajan*; Vu: *Vigna radiata*; Pm: *P. microcarpus*; Po: *P. oligospermus*; Pv: *P. vulgaris*; Ple: *P. leptoschyus* e Plu: *P. lunatus*.

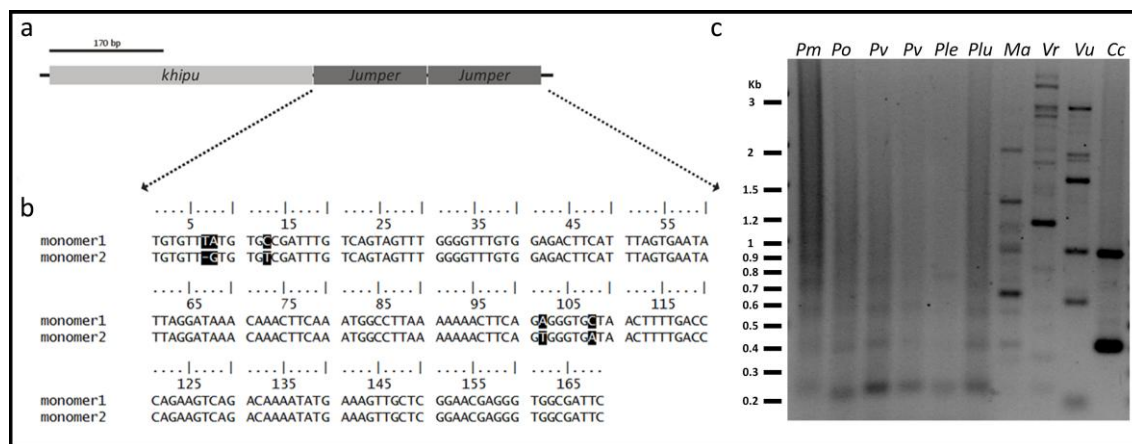
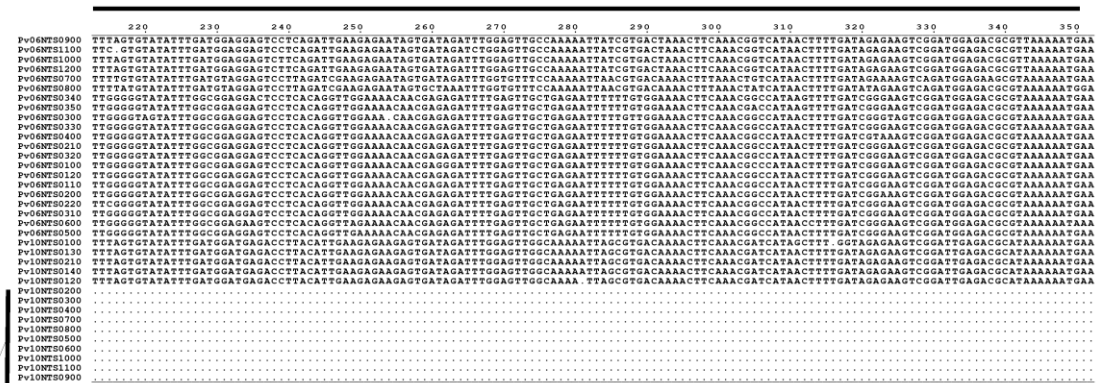


Fig. S3. (a) Representação esquemática do clone de *P. microcarpus* (*Pmkhipu2*) contendo *jumper*. Região em cinza claro mostra o fragmento de *khipu* (450 pb), enquanto a cinza escuro representa as duas unidades de *jumper*; (b) alinhamento dos dois monômeros de *jumper* identificados no presente clone. Nucleotídeos marcados em negrito representam polimorfismos ao longo da sequência; (c) Amplificação de *jumper* por PCR em diferentes espécies de *Phaseolus* (*Po*: *P. oligospermus*; *Pm*: *P. microcarpus*; *Pv*: *P. vulgaris* ‘BAT93’ e ‘JaloEPP’; *Ple*: *P. leptostachyus*; *Plu*: *P. lunatus*) e quatro espécies relacionadas (*Ma*: *Macroptilium atropurpureum*; *Vr*: *Vigna radiata*; *Vu*: *V. unguiculata* e *Cc*: *Cajanus cajan*). Barras na parte esquerda da figura representam os tamanhos dos fragmentos.



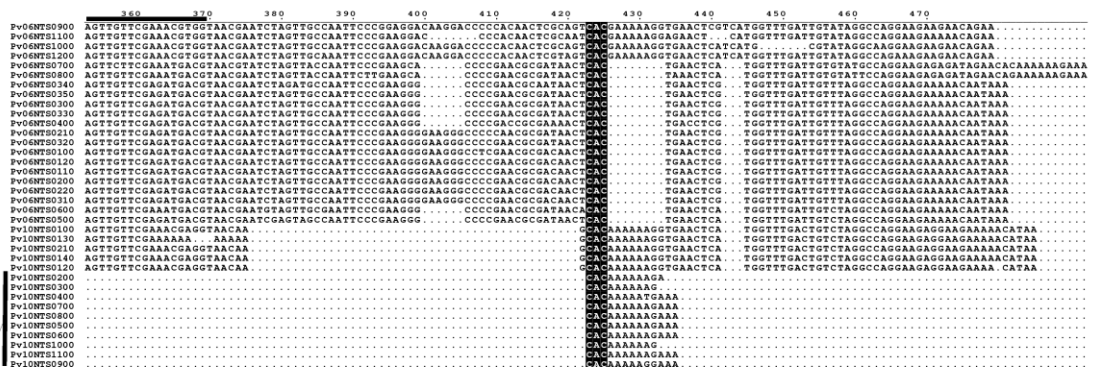
Continua...

Jumper

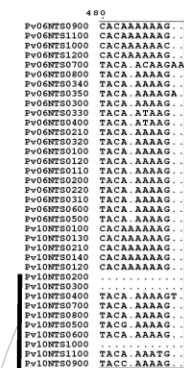


PvNTS without jumper

Jumper



PvNTS without jumper



PvNTS without jumper

Fig. S4. Alinhamento múltiplo de um conjunto de NTS extraídos do DNAr 5S de dados genômicos de *P. vulgaris*. Pv06: *Pv* cromossomo 6 e Pv10: *Pv* cromossomo 10. Linha horizontais na lateral esquerda marcam os NTS sem *jumper*.

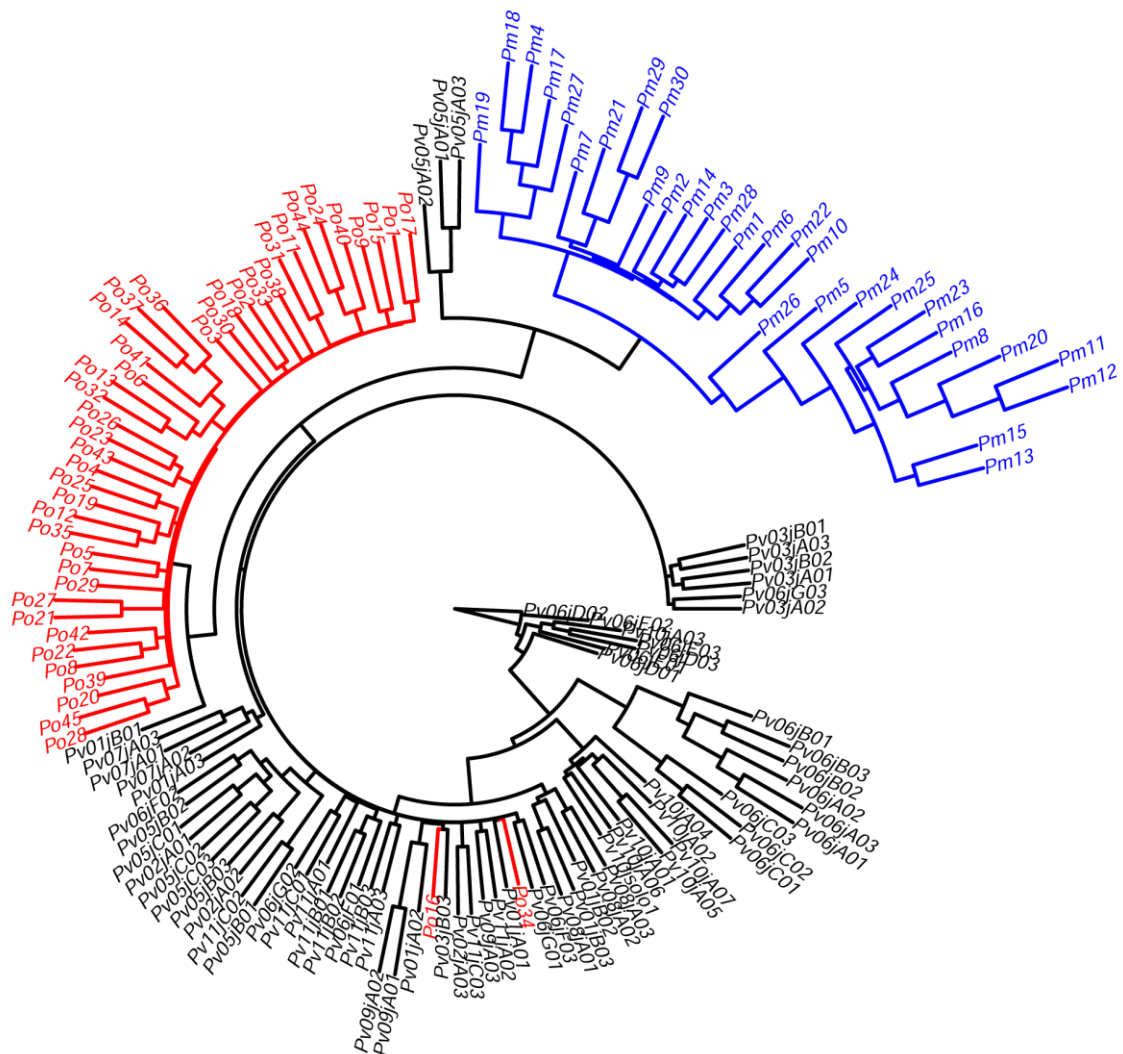


Fig. S5. Análise de máxima verossimilhança do satélite *jumper* de diferentes espécies de *Phaseolus*. *P. vulgaris* (Pv; preto), *P. oligospermus* (Po; vermelho) and *P. microcarpus* (Pm; azul).

5. Conclusões

5. Conclusões

- 1- *Khipu* foi confirmado como principal DNA satélite no genoma do feijão comum (*P. vulgaris*) e identificado também nos subtelômeros dos cromossomos das outras espécies de *Phaseolus* analisadas.
- 2- *Phaseolus lunatus* e *P. microcarpus*, entretanto, possuem um menor número de cópias de *khipu* em seus genomas.
- 3- Os diferentes padrões de hibridização observados por FISH para o *khipu mix* de *P. vulgaris* e os clones isolados das demais espécies (*P. leptostachyus*, *P. lunatus* e *P. microcarpus*) demonstram a presença de variabilidade intra- e interespecífica ao nível de sequência.
- 4- A organização genômica de *khipu* na forma de multímeros de 500 pb também foi observada para *P. leptostachyus*, *P. lunatus* e *P. microcarpus*, confirmando o tamanho desviante da unidade de repetição em comparação a maioria dos DNAs satélites.
- 5- Os subtelômeros em *P. microcarpus* apresentam uma complexidade de sequências maior em relação às outras espécies. Neste caso, *khipu* foi encontrado em íntima relação com outra sequência satélite, denominada *jumper*, em diferentes cromossomos.
- 6- A identificação de dois DNAs satélites nos subtelômeros de *P. microcarpus*, juntamente com a variação quantitativa encontrada para *khipu* na outras espécies do estudo, demonstram a ocorrência de amplificação diferencial dessas sequências em *Phaseolus* e suportam a hipótese de uma biblioteca de sequências (*Satellite DNA library*) como o principal mecanismo de evolução de DNAs satélites no gênero.

Normas da revista *Annals of Botany*

(http://www.oxfordjournals.org/our_journals/annbot/for_authors/)

INTRODUCTION

Scope of the journal

Annals of Botany is published for the Annals of Botany Company by Oxford University Press. Experimental, theoretical and applied papers on all aspects of plant science are welcome. The submitted manuscript or its essential content must not have been published previously or be under consideration for publication elsewhere. To merit publication in *Annals of Botany*, contributions should be substantial, written in clear English and combine originality of content with potential general interest. Submission of manuscripts that report small incremental advances or are of geographically local interest only is discouraged unless the implications of the findings are wide-reaching. Agronomic papers are expected to contain a substantial amount of basic plant biology. In general, a paper is unlikely to be accepted unless the referees and editors involved in its evaluation are enthusiastic about the science. The Covering Letter is an essential part of all submissions. It should include an ~60 word summary of the scientific strengths of the paper that the author(s) believe qualify it for consideration by *Annals of Botany*.

Language editing

If English is not your first language, you may wish to have your manuscript edited for language before submitting it. This is not a mandatory step, but may help to ensure that the academic content of your paper is fully understood by journal editors and reviewers. Language editing does not guarantee that your manuscript will be accepted for publication. There are many specialist language editing services available, including one offered by our publisher Oxford University Press, and you can find these easily by searching online. Authors are liable for all costs associated with such services.

Charges

Authors pay no fees or page charges unless electing for our Open Access scheme (see below for details). The corresponding author receives a free copy of the issue of the Journal in which their paper appears and a unique URL that gives access to a PDF (Portable Document Format) file of their article. The corresponding author is entitled to receive 25 printed offprints free of charge. These can be claimed using the Oxford Journals Authors Services site. Additional offprints can also be ordered using the Oxford Journals Authors Services site. Orders from the UK will be subject to the current UK VAT charge. For orders from elsewhere in the EU you or your institution should account for VAT by way of a reverse charge. Please provide us with your or your institution's VAT number. Colour photographs and graphics are also printed without charge where their use enhances scientific content or clarity.

Open Access

Annals of Botany authors have the option to publish their paper under the [Oxford Open](#) initiative; whereby, for a charge, their paper will be made freely available online immediately upon publication. *Annals of Botany* offers extremely competitive rates for Open Access publication, supported by its not-for-profit status aiming to promote botanical science. Corresponding authors/institutions can choose to pay a subsidized rate of 1000 GBP/1600 USD/1300 EUR, approximately 40% below the previous normal charge, and open access is further discounted or free for developing countries (click [here](#) for a list of qualifying countries). After your manuscript is accepted the corresponding author will be required to accept a mandatory licence to publish agreement. As part of the licensing process you will be asked to indicate whether or not you wish to pay for open access. If you do not select the open access option, your paper will be published with standard subscription-based access and you will not be charged.

Oxford Open articles are published under Creative Commons licences. Authors publishing in *Annals of Botany* can use the following Creative Commons licences for their articles:

- Creative Commons Attribution licence (CC-BY)

Please click [here](#) for more information about the Creative Commons licences.

If you choose the Open Access option you can pay the Open Access charges using our Author Services site. This will enable you to pay online with a credit/debit card, or request an invoice by email or post. Open access charges can be viewed [here](#) in detail.