



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE BIOCIÊNCIAS
DEPARTAMENTO DE BOTÂNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA VEGETAL

MARIANA ALEJANDRA BÁEZ

**EVOLUÇÃO DE SEQUÊNCIAS REPETITIVAS NO CARIÓTIPO BIMODAL
DE *Eleutherine bulbosa* (Miller) URBAN E ESPÉCIES DA SUBFAMÍLIA
IRIDOIDEAE (IRIDACEAE)**

Recife

2018

MARIANA ALEJANDRA BÁEZ

**EVOLUÇÃO DE SEQUÊNCIAS REPETITIVAS NO CARIÓTIPO BIMODAL
DE *Eleutherine bulbosa* (Miller) URBAN E ESPÉCIES DA SUBFAMÍLIA
IRIDOIDEAE (IRIDACEAE)**

Tese apresentada ao programa de Pós-Graduação em Biologia Vegetal da Universidade Federal de Pernambuco como requisito para a obtenção do título de Doutor em Biologia Vegetal na área de concentração Sistemática e Florística e na linha de pesquisa Citotaxonomia e Citogenética Vegetal.

Área de concentração: florística e sistemática.

Orientadora: Prof^ª. Dr^a Andrea Pedrosa-Harand

Recife

2018

Catálogo na Fonte:
Bibliotecário Bruno Márcio Gouveia, CRB-4/1788

Báez, Mariana Alejandra

Evolução de sequências repetitivas no cariótipo bimodal de *Eleutherine bulbosa* (Miller) urban e espécies da subfamília iridoideae (Iridaceae) / Mariana Alejandra Báez. – Recife, 2018.

96 f. : il.

Orientador: Prof^a. Dr^a. Andrea Pedrosa-Harand.

Tese (doutorado) – Universidade Federal de Pernambuco. Centro de Biociências. Programa de Pós-graduação Biologia Vegetal, Recife, 2018. Inclui referências e apêndices.

1. Genética vegetal 2. Botânica. 3. Angiosperma. I. Pedrosa-Harand, Andrea (Orientadora.) II. Título.

581.35

CDD (22.ed.)

UFPE/CB-2019-188

MARIANA ALEJANDRA BÁEZ

**EVOLUÇÃO DE SEQUÊNCIAS REPETITIVAS NO CARIÓTIPO BIMODAL
DE *Eleutherine bulbosa* (Miller) URBAN E ESPÉCIES DA SUBFAMÍLIA
IRIDOIDEAE (IRIDACEAE)**

Tese apresentada ao programa de Pós-Graduação em Biologia Vegetal da Universidade Federal de Pernambuco como requisito para a obtenção do título de Doutor em Biologia Vegetal na área de concentração Sistemática e Florística e na linha de pesquisa Citotaxonomia e Citogenética Vegetal.

Aprovada em: ____/____/____.

BANCA EXAMINADORA

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

Profa. Dra. Ana Cristina Brasileiro Vidal (Titular externo) Depto.de Genética, UFPE

Dr. Tiago Ribeiro Barros Dos Santos (Titular externo) Depto. de Botânica, UFPE

Dra. Fernanda de Oliveira Bustamante (Titular externo) Depto.de Genética, UFPE

Prof. Dr. Luiz Gustavo Rodrigues Souza (Titular Interno) Depto.de Botânica, UFPE

RESUMO

A maioria das angiospermas possuem cariótipos uniformes e simétricos mas algumas linhagens se caracterizam por possuírcariótipos assimétricos. Alguns cariótipos assimétricos pode estar marcado pela predominância de cromossomos com tamanhos altamente heterogêneos, conhecido como cariótipo bimodal. Estes cariótipos podem se originar por: rearranjos cromossômicos, aloploidia ou aumento em tamanho de um único conjunto de cromossomos. Na maioria dos grupos bimodais de plantas a organização cromossômica é pouco conhecida. A família Iridaceae, subfamília Iridoideae, se caracteriza por possuir espécies com cariótipos simétricos e assimétricos. Assim, o objetivo foi analisar a organização cromossômica das sequências repetitivas e a evolução dos cariótipos bimodais dentro da subfamília. Em *Eleutherinebulbosa*, ~40% do genoma é repetitivo, com abundância de LTR-retrotransposons e DNA satélite. A distribuição cromossômica destas mostrou que existe uma diferenciação cromossomos-específica, com uma alta acumulação nos cromossomos grandes e que se mantém na espécie irmã, *E. latifolia*, sugerindo diferenças estruturais e funcionais entre os conjuntos cromossômicos relacionadas à condição de cariótipo bimodal, no gênero. No entanto, *Alophiadrummondii* não mantém a distribuição cromossômica, sugerindo uma organização cromossômica dos cariótipos bimodais não conservada na subfamília. O cariótipo bimodal surgiu na tribo Tigridieae, tendo os cariótipos mais assimétricos na subtribo Tigridiniieae. Este evento não está relacionado com a evolução do tamanho do genoma. Apesar disto, pode ser sugerido uma evolução com amplificação e eliminação das sequências repetitivas cromossomo-específico, mas é possível que o mecanismo de amplificação/eliminação envolva as mesmas linhagens de sequências repetitivas mas que seria dependente do gênero.

Palavras-chave: Cariótipo Bimodal. Sequências repetitivas. Organização Cromossomo-Específica. Assimetria Cariotípica. Amplificação/Eliminação de Sequências.

ABSTRACT

Most angiosperms have uniform and symmetrical karyotypes but some lineages are characterized by asymmetric karyotypes. Some asymmetric karyotypes may be possess predominance of chromosomes with highly heterogeneous sizes, known as bimodal karyotype. These karyotypes can be originated by: chromosomal rearrangements, allopolyploidy or increase of size in a single set of chromosomes. In most plants bimodal groups the chromosome organization is poorly known. Iridaceae family, subfamily Iridoideae, is characterized to have species with symmetrical and asymmetrical karyotypes. Thus, the objective was to analyze the chromosomal organization of the repetitive sequences and the evolution of the bimodal karyotypes within the subfamily. In *Eleutherine bulbosa*, ~ 40% of the genome is repetitive, with abundance of LTR-retrotransposons and satellite DNA. Chromosomal distribution of these elements showed that there is a chromosome-specific differentiation, with a high accumulation in the large chromosomes and that is maintained in the sister species, *E. latifolia*, suggesting structural and functional differences between the chromosomal sets related to the condition of bimodal karyotype, in the genus. However, *Alophiadrummondii* does not maintain the same chromosomal distribution, suggesting a chromosomal organization of bimodal karyotypes not conserved within the subfamily. The bimodal karyotype appeared ones in the Tigridieae tribe, with the most asymmetrical karyotypes in the subtribe Tigridiniieae. This event is not co-related to the genome size evolution. Despite this, an evolution with amplification and elimination of the repetitive chromosome-specific sequences is suggested, but it is possible that the amplification/elimination mechanism involves the same lineages of repetitive sequences but it would be genus dependent.

Keywords: Bimodal Karyotype. Repetitive Sequences. Chromosome-Specific Organization. Karyotypic Asymmetry. Amplification / Elimination of Sequences.

SUMÁRIO

1	INTRODUÇÃO	8
2	FUNDAMENTAÇÃO TEÓRICA	11
2.1	EVOLUÇÃO CARIOTÍPICA.....	11
2.1.1	Cariótipos bimodais	11
2.1.2	Inversões cromossômicas.....	13
2.2	SEQUÊNCIAS REPETITIVAS.....	15
2.2.1	Sequências repetidas em tandem	16
2.2.1.1	DNA de Satélites	16
2.2.2	Sequências repetidas dispersas	19
2.2.2.1	Elementos transponíveis	19
2.3	CITOGENÔMICA COMPARATIVA	22
2.4	RECONSTRUÇÃO DE CARATERES ANCESTRAIS	23
2.5	FAMÍLIA <i>IRIDACEAE</i>	25
2.5.1	Relações filogenéticas na família <i>Iridaceae</i>	25
2.5.2	Citogenética e evolução cariotípica de <i>Iridaceae</i>	26
2.5.3	Gênero <i>Eleutherine</i> Herbert.....	29
3	EVOLUÇÃO DO CARIÓTIPO BIMODAL E ESTRUTURAÇÃO CROMOSSÔMICA NA SUBFAMÍLIA IRIDOIDEAE (<i>IRIDACEAE</i>)	31
4	CONCLUSÃO	63
	REFERÊNCIAS.....	64
	APÊNDICE A – ARTIGO PUBLICADO NA REVISTA FRONTIERS OF PLANT SCIENCE	70

1 INTRODUÇÃO

O cariótipo é o maior nível de organização estrutural e funcional do genoma nuclear. As características cariotípicas comumente usadas para as análises evolutivas são: número, tamanho e simetria, assim como posição e tipo de marcas cromossômicas (WEISS-SCHNEEWEISS e SCHNEEWEISS, 2013). A simetria cariotípica está determinada pelo tamanho dos cromossomos e/ou posição do centrômero. A maioria das angiospermas possuem cromossomos de tamanho similar, meta ou submetacêntricos e consequentemente com cariótipos uniformes e simétricos (STEBBINS, 1971; WEISS-SCHNEEWEISS e SCHNEEWEISS, 2013). No entanto, alguns grupos se caracterizam pela presença de cariótipos assimétricos. Este tipo de cariótipo é marcado pela predominância de cromossomos com centrômeros terminais/subterminais e/ou tamanhos cromossômicos altamente heterogêneos, este último conhecido como cariótipo bimodal (LEVITSKY, 1931, revisado por PERUZZI; EROĞLU, 2013; STEBBINS, 1971). A origem dos cariótipos bimodais é usualmente associada a três processos diferentes: - rearranjos cromossômicos envolvendo eventos de fusão-fissão (GEN-SHEN YIN et al., 2014; SCHUBERT; LYSACK, 2011), aloploidia (MCKAIN et al., 2012; SHIRAKAWA; NAGANO; HOSHI, 2012); e aumento em tamanho de um conjunto de cromossomos (DE LA HERRÁN et al., 2001). Foi postulado que cariótipos assimétricos seriam derivados e evolutivamente mais recentes que os cariótipos simétricos (MEDEIROS-NETO et al., 2017; STEBBINS, 1971). Cariótipos bimodais são comuns em vários grupos de animais, como em aves, répteis e anfíbios (NORONHA et al., 2016; SCHMID; STEINLEIN, 2017) e são bem conhecidos em monocotiledôneas como nos gêneros *Aloe* (FENTAW et al., 2013) e *Agave* (PALOMINO et al., 2012), entre outros. Na maioria dos grupos bimodais de plantas a organização cromossômica é pouco conhecida. No entanto, foi observado nos cromossomos grandes de várias espécies a presença de sequências altamente repetitivas, como regiões heterocromáticas bandas C positivas (*Cephalanthera*, *Epipactis* e *Limodorum* (Orchidaceae), (D'EMERICO et al., 1999); ou DNA satélites (*Ornithogalum longibracteatum*, (PEDROSA et al., 2001); *Muscari comosum* (DE LA HERRÁN et al., 2001, Hyacinthaceae). Em galinha (*Gallus gallus*), por exemplo, os microcromossomos constituem 30% do tamanho total do genoma, contêm mais genes que os macrocromossomos, e estes últimos estariam associados a sequências repetitivas em abundância (SMITH et al., 2000).

A família Iridaceae, Asparagales, com ~70 gêneros, se caracteriza por possuir espécies com cariótipos simétricos e assimétricos (GOLDBLATT et al., 2008). Na América do Sul estão representadas três das cinco tribos da subfamília Iridoideae: Sisyrinchieae, Trigridieae, Trimezieae (Goldblatt e Manning, 2008; Goldblatt et al., 2008). Para Iridoideae, o número cromossômico básico ancestral é $x = 8$, com números básicos secundários $x = 5, 6, 7, 9$ e 14 , consequência de dispoloidias ascendentes e descendentes, assim como, dispoloidia descendente e mais algum evento de poliploidia. A tribo Trigridieae apresenta espécies com cariótipos bimodais, em sua maioria com dois pares de cromossomos grandes e cinco pares de cromossomos pequenos, no entanto existem espécies que apresentam cariótipos simétricos (Alves et al., 2011; Moraes et al., 2015). Dentro desta tribo, está o gênero *Eleutherine* com duas espécies *E. bulbosa* e *E. latifolia*, apresentando cariótipos bimodais (GOLDBLATT; SNOW, 1991). *Eleutherine bulbosa*, apresenta número cromossômico $2n = 12$, com o par cromossômico I sendo 3 a 4 vezes maior que os demais pares. Este par maior apresenta uma heterozigose permanente devido a uma inversão pericêntrica assimétrica que inclui cerca de 70% do cromossomo, resultando em um cromossomo acrocêntrico e outro metacêntrico. Este par apresenta bandas DAPI⁺ e CMA⁺ e os sítios de DNAr 35S associados às bandas CMA⁺, dentro da região invertida, e sítios de DNAr 5S duplicados, fora da região de inversão (FEITOZA; GUERRA, 2011). Análises da distribuição de modificações de histonas têm mostrado que o par maior se diferencia do resto do complemento. Marcas que evidenciam atividade transcricional, marcaram fortemente cromossomos pequenos e o par maior marca unicamente na região terminal de um dos braços. As marcas associadas a regiões heterocromáticas, ricas em DNA repetitivo, mostram uma marcação mais intensa no par maior (FEITOZA; GUERRA, 2011). Dentro da mesma subtribo Tigridiineae, clado B, gêneros relacionados a *Eleutherine* também se caracterizam por possuir cariótipos bimodais. Estes gêneros possuem números cromossômicos diferentes, como *Gelasine* ($n = 7$), possuindo também espécies com cariótipos simétricos como *Gelasine coerulea*, e espécies como o gênero *Alophia* ($n = 14$), surgido por poliploidia (MORAES et al., 2015). No entanto, não é conhecida a evolução da assimetria cariotípica dentro da tribo e se existe uma estruturação cromossômica nestas espécies e sua relação com a condição bimodal.

Dentro deste contexto, o objetivo do trabalho foi entender a dinâmica de evolução do cariótipo bimodal. Para isso, foi analisada a fração repetitiva de espécie *Eleutherine bulbosa*, subfamília Iridoideae, com cariótipo bimodal para entender a distribuição cromossômica e possíveis mecanismo que levaram a essa distribuição. Assim também, como as sequências repetitivas estariam distribuídas em outras espécies com cariótipos bimodais relacionadas e qual seria a origem da assimetria cariotípica dentro da subfamília Iridoideae e entender os possíveis mecanismos que levaram ao surgimento do cariótipo bimodal dentro do grupo.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 EVOLUÇÃO CARIOTÍPICA

O cariótipo é o maior nível de organização estrutural e funcional do genoma nuclear. Variação nos cariótipos, tanto numéricas quanto estruturais, estão associadas a mudanças evolutivas. Os cariótipos podem ser alterados por meio de mudanças no número, estrutura e composição dos seus cromossomos (SCHUBERT, 2007). Alterações no nível de ploidia são classificadas em dois tipos: aloploidia, consequência da hibridização interespecífica, ou autopoliploidia dentro de uma mesma espécie. Ambas poliploidias podem ser consequência de erros na meiose (envolvendo gametas não reduzidos) ou mitose, que multiplicam o complemento cromossômico inteiro. Estas alterações numéricas podem ocasionalmente fornecer vantagens evolutivas aos indivíduos nos quais acontecem, sendo um dos mecanismos mais importantes para a especiação em plantas (HESLOP-HARRISON; SCHWARZACHER, 2011).

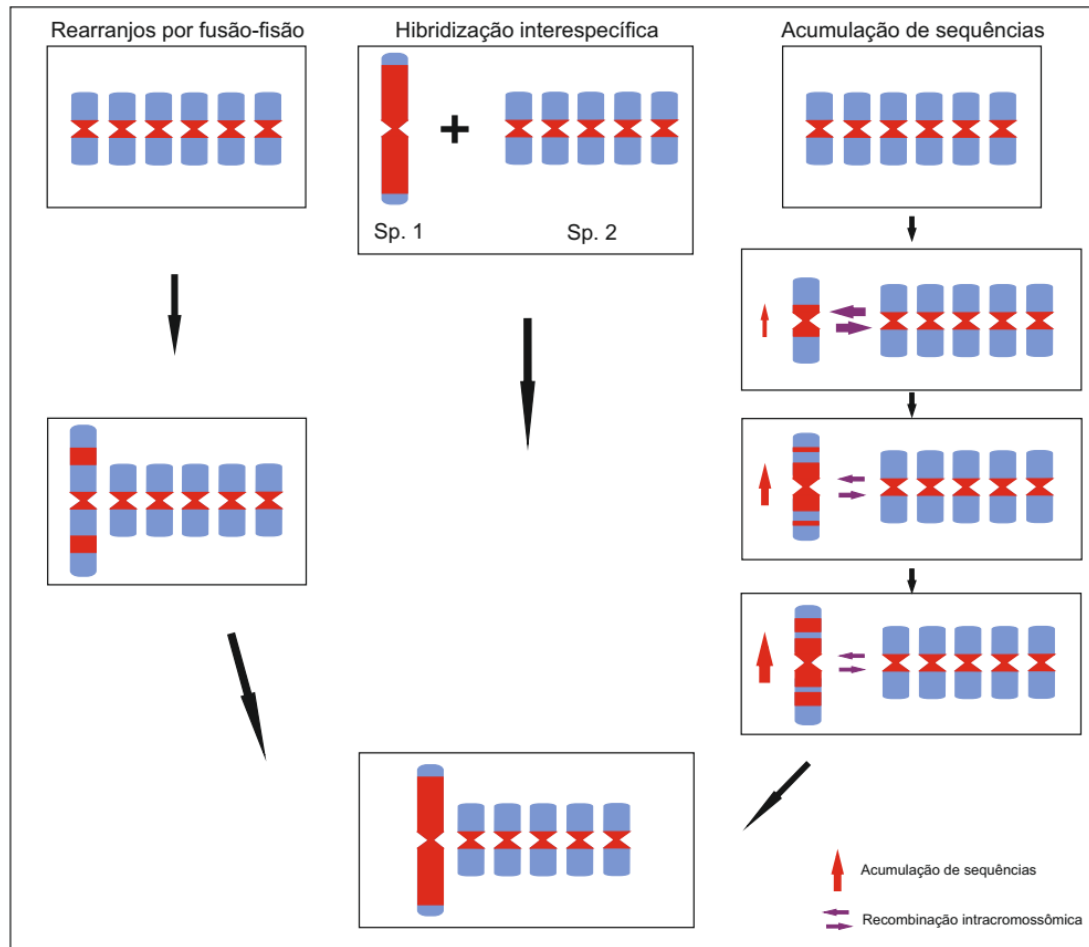
Outro tipo de mudança cromossômica é ao nível da composição de DNA dos cromossomos, que pode ser alterada evolutivamente por rearranjos primários e secundários. Rearranjos cromossômicos primários são o resultado da recombinação ilegítima durante o mecanismo de reparo das quebras de dupla fita do DNA, como as inserções, deleções, duplicações, inversões (para- e pericêntricas) e translocações recíprocas (intra- ou intercromossômicas). Rearranjos cromossômicos secundários podem surgir em organismos duplamente heterozigotos para dois rearranjos primários em um mesmo cromossomo. Recombinação homóloga entre as regiões rearranjadas do cromossomo resulta em gametas com novos cariótipos e gametas complementares exibindo o complemento cromossômico tipo selvagem restabelecido (SCHUBERT; LYSACK, 2011).

2.1.1 Cariótipos bimodais

A simetria cariotípica está determinada pelo tamanho dos cromossomos e/ou posição do centrômero. Mudanças no tamanho cromossômico podem afetar e ser afetadas pelo tamanho do genoma da espécie. A maioria das angiospermas possuem

cromossomos de tamanho pequeno, similar, meta ou submetacêntricos e consequentemente com cariótipos uniformes e simétricos(STEBBINS, 1971, WEISS-SCHNNEWEISS et al., 2013).Algumas linhagens se caracterizam pela presença de cariótipos assimétricos. Este tipo de cariótipo é marcado pela predominância de cromossomos com centrômeros terminais/subterminais (assimetria intracromossômica) e/ou tamanhos cromossômicos heterogêneos (assimetria intercromossômica). Um cariótipo com grande assimetria no tamanho cromossômico, constituído por dois conjuntos de tamanhos distintos, é conhecido como bimodal(PERUZZI; EROĞLU, 2013; STEBBINS, 1971). Os cariótipos bimodais possuem duas classes de tamanhos cromossômicos distintos(MCKAIN et al., 2012). Existem três hipóteses que propõem como os cariótipos bimodais surgem: a primeira propõe rearranjos cromossômicos envolvendo eventos de fusão-fissão, onde cromossomos maiores são produtos de fusão de cromossomos pequenos ou cromossomo pequenos são resultados da fissão dos cromossomos maiores(MCKAIN et al., 2012; SCHUBERT; LYSAK, 2011). O segundo mecanismo para o surgimento de cariótipos bimodais seria por aloploidia, envolvendo espécies parentais com diferentes tamanhos cromossômicos(MCKAIN et al., 2012). Por último, o aumento em tamanho de um *subset* de cromossomos causada pela acumulação diferencial de sequências repetitivas(DE LA HERRÁN et al., 2001).Em animais, os cariótipos bimodais são muito estudados. A galinha (*Gallus gallus domesticus*), por exemplo, possui micro- e macrocromossomos.Os microcromossomos constituem 30% do tamanho total do genoma, contêm duas vezes mais genes que os macrocromossomos e são associados a uma maior atividade transcricional. Foi sugerido que isto poderia estar associado a uma maior taxa de recombinação encontrada nos microcromossomos(SMITH et al., 2000) (Fig. 1).

Figura 1. Modelos dos diferentes possíveis mecanismos que originam cariótipos bimodais.



2.1.2 Inversões cromossômicas

As inversões cromossômicas são alguns dos principais mecanismos envolvidos na reorganização dos genomas (KIRKPATRICK, 2010). Uma inversão cromossômica é uma mudança estrutural, onde a consequência de duas quebras resulta na reinserção do segmento em orientação oposta dentro do mesmo cromossomo. Este rearranjo pode alterar o tamanho dos braços cromossômicos, padrão de bandas heterocromáticas e simetria do cromossomo. As inversões cromossômicas são classificadas em dois tipos, paracêntricas, não envolvendo o centrômero na região invertida, e pericêntricas, envolvendo o centrômero na região invertida. As inversões podem ser encontradas em estado homozigoto, quando os dois cromossomos de um par apresentam a inversão, ou em estado heterozigoto, quando só um dos cromossomos apresenta uma inversão cromossômica. Inversões em estado

heterozigoto, tanto pericêntricas como paracêntricas podem ser identificadas no indivíduo afetado pela formação de uma alça, denominada alça ou *loop* de inversão, possibilitando o pareamento em paquíteno durante a divisão meiótica. Alternativamente pode ocorrer ou não pareamento (sinapse) na área invertida, sem a formação de alça ou quiasma (recombinação) nesse segmento cromossômico. Quando a alça se forma e a recombinação acontece no interior da alça, as cromátides resultantes são anômalas, com deleções ou duplicações. Desse modo, inversões podem causar esterilidade pela produção de gametas não balanceados(LEVIN, 2002).

A principal consequência das inversões é a supressão da recombinação na região envolvida, podendo produzir isolamento genético entre indivíduos, populações e espécies. Assim, as inversões podem resultar em eventos de especiação, criando barreiras pós-zigóticas(KIRKPATRICK, 2006). Elas podem se fixar na população caso contenham diferenças adaptativas pré-existentes em alguns loci que conferem aos indivíduos portadores de vantagens adaptativas ao ambiente no qual se desenvolvem. Por outro lado, as regiões invertidas podem se converter em *hotspot* para a acumulação de diferenças selecionadas positivamente e genes que conferem incompatibilidade entre as espécies (KIRKPATRICK, 2010). As inversões são comumente encontradas como diferenças fixadas entre as espécies ou como polimorfismos dentro das espécies em muitos grupos de animais e plantas. Em *Mimulus guttatus*DC., (Scrophulariaceae), polimorfismos para inversões cromossômicas são distribuídos em diferentes regiões geográficas. Estes polimorfismos foram associados à divergência adaptativa e isolamento reprodutivo, assim como a características morfológicas, como tempo de floração, produção de flores e padrões de distribuição geográfica dependentes de condições ambientais(LOWRY; WILLIS, 2010). Em várias espécies do gênero *Drosophila*, a frequência das inversões podem variar clinalmente dependendo da latitude, isto é, variações graduais em uma mesma espécie por influência das condições ambientais latitudinais, sugerindo que os polimorfismos de inversões estão associados à adaptação climática. Estas clinas para polimorfismos de inversão podem se repetir em diferentes continentes, podendo ser mantidas ou não ao longo do tempo associadas a genes adaptados a um ambiente particular(ANDERSON et al., 2005).

As inversões também têm desempenhado um papel chave na evolução de cromossomos sexuais, uma vez que a seleção favorece a redução da recombinação entre genes sob seleção sexo-antagônica e entre loci que determinam o sexo (KIRKPATRICK, 2010). Em *Carica papaya* L., o mamoeiro, o sistema de determinação do sexo (X/Y) é controlado pela recente evolução de um par de cromossomos X/Y, no qual o cromossomo Y apresenta duas variantes. Esses dois tipos de cromossomos Y (Y e Y^h), são levemente diferentes com uma região específica para macho, MSY e com uma para o hermafrodita, HSY. A região HSY está associada a duas grandes inversões, ambas seguidas de vários rearranjos intracromossômicos (WANG et al., 2012). Ambos os cromossomos Y (Y e Y^h) acumularam mais sequências repetitivas de DNA e sequências de DNA de organelas em relação ao cromossomo X e ao genoma total (VANBUREN; MING, 2013; WANG et al., 2012). Todos os eventos evolutivos observados nos cromossomos sexuais de mamão foram atribuídos à falta de recombinação entre os cromossomos X e Y ou Y^h, o que teria levado à degeneração dos cromossomos Y (VANBUREN; MING, 2013; WANG et al., 2012).

2.2 SEQUÊNCIAS REPETITIVAS

O DNA nuclear de plantas é constituído por sequências codificadoras de cópia única ou pequeno número de cópias, *íntrons*, promotores e sequências de DNA reguladoras, assim como também por várias classes de sequências de DNA repetitivas. As sequências de DNA repetitivas formam uma grande parte dos genomas de plantas, estando constituídas por centenas a milhares de unidades dispostas em *tandem* ou dispersas no genoma e sujeitas a rápidas mudanças na sequência ou no número de cópias (HESLOP-HARRISON; SCHWARZACHER, 2011; RENNY-BYFIELD et al., 2011). Existem sequências repetitivas características nas regiões cromossômicas centroméricas e teloméricas e os DNA ribossomais (DNAr), que são os genes para RNAr e espaçadores intergênicos, nos loci 5S e 35S. Sequências repetidas em *tandem* ou DNA satélite consistem em motivos repetidos com várias cópias em uma ou mais localizações genômicas. Outra classe de sequências de DNA repetitivo são os elementos transponíveis, que incluem sequências que codificam

enzimas para sua própria replicação e integração no DNA nuclear. Estes elementos podem ser transposons de DNA ou retrotransposons (os mais abundantes e comuns) sendo os últimos mais relacionados à mudanças no conteúdo de DNA (HESLOP-HARRISON; SCHWARZACHER, 2011). A maioria destas sequências está envolvida em fenômenos de evolução genômica, causando diferenças nos tamanhos genômicos por inserções, deleções e amplificações, por exemplo (RENNY-BYFIELD et al., 2011; TAUTZ, 1993). As sequências mais relacionadas à evolução cariotípica em plantas são os elementos transponíveis e os DNA ribossomais (KOVARIK et al., 2008; TENAILLON; HOLLISTER; GAUT, 2010). Os elementos transponíveis têm a capacidade de se movimentar dentro do genoma causando expansão do mesmo, alterando a função dos genes ou atuando como matéria-prima para novos genes (TENAILLON; HOLLISTER; GAUT, 2010). Estes elementos promovem rearranjos cromossômicos mais facilmente que qualquer outro mecanismo. Em algumas espécies se encontram relacionados aos pontos de quebras de inversões cromossômicas, podendo estar envolvidos na origem ou manutenção do rearranjo na população (GRAY, 2000; HOFFMANN; RIESEBERG, 2008).

2.2.1 Sequências repetidas em tandem

2.2.1.1 DNA de Satélites

Sequências satélites (satDNA) são sequências de DNA repetidas, geralmente não codificadoras, abundantes, dispostas de forma adjacente em uma organização conhecida como *head-to-tail*, ou seja, estão orientadas na mesma direção (BISCOTTI; OLMO; HESLOP-HARRISON, 2015; PLOHL; MEŠTROVIĆ; MRAVINAC, 2012). Originalmente foram identificados por sua densidade diferencial, comparada com a fração principal do DNA genômico, na centrifugação com cloreto de cério (COONEY; MATTHEWS; BRADBURY, 1984). Também, foram definidos como os principais componentes das regiões cromossômicas heterocromáticas, pobres em genes e transcricionalmente suprimidas, que permanecem condensadas a longo do ciclo celular. Diferentes sequências satélites são encontradas nas regiões pericentroméricas, subteloméricas e intersticiais dos cromossomos (HESLOP-HARRISON; SCHWARZACHER, 2011; PLOHL; MEŠTROVIĆ; MRAVINAC,

2012). Assim, as regiões heterocromáticas podem estar constituídas por uma única família de satDNA, como foi observado em diferentes espécies dos gêneros *Citrus*, *Fortunella* e *Poncirus* para o satDNA CsSat, principal componente das bandas heterocromáticas CMA+ (BARROS E SILVA et al., 2010). Por outro lado, essas regiões podem estar constituídas por diferentes famílias, como observado em espécies do gênero *Phaseolus* com seu principal satDNA *kipu* das regiões subteloméricas e o satDNA *jumper*, presente tanto nas regiões subteloméricas e nas regiões pericentroméricas em algumas espécies (RIBEIRO et al., 2017a).

Os satDNA podem variar bastante quanto as suas características mesmo dentro da mesma espécie. Os monômeros destas diferem em comprimento, sequência de nucleotídeos, complexidade das sequências, unidade de repetição e abundância. Diversos mecanismos têm sido sugeridos para explicar sua evolução dessas sequências (PLOHL; MEŠTROVIĆ; MRVINAC, 2012). Os satDNA podem variar em tamanho por meio de amplificações/desamplificações via *crossing-over* desigual, replicação *slippage* e replicação saltatória ou em círculo rolante. O modelo de *crossing over* desigual sugere uma troca de segmentos entre homólogos ou entre cromátides irmãs de forma desequilibrada após pareamento errôneo, mudando quantitativamente a sequência (Charlesworth, 1994). De forma similar, a replicação *slippage*, que é o deslizamento da polimerase III durante a replicação de DNA, resulta em variações em tamanho nos satDNA (VIGUERA; CANCEILL; EHRLICH, 2001). Por outro lado, as sequências satélites podem ser amplificadas e se movimentar por meio de moléculas circulares extracromossômicas, ou seja, replicação em círculo rolante. Esta molécula circular, com parte das unidades de um monômero, são geradas por recombinação intracromossômica através de um loop. Quando se reintegrando no genoma por meio de recombinação homóloga com o loco original, expandem os *repeats* dessa região, mas quando há recombinação ectópica o processo formaria um novo loco da sequência (COHEN; SEGAL, 2009).

Cada satDNA, caracterizado por um monômero, forma um segmento genômico longo constituído por milhares de monômeros. Esta homogeneidade de sequências nos satDNAs é resultado da evolução não-independente das unidades de repetição. As mutações ao invés de acumular-se num único monômero, se espalham dentre as unidades de repetição do satDNA ou são eliminadas. Este fenômeno

conhecido como evolução em concerto é uma consequência de dois processos: homogeneização de sequência e fixação na população. As mutações são homogeneizadas por transferência de sequência não recíproca, por mecanismos como recombinação desigual, conversão gênica ou transposição. Após isto, a fixação na população é resultado da seleção aleatória do material genético na meiose e segregação cromossômica (FELINER; ROSSELLÓ, 2012; PLOHL; MEŠTROVIĆ; MRAVINAC, 2012). Supõe-se que a evolução em concerto homogeneiza as cópias por conversão gênica, processo de copiar e colar uma cópia genômica para outro locus sem importar se são ortólogos ou não, e/ou por recombinação desigual entre unidades homólogas (FELINER; ROSSELLÓ, 2012). O mecanismo de mudanças das unidades de repetição sustenta o processo conhecido como “molecular drive”. Neste processo as novas variantes de satDNA se espalham tanto dentro da mesma família de satDNA (homogeneização) como entre espécies com reprodução sexual (fixação), com o passar das gerações. Este processo originaria os padrões das famílias multigênicas dos DNAs repetitivos observados, homogeneidade dentro das espécies e diversidade entre as espécies, assim, espécies distintas apresentam famílias de satDNA diferentes (DAKAL et al., 2017). No entanto, a hipótese de “biblioteca de satélites de DNA” predisse que espécie relacionada compartilham um set ancestral de diferentes famílias de satDNA conservados que seriam amplificados diferencialmente em cada espécie devido a mecanismos estocásticos de evolução em concerto (RUIZ-RUANO et al., 2016). Isto foi observado em diferentes gêneros, por exemplo em *Rhynchospora* e *Phaseolus*, nos quais diferentes famílias de satDNA são compartilhadas pela maioria das espécies analisadas (RIBEIRO et al., 2017a, 2017b). Foi sugerido o termo “Satelitoma” para a coleção total de diferentes famílias de satDNA dentro de um genoma que também são compartilhados intraespecificamente, dentro de citótipos e populações distintas, e em espécies relacionadas (PALACIOS-GIMENEZ et al., 2017; RUIZ-RUANO et al., 2016).

Para assegurar a evolução em concerto, as taxas de recombinação precisam ser altas quando comparadas com as taxas de mutação. No caso em que a taxa de mutação exceda as taxas de recombinação, se esperaria variações intragenômicas significativa em diferentes regiões do genoma, sem restrições seletivas. Além disso, recombinação desigual entre cadeias de um homólogo (UHR) e recombinação

ilegítima (IR) foram sugeridos como poderosos mecanismos de remoção de sequências de DNA (TENAILLON; HOLLISTER; GAUT, 2010). Sendo assim, regiões cromossômicas que não sofrem recombinação, tais como regiões invertidas, tenderiam a acumular diferentes tipos de sequências repetitivas que evoluiriam diferencialmente do resto do genoma.

Em alguns casos, a homogeneização pode ser suprimida, levando a uma evolução que não é em concerto, onde as mutações se acumulam sem possibilidade de espalhamento ou eliminação. Por exemplo, a família de satDNA RET76 é altamente variável (não homogeneizada) e não fixada dentre as diferentes populações do grupo de cupins *Reticulitermes*, provavelmente como consequência do número limitado de reprodutores (ANDREA; MARINI; MANTOVANI, 2006). Regiões cromossômicas sem recombinação dentro de um mesmo genoma também podem mostrar diferenças na homogeneização. Em galinha, várias sequências centroméricas repetidas em *tandem* mostram uma distribuição cromossomo-específica, sendo características de macrocromossomos ou de microcromossomos, apresentando íntima relação entre domínios heterocromáticos diferentes em territórios cromossômicos específicos no núcleo interfásico (MASLOVA et al., 2015).

2.2.2 Sequências repetidas dispersas

2.2.2.1 Elementos transponíveis

As sequências móveis são capazes de auto inserir-se em novos locais dentro do genoma, geralmente com nenhum benefício aparente. São amplamente classificadas em dois grupos, de acordo com seu mecanismo de transposição. A classe I são os retroelementos, propagados via um intermediário de RNA, que é posteriormente transcrito reversamente ao DNA complementar, mecanismo conhecido como “copiar e colar”. Os mais abundantes em plantas são os retroelementos do tipo LTR, os quais possuem repetições terminais longas (*long terminal repeat*), mas também são representados pelos retrovírus endógenos (ERVs) e pelos retrotransposons não LTR, como os elementos LINE (*long interspersed nuclear elements*) (BISCOTTI; OLMO; HESLOP-HARRISON, 2015; WARREN et al., 2015). Os retrotransposons LTR são classificados com base na similaridade do

domínio da transcriptase reversa e os domínios das subunidades do genes pol. Um grupo pertencem aos Ty1/copia, onde as subunidades são a protease (PR), integrase (IN), transcriptase (RT) reversa e RNase H (RH), nessa ordem dentro de cada unidade de repetição. O outro grupo são os elementos Ty3/Gypsy onde as subunidades de cada unidade de repetição estão na ordem PR, RT, RH, e IN(ZHANG et al., 2014).

Na classe II estão incluídos os transposons de DNA que se caracterizam por movimentar-se dentro do genoma através de um mecanismo conhecido como “cortar e colar”, codificando uma enzima transposase e utilizando a maquinaria de replicação do hospedeiro, sem a presença de um intermediário de RNA. Alguns destes elementos, como os Polintons, se transpõem por meio da incisão de uma molécula de DNA de fita simples, que serve como base para a síntese do complementar; então a fita dupla se insere novamente no genoma. Por outro lado, os Helitrons se transpõem pelo mecanismo chamado “círculo rolante”(KAPITONOV; JURKA, 2001, 2006). Além destes tipos de elementos móveis chamados de elementos “autônomos”, por ter sequências codificadoras funcionais para as proteínas necessárias para sua transposição, existem os elementos “não autônomos” que não possuem sequências codificadoras para sua própria propagação. Estes elementos requerem a presença de um elemento autônomo no genoma hospedeiro que providencie *intrans* as proteínas necessárias para sua transposição. Dentro dos elementos da classe I os mais comuns são os SINE (*short interspersed nuclear elements*) e dentro dos elementos da classe II os mais conhecidos são os transposons de DNA do tipo MITE (WARREN et al., 2015).

O conteúdo de elementos transponíveis nos genomas é variável, influenciando o tamanho do genomas, os cariótipos e estabelecendo diferenças entre genomas de espécies relacionadas. A distribuição destes elementos nos genomas depende da especificidade de inserção e as taxas de remoção em diferentes regiões genômicas, geralmente via recombinação desigual e conversão gênica. Além disso, o silenciamento epigenético poderia resultar na inativação permanente dos elementos transponíveis pelo aumento da taxa de mutações pontuais neles. Este silenciamento epigenético também levaria a compactação na forma de heterocromatina diminuindo a expressão dos elementos transponíveis e também, inibindo os mecanismo de eliminação por recombinação (BENNETZEN; WANG, 2014). Assim, estes elementos

são os principais contribuintes na variação no tamanho dos genomas nas angiospermas, além da poliploidia (BENNETZEN; WANG, 2014). Genomas grandes seriam derivados de amplificações massivas de um pequeno número de famílias de LTR-retrotransposons, enquanto que genomas pequenos teriam tipo de estruturas similares desses elementos nas regiões heterocromáticas e as regiões gênicas estariam livres deles. Os efeitos dos elementos transponíveis no tamanho dos genomas pode ser altamente dinâmico e não previsível (BENNETZEN; WANG, 2014). O estudo comparativo de duas espécies do gênero *Zea* mostrou que ambas possuíam um aumento em algumas famílias de elementos enquanto que *Z. luxurians* apresentou aumento em numerosas famílias de elementos, resultando em um genoma 2x maior que o de *Z. mays*, em <2 milhões de anos (ESTEP; DEBARRY; BENNETZEN, 2013). Comparações do papel da expansão/contração do genoma entre espécies diploides e poliploides do gênero *Brachiaria*, através de dados de retrotransposons no transcriptoma, sugerem que a amplificação de diferentes retroelementos teve um impacto no comportamento e divergência no genoma que levaram à especiação (SANTOS et al., 2015).

Os elementos transponíveis podem afetar a estrutura dos genomas induzindo diferentes rearranjos cromossômicos através de inserções, recombinação ectópica, estando envolvidos nos pontos de quebra cromossômica (HU et al., 2011; WARREN et al., 2015). O transposição destes elementos pode alterar a estrutura ou expressão gênica inserindo-se no genoma dentro ou perto dos éxons, íntrons ou regiões reguladoras. Por outro lado, através de recombinação homóloga ou ectópica ilegítima, poderiam modificar ou deletar genes. Da mesma forma, podem induzir diferentes rearranjos como deleções, duplicações, inversões e translocações, afetando fortemente ao hospedeiro no nível individual bem como no nível populacional ou de espécie. Desta forma, em grande escala os rearranjos mediados por elementos transponíveis podem contribuir para a divergência das linhagens e na especiação (BENNETZEN; WANG, 2014; WARREN et al., 2015). Grandes rearranjos cromossômicos como inversões e deleções foram observados em cromossomos sexuais tanto de plantas como de animais. Nesse caso, os elementos transponíveis estariam participando na diferenciação dos cromossomos sexuais dos autossomos por meio da supressão da recombinação e através da expansão/contração dos

elementos transponíveis cromossomos sexuais se tornariam, em alguns casos, heteromórficos. Dentro destes cromossomos, as regiões com recombinação suprimida mostram diferente intensidade de espalhamento dos elementos quando comparada com regiões recombinantes do genoma (CHALOPIN et al., 2015; HOBZA et al., 2015). De forma similar, os elementos transponíveis contribuem na evolução de outros tipos de cromossomos especiais como os cromossomos B, supernumerários. Em *Secale cereale* L., por exemplo, contendo cromossomos Bs que não apresentam recombinação com o set de cromossomos A (autossomos), foi encontrado vários elementos transponíveis compartilhados entre ambos sets cromossômicos, indicando origem dos cromossomos Bs a partir dos cromossomos As. No entanto, ambos os tipos cromossômicos contêm elementos característicos, mostrando que sets cromossômicos sem recombinação podem apresentar ampliações diferenciadas de elementos repetitivos dentro de uma mesma espécie, seguindo vias evolutivas independentes (KLEMME et al., 2013).

2.3 CITOGENÔMICA COMPARATIVA

Uma parte importante das pesquisas genômicas se foca em entender a estrutura do genoma. O genoma das plantas é composto basicamente por sequências de DNA repetitivas de vários tipos distintos, alto número de cópias e grande variabilidade, o que os torna difíceis de analisar. Com o avanço de tecnologias de sequenciamento de nova geração (NGS), a realização de sequenciamento genômico de baixa cobertura para a identificação de sequências repetitivas ficou cada vez mais acessível (WEISS-SCHNEEWEISS et al., 2015). Dessa forma, um novo campo se abriu para o desenvolvimento de metodologias e programas que facilitem a análise desse montante de sequências geradas. Dentre essas, um programa que vem revelando grande potencial e uso na identificação de sequências repetitivas é o RepeatExplorer (Novak et al., 2010; Novak et al., 2013). Esse programa pode ser facilmente utilizado através de uma plataforma *online* (repeatexplorer.org) ou até mesmo por linha de comando em Linux oferecendo uma caracterização e identificação das principais sequências repetitivas presentes no genoma. Diversos trabalhos recentes demonstram a versatilidade e utilidade desse programa na identificação e

caracterização de sequências repetitivas com diferentes interesses (Heckmann et al., 2013; Klemme et al., 2013; Zhang et al., 2014, Marques et al., 2015, Ribeiro et al. 2016). Assim, essa ferramenta associadaa técnicas citogenéticas moleculares, principalmente a hibridização *in situ* fluorescente (FISH), fornece uma abordagem mais rápida e mais informativa chamada “citogenômica” (LEE et al., 2017; MACAS; NEUMANN; NAVRÁTILOVÁ, 2007; WEISS-SCHNEEWEISS et al., 2015).

No contexto genômico, o uso da FISH especialmente no mapeamento cromossômico de elementos repetitivos é particularmente importante na genômica estrutural e evolutiva porque permite entender a dinâmica dos diferentes elementos repetitivos assim como também o seu impacto filogenético (LEE et al., 2017; WEISS-SCHNEEWEISS et al., 2015). No gênero *Panax*, tanto espécies diploides como poliploides, possuem uma abundância, 39-52%, de elementos retrotransposons do tipo LTR, enquanto que uma espécie do gênero relacionado, *Aralia elata*, possui 17% destes elementos. Assim, pela comparação da distribuição cromossômica dos diferentes elementos retrotransposons tipo LTR, observaram que o tipo *PgDel* era o que mostrava uma distribuição particular com a maior proporção genômica em todas as espécies de *Panax*. Em conjunto, a abundâncias dos elementos e a distribuição desses elementos nos cromossomos levaram a evidenciar que os distintos genomas, em especial *P. quinquefolius*, evoluíram pela expansão do genoma pela rápida amplificação do elemento *PgDel* (LEE et al., 2017). Em *Rhynchospora pubera* foram identificadas *in silico* e confirmadas *in situ*, pela primeira vez, sequências específicas das regiões centroméricas de uma espécie com cromossomos holocêntricos. A família de DNA satélite *Tyba* e retroelementos centroméricos (CRRh) são as principais sequências de DNA associadas com os centrômeros funcionais dessa espécie (MARQUES et al., 2015). A presença e a distribuição centromérica desta família de DNA satélite e retroelementos é conservada dentro do gênero, assim como outras famílias de DNA satélites não centroméricas identificadas em outras espécies do gênero (RIBEIRO et al., 2017b).

2.4 RECONSTRUÇÃO DE CARATERES ANCESTRAIS

Abordagens empregando estudos filogenéticos tem permitido analisar e sugerir direções para mudanças cromossômicas. Por meio da combinação de dados cariotípicos e análises moleculares (COSTA et al., 2017; PELLICER et al., 2017; VAIO et al., 2013), os estudos filogenéticos comparativos permitem inferir de maneira mais clara, os eventos de hibridização ou mudanças nos números cromossômicos que levam à formação de novas espécies (VAIO et al., 2013, 2016). Estas análises combinadas têm demonstrado também que as disploidias descendentes são mais frequentes que as ascendentes na maioria dos grupos de plantas analisados. Embora, tanto casos de perdas como ganhos cromossômicos tenham ocorrido durante a evolução, como em *Melampodium* (Bakker et al. 2000; Stuessy et al. 2011; Cusimano et al. 2012). Além dos números cromossômicos básicos ancestrais, as análises filogenéticas podem ser utilizadas para estudar a evolução nos padrões de bandeamento cromossômico, os número e distribuição dos sítios de DNA ribossomais 5S e 45S, bem como as distribuições geográficas das espécies (COSTA et al., 2017; VAN-LUME et al., 2017). Em outros casos, quando combinadas com as análises de citogenética molecular, permitem identificar eventos de alopoliploidização e sugerir possíveis parentais. Mais recentemente, essas análises combinadas com estudos de tamanho do genoma em plantas sugeriram que o tamanho ancestral seria pequeno mas tanto ganhos como perdas de DNA ocorreram ao longo da evolução, permitindo reconhecer a direção das mudanças ocorridas (PELLICER et al., 2017). No caso da família Oxalidaceae há indícios de que o genoma ancestral fosse pequeno e que este, ao longo de sua evolução, teria sofrido grandes mudanças em alguns das seções, tanto no sentido de aumento quanto a redução do tamanho do genoma (VAIO et al., 2013). Na maioria das análises, os dados citogenéticos são plotados sobre as árvores filogenéticas (e.g. Peruzzi et al. 2009) ou ainda são analisados por métodos de reconstrução de caracteres ancestrais nos diferentes programas de análise filogenéticas (COSTA et al., 2017; MORAES et al., 2015). Mais recentemente, análises baseadas em modelos probabilísticos têm permitido testar diferentes hipóteses de evolução cromossômica numa determinada filogenia. Isso permite a reconstrução dos números cromossômicos ancestrais, assim como estima os eventos de poliploidização e disploidia ocorridos ao longo da evolução das plantas (MORAES et al., 2015). Embora este método apresente algumas limitações, como casos de múltiplos

números cromossômicos para uma mesma espécie, e inferência de eventos de poliploidia ímpar. Essa ferramenta representa uma forma rápida e relativamente simples de analisar números cromossômicos ancestrais e eventos cromossômicos ocorridos durante a evolução de espécies vegetais (ver GUERRA 2012).

2.5 FAMÍLIA *IRIDACEAE*

Esta família de monocotiledôneas é uma das maiores dentro da ordem Asparagales (CHASE; REVEAL, 2009). Possui entre 65-75 gêneros com mais de 2030 espécies, dentro de sete subfamílias monofiléticas (GOLDBLATT et al., 2008). Iridaceae compreende plantas com uma ampla diversidade de flores. Estas podem ser reconhecidas por sua corola; com três tépalas no verticilo exterior e três tépalas no verticilo interior usualmente parecidos em textura, forma e cor; e três estames (MANNING; GOLDBLATT, 2011). Esta diversidade de flores está associada a uma ampla diversidade de estratégias reprodutivas, incluindo associações a insetos ou aves polinizadoras. Nesta família, bem sendo particularmente estudadas as estruturas produtoras de óleo relacionadas com a polinização de abelhas dependente de óleo (CHAUVEAU et al., 2011).

2.5.1 Relações filogenéticas na família *Iridaceae*

Inicialmente, usando 52 caracteres morfológicos, a família Iridaceae foi classificada em quatro subfamílias: Isophysidoideae (1 gênero: 1 espécie), Nivenioideae (6:92), Iridoideae (20:~890) e Crocoideae (29:1032) (Goldblatt, 1990), confirmado por subsequentes estudos filogenéticos usando diferentes marcadores moleculares plastidiais (Reeves et al., 2001). Análises recentes ajudaram a evidenciar outras três subfamílias: Aristeoideae, Geosiridaceae e Patersonioideae (GOLDBLATT et al., 2008).

A única subfamília com representantes nos Neotrópicos (segundo maior centro de diversidade de Iridaceae) é a subfamília Iridoideae representada por 30 gêneros, atualmente reconhecidos, contendo ~900 espécies (MANNING; GOLDBLATT, 2011). Inicialmente foi agrupada por caracteres morfológicos e depois confirmada por

análises filogenéticas (GOLDBLATT et al., 2008). Na América do Sul estão representadas três das cinco tribos de Iridoideae: Sisyrinchieae, Trigrídieae, Trimezieae. No Brasil encontram-se 165 representantes das três tribos, pertencentes a 18 gêneros (GOLDBLATT et al., 2008; MANNING; GOLDBLATT, 2011). Trigrídieae, com maior número de representantes, está dividida em dois grandes clados correspondentes as duas subtribos estabelecidas previamente com base em dados morfológicos e citogenéticos: Cipurinae/Clado A e Trigrídiinae/Clado B (GOLDBLATT, 1982; MORAES et al., 2015a).

2.5.2 Citogenética e evolução cariotípica de *Iridaceae*

Existe uma grande variabilidade em número cromossômico, características cariotípicas e conteúdo de DNA dentre as espécies da família Iridaceae (MORAES et al., 2015a). Dentro da família o número cromossômicos varia de $2n = 6$ a $2n = 230$, sendo frequentes cariótipos bimodais e assimétricos, particularmente na subfamília Iridoideae (SOUZA-CHIES et al., 2012). Para Iridoideae, o número cromossômico básico ancestral sugerido é $x = 8$, com número básico secundários $x = 5, 6, 7, 8, 9$. Estes números teriam surgido por rearranjos cromossômicos dando em disploidias ascendentes, $x = 9$ em Sisyrinchieae, e descendentes $x = 5, 6, 7$ em Trigrídieae; já para em Trimezieae o número básico $x = 14$ teria se originado por disploidia descendente mais poliploidia (MORAES et al., 2015a). Tanto a diversidade nos números cromossômicos como análises de reconstrução de caracteres cromossômicos ancestrais mostram que a poliploidia assim como as disploidias foram fenômenos importantes na evolução das espécies de Iridaceae, particularmente na subfamília Iridoideae (MORAES et al., 2015a; SOUZA-CHIES et al., 2012). O tamanho do genoma dentro da família também mostra uma ampla variação, sendo observado para ~150 espécie entre 0,96 pg e 62,76 pg (Bennett e Leich, 2012).

Das tribos da subfamília Iridoideae, Trigrídieae contém espécies com cariótipos bimodais, geralmente com dois pares de cromossomos grandes e cinco pares de cromossomos pequenos. Dentro da subtribo Cipurinae o número básico é $x = 7$, apresentando diferentes níveis de ploidia sem eventos de disploidias (ALVES; LIMA; FELIX, 2011; MORAES et al., 2015a). Por outro lado, dentro da subtribo Trigrídiinae

presenta diversos números cromossômicos básicos, que provavelmente foram originados por poliploidia e eventos de disploidia descendentes em *Eleutherinex* = 6 e em *Phalocallisx* = 5 (MORAES et al., 2015a). Series poliploides são estados derivados e muito comuns dentro da tribo Trigrídieae e dentro de alguns gêneros, *Herbetia* por exemplo, considerando o valor 1Cx do tamanho do genoma, foi observado o que pareceria ser o fenômeno de diminuição do tamanho genômico seguido do evento de poliploidia (MORAES et al., 2015a). O tamanho do genoma dentro da subfamília Iridoideae mostra uma variação de quatro vezes, sendo observado valores 2C entre 3,90 pg a 15,43 pg (Bennett e Leitch, 2012) (MORAES et al., 2015a).

2.5.3 Gênero *Eleutherine* Herbert

Este gênero está constituído por plantas bulbosas, folhas lanceoladas e plissadas, com flores pequenas, brancas e de antese noturna, com distribuição na América Central e América do Sul, composta por apenas duas espécies: *E. bulbosa* (Miller) Urban e *E. latifolia* (Standley & Williams) Ravenna (GOLDBLATT; SNOW, 1991). As duas espécies são similares em relação às características morfológicas como as flores estrelas e a disposição dos estames. Citogeneticamente elas compartilham o número cromossômico básico $x = 6$, apresentando cariótipos bimodais com um par de cromossomos grandes e cinco pares de cromossomos pequeno (GOLDBLATT; SNOW, 1991).

Eleutherine bulbosa, distribuída amplamente na América do Sul, possui reprodução preferencialmente assexuada, apresenta um tamanho do genoma de 2,63 pg, número cromossômico $2n = 12$, e um cariótipo bimodal, com o par cromossômico I sendo 3 a 4 vezes maior que os demais pares. Este par maior apresenta uma heterozigose permanente devido a uma inversão pericêntrica assimétrica que inclui cerca de 70% do cromossomo, resultando em um cromossomo acrocêntrico e outro metacêntrico, ambos apresentando constrições secundárias adjacentes ao centrômero. A espécie apresenta características citológicas favoráveis para análises citogenéticas. O par maior apresenta tanto bandas DAPI⁺ como CMA⁺. O cromossomo acrocêntrico apresenta duas bandas DAPI⁺ intersticiais no braço longo, que no cromossomo metacêntrico são terminais no braço curto. As únicas bandas CMA⁺ se localizam neste par, associadas às contrações secundárias, com uma duplicação em *tandem* no metacêntrico. Guerra (1988) classificou as bandas CMA⁺ em bandas originais, que são as bandas CMA⁺ maiores e compartilhadas por ambos os cromossomos, e uma banda duplicada, que é de menor tamanho e está presente unicamente no cromossomo metacêntrico adjacente à banda original. Este segmento repetido não é idêntico ao original, apresentando apenas 40% do comprimento e uma maior atividade transcricional em comparação com os outros segmentos associados às contrações secundárias. Análises em meiose de *E. bulbosa* mostraram que as bandas CMA⁺ que correspondem à original estão sempre pareadas, o que significa que pareamento não homólogo nunca ocorre, e a banda CMA⁺ duplicada está sempre

não pareada. A formação de quiasmas ocorre na região fora da inversão e a região invertida forma uma alça sem quiasma (GUERRA, 1988). Sementes contendo variantes em homozigose não são viáveis, e na natureza só se observam indivíduos em heterozigose para a região invertida, mantendo esta espécie em heterozigose permanente e favorecendo a reprodução assexuada (M. Guerra, comunicação pessoal). A distribuição dos sítios de DNAr se encontra limitada ao par maior. Sítios de DNAr 35S estão associados às bandas CMA⁺, dentro da região invertida, e sítios de DNAr 5S estão duplicados na região terminal do braço longo do cromossomo acrocêntrico e na região terminal de um dos braços do cromossomo metacêntrico, fora da região de inversão (FEITOZA; GUERRA, 2011).

Análises da distribuição de modificações de histonas têm mostrado que o primeiro par de cromossomos de *E. bulbosa* se diferencia do resto do complemento. No caso de marcas que evidenciam atividade transcricional, todos os cromossomos pequenos marcaram fortemente e o par maior marca unicamente na região terminal de um dos braços de cada cromossomo do par. Além disso, a cromatina condensada em prófase, restrita ao par maior, é hipoacetilada e hipometilada em prófase e metáfase. As marcas associadas a regiões heterocromáticas, ricas em DNA repetitivo, mostram uma marcação mais intensa no par maior (FEITOZA; GUERRA, 2011). As características citogenéticas presentes em *E. bulbosa* a tornam um excelente material para estudar a evolução de inversões cromossômicas. Entre as características que a tornam um modelo se destacam a presença de poucos e grandes cromossomos, a condição de heterozigose permanente e a facilidade na identificação da grande inversão por marcas citológicas, tais como bandas CMA⁺ e DAPI⁺, constrições secundárias, sítios de DNAr e presença de uma duplicação em *tandem* dentro da região invertida.

3 EVOLUÇÃO DO CARIÓTIPO BIMODAL E ESTRUTURAÇÃO CROMOSSÔMICA NA SUBFAMÍLIA IRIDOIDEAE (IRIDACEAE)

Báez MA, Costa L, Pedrosa-Harand A

Laboratório de Citogenética e Evolução Vegetal, Departamento de Botânica, Universidade Federal de Pernambuco, Brasil

Autor para correspondência:

Mariana Báez

E-mail: marian_iaris@hotmail.com

Resumo

Cariótipo bimodais em plantas, caracterizados por dois grupos de cromossomos de tamanhos diferentes, representam o extremo da assimetria intercromossômica. Dentro da subfamília Iridoideae (Iridaceae), diversos gêneros apresentam cariótipos assimétricos e bimodais. Dentre eles, o gênero *Eleutherine* apresenta alta proporção de LTR-retrotransposons (LTR-RT) e sequências de DNA satélite com acumulação nos cromossomos grandes e distribuição diferenciada (uniforme vs. pericentromérica) entre cromossomos grandes e pequenos. Para avaliar a conservação dessa organização cromossômica particular em outros cariótipos bimodais do grupo e entender como os cariótipos bimodais evoluíram dentro da subfamília, foram amplificadas sequências repetitivas de *E. bulbosa* em *Gelasine coerulea*, *Phalocallis coelestis* e amplificadas e hibridizadas *in situ* em *Alophia drummondii*. Também foram reconstruídos os caracteres ancestrais de assimetria cariotípica e tamanho do genoma monoploide dentro da subfamília. Todas as espécies apresentaram amplificação das sequências repetitivas de *E. bulbosa*, indicando a presença destas famílias nos diferentes gêneros. *Alophia drummondii* mostrou distribuição dispersa uniforme ao longo de todos os cromossomos do complemento para os dois LTR-RT analisados, sugerindo que a organização cromossômica dos cariótipos bimodais na

subfamília não é conservada, apesar da característica bimodal compartilhada. O cariótipo bimodal surgiu na tribo Tigridaeae, estando os cariótipos bimodais mais assimétricos na subtribo Tigriniidae. O surgimento deste tipo de cariótipo não está diretamente relacionado à evolução do tamanho do genoma monoploide. Uma evolução com amplificação e eliminação diferencial de sequências repetitivas, envolvendo as mesmas linhagens de elementos mas com dinâmica gênero-específica, parece ser o mecanismo de evolução mais provável.

Palavras-chaves: assimetria cariotípica, organização cromossômica, reconstrução de caracteres ancestrais, retrotransposons, sequências repetitivas, tamanho do genoma, valor C

Introdução

O cariótipo é o maior nível de organização estrutural e funcional do genoma nuclear. Variações nos cariótipos, tanto numéricas quanto estruturais, estão associadas a mudanças evolutivas. As características cariotípicas comumente usadas para as análises evolutivas são número, tamanho e simetria, assim como distribuição de marcas cromossômicas, como a heterocromatina e sequências repetitivas (Weiss-Schneeweiss e Schneeweiss, 2013). A simetria cariotípica está determinada pelo tamanho dos cromossomos e/ou posição do centrômero. Mudanças no tamanho cromossômico podem afetar e ser afetadas pelo tamanho do genoma da espécie. A maioria das angiospermas possuem cromossomos de tamanho pequeno, similar, meta ou submetacêntricos e consequentemente com cariótipos uniformes e simétricos (Stebbins, 1971, Stace, 2000, Weiss-Schneeweiss e Schneeweiss, 2013).

Algumas linhagens se caracterizam pela presença de cariótipos assimétricos. Este tipo de cariótipo é marcado pela predominância de cromossomos com centrômeros terminais/subterminais (assimetria intracromossômica) e/ou tamanhos cromossômicos heterogêneos (assimetria intercromossômica). Um cariótipo com grande assimetria no tamanho cromossômico, constituído por dois conjuntos de tamanhos distintos, é conhecido como bimodal (Levitsky, 1931, revisado por Stebbins, 1971; Peruzzi e Eroğlu, 2013). Foi postulado que cariótipos assimétricos seriam derivados e evolutivamente mais recentes que os cariótipos simétricos (Stebbins 1971; Medeiros-Neto et al. 2017), no entanto também foi sugerido como um possível estado transitório ao invés de um simples estado evolutivo derivado final (Mandáková e Lysak 2008).

A origem dos cariótipos bimodais é usualmente associada a um de três processos diferentes. Rearranjos cromossômicos envolvendo eventos de fusão-fissão resultariam em cromossomos grandes como produtos da fusão de cromossomos pequenos ou cromossomos pequenos sendo resultado da fissão de cromossomos grandes (Burt, 2002; Schubert and Lysak, 2011; Gen-Shen Yin *et al.*, 2014). Aloploidia, envolvendo espécies parentais com diferentes tamanhos cromossômicos (Shirakawa et al. 2012; McKain et al. 2012), e aumento em tamanho de um conjunto de cromossomos causado pela acumulação diferencial de sequências

repetitivas (de la Herrán et al. 2001) seriam as outras duas causas possíveis. Cariótipos bimodais são bem conhecidos em monocotiledôneas como nos gêneros *Aloe* (Fentaw et al. 2013), *Agave* (Palomino et al. 2012), *Yucca* (Vosa 2005), e no gênero *Hypochaeris* de Asteraceae (Fiorin et al. 2013). Na maioria dos grupos bimodais de plantas, a organização cromossômica é pouco conhecida. No entanto, foi observado nos cromossomos grandes de várias espécies a presença de sequências altamente repetitivas, como regiões heterocromáticas C positivas em *Cephalanthera*, *Epipactis* e *Limodorum* (Orchidaceae, D'Emérico et al., 1999); ou DNA satélites (DNA_{sat}) em *Ornithogalum longibracteatum* (Pedrosa et al., 2001) e *Muscari comosum* (de la Herrán et al., 2001, Hyacinthaceae), possivelmente relacionados com o aumento na assimetria destes cariótipos.

A família Iridaceae, Asparagales, com ~70 gêneros, dentro de sete subfamílias monofiléticas, se caracteriza por possuir espécies com cariótipos simétricos e assimétricos (Goldblatt et al. 2008). A única subfamília com representantes nos Neotrópicos (segundo maior centro de diversidade de Iridaceae) é a subfamília Iridoideae com 30 gêneros e ~900 espécies (Goldblatt e Manning, 2008). Na América do Sul estão representadas três das cinco tribos da subfamília Iridoideae: Sisyrinchieae, Trigrídieae, Trimezieae (Goldblatt e Manning, 2008; Goldblatt et al., 2008). A tribo Trigrídieae, com maior número de representantes, está dividida em dois grandes clados correspondentes às duas subtribos estabelecidas previamente com base em dados morfológicos e citogenéticos: Cipurinae/Clado A e Trigrídiinae/Clado B (Goldblatt 1982; Moraes et al. 2015).

Para Iridoideae, o número cromossômico básico ancestral é $x = 8$, com números básicos secundários $x = 5, 6, 7, 9$ e 14 , consequência de disploidias ascendentes ($x = 9$) dentro da tribo Sisyrinchieae e descendentes ($x = 5, 6, 7$) dentro da tribo Trigrídieae, assim como, disploidia descendente e poliploidia ($x = 14$) na tribo Trimezieae. Eventos de poliploidia também teriam ocorrido dentro de Trigrídieae: na subtribo Trigrídiinae, no gênero *Alophia* e no clado mais derivado dentro dele composto por *Cardiostigma* Baker/ *Tigridia* Juss./ *Sessilanthera* Molseed & Cruden/ *Cobana* Ravenna (Moraes et al. 2015). Dentro da subtribo Cipurinae, clado A, o número básico é $x = 7$, apresentando diferentes níveis de ploidias e eventos de disploidias (Alves et al. 2011; Moraes et al. 2015). A tribo Trigrídieae é a única da

subfamília que apresentam espécies com cariótipos bimodais, em sua maioria com um conjunto de cromossomos grandes e um conjunto de cromossomos pequenos, independente dos diferentes números cromossômicos básicos. No entanto existem espécies que apresentam cariótipos não bimodais, como *Gelasine coerulea* (Alves et al., 2011; Moraes et al., 2015).

Dentro da tribo Trigidieae está o gênero *Eleutherine*, com duas espécies *E. bulbosa* e *E. latifolia* apresentando cariótipos fortemente bimodais ($n = 6$, 1L+5S) (Goldblatt and Snow 1991). Em *E. bulbosa*, a fração repetitiva representa ~40% do genoma, com alta proporção de elementos transponíveis do tipo LTR-retrotransposons (LTR-RT) e sequências de DNA satélite (satDNA). A distribuição cromossômica destas sequências repetitivas mostrou que existe uma diferenciação entre cromossomos grandes e cromossomos pequenos nesta espécie, com uma alta acumulação de LTR-RT de forma dispersa nos cromossomos grandes e uma distribuição pericentromérica das mesmas sequências nos cromossomos pequenos. A distribuição da marca para eucromatina H3K4me3 também mostrou diferenciação entre os tipos cromossômicos, sendo mais forte nos cromossomos pequeno e quase ausente nos cromossomos grandes. Outras marcas epigenéticas e o tempo de replicação do DNA na fase S também mostrou clara diferenciação entre os tipos cromossômicos. A distribuição dos LTR-RT específica dos tipos cromossômicos em *E. bulbosa* se mantém na espécie irmã, *E. latifolia* (Báez et al., em preparação). Sendo assim, no gênero *Eleutherine*, cromossomos grandes e pequenos mantem uma composição de sequências repetitivas e estruturação da cromatina diferencial, sugerindo diferenças estruturais e funcionais entre estes conjuntos cromossômicos relacionadas à condição de cariótipo bimodal. Dentro da mesma subtribo Tigridiineae, clado B, gêneros relacionados a *Eleutherine* também se caracterizam por possuir cariótipos bimodais. Estes gêneros possuem números cromossômicos diferentes, como *Gelasine* ($n = 7$), possuindo também espécies com cariótipos simétricos como *Gelasine coerulea*, e espécies como o gênero *Alophia* ($n = 14$), surgido por poliploidia (Moraes et al. 2015). No entanto, não é conhecida a evolução da assimetria cariotípica dentro da tribo e se existe uma estruturação cromossômica peculiar nestas espécies associada à sua condição bimodal.

Dentro deste contexto de distribuição diferencial de sequências repetitivas entre cromossomos grandes e pequenos no gênero *Eleutherine* e a presença de cariótipos bimodais em espécies da tribo Trigidieae com diferentes números cromossômicos básicos, levantamos as seguintes questões: Como as sequências repetitivas estão distribuídas entre cromossomos grandes e pequenos em outras espécies com cariótipos bimodais relacionadas ao gênero *Eleutherine*? As mesmas sequências repetitivas que apresentam distribuição diferencial estão presentes em espécies que não apresentam cariótipos bimodais dentro da tribo Trigidieae? A diferenciação estrutural cromossômica característica do cariótipo bimodal de *Eleutherine* é mantida dentro da tribo? Qual é a origem da assimetria cariotípica dentro da subfamília Iridoideae? Existe alguma correlação entre a evolução da assimetria cariotípica com a evolução do tamanho genômico dentro da subfamília Iridoideae?

Assim, o objetivo do trabalho foi entender se a distribuição cromossomo-específica das sequências repetitivas no gênero *Eleutherine* é mantida em outras espécies dentro da tribo Trigidieae e como evoluiu a assimetria cariotípica dentro da subfamília Iridoideae.

Materiais e Métodos

Material vegetal

Para as análises cromossômicas foram utilizadas sementes de *Gelasine coerulea* (Vell.) Ravenna e *Phallocallis coelestis* (Lehm.) Ravenna, cedidas pela Dra. Eliane Kaltchuk dos Santos da Universidade Federal de Rio Grande do Sul, e sementes de *E. latifolia*, cedidas pela Dra. Guadalupe Munguía Linno da Universidade de Guadalajara, México. Todas as sementes foram germinadas em câmara úmida (3 - 4 meses) e as plântulas foram mantidas em substrato numa sala de germinação a 24°C e depois a temperatura ambiente. Indivíduos de *Alophia drummondii* (Graham) R.C. Foster foram coletados em Buíque, Pernambuco, Brasil e foram cedidos pelo Dr. Marcelo Guerra da Universidade Federal de Pernambuco. Indivíduos de *E. bulbosa* foram coletados em Aracuara, Bahia, Brasil e foram cedidos pelo Dr. Luiz Gustavo Souza da Universidade Federal de Pernambuco (voucher número UFP 82763). As

plantas crescidas de sementes e os indivíduos coletados foram mantidos em cultivo no Jardim Experimental do Laboratório de Citogenética e Evolução Vegetal da Universidade Federal de Pernambuco, Brasil.

Preparações cromossômicas

Pontas de raízes obtidas a partir das sementes ou bulbos foram pre-tratadas em 8-hidroxiquinoleína 0,02 mM a 10 °C por 24 h, fixadas em etanol: ácido acético (3:1 v/v) de 2 a 24 h a temperatura ambiente e estocadas -20 °C. As raízes fixadas foram digeridas com 2% celulase (Onozuka) - 20% pectinase (Sigma) por 60 min a 37 °C, esmagadas numa gota de ácido acético 45% e a lamínula foi retirada em nitrogênio líquido.

Extração de DNA e amplificação por PCR

O DNA genômico de *A. drummondii*, *G. coerula* e *P. coelestis* foi extraído seguindo o protocolo de Doyle e Doyle (1987), a partir de folhas frescas nas duas primeiras espécies e a partir de sementes na última espécie. O DNA da espécie controle *E. bulbosa* foi extraído, a partir de folhas frescas, usando o DNAeasy Plant Mini Kit (Qiagen) seguindo protocolo do fabricante. As sequências repetitivas mais abundantes de *E. bulbosa*, o DNA satélite, Ebusat2, e quarto LTR-retrotransposons (LTR-RT) Ty1/Copia Maximus e Tork e Ty3/Gypsy Tat e Chromovirus, foram amplificadas por PCR nas três espécies relacionadas, utilizando *primers* desenhados para *E. bulbosa* (Báez *et al.*, em preparação). 40 ng de DNA genômico foi usado para todas as reações de PCR, contendo também 1x PCR buffer (Invitrogen), 20 µM MgCl, 0.1 µM de cada dNTP, 0.5 µM de cada *primer*, 0.024 U *Taq* polymerase (*Taq* DNA polymerase Platinum, Invitrogen-Brasil) e água para volume total de 50 µl. As condições da PCR foram 1× 94°C 3 min, 30× 94°C 1 min, 55°C 1 min, 72°C 1 min e 1× 72°C 10 min. Os produtos foram checados em gel de agarose 1% e fotografados.

Hibridização in situ fluorescente e estimativa da intensidade do sinal fluorescente

A hibridização *in situ* fluorescente (FISH) foi feita nas espécies *Alophia drummondii* e *Gelasine coerulea* utilizando as sondas dos DNA satélites Ebusta1, Ebusat2 e de quatro LTR-RT, Ty1/Copia Maximus e Tork e Ty3/Gypsy Tat e Chromovirus de *E. bulbosa* (Baéz *et al.*, em preparação). Adicionalmente, foram selecionadas três sequências repetitivas para amplificação e localização cromossômica na espécie *Alophia drummondii*: o DNA satélite Ebusat2, e dois LTR-RT Ty1/Copia Tork e Ty3/Gypsy Chromovirus. Estas foram marcadas com Cy3-dUTP por *nick translation* usando o kit Nick translation mix (Roche) a partir dos produtos de PCR amplificados da própria espécie.

A FISH foi feita como descrito por Pedrosa *et al.* (2002), com modificações na stringência de hibridização final para 42%, com o mix hibridização contendo 10% (w/v) dextran sulfato, 6× SSC e 5 ng/μl de cada sonda e lavagens pós-hibridização 5 vezes em 2xSSC. As imagens foram capturadas num microscópio Leica DM5500 B com uma câmera acoplada Leica DFC345 FX e o software LAS AF. As imagens foram editadas com o programa Adobe Photoshop CS5.

Para quantificar o acúmulo de sequências repetitivas ao longo dos tipos cromossômicos, foram comparadas as intensidades entre o sinal do LTR-RT Ty1/Copia Tork e o contracolorante DAPI. Foi estimada a intensidade de fluorescência do sinal para *E. bulbosa* e *A. drummondii* ao longo dos cromossomos grandes (ambos homólogos por separado para *E. bulbosa*) e pequenos e na área total dos cromossomos grandes (ambos homólogos por separado para *E. bulbosa*) e pequenos. Estas medições foram realizadas no software ImageJ (Abramoff *et al.*, 2004). A intensidade do sinal foi quantificada em 10 ou oito pontos ao longo dos cromossomos grandes e pequenos, respectivamente, em cinco metáfases para cada espécie. Para o cálculo da razão entre as intensidades do sinal do LTR-RT Ty1/Copia Tork e do DAPI, a intensidade foi quantificada na área total de ambos os tipos cromossômicos em ambas as espécies. Porcentagens nos gráficos mostram a proporção do conteúdo de DNA de cada cromossômico em relação ao complemento diploide da espécie.

Estimativa dos índices de assimetria cariotípica e tamanho do genoma haploide

Para a estimativa dos índices de assimetria foram utilizadas imagens das células metafásicas de espécies de diferentes gêneros da subfamília Iridoideae (Iridaceae), listadas na Tabela 1. Cada cromossomo de todas as metáfases foram medidos usando a ferramenta 'Ruler' do software Adobe Photoshop CS3 e o comprimento em μm foi calculado baseado na barra de escala das respectivas imagens. Para cada espécie, o índice de assimetria foi estimado segundo Zarco (1986), usando a fórmula para o índice A_2 que calcula a relação entre o tamanho dos diferentes cromossomos, sendo a relação entre o desvio padrão (S) e a média do comprimento cromossômico ($\bar{X}$) para cada amostra:

$$A_2 = \frac{S}{\bar{X}}$$

Os valores para o tamanho do genoma monoploide foram usados para reconstrução do tamanho do genoma devido aos diferentes níveis de ploidia das espécies. Estes valores foram compilados de Moraes *et al.* (2015). Para *Alophia silvestris*, *Ennealophus euryandrus*, *Libertia sessiliflora*, *Nemastylis tenuis*, *Neomarica northiana*, *Orthrosanthus chimboracensis* *Trimezia martinicensis*, os valores do tamanho do genoma monoploide foram estimados baseados no comprimento do complemento cromossômico total e o comprimento total e tamanho do genoma das espécies filogeneticamente próximas.

Métodos Filogenéticos Comparativos

Para entender a correlação entre o índice de assimetria cariotípica e a evolução do cariótipo bimodal, assim como sua relação com a evolução do tamanho do genoma monoploide, foram empregados Métodos Filogenéticos Comparativos (PCMs). Os índices de assimetria e tamanhos do genoma haploide estimados neste trabalho junto com os estimados na literatura previamente (Tabela 1), foram compilados para as análises do grupo. Foi usada a topologia da árvore filogenética previamente publicada para a subfamília Iridoideae (Moraes *et al.*, 2015). Como os índices de assimetria foram estimados unicamente nas espécies com imagens das metáfases disponíveis na literatura, não foi possível estimar este índice para todas as espécies incluídas na filogenia. Por isto, foi utilizado o algoritmo 'drop.tip' incluído no pacote 'phytools' v.

5.30 (Revell 2012) no software R v. 3.0.1 (R Core Team 2015) para excluir os terminais da árvore correspondente às espécies sem índice de assimetria estimado. Também foi inferido o sinal filogenético para a assimetria cariotípica dentro da subfamília Iridoideae usando o índice λ de Pagel (Pagel, 1999). O sinal filogenético é um importante fator nas análises PCM, porque espécies com características similares frequentemente permanecem relacionadas (Münkemüller, 2012). O estado ancestral da assimetria cariotípica e tamanho do genoma monoploide foi estimado usando a abordagem de máxima verossimilhança como implementado no pacote 'phytools' v. 5.30 (Revell 2012) no software R v. 3.0.1 (R Core Team 2011). Para a árvore contendo apenas as espécies com Índice de Assimetria Cariotípica estimado, foi utilizada a ferramenta 'contMap' implementada no pacote 'phytools' v. 5.30 (Revell 2012) no software R v. 3.0.1 (R Core Team 2011). Para evidenciar a existência de correlação entre tamanho do genoma monoploide e a assimetria cariotípica na subfamília Iridoideae foram usadas duas diferentes abordagens comparativas: 1) contrastes filogenéticos independentes (phylogenetic independent contrasts: PICs) (Felsenstein 1985); 2) mínimos quadrados filogenéticos generalizados (Phylogenetic generalized least squares: PGLS) (Butler and King, 2004). Estas análises foram feitas usando os pacotes 'ape' v. 4.0 (Paradis 2017) e 'caper' v. 0.5.2 (Orme *et al.*, 2013) no software R v. 3.0.1 (R Core Team 2015).

Resultados

Presença das famílias de sequências repetitivas em espécies relacionadas a Eleutherine

As espécies de *Eleutherine* apresentam cariótipo bimodal com acúmulo de elementos repetitivos nos cromossomos grandes (Baéz *et al.* em preparação). Para determinar a presença desses elementos em espécies relacionadas da subfamília, *primers* desenvolvidos em *E. bulbosa* para as seis famílias mais abundantes de sequências repetitivas foram utilizados para amplificação por PCR e confirmaram a presença destas nas espécies *Alophia drummondii*, *Gelasine coerulea* e *Phalocallis coelestis*. Os dois satDNA, Ebusat1 (característico dos cromossomos grandes) e

Ebusat2 (característico da região centromérica), mostraram o mesmo tamanho de fragmentos amplificados que os observados para *E. bulbosa*. Para as duas linhagens de LTR-RT Ty1/Copia, Tork e Maximus, e as duas LTR-RT Ty3/Gypsy, Chromovirus e Tat, o domínio da Integrase mostrou ampliações com mesmos tamanhos de fragmentos que em *E. bulbosa*, exceto em *P. coelestis* para o Chromovirus que não mostrou amplificação (Fig. 1).

Distribuição cromossômica das sequências de DNA repetitivo em Alophia drummondii

Nas espécies de *Eleutherineos* cromossomos grandes apresentam acúmulo de retroelementos com distribuição uniforme, típica de genomas grandes, enquanto os cromossomos pequenos apresentam distribuição pericentromérica, típica de genomas pequenos (Baéz et al. em preparação). Para evidenciar se essa distribuição bimodal é mantida no grupo, foram escolhidas e amplificadas a partir de DNA genômico de *A. drummondii*, das três espécies utilizadas, a filogeneticamente mais distante, três famílias repetitivas características de *E. bulbosa*, Ebusat2, característico da região pericentromérica, e dois LTR-RT, Ty1/Copia-Tork e Ty3/Gypsy-Chromovirus. Para o satDNA Ebusat2 não foi observado nenhum sinal de hibridização nos cromossomos de *A. drummondii* (dado não mostrado). Ambos os LTR-RT mostraram sinais de hibridização com distribuição cromossômica similar. Estas famílias repetitivas apresentam distribuição uniforme e dispersa ao longo de todos os cromossomos do complemento, tanto grandes quanto pequenos (Fig. 2).

Para confirmar a distribuição uniforme das sequências repetitivas em ambos os tipos cromossômicos de *A. drummondii*, a intensidade dos sinais de LTR-RT Ty1/Copia-Tork foi medida nesta espécie e em *E. bulbosa*, para comparação. Em *A. drummondii* foi observado que este LTR-RT se distribui de forma uniforme ao longo do cromossomo e de forma similar em ambos os tipos cromossômicos, sem nenhuma região evidenciando uma aparente acumulação. A razão entre a intensidade de fluorescência do sinal da sonda LTR-RT e do contracolorante de DNA DAPI foi similar entre cromossômicos grandes e pequenos, o que é evidenciado pela razão de 0,99 entre eles. Isso sugere que o sinal de hibridização do LTR-RT é proporcional à quantidade de DNA dos cromossomos, tanto grande como pequeno. Por outro lado,

em *E. bulbosa*, o sinal de LTR-RT se distribuiu de forma uniforme ao longo dos cromossomos grandes, porém, nos cromossomos pequenos, foi mais intenso na região pericentromérica, diminuindo nas regiões distais dos braços cromossômicos. A razão entre as intensidades relativas do LTR-RT de cromossomos pequenos e grandes foi de 0,83 e 0,85 em relação ao acrocêntrico e ao metacêntrico respectivamente. Isso mostra que, diferentemente de *Alophia*, existe uma maior abundância relativa do LTR-RT em cromossomos grandes de *Eleutherine* quando comparada com os cromossomos pequenos, provavelmente devido à distribuição restrita à região pericentromérica dos mesmos nos cromossomos pequenos (Fig. 3).

Reconstrução da assimetria cariotípica e tamanho do genoma haploide na subfamília Iridoideae

Uma vez que a amplificação de sequências repetidas parece ter tido papel importante para a bimodalidade em *Eleutherine*, foram reconstruídos a assimetria cromossômica e o tamanho do genoma ancestral e avaliada sua evolução na subfamília Iridoideae usando 27 espécies. O índice de assimetria cariotípica variou entre 0,14 em *Libertia sessiliflora* e 0,83 em *Ennealophus euryandrus* (Tabela 1). Das espécies com cariótipo assimétrico, que variaram entre 0,26 - 0,83, todos foram previamente descritos como cariótipos bimodais, pertencendo aos gêneros *Calydorea*, *Cipura*, *Cypella*, *Herbertia*, *Kelissa*, *Nemastylis*, *Onira*, *Alophia*, *Eleutherine*, *Gelasine* (só uma das espécies analisadas), *Tigridia*, *Ennealophus*. Apesar disto, o índice de assimetria de *Trimezia spatha* foi 0,27 mas o cariótipo não bimodal. Considerando apenas as estimativas diretas de tamanho de genoma, o tamanho do genoma monoploide 1Cx variou de 0,48 pg em *Sisyrinchium convolutum* e 4,35 pg em *Tigridia pavonia* e *Trimezia spathata* (Tabela 1). Os índices lambda de Pagel encontrados foram 0,99 ($p < 0.01$) para assimetria cariotípica e 0,98 ($p < 0.01$) para tamanho do genoma monoploide, indicando alto sinal filogenético para ambas as características dentro da subfamília Iridoideae. Isto confirma que espécies mais intimamente relacionadas apresentam valores similares para estes caracteres.

A reconstrução do caráter ancestral para a assimetria cariotípica indicou que o cariótipo ancestral da subfamília possuía um certo grau de assimetria. Na primeira

linhagem a divergir (clado 1, tribo Sisyrinchieae), o cariótipo se tornou mais simétrico. No clado 2 (tribo Trimezieae), os cariótipos também apresentaram um aumento na simetria. Por outro lado, no clado 3 (tribo Tigridaeae), os cariótipos se tornaram mais assimétricos (na sua maioria bimodais), especialmente no clado B (subtribo Tigridiinae), com os extremos nos gêneros *Eleutherine* e *Ennealophus* (Tabela 1 e Fig. 4).

A reconstrução do caráter ancestral para o tamanho do genoma monoploide mostrou que o genoma ancestral era pequeno (segundo classificação em (Bennett e Leitch 2005), com valor 1Cx intermediário entre os tamanhos observados dentro da subfamília. No clado 1, houve uma diminuição do tamanho do genoma nas espécies analisadas. No clado 2, os genomas aumentaram, com algumas espécies do gênero *Trimezia* e *Neomerica* apresentando os maiores genomas da subfamília. No clado 3, os genomas mantiveram, em sua maioria, o estado ancestral. No entanto, no clado A os genomas tiveram um aumento dentro dos gêneros *Calydorea* e *Cipura*. Já no clado B, os genomas diminuíram nos gêneros *Eleutherine* e *Ennealophus* e aumentaram nas espécies de *Tigridia* analisadas (Fig. 4).

As análises de PICs e PGLs mostraram que não existe correlação entre a evolução do índice de assimetria cariotípica e do tamanho do genoma monoploide ($p = 0,50$ para ambas análises) dentro da subfamília Iridoideae (Fig 5). A comparação das árvores da reconstrução do caráter ancestral para a assimetria cariotípica e o tamanho do genoma haploide também evidenciaram que não existe uma correlação entre o aumento ou diminuição do tamanho do genoma haploide e o aumento ou diminuição da assimetria cariotípica. Espécies de clados com cariótipos simétricos mostraram tanto diminuição do tamanho do genoma (clado 1) como aumento do tamanho do genoma (clado 2). Espécies do clado 3 (com aumento da assimetria e cariótipos bimodais) mantiveram o tamanho do genoma monoplóide, mas em alguns casos foi possível observar aumento e nos cariótipos mais assimétricos os genomas diminuíram (clado 3B) (Fig. 4).

Discussão

Organização cromossômica do DNA repetitivo em espécies com cariótipos bimodais na subfamília Iridoideae

As espécies de *Eleutherine* apresentam cariótipo bimodal com cromossomos grandes apresentando acúmulo de retroelementos com distribuição uniforme, típica de genomas grandes, e cromossomos pequenos com distribuição pericentromérica, típica de genomas pequenos (Baéz et al. em preparação). Essa organização não foi observada na espécie *Alophia drummondii*, outra espécie com cariótipo bimodal do clado B da tribo Tigridaeae. Esses dados evidenciam que existe divergência na estruturação cromossômica dos diferentes cariótipos bimodais dentro da tribo Tigridaeae. Esta situação é diferente da observada em outros grupos de espécies com cariótipos bimodais. Acumulação de sequências repetitivas do tipo retrotransposons não-LTR cromossomo-específica foi observada em diferentes espécies de Galliformes. O retrotransposon CR1, caracterizado em galinha, se distribui em todos os cromossomos com uma alta acumulação nos cromossomos grandes formando um padrão de bandas não correspondente com bandas R ou G. Esta acumulação e distribuição se mantêm em Anseriformes, Passeriformes e Falconiformes (Coullin et al. 2005). Por isso, os mecanismos de evolução das sequências repetitivas em cromossomos grandes e pequenos dentro da subfamília Iridoideae parecem ser mais dinâmicos e diferenciados a depender do gênero. Essa diferença pode estar relacionada às diferenças entre os cariótipos bimodais de Iridoideae, que apresentam uma gradação de assimetria, com cariótipos desde simétricos até claramente bimodais, enquanto nesses grupos de aves os cariótipos são fortemente bimodais, com macro e microcromossomos, revelando uma organização diferenciada.

Assim como proposto para *Eleutherine* (Baéz et al, em preparação), as mesmas sequências repetitivas poderiam estar envolvidas na evolução dos cariótipos bimodais de vários dos gêneros da subfamília Iridoideae, uma vez que, em quatro dos gêneros analisados (*Eleutherine*, *Alophia*, *Gelasine* e *Phalocallis*), os mesmos elementos LTR-RT e satDNA foram detectados. Assim, as sequências características de *E. bulbosa* tiveram provavelmente uma origem comum dentro do clado B da subfamília Iridoideae. Para a sequência Ebusat2, centromérica em *Eleutherine*, não foi possível observar sinais de hibridização em nenhum cromossomo de *A. drummondii*. Essa falta de sinais

pode indicar que a abundância desta família repetitiva na espécie é menor que a necessária para a visualização por FISH. Dentro do gênero *Prospero*, o número de cópias do satDNA centromérico *PaB6* possui grandes diferenças tanto dentro de citótipos da mesma espécie, *P. autumnale*, como em espécies relacionadas. Nas espécies relacionadas *P. hanburyi* e *P. obtusifolium*, o número de cópias foi sempre menor, coincidindo com sinais muito fracos nas hibridizações *in situ* nos cromossomos de cada espécie (Emadzade et al. 2014). No feijão comum, *Phaseolus vulgaris*, a sequências repetitiva *PvMeso* é característica dos acessos mesoamericanos e estão presentes nos acessos andinos e nas espécies próximas num baixo número de cópias unicamente visível por Southern blot com longa exposição. A origem de *PvMeso* seria recente no gênero e teria ocorrido unicamente no grupo Mesoamericano (Ribeiro et al. 2011).

Evolução da assimetria cariotípica na subfamília Iridoideae

A reconstrução do cariótipo ancestral da subfamília Iridoideae indicou que cariótipo ancestral seria simétrico, não bimodal, embora possuísse um certo grau de assimetria. Isto concorda com a hipótese de que cariótipos assimétricos seriam derivados de cariótipos simétricos (Stebbins 1971; Medeiros-Neto et al. 2017). O cariótipo teria se mantido simétrico, porém aumentado a simetria, nos dois primeiros clados que divergiram na evolução desta subfamília (tanto no clado 1, específico de espécies da tribo Sysirinchieae, quanto no clado 2, específico de espécies da tribo Trimezieae). A assimetria cariotípica aumentou no clado 3 (tribo Tigridieae) e o cariótipo bimodal surgiu também neste clado. Dentro deste clado, a assimetria cariotípica foi acentuada nas espécies do clado B/Tigridiinae e possivelmente houve uma reversão no ramo que derivou na espécie *Gelasine coerulea*. Em Agavaceae o surgimento do cariótipo bimodal ocorreu uma única vez no clado *Yucca-Agave*, incluindo também o gênero *Hosta*, conhecido como clado ABK (Agavoideae Bimodal Karyotype, McKain et al. 2012, 2016). Isto é recorrente dentro de Asparagales, onde o cariótipo bimodal tem surgido várias vezes (Pires et al. 2006).

Evolução da assimetria vs. evolução do tamanho do genoma monoploide: mecanismo de evolução do cariótipo bimodal

O tamanho do genoma monoploide ancestral da subfamília foi um tamanho considerado pequeno (segundo Bennett e Leitch 2005), intermediário entre os encontrados na subfamília. Foram identificados tanto eventos de diminuição, clado 1, como de aumento, clado 2 e alguns ramos no clado 3, do tamanho do genoma monoploide. No entanto, estes eventos não estão relacionados com o aumento ou diminuição da assimetria cariotípica, como confirmado pelas análises de PICs/PGLS. Em outros grupos com cariótipos bimodais como em Agavaceae o surgimento da bimodalidade sim esteve correlacionado a aumento no tamanho do genoma. Para o clado ABK, Agavaceae, os eventos de paleopoliploidia foram correlacionados com o surgimento do cariótipo bimodal deste grupo de espécies na família (McKain et al. 2012).

A manutenção da simetria cariotípica está relacionada tanto com aumento como com a diminuição do tamanho do genoma monoploide. Isto sugere uma possível amplificação e eliminação, respectivamente, relativamente uniforme para todos os cromossomos do complemento levando a cariótipos simétricos em ambos os casos. Por outro lado, no clado onde a assimetria cariotípica aumenta, surgindo inclusive os cariótipos bimodais, o tamanho do genoma se mantém estável de forma geral, evidenciando uma possível amplificação/eliminação diferencial nos conjuntos de cromossomos grandes e pequenos, respectivamente, levando à assimetria cariotípica. Dentro deste clado, os cariótipos mais assimétricos diminuíram o tamanho do genoma monoploide. No caso do gênero *Eleutherine*, a diminuição do tamanho do genoma poderia estar relacionado à diminuição do número cromossômico básico de $n = 7$ para $n = 6$ (Moraes et al. 2015), com menos um par de cromossomos grandes em relação aos outros cariótipos bimodais, no entanto, não tem até o momento indícios de fusão cromossômica para esta espécie nem para a subfamília. No caso de *Ennealophus*, pode estar relacionada a uma redução no tamanho dos cromossomos pequenos, já que são os menores dentre os cromossomos pequenos das demais espécies bimodais.

De forma geral, observando os comprimentos dos pares cromossômicos dentro dos cariótipos bimodais, os cromossomos pequenos mantem, com algumas exceções, um comprimento em torno dos 2 μm , similar aos comprimentos dos cromossomos dos cariótipos simétricos na tribo Sysirinchieae. Por outro lado, os maiores cromossomos grandes variam entre 5- 12 μm . Esta grande variação nos comprimentos cromossômicos, com aumento dos cromossomos grandes e relativa estabilidade nos cromossomos pequenos, poderia indicar um equilíbrio entre a amplificação de sequências nos cromossomos grandes e eliminação preferencial nos cromossomos pequenos. Isto estaria em concordância com o observado na distribuição cromossomo-específico de sequências repetitivas no gênero *Eleutherine*, com acúmulo de diferentes sequências LTR-RT e satélites de DNA nos cromossomos grandes (Báez et al., em preparação). Isto é observado também em outros grupos como as sequências de satDNA caracterizadas em *Muscari comosum*, MCSAT, que estão amplamente acumuladas nos cromossomos grandes desta espécie, sugerindo o aumento do tamanho cromossômico como consequência. Análises deste satDNA em *M. matritensis* e *M. dionysicum* mostraram amplificação e localização cromossômica similar, no entanto a proporção de acumulação nestas espécies é bem menor (de la Herrán et al. 2001). Por outro lado, considerando os cromossomos do gênero *Neomerica*, com cariótipo simétrico e cromossomos em torno de 5 μm , o mecanismo que levaria a uma maior eliminação nos cromossomos pequenos também seria possível. De fato, é provável que tanto aumento quanto diminuição de cromossomos grandes e pequenos esteja acontecendo na subfamília Iridoideae, uma vez que a distribuição das mesmas sequências em *Eleutherine* e *Alophia drummondii* não é similar. Outros fatores podem estar influenciando a evolução dos cariótipos bimodais nesta família. Isto foi observado para outras famílias com cariótipos bimodais dentro de Asparagales, como no clado ABK de Agavaceae. Neste grupo, certos fatores ecológicos como tipo de solo e variação floral com múltiplas estratégias reprodutivas se mostraram particularmente envolvidas na evolução do cariótipo bimodal (McKain et al. 2016).

Na subfamília Iridoideae, a assimetria cariotípica aumentou e o cariótipo bimodal teria surgido uma vez na tribo Tigridieae. Este evento não estaria diretamente relacionado com o tamanho do genoma monoploide. Apesar disso, comparando a

evolução de ambos os caracteres analisados, assimetria cariotípica e tamanho do genoma monoploide, pode ser sugerido uma evolução com amplificação e eliminação das sequências repetitivas cromossomo-específico, como sugerido pelas organizações cromossômicas de sequências repetitivas em algumas espécies. A distribuição cromossômica das sequências repetitivas é diferente em gêneros diferentes dentro da tribo Tigridieae, não mantem a estruturação cromossômica bimodal observada em algumas espécies da tribo. Assim, é possível que o mecanismo de amplificação/eliminação seja dependente do gênero e talvez às próprias particularidades ecológicas.

Agradecimentos

Agrademos a Prof. Dra. Eliane Kaltchuk dos Santos da Universidade Federal de Rio Grande do Sul por ceder as sementes de *Gelasine coerulea* e *Phallocallis coelestis* e ao Prof. Dr. Marcelo Guerra por ceder os indivíduos de *Alophia drummondii*. Também agradecemos a Dra. Ana Paula Moraes da Universidade Estadual de São Paulo pela disponibilização da topografia da árvore da subfamília Iridoideae. Também agradecemos à Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brasil, pelo suporte financeiro.

Referências

- Aaron R, Ortiz-Catedral O (2002) Cytology and pollen morphology of *Ainea conzatii* and *Cardiostigma hintonii* (Tigridieae: Iridaceae) and their phylogenetic consequences. *Botany in the Curriculum: Integrating Research and Teaching*. Available at: <http://2002.botanyconference.org/section12/abstracts>
- Alves LIF, Lima SAA, Felix LP (2011) Chromosome characterization and variability in some Iridaceae from Northeastern Brazil. *Genet Mol Biol* 34:259–267. doi: 10.1590/S1415-47572011000200016

- Arroyo Martínez HA, Arzate Fernández AM, Barba González R, Piña Escutia JL (2017) Karyotype determination of three *Tigridia* species (Asparagales, Iridaceae). *Caryologia* 70:211–215. doi: 10.1080/00087114.2017.1321312
- Bennett MD, Leitch IJ (2005) Genome Size Evolution in Plants. In: *The Evolution of the Genome*. Elsevier, pp 89–162
- Bennett MD, Leitch IJ (2012) *Plant DNA C-values database*. Kew: Royal Botanic Gardens. Available at: <http://data.kew.org/cvalues/>
- Blanchon DJ, Murray BG, Braggins JE (2000) Chromosome numbers in the genus *Liberlia* (Iridaceae). *N Z J Bot* 38:245–250. doi: 10.1080/0028825X.2000.951268
- Burt DW (2002) Origin and evolution of avian microchromosomes. *Cytogenet Genome Res* 96:97–112
- Chukr NS, Giulietti AM (2008) Revisão de *Trimezia* Salisb. ex. Herb. (Iridaceae) para o Brasil. *Sitientibus* 8: 15–58
- Coullin P, Bed'Hom B, Candelier JJ, et al (2005) Cytogenetic repartition of chicken CR1 sequences evidenced by PRINS in Galliformes and some other birds. *Chromosome Res* 13:665–673. doi: 10.1007/s10577-005-1004-7
- de la Herrán R, Robles F, Cuñado N, et al (2001) A heterochromatic satellite DNA is highly amplified in a single chromosome of *Muscari* (Hyacinthaceae). *Chromosoma* 110:197–202. doi: 10.1007/s004120000115
- D'Emerico S, Grünanger P, Scrugli A, Pignone D (1999) Karyomorphological parameters and C-band distribution suggest phyletic relationships within the subtribe Limodorinae (Orchidaceae). *Plant Syst Evol* 217:147–161
- De Tullio L, Roitman G, Bernardello G, Tullio LD (2008) *Tamia* (Iridaceae), a synonym of *Calydorea*: cytological and morphological evidence. *Systematic Botany* 33: 509–513
- Emadzade K, Jang T-S, Macas J, et al (2014) Differential amplification of satellite PaB6 in chromosomally hypervariable *Prospero autumnale* complex (Hyacinthaceae). *Ann Bot* 114:1597–1608. doi: 10.1093/aob/mcu178

Fentaw E, Dagne K, Rønsted N, et al (2013) Karyotypes in Ethiopian Aloe species (Xanthorrhoeaceae: Asphodeloideae). *Kew Bull* 68:599–607. doi: 10.1007/s12225-013-9475-8

Fiorin FG, Ruas PM, Ortiz MA, et al (2013) Karyotype studies on populations of two *Hypochaeris* species (*H. catharinensis* and *H. lutea*), Asteraceae, endemics to southern Brazil. *Genet Mol Res* 12:1849–1858. doi: 10.4238/2013.January.4.4

Gen-Shen Yin, Yang Zhi-Yun, Chiang Tzen-Yuh, Gong Xun (2014) Tracing the origin of the Bimodal karyotype of the tribe Lilieae (Liliaceae) based on comparative karyotypes analyses. *7 3 7 - 7 4 6*

Goldblatt P (1982) Chromosome Cytology in Relation to Suprageneric Systematics of Neotropical Iridaceae. *Syst Bot* 7:186–198. doi: 10.2307/2418327

Goldblatt P, Takei M (1997) Chromosome Cytology of Iridaceae-Patterns of Variation, Determination of Ancestral Base Numbers, and Modes of Karyotype Change. *Ann Mo Bot Gard* 84:285. doi: 10.2307/2400005

Goldblatt P, Rodriguez A, Powell MP, et al (2008) Iridaceae ‘Out of Australasia’? Phylogeny, biogeography, and divergence time based on plastid DNA sequences. *Syst Bot* 33:495–508

Goldblatt P, Snow N (1991) Systematics and Chromosome Cytology of Eleutherine Herbert (Iridaceae). *Ann Mo Bot Gard* 78:942–949. doi: 10.2307/2399735

Kenton A, Heywood CA (1984) Cytological studies in South American Iridaceae. *Plant Syst Evol* 146:87–104

Kenton AY, Rudall PJ, Johnson AR. 1986. Genome size variation in *Sisyrinchium* L. (Iridaceae) and its relationship to phenotype and habitat. *Botanical Gazette* 147: 342–354

Kenton A, Dickie JB, Langton DH, Bennett MD (1990) Nuclear DNA amount and karyotype symmetry in *Cypella* and *Hesperoxiphion* (Tigridieae: Iridaceae). *Evolutionary Trends in Plants* 4: 59–69

- Levitsky, GA (1931) The karyotype in systematics. *Bulletin of Applied Botany, Genetics and Plant Breeding* 27: 220–240.
- Mandáková T, Lysak MA (2008) Chromosomal Phylogeny and Karyotype Evolution in $x=7$ Crucifer Species (Brassicaceae). *Plant Cell* 20:2559–2570. doi: 10.1105/tpc.108.062166
- McKain MR, McNeal JR, Kellar PR, et al (2016) Timing of rapid diversification and convergent origins of active pollination within Agavoideae (Asparagaceae). *Am J Bot* 103:1717–1729. doi: 10.3732/ajb.1600198
- McKain MR, Wickett N, Zhang Y, et al (2012) Phylogenomic analysis of transcriptome data elucidates co-occurrence of a paleopolyploid event and the origin of bimodal karyotypes in Agavoideae (Asparagaceae). *Am J Bot* 99:397–406. doi: 10.3732/ajb.1100537
- Medeiros-Neto E, Nollet F, Moraes AP, Felix LP (2017) Intrachromosomal karyotype asymmetry in Orchidaceae. *Genet Mol Biol* 40:610–619. doi: 10.1590/1678-4685-GMB-2016-0264
- Moraes AP, Souza-Chies TT, Stiehl-Alves EM, et al (2015) Evolutionary trends in Iridaceae: new cytogenetic findings from the New World. *Bot J Linn Soc* 177:27–49
- Murin A, Majovsky J (1976) IOPB Chromosome number reports LIII. *Taxon* 25: 483–500
- Oliver RL, Lewis WH. 1962. Chromosome numbers of *Sisyrinchium* (Iridaceae) in eastern North America. *Sida* 1:43–48
- Palomino G, Martínez J, Cepeda-Cornejo V, Pimienta-Barrios E (2012) Nuclear genome size and cytotype analysis in *Agave cupreata* Trel. & Berger (Agavaceae). *Caryologia* 65:281–294. doi: 10.1080/00087114.2012.752915
- Pedrosa A, Jantsch MF, Moscone EA, et al (2001) Characterisation of pericentromeric and sticky intercalary heterochromatin in *Ornithogalum longibracteatum* (Hyacinthaceae). *Chromosoma* 110:203–213. doi: 10.1007/s004120000125

Pedrosa Andrea, Sandal Niels, Stougaard Jens, et al (2002) Chromosomal Map of the Model Legume *Lotus japonicus*. *Genetics* 161:1661–1672

Peruzzi L, Eroğlu HE (2013) Karyotype asymmetry: again, how to measure and what to measure? *Comp Cytogenet* 7:1–9. doi: 10.3897/CompCytogen.v7i1.4431

Pires C, Maureira I, Givnish T, et al (2006) Phylogeny, Genome Size, and Chromosome Evolution of Asparagales. *Aliso* 22:287–304. doi: 10.5642/aliso.20062201.24

Ribeiro T, dos 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. doi: 10.1007/s10709-011-9615-8

Rudall (1986) *An Anatomical and Chromosomal Investigation of Sisyrinchium and Allied Genera* Author(s): Paula Rudall, A. Y. Kenton and T. J. Lawrence Reviewed work(s): *Bot Gaz* 147:466–477

Schubert I, Lysak MA (2011) Interpretation of karyotype evolution should consider chromosome structural constraints. *Trends Genet* 27:207–216. doi: 10.1016/j.tig.2011.03.004

Shirakawa J, Nagano K, Hoshi Y (2012) Polyploid genome structure of *Drosera spatulata* complex (Droseraceae). *Cytologia (Tokyo)* 77:97–106

Stace CA (2000) Cytology and Cytogenetics as a Fundamental Taxonomic Resource for the 20th and 21st Centuries. *Taxon* 49:451–477. doi: 10.2307/1224344

Stebbins GL (1971) Chromosomal evolution in higher plants. *Chromosom Evol High Plants*

Vosa CG (2005) On chromosome uniformity, bimodality and evolution in the tribe Aloineae (Asphodelaceae). *Caryologia* 58:83–85. doi: 10.1080/00087114.2005.10589437

Weiss-Schneeweiss, H, Schneeweiss, GM (2013) Karyotype Diversity and Evolutionary Trends in Angiosperms. In: Greilhuber, J., Dolezel, J., Wendel, J.F. (eds) Plant Genome Diversity Volume 2, Physical Structure, Behaviour and Evolution of Plant Genomes. Springer, Vienna, pp 209-230.

Zarco CR (1986) A New Method for Estimating Karyotype Asymmetry. *Taxon* 35:526–530. doi: 10.2307/1221906

Tabela 1. Espécies da subfamília Iridoideae incluídas na árvore filogenética de Moraes et al. (2015), número cromossômico e nível de ploidia, valor 2C do tamanho do genoma, tamanho do genomas haploide (1Cx), índice de assimetria cariotípica (A₂) e presença de cariótipo bimodal. Em negrito os dados novos.

Espécie	2n (nível ploidia)	2C	1Cx	A ₂	Bimodal	Referência
Tigridieae Clado A						
<i>Ainea konzattii</i> (R.C.Foster) Ravenna	28 (4x)			-		AOC02
<i>Calydorea</i> Herb.						
<i>C. alba</i> Roitman & A. Castillo	14 (2x)	8,34	4,17	-	-	M15
<i>C. approximata</i> R.C.Foster	14 (2x)	4,58	2,29	0,26	Sim	M15/PT
<i>C. azurea</i> Klatt	28 (4x)	-	-	-	-	GT97
<i>C. basáltica</i> Ravenna	14 (2x)	7,72	3,86	-	-	M15
<i>C. campestris</i> Baker	14 (2x)	6,96	3,48	0,28	Sim	M15/PT
<i>C. crocoides</i> Ravenna	14 (2x)	6,90	3,45	-	-	M15
<i>C. minima</i> Roitman & A. Castillo	-	-	-	-	-	-
<i>C. pallens</i> Griseb.	28 (4x)	-	-	-	-	GT97
<i>C. undulata</i> Ravenna	14 (2x)	-	-	-	-	TRB08
<i>C. xiphioides</i> (Poepp.) Espinosa	42 (6x)			-		GT97
<i>Catila amabilis</i> Ravenna	14 (2x)	4,08	2,04	-	-	GT97
<i>Cipura</i> Aubl.						
<i>C. paludosa</i> Aubl.	14 (2x)	16.9	4.22	0.50	Sim	A11
<i>C. campanulata</i> Ravenna	-	-	-	-	-	-
<i>Cypella</i> Herb						
<i>C. aquatilis</i> Ravenna	14 (2x)	4,78	2,39	-	-	K90
<i>C. hauthallii</i> Ravenna	14 (2x)	4,18	2,09	0,27	Sim	M15/PT
<i>C. herbertii</i> Herb.	14 (2x)	3,90	1,95	0,30	Sim	M15/PT
<i>C. osteniana</i> Beauverd	-	-	-	-	-	-
<i>C. pusilla</i> (Link & Otto) Benth. &Hook.f. ex. B.D. Jacks	-	-	-	-	-	-
<i>Herbertia</i> Sweet						

<i>H. darwinii</i> Roitman & A. Castillo	14 (2x)	4,60	2,30	0,35	Sim	M15/PT
<i>H. lahue</i> (Molina) Goldblatt	56 (8x)	15,43	1,93	-	Sim	M15
<i>H. aff. Lahue</i> Molina) Goldblatt	-	-	-	-	-	-
<i>H. pulchella</i> Sweet	28 (4x)	8,60	2,15	0,29	Sim	GT97/PT
<i>H. quareimana</i> Ravenna	28 (4x)	8,70	2,17	0,28	Sim	M15/PT
<i>H. tigridioides</i> Goldblatt	-	-	-	-	-	-
<i>H. zebrina</i> Deble	14 (2x)	4,69	2,34	0,30	Sim	M15/PT
<i>Kelissa brasiliensis</i> (Baker) Ravenna	14 (2x)	4,73	2,36	0,35	Sim	M15/PT
<i>Larentia</i> Klatt						
<i>L. mexicana</i> (C.V.Morton & R.C.Foster) Goldblatt	-	-	-	-	-	-
<i>L. rosei</i> (R.C. Foster) Ravenna	-	-	-	-	-	-
<i>Nemastylis</i> Nutt.						
<i>N. geminiflora</i> Small	56 (8x)	-	-	-	-	G82
<i>N. tenuis</i> (S.Watson)	14 (2x)	-	2,17*	0,28	Sim	G82/PT
<i>Onira unguiculata</i> (Baker) Ravenna	14 (2x)	4,57	2,28	0,28	Sim	M15/PT
Tigridieae CladoB						
<i>Alophia</i> Herb.						
<i>A. sylvestris</i> (Loes) Goldblatt	28 (2x)	-	2,54*	0,50	Sim	M15/PT
<i>A. veracruzana</i> Goldblatt&Howard	28 (2x)	-	-	-	-	-
<i>Cardenanthus vargasii</i> R.C. Foster	-	-	-	-	-	-
<i>Cardiostigma</i> Baker						
<i>C. longispathum</i> (Herb.) Baker	-	-	-	-	-	-
<i>C. mexicanum</i> (R.C.Foster) Ravenna	-	-	-	-	-	-
<i>Cobana guatemalensis</i> (Standl.) Ravenna	28 (2x)	-	-	-	-	GT97
<i>Eleutherine</i> Herb.						
<i>E. bulbosa</i> (Mill.) Urb.	12 (2x)	2,50	1,25*	0,63	Sim	PT/A11

<i>E. latifolia</i> (Standl & L.O.Williams) Ravenna	12 (2x)	2,80	1,40*	0,55	Sim	PT
<i>Ennealophus euryandrus</i> (Griseb.) Ravenna	14 (2x)	-	0,95*	0,83	Sim	GT97
<i>Gelasine</i> Herb.						
<i>G. coerulea</i> Ravenna	14 (2x)	5,05	2,53	0,18	Não	M15
<i>G. elongata</i> (Graham) Ravenna	12 (2x)	3.50	1.74	0.37	Sim	BL12
<i>Hesperoxiphion peruvianum</i> (Baker) Baker	14 (2x)	2,70	1,35	-	-	K90
<i>Mastigostyla</i> I.M Johnst.						
<i>M. cyrtophylla</i> I.M Johnst.	-	-	-	-	-	-
<i>Mastigostyla sp01</i> I.M Johnst.	-	-	-	-	-	-
<i>Phalocallis coelestis</i> (Lehm.) Ravenna	10 (2x)	7,21	3,60	-	Sim	GT97/M15
<i>Sessilanthera</i> Molseed & Cruden						
<i>S. heliantha</i>	-	-	-	-	-	-
<i>S. latifolia</i> (Weath.) Molseed & Cruden	28 (2x)	-	-	-	-	G82
<i>Tigridia</i> Juss.						
<i>T. alpestris subsp. alpestris</i> Molseed	28 (2x)	-	-	-	-	-
<i>T. alpestris subsp. Obtusa</i> Molseed	28 (2x)	7,75	3,88	0,47	Sim	BL12/AM17
<i>T. convoluta</i>	28 (2x)	-	-	-	-	-
<i>T. chiapensis</i> Molseed	28 (2x)	-	-	-	-	-
<i>T. flammea</i> (Lindl.) Ravenna	28 (2x)	-	-	-	-	-
<i>T. immaculata</i> Ravenna	28 (2x)	-	-	-	-	-
<i>T. meleagris</i> (Lindl.) G. Nicholson	28 (2x)	-	-	-	-	G82
<i>T. oaxacana</i> Johnst	28 (2x)	-	-	-	-	-
<i>T. orthantha</i> (Lem.) Ravenna	28 (2x)	-	-	-	-	G82
<i>T. pavonia</i> (L.f.) DC.	28 (2x)	8,70	4,35	0,42	Sim	BL12/AM17
<i>T. philippiana</i> Johnst	28 (2x)	-	-	-	-	-

<i>O. chimboracensis</i> (Kunth) Baker	54 (5x)	-	0,24*	0,19	Não	GT97/PT
<i>O. monadelphus</i> Ravenna	-	-	-	-	-	-
<i>O. multiflorus</i> Sweet	-	-	-	-	-	-
<i>Sisyrinchium</i> L.						
<i>S. aff. pusillum</i> OC.	-	-	-	-	-	-
<i>S. angustifolium</i> Miller	96 (12x)	-	-	-	-	MM76
<i>S. californicum</i> (Ker. Gaswl.) Dryand.	32 (4x)	1,40	0,35	-	-	KRJ86
<i>S. chilensis</i> Hook.	16 (2x)	1,46	0,73	-	-	KH84
<i>S. convolutum</i> Nooca	36 (4x)	1,90	0,48	0,17	Não	KH84/PT
<i>S. micranthum</i> Cav.	32 (4x)	-	-	-	-	G82
<i>S. minus</i> Engelm. & A.Gray	10 (x?)	-	-	-	-	OL62
<i>S. palmifolium</i> L.	-	-	-	-	-	-
<i>S. sellowianum</i> Klatt	-	-	-	-	-	-
<i>S. striatum</i> Sm.	18 (2x)	2,50	1,25	-	-	KRJ86
<i>S. uliginosum</i> Ravenna	-	-	-	-	-	-
<i>Solenomelus</i> Miers.						
<i>S. pedunculatus</i>	-	-	-	-	-	-
<i>S. segethi</i> Kuntze	48 (2x)	-	-	-	-	RKL86

* valores do tamanho do genoma monoploide estimados baseados no comprimento do complemento cromossômico total e o comprimento total e tamanho do genoma das espécies filogeneticamente próximas.

Referências: A11: Alves et al. (2011), AM17: Arroyo Martinez et al. (2017), AOC02: Aaron e Ortiz-Catedral (2002), B00: Blanchon et al. (2000), BL12: Bennett e Leich (2012), CG08: Chukr e Giulietti (2008), G82: Goldblatt (1982), GT97: Goldblatt e Takei (1997), K90: Kenton et al. (1990), KH84: Kenton e Heywood (1984), KRJ86: Kenton et al. (1986), MM76: Murin e Majovsky (1976), M15: Moraes et al. (2015), OL62: Oliver e Lewis (1962), RKL86: Rudall et al. (1986), TRB08: de Tullio et al. (2008), PT: presente trabalho.

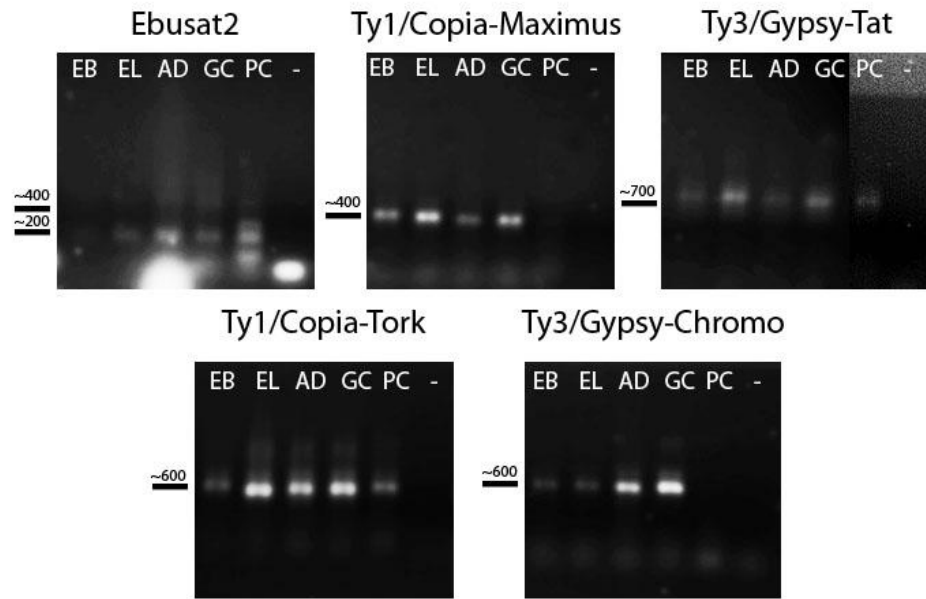


Figura 1. Produtos de PCR das sequências repetitivas características de *E. bulbosa* em espécies relacionadas. EB = *Eleutherine bulbosa* (controle positivo), EL = *Eleutherinelatifolia*, AD= *Alophia drummondii*, GC = *Gelasine coerulea*, PC = *Phalocalliscoeletis*. Barras pretas correspondem ao tamanho do fragmento específico para cada família repetitiva em pares de bases.

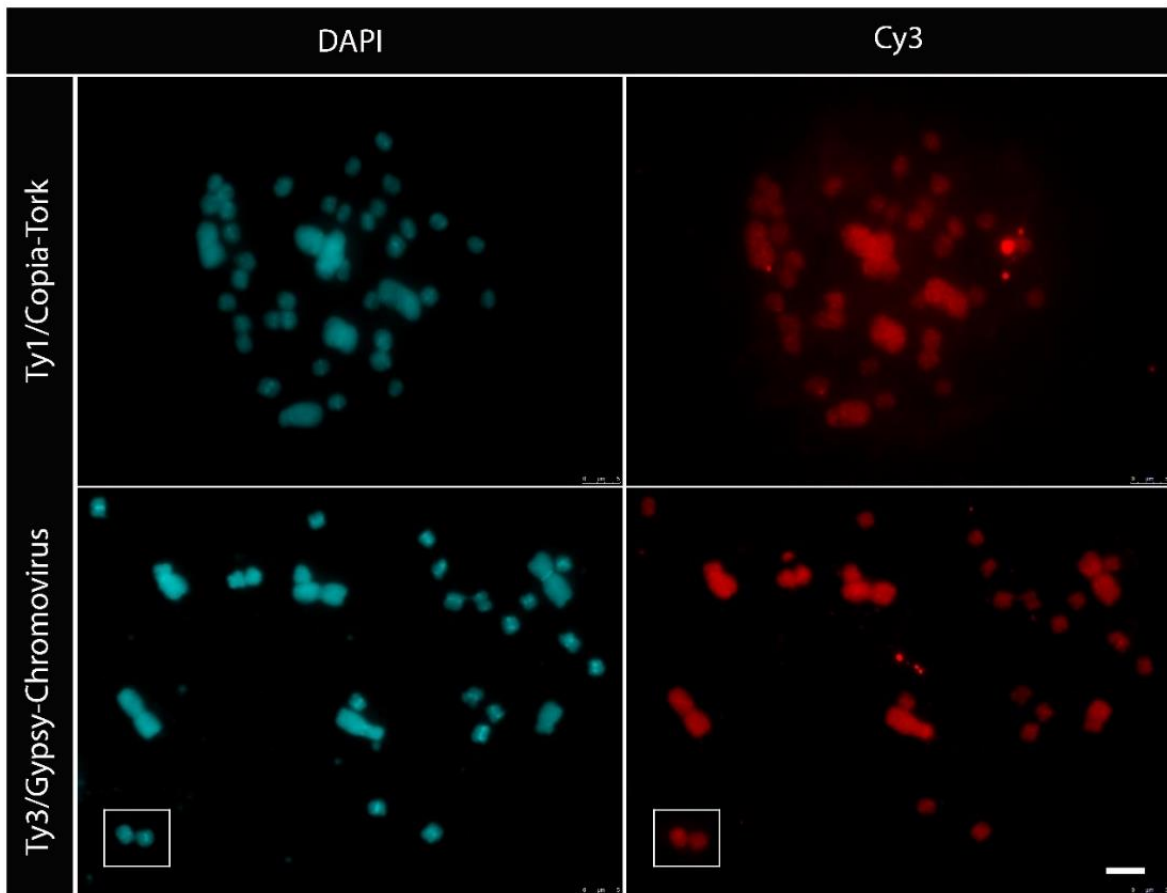


Figura 2.Distribuição cromossômica dispersa uniforme de duas famílias de LTR-RE em cromossomos metafásicos de *Alophia drummondii*. Insertos mostram cromossomos fora do campo central da célula e a barra corresponde a 5 µm.

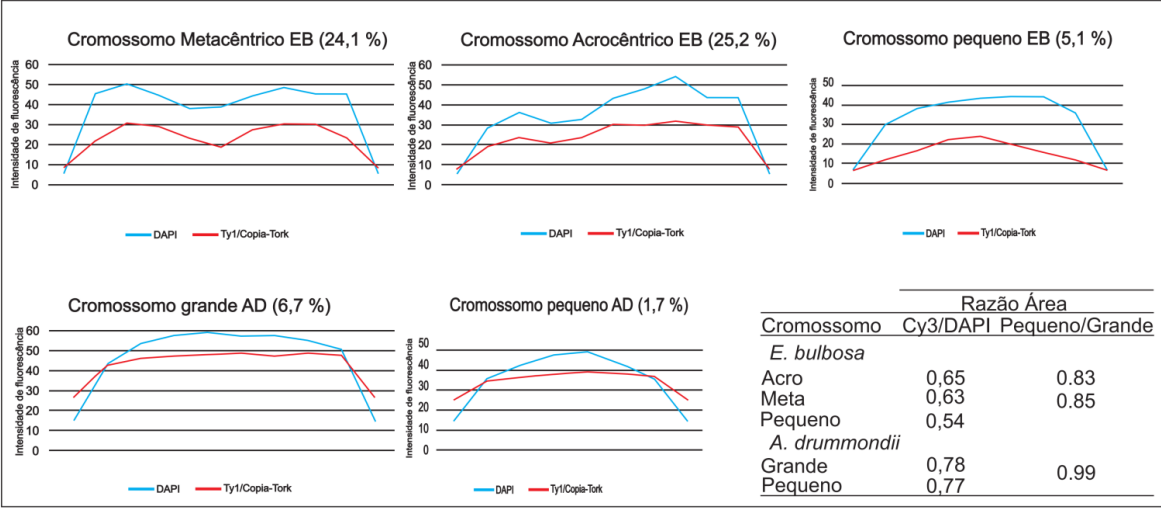


Figura 3.Distribuição cromossômica da intensidade de fluorescência do LTR-retrotransposon Ty1/Tork ao longo dos cromossomos grandes e pequenos de *E. bulbosa*(EB) e *A. drummondii*(AD). Gráficos mostram medidas lineares ao longo de cada cromossomo. Porcentagens nos gráficos mostram a proporção do conteúdo de DNA de cada cromossômico em relação ao complemento diploide da espécie. Tabela mostra valores médios das razões entre a intensidade do sinal (Cy3) e do contracorante (DAPI) e entre essas razões para os tipos cromossômicos (grande e pequeno), considerando a área dos cromossomos medidos.

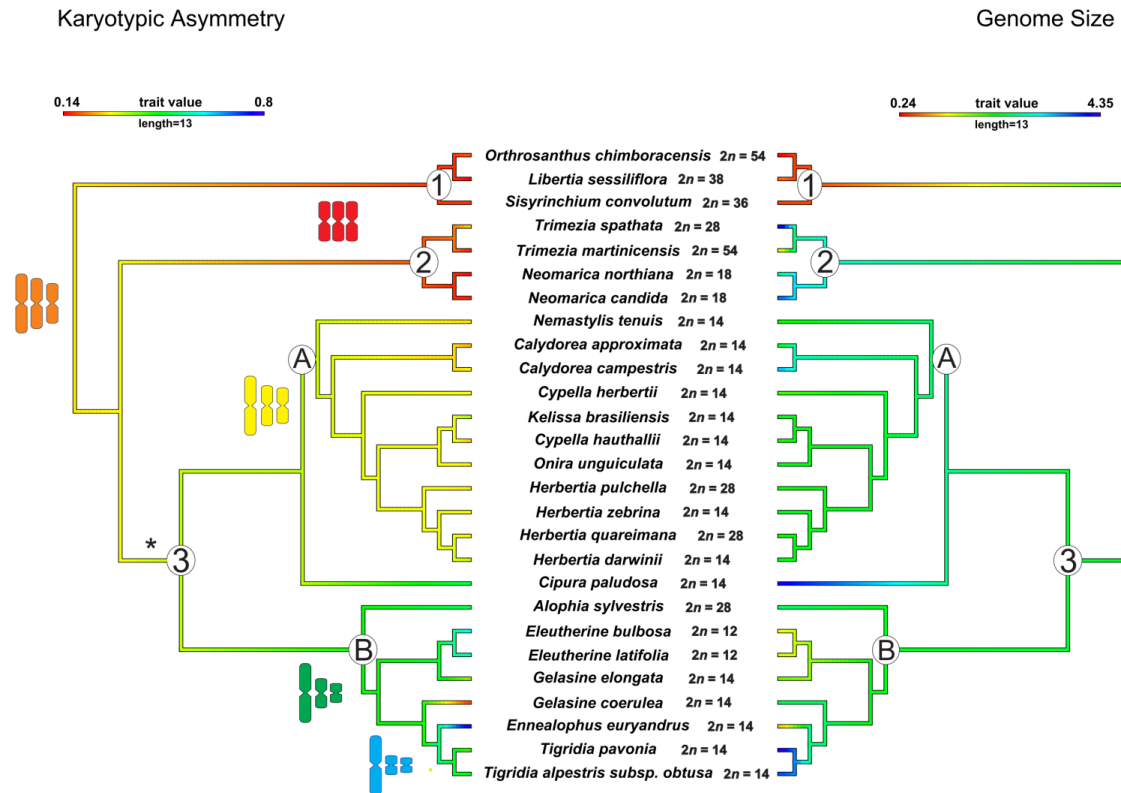


Figura 4. Evolução da assimetria cariotípica e o tamanho do genoma haploide (1Cx) na subfamília Iridoideae (Iridaceae) com destaque para a tribo Trigrídieae. Reconstrução do caráter ancestral para a assimetria cariotípica (direita) e o tamanho do genoma haploide (esquerda) por máxima parcimônia baseada na filogenia de Moraes *et al.* (2015). Clado 1 corresponde à tribo Sisyrinchieae, clado 2 à tribo Trimezieae, clado 3 à tribo Tigrídieae com seus subclados A da subtribo Cipurinae e o subclado B da subtribo Tigrídiinae. Representações esquemáticas dos cromossomos dos cariótipos simétricos (vermelho, laranja) e dos diferentes níveis de assimetria (amarelo, verde e azul) se mostram nos respectivos clados onde ocorrem esses tipos cariotípicos. Asterisco se mostra o aumento da assimetria cariotípica com surgimento do cariótipo biomdial.

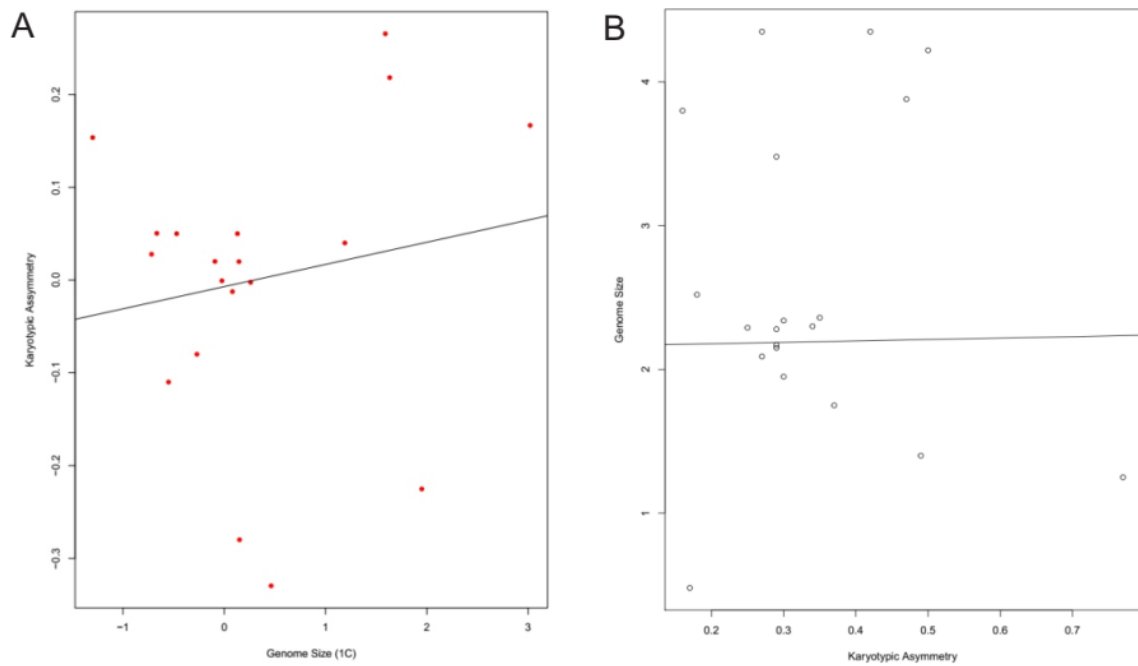


Figura 5. Gráficos de correlação estatística da evolução dos caracteres assimetria cariotípica e tamanho do genoma haploide. A, análises de PICs mostrando no eixo y os valores de contraste da assimetria cariotípica e no eixo x os valores de contraste do tamanho do genoma monoploide. B, análises de PGLS mostrando no eixo y os valores do tamanho do genoma monoploide e no eixo x os valores da assimetria cariotípica.

4 CONCLUSÃO

- Distribuição das sequências repetitivas é cromossomo-específica com acúmulo de sequências nos cromossomos grandes do complemento
- Distribuição da cromatina de marcas eucromáticas e da região pericentromérica, mostra uma organização diferencial entre cromossomos grande e pequenos
- Organização cromossômica de sequências e da cromatina evidenciam uma estruturação relacionada à condição bimodal do cariótipo, mostrando comportamento de subgenomas dentro de uma mesma espécie
- Estruturação bimodal em subgenomas se mantém nas espécies dentro do gênero *Eleutherine*
- Estruturação bimodal, com acúmulo nos cromossomos grandes, não se mantém no gênero *Alophia*, mas distante filogeneticamente
- O cariótipo bimodal surgiu na tribo Tigridaeae dentro da subfamília Iridoideae
- Evolução da assimetria cariotípica não está correlacionada com a evolução do tamanho do genoma monoploide
- Evolução do cariótipo bimodal estaria relacionada com um equilíbrio entre amplificação de sequências nos cromossomos grandes e eliminação delas nos cromossomos pequenos
- Evolução do cariótipo bimodal, envolvendo o mecanismo de amplificação/eliminação de sequências, seria diferentes dentro dos diferentes gênero na subfamília Iridoideae

REFERÊNCIAS

- ALVES, L. I. F.; LIMA, S. A. A.; FELIX, L. P. Chromosome characterization and variability in some Iridaceae from Northeastern Brazil. **Genetics and Molecular Biology**, v. 34, n. 2, p. 259–267, 2011.
- ANDERSON, A. R. et al. The latitudinal cline in the In(3R)Payne inversion polymorphism has shifted in the last 20 years in Australian *Drosophila melanogaster* populations: CLINAL VARIATION FOR IN(3R)PAYNE IN DROSOPHILA. **Molecular Ecology**, v. 14, n. 3, p. 851–858, 7 fev. 2005.
- ANDREA, L.; MARINI, M.; MANTOVANI, B. Non-concerted evolution of the RET76 satellite DNA family in Reticulitermes taxa (Insecta, Isoptera). **Genetica**, v. 128, n. 1–3, p. 123–132, 1 set. 2006.
- BARROS E SILVA, A. E. et al. The evolution of CMA bands in Citrus and related genera. **Chromosome Research**, v. 18, n. 4, p. 503–514, jun. 2010.
- BENNETZEN, J. L.; WANG, H. The Contributions of Transposable Elements to the Structure, Function, and Evolution of Plant Genomes. **Annual Review of Plant Biology**, v. 65, n. 1, p. 505–530, 29 abr. 2014.
- BISCOTTI, M. A.; OLMO, E.; HESLOP-HARRISON, J. S. (PAT). Repetitive DNA in eukaryotic genomes. **Chromosome Research**, v. 23, n. 3, p. 415–420, 1 set. 2015.
- CHALOPIN, D. et al. Comparative Analysis of Transposable Elements Highlights Mobilome Diversity and Evolution in Vertebrates. **Genome Biology and Evolution**, v. 7, n. 2, p. 567–580, 1 fev. 2015.
- CHASE, M. W.; REVEAL, J. L. A phylogenetic classification of the land plants to accompany APG III. **Botanical Journal of the Linnean Society**, v. 161, n. 2, p. 122–127, 1 out. 2009.
- CHAUVEAU, O. et al. Evolution of oil-producing trichomes in *Sisyrinchium* (Iridaceae): insights from the first comprehensive phylogenetic analysis of the genus. **Annals of Botany**, v. 107, n. 8, p. 1287–1312, 1 jun. 2011.
- COHEN, S.; SEGAL, D. Extrachromosomal Circular DNA in Eukaryotes: Possible Involvement in the Plasticity of Tandem Repeats. **Cytogenetic and Genome Research**, v. 124, n. 3–4, p. 327–338, 2009.
- COONEY, C. A.; MATTHEWS, H. R.; BRADBURY, E. M. 5-Methyldeoxycytidine in the Physarum minichromosome containing the ribosomal RNA genes. **Nucleic Acids Research**, v. 12, n. 3, p. 1501–1516, 10 fev. 1984.
- COSTA, L. et al. Comparative cytomolecular analyses reveal karyotype variability related to biogeographic and species richness patterns in Bombacoideae (Malvaceae). **Plant Systematics and Evolution**, v. 303, n. 9, p. 1131–1144, nov. 2017.

DE LA HERRÁN, R. et al. A heterochromatic satellite DNA is highly amplified in a single chromosome of *Muscari* (Hyacinthaceae). **Chromosoma**, v. 110, n. 3, p. 197–202, jul. 2001.

ESTEP, M. C.; DEBARRY, J. D.; BENNETZEN, J. L. The dynamics of LTR retrotransposon accumulation across 25 million years of panicoid grass evolution. **Heredity**, v. 110, n. 2, p. 194, fev. 2013.

FEITOZA, L.; GUERRA, M. Different types of plant chromatin associated with modified histones H3 and H4 and methylated DNA. **Genetica**, v. 139, n. 3, p. 305–314, mar. 2011.

FELINER, G. N.; ROSSELLÓ, J. A. Concerted Evolution of Multigene Families and Homoeologous Recombination. In: **Plant Genome Diversity Volume 1**. [s.l.] Springer, Vienna, 2012. p. 171–193.

GOLDBLATT, P. Chromosome Cytology in Relation to Suprageneric Systematics of Neotropical Iridaceae. **Systematic Botany**, v. 7, n. 2, p. 186–198, 1982.

GOLDBLATT, P. et al. Iridaceae ‘Out of Australasia’? Phylogeny, biogeography, and divergence time based on plastid DNA sequences. **Systematic Botany**, v. 33, n. 3, p. 495–508, 2008.

GOLDBLATT, P.; SNOW, N. Systematics and Chromosome Cytology of Eleutherine Herbert (Iridaceae). **Annals of the Missouri Botanical Garden**, v. 78, n. 4, p. 942–949, 1991.

GRAY, Y. H. It takes two transposons to tango: transposable-element-mediated chromosomal rearrangements. **Trends in Genetics**, v. 16, n. 10, p. 461–468, 2000.

GUERRA, M. DOS S. Mitotic and meiotic analysis of a pericentric inversion associated with a tandem duplication in *Eleutherine bulbosa*. **Chromosoma**, v. 97, n. 1, p. 80–87, 1 set. 1988.

HESLOP-HARRISON, J. S. P.; SCHWARZACHER, T. Organisation of the plant genome in chromosomes: Organisation of the plant genome in chromosomes. **The Plant Journal**, v. 66, n. 1, p. 18–33, abr. 2011.

HOBZA, R. et al. Impact of repetitive DNA on sex chromosome evolution in plants. **Chromosome Research**, v. 23, n. 3, p. 561–570, 1 set. 2015.

HOFFMANN, A. A.; RIESEBERG, L. H. Revisiting the Impact of Inversions in Evolution: From Population Genetic Markers to Drivers of Adaptive Shifts and Speciation? **Annual Review of Ecology, Evolution, and Systematics**, v. 39, n. 1, p. 21–42, dez. 2008.

HU, T. T. et al. The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. **Nature Genetics**, v. 43, n. 5, p. 476, maio 2011.

- KAPITONOV, V. V.; JURKA, J. Rolling-circle transposons in eukaryotes. **Proceedings of the National Academy of Sciences**, v. 98, n. 15, p. 8714–8719, 17 jul. 2001.
- KAPITONOV, V. V.; JURKA, J. Self-synthesizing DNA transposons in eukaryotes. **Proceedings of the National Academy of Sciences of the United States of America**, v. 103, n. 12, p. 4540–4545, 2006.
- KIRKPATRICK, M. Chromosome Inversions, Local Adaptation and Speciation. **Genetics**, v. 173, n. 1, p. 419–434, 17 mar. 2006.
- KIRKPATRICK, M. How and Why Chromosome Inversions Evolve. **PLoS Biology**, v. 8, n. 9, p. e1000501, 28 set. 2010.
- KLEMME, S. et al. High-copy sequences reveal distinct evolution of the rye B chromosome. **New Phytologist**, v. 199, n. 2, p. 550–558, 1 jul. 2013.
- KOVARIK, A. et al. Evolution of rDNA in Nicotiana Allopolyploids: A Potential Link between rDNA Homogenization and Epigenetics. **Annals of Botany**, v. 101, n. 6, p. 815–823, 1 abr. 2008.
- LEE, J. et al. Rapid amplification of four retrotransposon families promoted speciation and genome size expansion in the genus Panax. **Scientific Reports**, v. 7, n. 1, dez. 2017.
- LEVIN, D. A. **The role of chromosomal change in plant evolution**. Oxford ; New York: Oxford University Press, 2002.
- LOWRY, D. B.; WILLIS, J. H. A Widespread Chromosomal Inversion Polymorphism Contributes to a Major Life-History Transition, Local Adaptation, and Reproductive Isolation. **PLOS Biology**, v. 8, n. 9, p. e1000500, 28 set. 2010.
- MACAS, J.; NEUMANN, P.; NAVRÁTILOVÁ, A. Repetitive DNA in the pea (*Pisum sativum* L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and *Medicago truncatula*. **BMC Genomics**, v. 8, n. 1, p. 427, 2007.
- MANNING, J. C.; GOLDBLATT, P. Iridaceae. **Bothalia**, v. 41, n. 1, p. 171–175, 13 dez. 2011.
- MARQUES, A. et al. Holocentromeres in *Rhynchospora* are associated with genome-wide centromere-specific repeat arrays interspersed among euchromatin. **Proceedings of the National Academy of Sciences**, v. 112, n. 44, p. 13633–13638, 3 nov. 2015.
- MASLOVA, A. et al. Three-dimensional architecture of tandem repeats in chicken interphase nucleus. **Chromosome Research**, v. 23, n. 3, p. 625–639, 1 set. 2015.
- MCKAIN, M. R. et al. Phylogenomic analysis of transcriptome data elucidates co-occurrence of a paleopolyploid event and the origin of bimodal karyotypes in

Agavoideae (Asparagaceae). **American Journal of Botany**, v. 99, n. 2, p. 397–406, 1 fev. 2012.

MORAES, A. P. et al. Evolutionary trends in Iridaceae: new cytogenetic findings from the New World. **Botanical Journal of the Linnean Society**, v. 177, n. 1, p. 27–49, 2015a.

MORAES, A. P. et al. Evolutionary trends in Iridaceae: new cytogenetic findings from the New World: Evolutionary trends in Iridoideae. **Botanical Journal of the Linnean Society**, v. 177, n. 1, p. 27–49, jan. 2015b.

PALACIOS-GIMENEZ, O. M. et al. High-throughput analysis of the satellitome revealed enormous diversity of satellite DNAs in the neo-Y chromosome of the cricket *Eneoptera surinamensis*. **Scientific Reports**, v. 7, n. 1, dez. 2017.

PELLICER, J. et al. Genome size dynamics in tribe Gilliesieae (Amaryllidaceae, subfamily Allioideae) in the context of polyploidy and unusual incidence of Robertsonian translocations. **Botanical Journal of the Linnean Society**, v. 184, n. 1, p. 16–31, 1 maio 2017.

PERUZZI, L.; EROĞLU, H. E. Karyotype asymmetry: again, how to measure and what to measure? **Comparative Cytogenetics**, v. 7, n. 1, p. 1–9, 14 mar. 2013.

PLOHL, M.; MEŠTROVIĆ, N.; MRAVINAC, B. Satellite DNA Evolution. v. 7, p. 126–152, 2012.

RENNY-BYFIELD, S. et al. Next Generation Sequencing Reveals Genome Downsizing in Allotetraploid *Nicotiana tabacum*, Predominantly through the Elimination of Paternally Derived Repetitive DNAs. **Molecular Biology and Evolution**, v. 28, n. 10, p. 2843–2854, 1 out. 2011.

RIBEIRO, T. et al. Evolutionary dynamics of satellite DNA repeats from *Phaseolus* beans. **Protoplasma**, v. 254, n. 2, p. 791–801, mar. 2017a.

RIBEIRO, T. et al. Centromeric and non-centromeric satellite DNA organisation differs in holocentric *Rhynchospora* species. **Chromosoma**, v. 126, n. 2, p. 325–335, mar. 2017b.

RUIZ-RUANO, F. J. et al. High-throughput analysis of the satellitome illuminates satellite DNA evolution. **Scientific Reports**, v. 6, n. 1, set. 2016.

SANTOS, F. C. et al. Chromosomal distribution and evolution of abundant retrotransposons in plants: gypsy elements in diploid and polyploid *Brachiaria* forage grasses. **Chromosome Research**, v. 23, n. 3, p. 571–582, set. 2015.

SCHUBERT, I. Chromosome evolution. **Current Opinion in Plant Biology**, Genome Studies and Molecular Genetics / Edited by Stefan Jansson and Edward S Buckler. v. 10, n. 2, p. 109–115, 1 abr. 2007.

SCHUBERT, I.; LYSAK, M. A. Interpretation of karyotype evolution should consider chromosome structural constraints. **Trends in Genetics**, v. 27, n. 6, p. 207–216, 1 jun. 2011.

SMITH, J. et al. Differences in gene density on chicken macrochromosomes and microchromosomes. **Animal Genetics**, v. 31, n. 2, p. 96–103, 1 mar. 2000.

SOUZA-CHIES, T. T. et al. Studies on diversity and evolution of Iridaceae species in southern Brazil. **Genetics and Molecular Biology**, v. 35, n. 4, p. 1027–1035, 2012.

STEBBINS, G. L. Chromosomal evolution in higher plants. **Chromosomal evolution in higher plants.**, 1971.

TAUTZ, D. Notes on the definition and nomenclature of tandemly repetitive DNA sequences. In: **DNA Fingerprinting: State of the Science**. Progress in Systems and Control Theory. [s.l.] Birkhäuser, Basel, 1993. p. 21–28.

TENAILLON, M. I.; HOLLISTER, J. D.; GAUT, B. S. A triptych of the evolution of plant transposable elements. **Trends in Plant Science**, v. 15, n. 8, p. 471–478, 1 ago. 2010.

VAIO, M. et al. Molecular phylogeny and chromosome evolution among the creeping herbaceous *Oxalis* species of sections Corniculatae and Ripariae (Oxalidaceae). **Molecular Phylogenetics and Evolution**, v. 68, n. 2, p. 199–211, 1 ago. 2013.

VAIO, M. et al. Phylogenetic and cytogenetic relationships among species of *Oxalis* section Articulatae (Oxalidaceae). **Plant Systematics and Evolution**, v. 302, n. 9, p. 1253–1265, nov. 2016.

VANBUREN, R.; MING, R. Dynamic transposable element accumulation in the nascent sex chromosomes of papaya. **Mobile Genetic Elements**, v. 3, n. 1, p. e23462, jan. 2013.

VAN-LUME, B. et al. Heterochromatic and cytomolecular diversification in the Caesalpinia group (Leguminosae): Relationships between phylogenetic and cytogeographical data. **Perspectives in Plant Ecology, Evolution and Systematics**, v. 29, p. 51–63, dez. 2017.

VIGUERA, E.; CANCEILL, D.; EHRLICH, S. D. Replication slippage involves DNA polymerase pausing and dissociation. **The EMBO Journal**, v. 20, n. 10, p. 2587–2595, 15 maio 2001.

WANG, J. et al. Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. **Proceedings of the National Academy of Sciences**, v. 109, n. 34, p. 13710–13715, 21 ago. 2012.

WARREN, I. A. et al. Evolutionary impact of transposable elements on genomic diversity and lineage-specific innovation in vertebrates. **Chromosome Research**, v. 23, n. 3, p. 505–531, 1 set. 2015.

WEISS-SCHNEEWEISS, H. et al. **Employing next generation sequencing to explore the repeat landscape of the plant genome.** [s.l: s.n.].

ZHANG, L. et al. The structure and retrotransposition mechanism of LTR-retrotransposons in the asexual yeast *Candida albicans*. **Virulence**, v. 5, n. 6, p. 655–664, 15 ago. 2014.

APÊNDICE A – ARTIGO PUBLICADO NA REVISTA FRONTIERS OF PLANT
SCIENCE

**Together but different: the subgenomes of the bimodal
Eleutherine karyotypes are differentially organized**

**Mariana Báez^{1*}, Magdalena Vaio², Steven Dreissig³, Veit Schubert³, Andreas Houben³,
Andrea Pedrosa-Harand¹**

¹Laboratory of Plant Cytogenetics and Evolution, Department of Botany, Federal University of Pernambuco, Brazil

²Laboratory of Genetics, Department of Plant Biology, College of Agronomy, University of the Republic, Uruguay

³Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany

***Corresponding Author:**

marian_iaris@hotmail.com

ORCID: 0000-0002-7874-6385

Key words: retrotransposons, satellite DNA, repetitive sequences accumulation, DNA replication, histone modification, inversion

Summary

Bimodal karyotypes are characterized by the presence of two sets of chromosomes of contrasting size. *Eleutherine bulbosa* ($2n = 12$) presents a bimodal karyotype with a large chromosome pair, which has a pericentric inversion in permanent heterozygosity with suppressed recombination, and five pairs of 3-4 times smaller chromosomes. Aiming to understand whether high copy number sequence composition differs between both chromosome sets, we investigated the repetitive DNA fraction of *E. bulbosa* and compared it to the chromosomal organization of the related *E. latifolia* species, not containing the pericentric inversion. We also compared the repetitive sequence proportions between the heteromorphic large chromosomes of *E. bulbosa* and between *E. bulbosa* and *E. latifolia* to understand the influence of the chromosome inversion on the dynamics of repetitive sequences. The most abundant repetitive families of the genome showed a similar chromosomal distribution in both homologues of the large pair and in both species, apparently not influenced by the species-specific inversions. The repeat families Ebusat1 and Ebusat4 are localized interstitially only on the large chromosome pair, while Ebusat2 is located in the centromeric region of all chromosomes. The four most abundant retrotransposon lineages are accumulated in the large chromosome pair. Replication timing and distribution of epigenetic and transcriptional marks differ between large and small chromosomes. The differential distribution of retroelements appears to be related to the bimodal condition and is not influenced by the nonrecombining chromosome inversions in these species. Thus, the large and small chromosome subgenomes of the bimodal *Eleutherine* karyotype are differentially organized and provably evolved by repetitive sequences accumulation on the large chromosome set.

Key words: retrotransposons, satellite DNA, repetitive sequences accumulation, DNA replication, histone modification, inversion

Introduction

Bimodal karyotypes are characterized by the presence of two sets of chromosomes of contrasting size. The origin of bimodal karyotypes is usually associated with one of the following processes; (i) chromosomal rearrangements involving fusion-fission events generate large chromosomes as fusion products of small chromosomes, or small chromosomes result from the fission of large chromosomes (Burt, 2002; Gen-Shen et al., 2014; Schubert and Lysak, 2011). (ii) The combination of different parental species (allopolyploidization) may combine species with different chromosome sizes (McKain et al., 2012; Shirakawa et al., 2012). (iii) The differential accumulation of repetitive sequences may increase the size of a subset of chromosomes (de la Herrán et al., 2001). Bimodal karyotypes are common within several animals groups, such as birds, reptiles and amphibians (Noronha et al., 2016; Masabanda et al., 2004; Stock and Mengden, 1975). Also, many plant genera, such as *Agave*, *Yucca*, *Hosta* (Akemine, 1935; Palomino et al., 2012; Watkins, 1936), *Aloe* (Brandham, 1998; Fentaw et al., 2013), and *Hypochaeris* (Fiorin et al., 2013) show bimodal karyotypes.

In animal bimodal karyotypes, gene content, the abundance of heterochromatic repetitive sequences, as well as the replication behaviour differ between both chromosome sets (McQueen et al., 1998; Smith et al., 2000). For instance, chicken microchromosomes are early replicating, harbour twice as many genes as macrochromosomes, and are associated with an increased gene transcriptional activity (McQueen et al., 1998). In contrast, in most bimodal plant groups the chromosome organization is largely unknown. The bimodal karyotypes of some Orchidaceae species contain large chromosomes with a higher proportion of C-banding-positive heterochromatin (D'Emerico et al., 1999). In *Ornithogalum longibracteatum* (Hyacinthaceae), one satellite DNA sequence (satDNA) is the major constituent of the heterochromatin of the large chromosomes (Pedrosa et al., 2001). A specific satDNA, found in *Muscari comosum* (Hyacinthaceae) is related to the heterochromatic bands of the large chromosomes, and it has been suggested to cause the increase of asymmetry of the karyotypes within this genus (de la Herrán et al., 2001). Independent of composition and origin, in both animal and plant bimodal species, it was suggested that the maintenance of these chromosome size differences could be related to the genome structure and function (Coullin et al., 2005; Griffin et al., 2015; Vosa, 2005).

Eleutherine (Iridaceae) is a Neotropical genus of the subfamily Iridoideae and comprises two species, both with bimodal karyotypes (Goldblatt and Snow, 1991). *Eleutherine bulbosa* ($2n = 12$) has one chromosome pair (chromosome I) which is 3-4 times larger than the other pairs. The large chromosome pair is heteromorphic due to an asymmetric pericentric inversion in heterozygosity, encompassing about 70% of the chromosome and resulting in one acrocentric and one metacentric homologue (Guerra, 1988). This pair contains two DAPI-positive heterochromatic bands. They are located interstitially in the long arm of the acrocentric and terminally in the short arm of the metacentric homologue. CMA-positive bands are located in the pericentromeric region of both homologues. The presence of rDNA sites is limited to chromosome pair I. While the 35S rDNA sites are located inside of the chromosomal inversion, the 5S rDNA sites are duplicated in the terminal region of the long arm of both chromosomes, outside of the inversion (Feitoza and Guerra, 2011). The second species of the genus, *E. latifolia*, has also a $2n = 12$ bimodal karyotype with a pair of large acrocentric chromosomes, but without an inversion (Goldblatt and Snow, 1991).

All small chromosomes of *E. bulbosa* are enriched in euchromatin marks, like acetylated histone H4K5 and dimethylated H3K4. In contrast, the large chromosome pair is 5-mC

hypermethylated (Feitoza and Guerra, 2011), showing a chromatin differentiation between both chromosome sets. Meiotic analysis showed that the inverted region of the large chromosome pair was devoid of recombination, with chiasmata observed only outside the inversion loop (Guerra, 1991). All analysed individuals and populations of *E. bulbosa* were heterozygous, and this heterozygosity is supposed to be fixed preferentially by asexual reproduction (Guerra, 1988; 1991).

The process of recombination is linked to the evolution of repetitive sequences, as observed for satellite DNA homogenization via gene conversion (Feliner and Rosselló, 2012). Furthermore, unequal recombination between homologous chromatids or illegitimate recombination were proposed as powerful mechanisms for removing repetitive sequences (Tenaillon et al., 2010) and decreasing genome size (Renny-Byfield et al., 2011). Thus, chromosomal regions devoid of recombination, such as inverted regions, could tend to accumulate different types of repetitive sequences, which may evolve differentially from the rest of the genome. Within this context, the 5-mC hypermethylation of *E. bulbosa* chromosome pair I, which contrasts to the small chromosome pairs, and the lack of recombination between the homologues of chromosome pair I in a large segment led to the following questions. Does the distribution of repetitive, epigenetics histone marks and timing of DNA replication differ between large and small chromosomes within a bimodal karyotype? Does the repetitive composition differ in the non-recombining, inverted region between the homologues of chromosome pair I? Is the distribution of repeats conserved between large and small chromosomes within the *Eleutherine* species?

Therefore, we describe the repeat composition and chromosome organization of *E. bulbosa*. The findings were compared to the sister species *E. latifolia*, also showing a bimodal karyotype but lacking the large chromosome inversion.

Materials and Methods

Materials

Plants of *E. bulbosa* were collected in Aracuara, Bahia State, Brazil (voucher number UFP 82763), and cultivated in the experimental garden of the Laboratory of Plant Cytogenetic and Evolution from the Federal University of Pernambuco, Recife, Brazil. Seeds of *E. latifolia* were kindly provided by Dr Guadalupe Munguía Linno from Guadalajara University, Mexico. Seeds were germinated in a wet chamber (3 - 4 months), the seedlings were transferred into soil and cultivated in a germination room at 24°C.

Genome size estimation

Samples were prepared from 40-50 mg of young leaves of *E. bulbosa* (Miller) Urban or *E. latifolia* (Standl & L.O.Williams) Ravenna in 1 ml of LB nuclear isolation buffer and filtered through a 30 µm nylon filter (Doležel et al., 2007). *Solanum lycopersicum* L. (2C = 1.96 pg) served as standard. Nuclei were stained with propidium iodide (50 µg mL⁻¹) and RNase (50 mg mL⁻¹) was added to prevent staining of double-stranded RNA. The nuclear DNA content was determined with a Partec CyFlow SL (Partec) flow cytometer, and results were analysed with Flomax program. For genome size estimations three replicates were analysed and the nuclear DNA content for each species was calculated according to the formula:

2C nuclear DNA content of the sample (pg)

= $\frac{\text{sample G0/G1}}{\text{reference standard G0/G1}}$ x 2C nuclear DNA content of the reference standard

reference standard G0/G1

Extraction of genomic DNA and DNA isolation from microdissected chromosomes

Total genomic DNA was extracted from young leaves of one individual of *E. bulbosa* using the DNAeasy Plant Mini Kit (Qiagen) according to manufacturer's instruction.

For microdissection of the two large homologous chromosomes, root tips were collected from bulbs, pre-treated in 8-hydroxyquinoline at 10°C for 24 h and fixed in 2 % formaldehyde for 15 min under vacuum. Fixed root tips were chopped in a nuclei isolation buffer (Doležel et al., 2007) and filtered through a 30 µm nylon membrane. The cell solution was centrifuged onto a microscopic slide at 2,000 rpm for 10 minutes (Shandon, CytoSpin3). Slides, mounted in 4', 6-diamidino-2-phenylindole (DAPI), were hit with a metal needle to physical separate chromosomes from broken cells. Ten chromosomes of each large chromosome homologue were isolated by microdissection with a glass needle, using a Zeiss Axio Zoom.V16 microscope coupled to an AxioCam 289 MRc5 digital camera (Zeiss) and the aureka® microsampling platform (aura-optik). The chromosomal DNA was amplified by multiple displacement amplification according to the protocol described in Dreissig et al., (2015) with minor modifications. Briefly, chromosomes were collected in 0.5 µl H₂O and 1 µl sample buffer (GE Healthcare, Genomiphi V2), then were incubated in alkaline lysis buffer (Gole et al., 2013) and 0.1 µg/µl of proteinase K (Sigma) at 37°C for 1 h, followed by heat inactivation at 65°C for 10 min. After incubation, 0.5 µl of neutralizing buffer (Gole et al., 2013) was added and the samples were left on ice while a master mix was prepared (3.5 µl of sample buffer, 4.5 µl reaction buffer and 0.5 µl of enzyme mix, Genomiphi V2 GE Healthcare). The samples were incubated at 30°C for 8 h followed by heat inactivation at 65°C for 10 min, and then cooled down to 4°C and kept at -20°C. A primer pair specific for CL29 of *E. bulbosa*, an LTR Ty3/Gypsy-Tat repeat (Online Resource 1), was used to check whether the generation of chromosome-derived DNA was successful.

Next Generation Sequencing and sequences analysis

Genomic DNA and chromosome microdissection-derived DNA were used for paired-end, 100 bp reads, and single-end, 250 bp reads, Illumina sequencing, respectively (Genbank Bioproject PRJNA549830). The repetitive fraction analysis was performed with 400 Mbp of reads of the genomic DNA (0.32x genome coverage) and 2,160 Mbp for each homologue of the large chromosome pair (~8x genome coverage). Sequenced reads were analyzed with the similarity-based read clustering method, implemented in the RepeatExplorer pipeline (Novak et al., 2013). Reads were filtered by quality with the default sets (quality cut-off value= 10, within a 95% of the bases in the sequence) and genomic paired-end reads were joined with the interlaced tool. For single-end reads, datasets from both chromosome types, were assigned a unique identifier and joined into a single dataset with the concatenate tool. For both datasets (genomic and chromosomes) clustering was performed with a minimum overlap of 55% and a similarity of 90%. For sequences of microdissected chromosomes, three independent analyses were performed, using a different dataset of reads of the same sequencing, to confirm the

proportions of each cluster on the two different homologues chromosomes. Repeat annotation and classification was performed for those clusters with an abundance >0.01%. For basic repeat classification, protein domains were identified using the tool 'Find RT Domains' in RepeatExplorer (Novak et al., 2013). Searches for sequence similarity, using different databases (GenBank and TIGR) were performed and graph layouts of individual clusters were examined using the SeqGrapheR program (Novak et al., 2013). Satellite DNAs were identified based on the graph layout and further examined using DOTTER (Sonnhammer and Durbin, 1995).

Amplification, cloning and sequencing

The seven most abundant repeats of the total genome, three satellite DNAs (satDNA: Ebusat1, Ebusat2, Ebusat3) and four LTR-retrotransposons (LTR-RT) (Ty1/Copia-Maximus and -Tork and Ty3/Gypsy-Tat and -Chromovirus), were PCR amplified. In addition, one satDNAs (Ebusat4) from microdissected acrocentric chromosome DNA were also PCR amplified. For satellite DNAs, primers were designed, facing outwards of the repeat unit, for the consensus sequences and from the region where most of the reads were conserved. LTR-RT-specific primers were designed to amplify the Integrase domain, commonly used in chromosome analyses with repetitive sequences, for being suggested as the most conserved domain within the domains of the retrotransposons (Table S1). The conserved region of the integrase domain was identified using the SeqGrapheR program (Novak et al., 2013). 40 ng of genomic DNA was used for all PCR reactions with 1× PCR buffer (Invitrogen), 2 mM MgCl₂, 0.1 mM of each dNTP, 0.4 μM each primer, 0.025 U *Taq* polymerase (Platinum *Taq* DNA polymerase, Invitrogen) and water. PCR conditions were: 94°C 3 min, 30× (94°C 1 min, 55°C 1 min, 72°C 1 min) and 72°C 10 min. PCR fragments were purified from a 1% agarose gel using the AxyPrep DNA gel extraction kit (Axygen Biosciences) and cloned with the pGEM®-T Vector cloning system (Promega) using JM109 *E. coli* high-efficiency competent cells (Promega), following manufacturer's instructions. One positive clone of each repetitive element was sequenced with a 3500 Genetic Analyzer Sanger sequencing platform at the Biosciences Center of the Federal University of Pernambuco for confirming its identity. Sequences were deposited in the Genbank database as MK228130-MK228135.

Chromosome preparation and fluorescence *in situ* hybridization

Cloned satellite DNAs, rDNAs and retrotransposons were labelled with either Cy3-dUTP, Cy5-dUTP or **digoxigenin-11-dUTP** by nick translation using a Nick translation mix (Roche, Brazil) or with DNase I (0.002 U) and DNA polymerase (4 U) enzymes following Kato et al., (2004). 35S rDNA sites were detected with the pTa71 clone from *Triticum aestivum* (Gerlach and Bedbrook, 1979). Clone D2 from *Lotus japonicus* (Pedrosa et al., 2002) was used to detect the 5S rDNA.

Chromosomes were prepared from root tips collected from bulbs, pre-treated in 0.02M 8-hydroxyquinoline at 10°C for 24 h and fixed in ethanol: acetic acid (3:1 v/v) for 2 to 24 h at room temperature and stored at -20°C. Fixed root tips were digested with 2 % cellulase-20 % pectinase for 90 min at 37°C, and squashed in a drop of 45 % acetic acid. Fluorescent *in situ* hybridization was performed as described by Pedrosa et al. (2002). The hybridization mix contained 50% (v/v) formamide, 10% (w/v) dextran sulfate, 2× SSC, and 5 ng/μl of each probe. Slides were denatured at 75°C for 5 min, and the final stringency of hybridization was 76%.

Images were captured using a Leica DM5500 B microscope with a Leica DFC345 FX coupled camera and the LAS AF software. Images were edited with Adobe Photoshop CS5.

Immunodetection of histone modifications and active RNA polymerase II

Antibodies for three different histone modifications were used: one euchromatic mark, rabbit anti-histone H3K4me3 (Abcam1012, diluted 1:300) and two pericentromeric chromatin marks: mouse anti-H3S10ph (Abcam 14955, diluted 1:2,000) and rabbit anti-H2AT120ph (Demidov et al., 2014, diluted 1:500). The latter antibody was developed for the same peptide as described in Dong and Han (2012). A mark for transcriptional activity was also applied: rat anti-RNAPIISer2ph (Millipore 04-1571, diluted 1:100). For immunostaining, root tips were pretreated with 2 mM 8-hydroxyquinoline for 24 h at 10°C. For RNAPIISer2 detection, nuclei were isolated from leaves. Both were fixed in freshly prepared 4% paraformaldehyde (dissolved in 1× PBS) for 30 min on ice and then washed 3-times for 15 min in 1× PBS on ice. Fixed root tips and leaves were chopped in a nuclei isolation buffer (Doležel et al., 2007) and filtered through a 30 µm nylon membrane. The cell suspension was used to prepare slides by centrifugation onto a microscopic slide at 2,000 rpm for 3 min (Shandon, CytoSpin3). Slides were incubated in 3% BSA for 30 min at 37°C. Primary antibodies, diluted in 1 % BSA, were incubated overnight at 4°C and detected with Alexa 488-conjugated anti-rabbit (Dianova 711-545-152, diluted 1:200), Alexa 488-conjugated anti-mouse (Molecular probes A11001, diluted 1:200), Alexa 488-conjugated anti-rat (Dianova 112-545-167, diluted 1:200) or goat Cy3-conjugated anti-mouse (Dianova 115-165-062, diluted 1:300) antibodies in 1% BSA and incubated for 1 h at 37°C.

After immunostaining with H2AThr120ph and RNAPIISer2ph, FISH was performed subsequently to analyse the colocalization with the Ebusat2 satellite and the Ty3/Gypsy-Tat LTR-retrotransposon, respectively. Therefore, the slides were washed twice in 1× PBS, fixed in ethanol:acetic acid (3:1 v/v) for at least 24 h at room temperature in the dark, dehydrated and prehybridized in 15 µl of DS20 (50% formamide, 10% dextran sulfate, 2× SSC) overnight at 37°C. Slides were washed in 2× SSC, dehydrated and denatured in 0.2 N NaOH in 70% ethanol for 10 min at room temperature. Afterwards, additional dehydration was performed and the slides were hybridized with 50 ng of the probe in DS20 overnight at 37°C.

Images for histone modifications were captured using an epifluorescence microscope BX61 (Olympus) equipped with a cooled CCD camera (Orca ER, Hamamatsu). To achieve super-resolution for RNAPIISer2ph imaging, 3D structured illumination microscopy (3D-SIM) was applied using a 63×1.4NA Oil Plan-Apochromat objective of an Elyra PS.1 microscope system and the software ZEN (Carl Zeiss, Germany) (Weisshart et al., 2016). For histone modification marks, the fluorescence intensity was estimated along the chromosomes using the ImageJ software (Abramoff et al. 2004). Intensity measurements were done at 10 or eight consecutive circles of ~50-60 pixels each, along with both homologues of the large chromosomes pair and one small chromosome pair, respectively. Five metaphases per mark were measured and a mean of the measurements of each position along the chromosomes was calculated. We defined a ratio between the intensity of DAPI and the histone modification fluorescence along the chromosomes.

DNA replication analysis

DNA replication analysis was performed with the EdU kit (BCK-EdU 594-1, baseclick GmbH, Germany) following the manufacturer's protocol. Root tips were collected and incubated in a humid chamber with a filter paper embedded in an EdU solution for 3 h at room temperature for the incorporation of the dNTP analogue. After incorporation, root tips

recovered in water for 30 min, were pretreated in 8-hydroxyquinoline for 24 h at 10°C and fixed in ethanol:acetic acid (3:1 v/v). The cell walls were digested by treating the root tips with an enzyme mix of 0.7 % cellulase R10 (Duchefa C8001), 1 % pectolyase (Sigma P3026), 1% cytohelicase (Sigma C8274) for 90 min at 37°C. The squashing of chromosomes was performed in a drop of 45 % acetic acid. Then, the slides were immersed into liquid nitrogen to remove the coverslips. The slides were incubated with 3 % BSA for 20 min at room temperature, and then with the detection mixture for 30 min at room temperature in the dark. Sequential FISH with the Ebusat 1 and Ty3/Gypsy-Tat LTR-RT probes was performed as described above. Images were captured as described above.

Results

Bimodal karyotype of *E. bulbosa* is characterized by a differential chromosome-type specific repeat distribution

To characterize the repetitive DNA fraction of *E. bulbosa* (1C = 1.25 Gbp), its genome was sequenced at 0.32× genome coverage. Reads, comprising in total 400 Mbp, were grouped into 118,468 clusters containing from 2 to 90,876 reads. Clusters included 49.7% of all reads, with the major 257 clusters representing at least 0.01% of the genome each. The analysis revealed three major satellite DNA families (satDNAs), 11 transposable element families (LTR-retrotransposons and LINE), DNA transposons and ribosomal DNA sequences (Table 1). The largest clusters were identified as satDNAs: CL1 (Ebusat1) representing 4.99% of the genome with a complex repeat unit with variable length; CL2 (Ebusat2) representing 2.16% of the genome with a 261-bp repeat unit; and CL3 (Ebusat3) represented 1.4% of the genome. The LTR-like retrotransposons constituted approximately 28% of the genome, with the Ty3/gypsy superfamily exceeding 2.15 fold the genome proportion of the Ty1/copia superfamily. Within the former, Tat and Chromovirus were the only highly abundant lineages. Within Ty1/copia retrotransposons, eight lineages were identified, with Maximus and Tork being the most abundant.

The most abundant tandem repeat Ebusat1 localized at the large chromosome pair, mostly forming two interstitial blocks within the chromosome inversion: in the long arm of the acrocentric and in the short arm of the metacentric chromosome. A third weaker signal was observed outside of the inversion, in a distal position on the long arm of both chromosomes (Fig. 1). Ebusat2 was located in the centromeric region of all chromosomes, showing similar hybridization intensities among all chromosomes (Fig. 1). Both satDNAs co-localized with DAPI⁺ bands, except for the third smaller band of Ebusat1. For Ebusat3, no hybridization signals were detected, although the tandem repeat nature was confirmed by Dotter and the amplified PCR fragments showed a ladder-like pattern.

All four most abundant LTR-retrotransposon lineages showed a high accumulation on the large chromosome pair, with dispersed labelling along the entire chromosome, except for the proximal region in chromosome I corresponding to the 35S ribosomal DNA (Fig. 1, Fig. S1). Three of them (Tat, Chromovirus and Tork) showed pericentromeric labelling on the small chromosomes, while Maximus showed a more scattered distribution, enriched proximally on small chromosomes (Fig. 1, Fig. S1). These results demonstrate that both large and small chromosome types share the same repetitive DNA sequences, except for Ebusat1, which is present only in the large pair. However, except for Ebusat2, each chromosome type displays a specific chromosomal distribution for these repetitive sequences, with a higher abundance of repeats in the large chromosome pair.

In order to compare the DNA composition of the homologues of the large chromosome pair, microdissection was performed, and the isolated chromosome-specific DNA was sequenced. About 2,160 Mbp of the sequence reads from each homologue (within and outside the inversion) were used for comparative *in silico* analyses, representing $\sim 8\times$ coverage for each chromosome. The reads were grouped into 13,279 clusters containing from 2 to 80,108 reads. The major 267 clusters, with a minimum of 0.01% of the chromosome DNA proportion, represented $\sim 27\%$ of the chromosome DNA. Further, two different clustering analyses were run using two different sets of sequencing data each, showing similar results. The analysis revealed four satDNA families, Ebusat1, Ebusat2 and Ebusat3, as well as one previously undetected ones, Ebusat4, seven LTR-RT lineages previously characterized in the genomic analysis, LINEs, DNA transposons and ribosomal DNA sequences. Except for Ebusat2, the amount of all satDNAs seems to be higher in the larger chromosomes, in agreement with the distribution of Ebusat1 in this chromosome pair. In contrast, the LTR-RT lineages, except TAR in the metacentric chromosome, showed a similar or lower proportion within chromosome I than the total genome proportions. This may indicate a preferential amplification of the most abundant tandem repeats during the process of whole genome amplification after microdissection (Table S2). Ebusat4 was mostly localized within the chromosome inversion, associated with the interstitial bands of the large chromosome pair as Ebusat1, in the long arm of the acrocentric and in the short arm of the metacentric chromosome (Fig.1). The hybridization signals were similar in both chromosomes of the large pair in disagreement with the differential proportion of Ebusat4 observed after analyzing the DNA composition of microdissected chromosomes. This difference could be explained by the uneven amplification of microdissected DNA (Table S2).

These results confirmed the presence of the major repetitive families in the large chromosome pair, as well as the presence of one new satDNA, Ebusat4, clearly enriched in both homologues of this pair. This distribution of repeats was not influenced by the occurrence of the nonrecombining inverted region of the large chromosome pair.

The chromatin composition differs between large and small chromosomes

As the identified repetitive DNA families showed a distinct chromosome-type specific distribution, we investigated the distribution of a subset of posttranslational histone modifications to understand whether the observed DNA composition differences are associated with a different chromatin organization.

The small chromosomes of *E. bulbosa* showed strong labelling along the chromosome arms by the euchromatin histone mark H3K4me3, with weaker labelling in the pericentromeric regions. In contrast, both large homologues showed weaker labelling along the entire chromosomes (Fig. 2). Histone modification marks for pericentromeric regions displayed contrasting results. While all chromosomes showed a pericentromeric distribution for phosphorylated H3 at serine 10, the large chromosome pair was strongly phosphorylated with H2AThr120ph at the proximal chromosome regions, flanking the pericentromeres and including the rDNA site. Towards distal arm regions, the phosphorylation of H2AThr120 gradually decreased. All small chromosomes showed a weak phosphorylation at the pericentromeres, and weaker signals on the proximal and distal regions (Fig. 2 and Movie S1, S2). The different distribution of both histone modifications and between the small and large chromosomes was confirmed by fluorescence intensity measurements. Subsequent FISH with the centromeric Ebusat2 satellite repeats confirmed its localization between both proximal H2AThr120ph positive regions (Fig. 2).

We also analyzed whether different transcriptional activities between both chromosome types exist. Therefore, we applied specific antibodies against RNA polymerase II phosphorylated at serine 2 (RNAPIISer2ph) as a mark for transcriptional activity. Antibodies specific for the phosphorylation state of a peptide allow the discrimination between active and inactive RNAPII (Bourdon et al., 2012). For the elongation step of transcription, phosphorylation at serine 2 is required (Ni et al. 2004). RNAPIISer2 displayed a dispersed distribution in interphase nuclei, with a certain accumulation within the nucleus interior (Fig. 3). Subsequent FISH with LTR-RT Ty3/Gypsy-Tat showed that the regions enriched with these repetitive elements, which are more abundant in the large chromosome pair and mostly present at the periphery of interphase nuclei, were only seldom associated with active RNAPII. In contrast, regions free of these repetitive elements, mostly in the central nuclear region, were enriched with RNAPIISer2ph (Fig. 3).

Altogether, these results demonstrate that chromatin composition and transcriptional activity differ between large and small chromosomes. These differences are possibly a consequence of the different repetitive sequence composition in both chromosome types.

Replication dynamics differ between large and small chromosomes

As large and small chromosomes present a different repeat and chromatin composition, replication analysis was performed to uncover a potentially different replication behaviour of both chromosome types. Compared to the small chromosomes, the large chromosome pair incorporated EdU at a different time. Furthermore, incorporation into the large chromosomes was observed as bands along the chromosome arms (Fig. 4a). Some cells showed this banding pattern together with the incorporation at the pericentromeric regions on all small chromosomes. Subsequent FISH with Ebusat1 and the Ty3/Gypsy-Tat element showed a partial colocalization between the EdU incorporation and these repeats in metaphase chromosomes as well as in interphase nuclei, mostly at the nucleus periphery (Fig. 4b, c). These data indicate that the different chromosome structures also influence the replication timing, likely as a consequence of the presence of distinct repetitive sequences and chromatin compositions in the large and small chromosomes.

Chromosomal distribution of repeats is similar in *Eleutherine latifolia*

To analyze whether the subgenome repeat distribution is a specific feature of *E. bulbosa*, or a bimodal karyotype feature in the genus, we hybridized the major repeat sequences of *E. bulbosa* to the chromosomes of the sister species *E. latifolia* (2C = 1.4 pg). Both 35S and 5S rDNA sites were located on both of the large acrocentric homologues, with the 35S sites at the pericentromeric regions and the 5S sites at the interstitial regions of the long arms (Fig. 5). The position of the 35S rDNA is similar to that in the acrocentric chromosome of *E. bulbosa*. The 5S rDNA sites are also duplicated in closer proximity as in *E. bulbosa*. However, in one of these homologues both signals are located more distally, suggesting a paracentric inversion present in heterozygosity. The most abundant repetitive elements of *E. bulbosa*, both satDNAs (Ebusat1 and Ebusat2), the Ty3/Gypsy-Tat and Ty1/Copia-Maximus LTR-RT lineages, showed a similar distribution in *E. latifolia* as it was found in *E. bulbosa* chromosomes. Ebusat1 showed two distal bands at the long arms of the large acrocentric pair, one stronger and one weaker band, with inverted orientation between the homologues, confirming the paracentric inversion (brackets in Fig. 5). Ebusat2 displayed a centromeric distribution in all chromosomes of the complement. Ty3/Gypsy-Tat and Ty1/Copia-Maximus LTR-RT exhibited an accumulation on the large acrocentric pair, similar as observed in *E. bulbosa*, irrespective of the position of the

inversion. Ty1/Copia-Maximus was uniformly dispersed along the largest acrocentric chromosome pair. It was proximally distributed on all small chromosomes, but denser at the pericentromeric regions. Similarly, Ty3/Gypsy-Tat was uniformly dispersed along the large pair and showed a pericentromeric distribution in the small chromosomes (Fig. 5). Chromosome signals from both satDNAs and LTR-RT were weaker in *E. latifolia* than in *E. bulbosa*. This suggests differences of these repeats families between both species, likely due to differences in sequence similarity and/or abundance. However, both species have the same distribution of repeat sequences, indicating that the chromosome set-specific repeat distributions are conserved in the genus and was not influenced by species-specific chromosome rearrangements. We suggest that this type of chromosome set-repeat distribution is a characteristic of bimodal karyotypes.

Discussion

Chromosomal distribution of repetitive DNA in *E. bulbosa* reflects its bimodal karyotype

Although repetitive sequences are enriched in the large chromosomes of several bimodal plant species (de la Herrán et al., 2001; Pedrosa et al., 2001), none of them showed a high degree of accumulation of LTR retroelements, together with satDNA, as seen in *Eleutherine*. The repetitive genome fraction of *E. bulbosa* is mainly composed of LTR retroelements, with more abundant Ty3/Gypsy-like than Ty1/Copia-like elements. Within eukaryote genomes, LTR-RT Ty3/Gypsy and Ty1/Copia are most abundant in plants (Kumar and Bennetzen, 1999), with the Ty3/Gypsy elements being the most abundant in the majority of angiosperm families (Weiss-Schneeweiss et al., 2015). All LTR retrotransposons were highly accumulated on the large chromosome pair of *Eleutherine*, showing a uniformly dispersed distribution, similar to the distribution observed for several mobile elements in large genomes (Lamb, 2006; Lamb et al., 2007). In contrast, the small chromosomes of *Eleutherine* accumulated the LTR retroelements only at the pericentromeric regions. This restricted distribution of repetitive sequences is typical for small genome species, e.g. in *Arabidopsis* (Heslop-Harrison et al., 1997), *Phaseolus* (Padeken et al., 2015) and *Brachiaria* (Santos et al., 2015). Thus, the differential distribution of retroelements appears to be related to the bimodal condition and is not influenced by the nonrecombining chromosome inversions in these species. The large and the small chromosome sets of *Eleutherine* constitute two distinct subgenomes with respect to the retroelement distribution.

The chromatin organization is partly similar in *Eleutherine* and bird species. The microchromosomes of birds are enriched in genes and hyperacetylated at histone H4K5, while macrochromosomes are gene poor and hypoacetylated at histone H4K5 (McQueen et al., 1998). However, unlike the large chromosomes of *E. bulbosa*, bird macrochromosomes are poor in heterochromatin, with less 5-methylcytosine-rich regions than microchromosomes, probably related to CpG islands associated with the high gene content of the microchromosomes (Schmid and Steinlein, 2017). Nevertheless, some transposable elements show the differential chromosomal distribution in bimodal karyotypes of different animal groups. In birds, CR1-like retroelements are spread over nearly all chromosomes but have a higher density on macrochromosomes with a particular banding pattern (Coullin et al., 2005). The Rex6 transposable element is also densely distributed in the largest chromosome pairs in several species of *Podocnemis* turtles. It was suggested that Rex6 may influence the genomic structure, interfering with gene regulation (Noronha et al., 2016).

Retrotransposons are abundant components of large chromosomes in some plants with bimodal karyotypes, but differentiation of chromosome sets is less evident. The large chromosomes of South American *Hypochaeris* species are enriched by Ty1-Copia LTR-RT along their entire lengths, while small chromosomes lacked those elements in most of the long arms. However, this pattern is not maintained in other species of the genus, mainly not in Old world species (Morocco and Croatia; Ruas et al., 2008). On the other hand, in three species of the genus *Alstroemeria* with asymmetric chromosomes, the distribution of a Ty1-Copia like LTR-RT was equally dispersed over all chromosomes (Kuipers et al., 1998).

Here, two major satellite DNAs represent a large proportion of the genome (~7 %). Such a high proportion of one or two satDNA families is unusual in most plant species. Usually, many satDNA families with a low abundance or few satDNA families with slightly larger genome abundance were observed (Hemleben et al., 2007; Steflova et al., 2013). Contrary to what was suggested (Melters et al., 2013), the most abundant satDNA from *E. bulbosa* is not centromeric but accumulated interstitially exclusively in the large chromosome pair. The second most abundant satDNA is present in all centromeric regions. However, whether this location is associated with centromere functions has not yet been clarified. As in *Eleutherine*, large chromosomes of bimodal karyotypes in some plants accumulate satDNA. In *Muscari* species, a large chromosome specific satDNA found in *M. comosum* is conserved within the genus and has been proposed to mediate the increase of karyotype asymmetry (de la Herrán et al., 2001). One of the major components of intercalary heterochromatin on the large chromosomes of *Ornithogalum longibracteatum* was also a specific satDNA (Pedrosa et al., 2001). Therefore, as clearly observed in *Eleutherine*, large and small chromosomes within a bimodal karyotype can maintain a differential DNA composition. This can cause structural and functional differences between these subgenomes.

Large and small chromosomes differ in heterochromatin, gene content and replication timing

Chromatin-associated with H3K4me3 and H2AThr120ph histone marks differentiate the large and small chromosomes in *E. bulbosa*. The dense distribution of H3K4me3 along all small chromosomes is associated with euchromatin, as it was observed also for other histone modification marks indicating euchromatin in the small chromosomes of this and other species with small genomes (Feitoza and Guerra, 2011). The less intense distribution of H3K4me3 on the large chromosomes is likely associated with a higher proportion of heterochromatin, as revealed by the repeat distribution of in the present work, as well as by 5-methylcytosine and H3K9me2 localization (Feitoza and Guerra, 2011).

The so-called universal histone modification mark for the pericentromeric region, H2AThr120ph (Dong and Han, 2012; Demidov et al., 2014), did not exclusively label the pericentromeric region in *E. bulbosa* as defined by the H3S10ph mark. The large chromosomes also showed intense H2AThr120ph signal along their lengths. This suggests that the proximal chromatin composition may differ between both chromosomes sets and that it is rather related to the chromosome structure than to centromere function. A comparable atypical distribution of this histone mark was found for the large Y chromosome of *Coccinia grandis* (Sousa et al., 2016).

Large and small subgenomes of *Eleutherine* also differ in their replication timing. Large chromosomes showed a banding pattern after EdU incorporation. This suggests that, within these chromosomes, there are regions which perform replication at a different time than those of the small chromosomes. This could be due to the high repeat composition within large chromosomes, as confirmed by the colocalization of some repetitive elements with EdU bands. Differential replication timing was also reported for chicken macrochromosomes displaying late replication, and microchromosomes showing early replication, both related to their different gene content (McQueen et al., 1998). This assumption could also be valid for *E. bulbosa*, because the distribution of active RNAPII in interphase nuclei suggests chromosome regions with different transcription activities. Chromatin enriched in active RNAPII did not colocalize to repeat-rich nuclei regions and is possibly associated with small chromosomes. In addition, a weaker RNAPIISer2ph labelling and more repetitive DNA were observed at the nuclear periphery, possibly associated with large chromosomes. Together, this data suggests that the large chromosomes of *Eleutherine* are mainly composed of heterochromatin and heterochromatin-like, early-condensing euchromatin (Feitoza and Guerra, 2011; Guerra et al., 1988). They have a lower gene density and partially replicate later. In contrast, small chromosomes are composed mainly of euchromatin, are gene-rich and replicate earlier. Thus, the large and small chromosomes represent structurally and functionally differentiated subgenomes within the same species.

The bimodal chromosomal organization is maintained within the sister species *E. latifolia*

Both investigated *Eleutherine* species display a similar repeat distribution, indicating that these repeats originated and underwent a chromosome-type specific accumulation before the separation of these two species. The newly discovered chromosome inversion in *E. latifolia*, a paracentric inversion in the long arm of one homologue of the large chromosome pair, also involved Ebusat1. One breakpoint is apparently close to the breakpoint of the *E. bulbosa* pericentric inversion. Both events may suggest that this chromosome pair is prone to chromosome rearrangements, possibly due to its highly repetitive sequence content. Repetitive sequences are potential sites for chromosome rearrangements through homology-directed recombination repair using ectopic homologous repeats as a template (Charlesworth et al., 2005). Although vegetative reproduction in *Eleutherine* could be responsible for the maintenance of both inversions in heterozygous condition, the recurrent rearrangements in this chromosome pair are possibly associated with a strong purifying selection, with lethality under homozygous conditions. This may lead to permanent heterozygosity of this chromosome pair, even in individuals propagated via seeds. First, we hypothesized that the chromosomal inversion in *E. bulbosa* could have lead to repetitive sequence accumulation on the large chromosome pair. However, since the repeat distribution is uniform along the entire large chromosome pair, even outside the inverted region, and is similar in both *Eleutherine* species, it is more likely that it is related to the bimodal structuration and function of these karyotypes, rather than the consequence of chromosome inversions.

How could a bimodal karyotype evolve in *Eleutherine*?

Different types of bimodal karyotypes exist, since not in all bimodal species different chromatin and sequence composition between large and small chromosomes exist, as presented for *Eleutherine*. There are different hypotheses for the origin of bimodal karyotypes.

Interspecific hybridization appears not to apply to *Eleutherine*, since the bimodal karyotype is characteristic for the whole tribe Tigridieae, with no evidence of allopolyploidy for its origin (Moraes et al., 2015).

Although paleo-allopoloidization or ancestral chromosomal rearrangements cannot be completely ruled out, our data indicate an increase of the size of one set of chromosomes due to a differential repetitive sequences accumulation. This phenomenon was suggested for species of the *Muscari* genus, where differential accumulation of one satDNA was observed to be associated to the increase in the size of a subset of chromosomes (de la Herrán et al., 2001). It is possible that an initial random accumulation of repetitive sequences in one chromosome pair in *Eleutherine* ancestral gradually increased its size and led to its chromatin differentiation, differential replication and transcription behaviour. Consequently, this may lead to a higher repetitive sequence accumulation and divergence of different repetitive families in the large chromosomes, which is less deleterious than in the more gene-rich small chromosomes (Vosa, 2005; see Fig. 6). In both *Eleutherine* species, this repetitive environment provided a special background for chromosomes inversion.

In chicken, the microchromosomes present an increased rate of meiotic recombination compared to the macrochromosomes, perhaps due to a gene composition that favours meiotic recombination. This could lead to evolutionary pressure for an increase of gene density on small chromosomes (Rodionov et al., 2002; Smith et al., 2000). The negative feedback between repeats and recombination may be intensified by intrachromosomal rearrangements, also contributing to suppress recombination and subgenome differentiation in bimodal karyotypes, as our results suggested for *Eleutherine*. The possible gradual increase of meiotic recombination within the small chromosome set and, in consequence, a decrease of recombination between small and large chromosome sets could also favour the differentiation of the structure and function of both chromosome set and the maintenance of the bimodal karyotypes over time.

Author Contribution Statement

MB, AH and AP designed the study. MB and MV carried out the bioinformatic studies. SD carried out the chromosome microdissection with the help of MB. VS carried out the super-resolution microscopy. MB carried out the cytogenetic, molecular and flow cytometry experiments, interpreted the results, and wrote the manuscript with the help of AH, and AP. All authors read and approved the final manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

We are indebted to Dr Gustavo Souza, from the Federal University of Pernambuco, Brazil for providing *E. bulbosa* plants, to Dr Guadalupe Munguía Linno from Guadalajara University, Mexico, for providing *E. latifolia* seeds, to Prof Ingo Schubert for discussion and to Dr Axel Himmelbach (IPK, Gatersleben) for Illumina sequencing.

Funding

We thank the Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Germany, for financial support. This study was supported in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES, Finance Code 001).

Supplementary Material

Table S1. List of primer to amplify *E. bulbosa* repetitive elements

Table S2. Proportions repetitive element (%) in the total genome of *E. bulbosa*. Estimated proportion of each repeat in the metacentric and acrocentric homologue of pair I, with a preferential amplification of the most abundant repetitive families likely caused by the whole genome amplification of microdissected DNA.

Figure S1. Distribution of LTR-retrotransposons Ty1/Copia-Tork and Ty3/Gypsy-Chromovirus elements in *E. bulbosa*. Both types of repeats are highly enriched along the large chromosome pair and the at the pericentromeric region in all small chromosomes.

Movie S1. Movie showing the differential chromatin organization within the centromeric region of *E. bulbosa* chromosomes mitotic complement. Pericentromeric histone modification marks H3S10ph (red) and H2ATh120ph (green), on both chromosome types. Chromosomes are stained with DAPI (blue).

Movie S2. Movie showing the differential chromatin organization within the centromeric region of *E. bulbosa* metacentric chromosome. Pericentromeric histone modification marks H3S10ph (red) and H2ATh120ph (green). Chromosomes are stained with DAPI (blue).

References

- Akemine, T. (1935) Chromosome studies on *Hosta* I. The chromosome numbers in various species of *Hosta*. J. Fac. Sci. Hokkaido Imp Univ Ser 5 Bot 5, 25–32.
- Brandham, P.E. and Doherty, M.J. (1998) Genome Size Variation in the Aloaceae, an Angiosperm Family Displaying Karyotypic Orthoselection. *Annals of Botany* 82, 67–73.
- Bourdon, M., Pirrello, J., Cheniclet, C., et al. (2012) Evidence for karyoplasmic homeostasis during endoreduplication and a ploidydependent increase in gene transcription during tomato fruit growth. *Development* 139, 3817–3826.
- Burt, D.W. (2002) Origin and evolution of avian microchromosomes. *Cytogenet. Genome Res* 96, 97–112.
- Charlesworth, D., Charlesworth, B. and Marais, G. (2005) Steps in the evolution of heteromorphic sex chromosomes. *Heredity* 95, 118–128.
- Coullin, P., Bed'Hom, B., Candelier, J.J., Vettese, D., Maucolin, S., Moulin, S., Galkina, S.A., Bernheim, A. and Volobouev, V. (2005) Cytogenetic repartition of chicken CR1 sequences evidenced by PRINS in Galliformes and some other birds. *Chromosome Res* 13, 665–673.

- de la Herrán, R., Robles, F., Cuñado, N., Santos, J.L., Ruiz Rejón, M., Garrido-Ramos, M.A. and Ruiz Rejón, C. (2001) A heterochromatic satellite DNA is highly amplified in a single chromosome of *Muscari* (Hyacinthaceae). *Chromosoma* 110, 197–202.
- D’Emerico, S., Grünanger, P., Scrugli, A. and Pignone, D. (1999) Karyomorphological parameters and C-band distribution suggest phyletic relationships within the subtribe Limodorinae (Orchidaceae). *Plant Syst. Evol* 217, 147–161.
- Demidov, D., Schubert, V., Kumke, K., et al. (2014) Anti-Phosphorylated Histone H2AThr120: A Universal Microscopic Marker for Centromeric Chromatin of Mono- and Holocentric Plant Species. *Cytogenet. Genome Res* 143, 150–156.
- Doležel, J., Greilhuber, J. and Suda, J. (2007) Estimation of nuclear DNA content in plants using flow cytometry. *Nat. Protoc* 2, 2233.
- Dong, Q. and Han, F. (2012) Phosphorylation of histone H2A is associated with centromere function and maintenance in meiosis. *Plant J* 71, 800–809.
- Dreissig, S., Fuchs, J., Cápál, P., Kettles, N., Byrne, E. and Houben, A. (2015) Measuring Meiotic Crossovers via Multi-Locus Genotyping of Single Pollen Grains in Barley. *PLoS One* 10(9):e0137677. doi: 10.1371/journal.pone.0137677.
- Feitoza, L. and Guerra, M. (2011) Different types of plant chromatin associated with modified histones H3 and H4 and methylated DNA. *Genetica* 139, 305–314.
- Feliner, G.N. and Rosselló, J.A. (2012) Concerted Evolution of Multigene Families and Homoeologous Recombination. In: *Plant Genome Diversity Volume 1*. Springer, Vienna, pp. 171–193.
- Fentaw, E., Dagne, K., Rønsted, N., Demissew, S. and Grace, O.M. (2013) Karyotypes in Ethiopian *Aloe* species (Xanthorrhoeaceae: Asphodeloideae). *Kew Bull* 68, 599–607.
- Fiorin, F.G., Ruas, P.M., Ortiz, M.A., Urtubey, E., Matzenbacher, N.I. and Ruas, C.F. (2013) Karyotype studies on populations of two *Hypochaeris* species (*H. catharinensis* and *H. lutea*), Asteraceae, endemics to southern Brazil. *Genet Mol Res* 12, 1849–1858.
- Gen-Shen, Y., Yang, Z., Chiang, T., and Gong, X. (2014) Tracing the origin of the Bimodal karyotype of the tribe Liliaceae (Liliaceae) based on comparative karyotypes analyses. *Plant Div Res* 36, 737–746.
- Gerlach, W.L. and Bedbrook, J.R. (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. *Nucleic Acids Res* 7, 1869–1885.
- Goldblatt, P. and Snow, N. (1991) Systematics and Chromosome Cytology of *Eleutherine* Herbert (Iridaceae). *Ann. Mo Bot Gard* 78, 942–949.
- Griffin, D.K., Romanov, M.N., O’Connor, R., Fowler, K.E. and Larkin, D.M. (2015) Avian cytogenetics goes functional, in: *Third Report on Chicken Genes and Chromosomes 2015* M. Schmid, J. Smith, and D. W. Burt, eds. *Cytogenet. Genome Res* 145, 100–105.

- Guerra, M. (1988) Mitotic and meiotic analysis of a pericentric inversion associated with a tandem duplication in *Eleutherine bulbosa*. *Chromosoma* 97, 80–87.
- Guerra, M. (1991) Cis-acting regulation of NOR cistrons in *Eleutherine bulbosa* (Iridaceae). *Genetica* 83, 235–241.
- Hemleben, V., Kovarik, A., Torres-Ruiz, R.A., Volkov RA and Beridze T (2007) Plant highly repeated satellite DNA: Molecular evolution, distribution and use for identification of hybrids. *Syst. Biodivers* 5, 277–289.
- Heslop-Harrison, J.S., Brandes, A., Taketa, S., Schmidt, T., Vershinin, A.V., Alkhimova, E.G., Kamm, A., Doudrick, R.L., Schwarzacher, T., Katsiotis, A., Kubis, S., Kumar, A., Pearce, S.R., Flavell, A.J. and Harrison, G.E. (1997) The chromosomal distributions of Ty1-*copia* group retrotransposable elements in higher plants and their implications for genome evolution. *Genetica* 100, 197–204.
- Kuipers, A.G., Heslop-Harrison, P.J. and Jacobsen, E. (1998) Characterisation and physical localisation of Ty1-copia-like retrotransposons in four *Alstroemeria* species. *Genome* 41, 357–367.
- Kumar, A. and Bennetzen, J.L. (1999) Plant Retrotransposons. *Annu Rev Genet* 33, 479–532.
- Lamb, J.C. (2006) Retroelement Genome Painting: Cytological Visualization of Retroelement Expansions in the Genera *Zea* and *Tripsacum*. *Genetics* 173, 1007–1021.
- Lamb, J.C., Meyer, J.M., Corcoran, B., Kato, A., Han, F. and Birchler, J.A. (2007) Distinct chromosomal distributions of highly repetitive sequences in maize. *Chromosome Res* 15, 33–49.
- Masabanda, J.S., Burt, D.W., O'Brien, P.C.M., et al., (2004) Molecular Cytogenetic Definition of the Chicken Genome: The First Complete Avian Karyotype. *Genetics* 166, 1367–1373.
- McKain, M.R., Wickett, N., Zhang, Y., Ayyampalayam, S., McCombie, W.R., Chase, M.W., Pires, J.C., dePamphilis, C.W. and Leebens-Mack, J. (2012) Phylogenomic analysis of transcriptome data elucidates co-occurrence of a paleopolyploid event and the origin of bimodal karyotypes in Agavoideae (Asparagaceae). *Am J Bot.* 99, 397–406.
- McQueen, H.A., Siriaco, G. and Bird, A.P. (1998) Chicken microchromosomes are hyperacetylated, early replicating, and gene rich. *Genome Res* 8, 621–630.
- Melters, D.P., Bradnam, K.R., Young, H.A., et al., (2013) Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biol* 14, R10.
- Moraes, A.P., Souza-Chies, T.T., Stiehl-Alves, E.M., et al., (2015) Evolutionary trends in Iridaceae: new cytogenetic findings from the New World. *Bot J Linn. Soc* 177, 27–49.
- Ni, Z., Schwartz, B.E., Werner, J., Suarez, J.R. and Lis, J.T. (2004) Coordination of transcription, RNA processing, and surveillance by P-TEFb kinase on heat shock genes. *Molecular Cell* 13, 55–65.

- Noronha, R.C.R., Barros, L.M.R., Araújo, R.E.F., Marques, D.F., Nagamachi, C.Y., Martins, C. and Pieczarka, J.C. (2016) New insights of karyoevolution in the Amazonian turtles *Podocnemis expansa* and *Podocnemis unifilis* (Testudines, Podocnemidae). *Mol Cytogenet* 9, 73.
- Novak, P., Neumann, P., Pech, J., Steinhaisl, J. and Macas, J. (2013) RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* 29, 792–793.
- Padeken, J, Zeller, P. and Gasser, S.M. (2015) Repeat DNA in genome organization and stability. *Curr. Opin. Genet Dev* 31, 12–19.
- Pedrosa, A., Jantsch, M.F., Moscone, E.A., Ambros, P.F. and Schweizer, D. (2001) Characterisation of pericentromeric and sticky intercalary heterochromatin in *Ornithogalum longibracteatum* (Hyacinthaceae). *Chromosoma* 110, 203–213.
- Pedrosa, A., Sandal, N., Stougaard, J., Schweizer, D., and Bachmair, A., (2002) Chromosomal Map of the Model Legume *Lotus japonicus*. *Genetics* 161, 1661–1672.
- Renny-Byfield, S., Chester, M., Kovařík, A., et al., (2011) Next Generation Sequencing Reveals Genome Downsizing in Allotetraploid *Nicotiana tabacum*, Predominantly through the Elimination of Paternally Derived Repetitive DNAs. *Mol Biol Evol* 28, 2843–2854.
- Rodionov, A.V., Lukina, N.A, Galkina, S.A., Solovei, I. and Saccone, S. (2002) Crossing Over in Chicken Oogenesis: Cytological and Chiasma-Based Genetic Maps of the Chicken Lampbrush Chromosome 1. *J Hered* 93, 125–129.
- Ruas, C.F., Weiss-Schneeweiss, H., Stuessy, T.F., et al. (2008) Characterization, genomic organization and chromosomal distribution of Ty1-copia retrotransposons in species of *Hypochoeris* (Asteraceae). *Gene* 412, 39–49.
- Santos, F.C., Guyot, R., do Valle, C.B., Chiari, L., Techio, V.H., Heslop-Harrison, P. and Vanzela, A.L.L. (2015) Chromosomal distribution and evolution of abundant retrotransposons in plants: gypsy elements in diploid and polyploid *Brachiaria* forage grasses. *Chromosome Res* 23, 571–582.
- Schmid, M. and Steinlein, C. (2017) The Hypermethylated Regions in Avian Chromosomes. *Cytogenet. Genome Res* 151, 216–227.
- Schubert, I. and Lysak, M.A. (2011) Interpretation of karyotype evolution should consider chromosome structural constraints. *Trends Genet* 27, 207–216.
- Shirakawa, J., Nagano, K. and Hoshi, Y. (2012) Polyploid genome structure of *Drosera spatulata* complex (Droseraceae). *Cytologia* 77, 97–106.
- Smith, J., Bruley, C.K., Paton, I.R., et al., (2000) Differences in gene density on chicken macrochromosomes and microchromosomes. *Anim Genet* 31, 96–103.
- Sonnhammer, E.L.L. and Durbin, R. (1995) A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene* 167, GC1–GC10.

- Sousa, A., Bellot, S., Fuchs, J., Houben, A. and Renner, S.S. (2016) Analysis of transposable elements and organellar DNA in male and female genomes of a species with a huge Y chromosome reveals distinct Y centromeres. *Plant J* 88, 387–396.
- Steflova, P., Tokan, V., Vogel, I., Lexa, M., Macas, J., Novak, P., Hobza, R., Vyskot, B. and Kejnovsky, E. (2013) Contrasting Patterns of Transposable Element and Satellite Distribution on Sex Chromosomes (XY1Y2) in the Dioecious Plant *Rumex acetosa*. *Genome Biol. Evol* 5, 769–782.
- Stock, A.D. and Mengden, G.A. (1975) Chromosome banding pattern conservatism in birds and nonhomology of chromosome banding patterns between birds, turtles, snakes and amphibians. *Chromosoma* 50, 69–77.
- Tenaillon, M.I., Hollister, J.D. and Gaut, B.S. (2010) A triptych of the evolution of plant transposable elements. *Trends Plant Sci* 15, 471–478.
- Vosa, C.G. (2005) On chromosome uniformity, bimodality and evolution in the tribe Aloineae (Asphodelaceae). *Caryologia* 58, 83–85.
- Watkins, G.M. (1936) Chromosome Numbers and Species Characters in *Yucca*. *Am. J. Bot* 23, 328–333.
- Weiss-Schneeweiss, H., Leitch, A., McCann, J., Jang, T.S. and Macas, J. (2015) Employing next generation sequencing to explore the repeat landscape of the plant genome. In: Hörandl, E. & Appelhans, M.S. *Next-Generation Sequencing in Plant Systematics*. International Association for Plant Taxonomy (IAPT).
- Weisshart, K., Fuchs, J. and Schubert, V. (2016). Structured illumination microscopy (SIM) and photoactivated localization microscopy (PALM) to analyze the abundance and distribution of RNA polymerase II molecules in flow sorted *Arabidopsis* nuclei. *Bio Protoc* 6:e1725.

Table 1. Proportion (%) of repetitive elements present in the total genome of *E. bulbosa*.

Repetitive element		Total genome (%)
Satellite	Ebusat 1	4.99
	Ebusat 2	2.16
	Ebusat 3	1.40
	Ebusat 4	-*
LTR-Ty3/Gypsy	Tat	14.26
	Chromovirus	1.76
LTR-Ty1/Copia	Maximus	1.78
	Tork	1.46
	TAR	1.23
	AleII	1.19
	Angela	0.98
	Ivana/Oryco	0.31
	Bianca	0.48
	AleI.Retrofit	0.031
Unclassified LTR		6.39
LINE		0.19
DNA Transposons		2.08
rDNA		0.53
Microsatellite		0.33
Unclassified		1.88
Total		43.04

*Found only in the acrocentric chromosome via *in silico* analysis

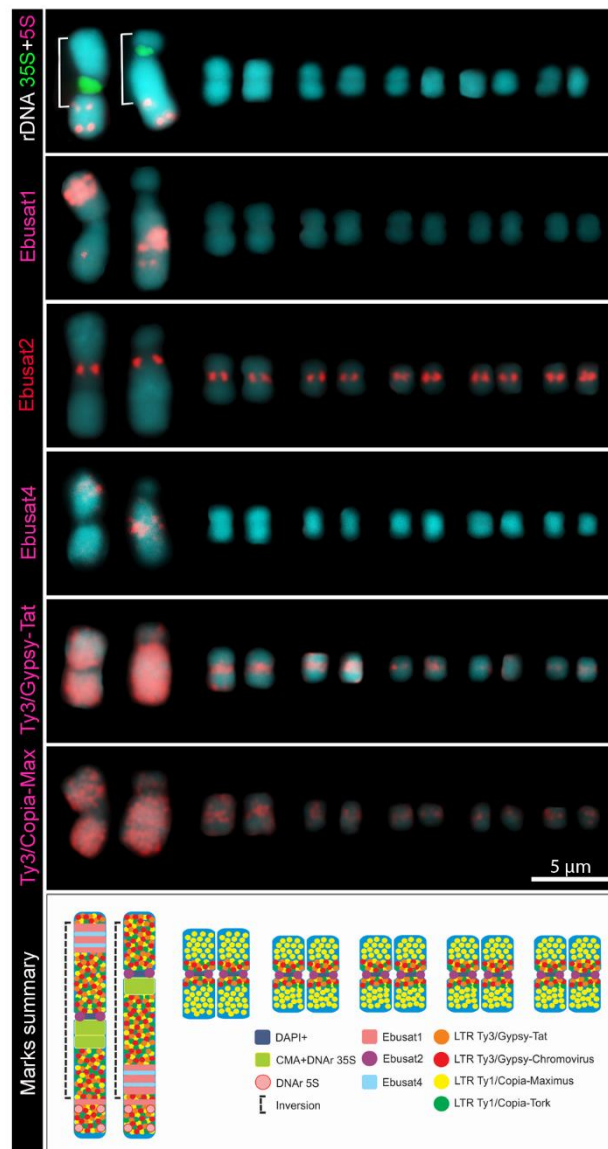


Figure 1 Comparative karyograms showing the distribution of repetitive elements in the *E. bulbosa* chromosomes. The brackets indicate the inverted chromosome region. The scheme below summarizes the repeat distribution.

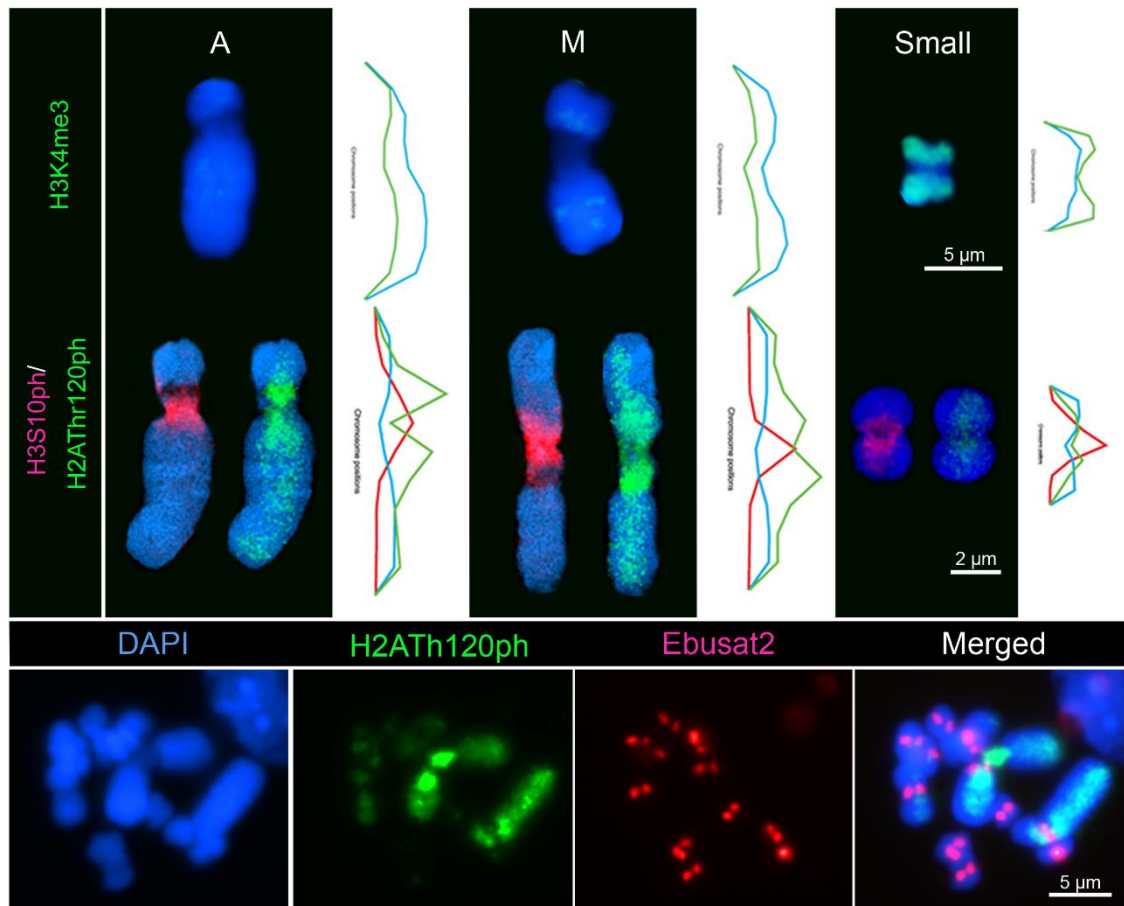


Figure 2 Chromatin organization differs between the large and small chromosomes of *E. bulbosa*. Top: Distribution of the euchromatin mark H3K4me3 on the large acrocentric (A), metacentric (M) and a small chromosome, each with a fluorescence intensity histogram of whole chromatin (DAPI) and H3K4me3 staining. Middle: Distribution of the pericentromeric histone modification marks H3S10ph and H2AThr108ph on the three chromosome types, with their respective fluorescence intensity histograms. Bottom: H2AThr108ph distribution along the chromosome arms and the Ebusat2 DNA localization at the centromeres.

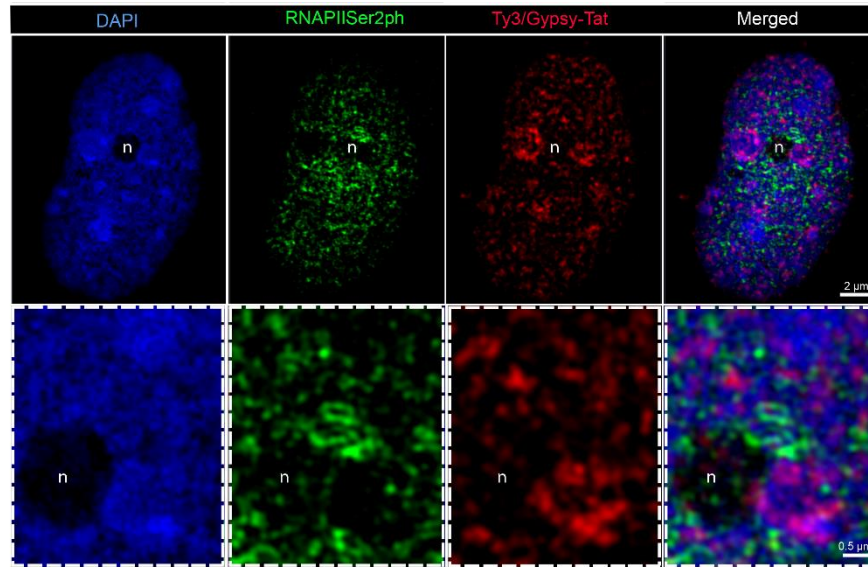


Figure 3 Distribution of RNAPIISer2ph and Ty3/Gypsy-Tat repetitive elements in *E. bulbosa* interphase nuclei, analyzed by super-resolution microscopy (SIM). Active RNAPII phosphorylated at serine 2 (RNAPIISer2ph) accumulates especially at the nucleus interior. Regions enriched in LTR Ty3/Gypsy-Tat elements, mostly at the periphery of interphase nuclei, were only seldom associated with active RNAPII, localized around the nucleolus (n, enlarged region).

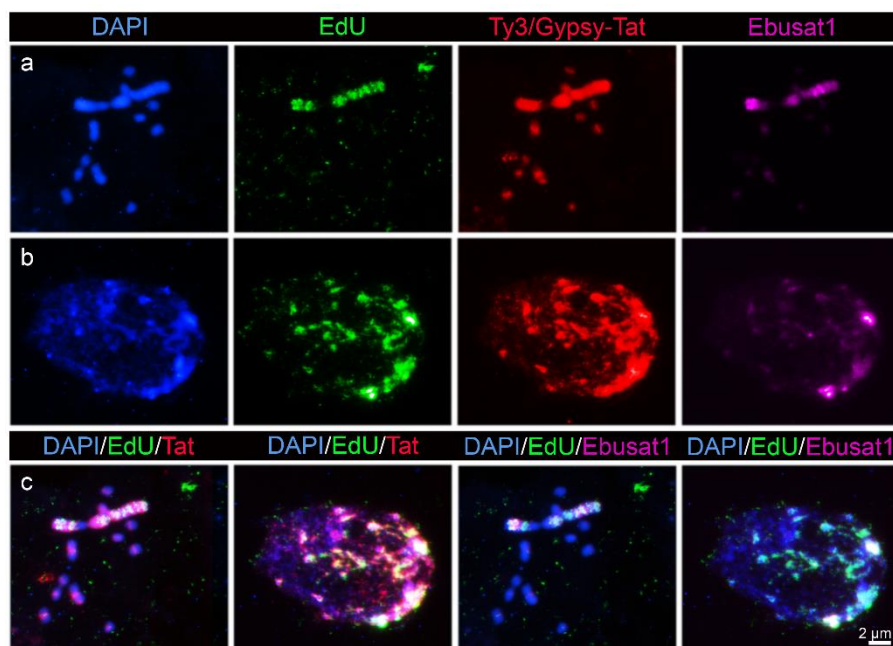


Figure 4 Replication dynamics differs between the large and small chromosomes of *E. bulbosa*. Metaphase chromosomes (a) and interphase nuclei (b) after EdU incorporation followed by LTR Ty3/Gypsy-Tat and Ebusat1 DNA localization. Co-localization is shown in c.

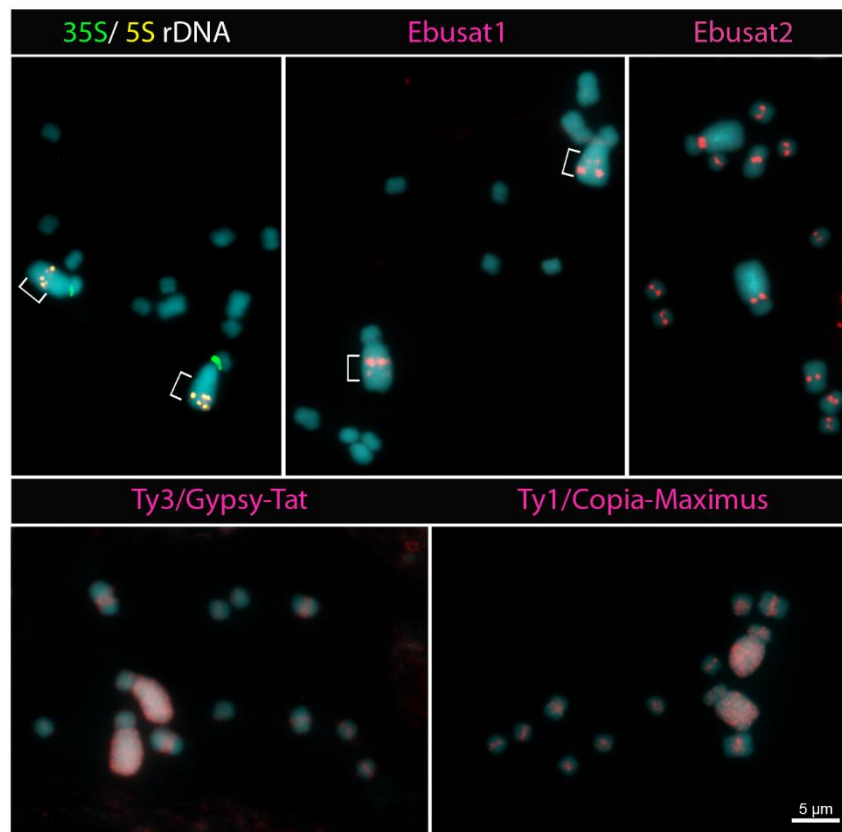


Figure 5 Distribution of repetitive elements in *E. latifolia*. The localization of the 5S, 35S ribosomal DNA (rDNA) sites and the Ebusat1 repeat indicate a paracentric inversion at the long arm of the large chromosome pair (brackets).

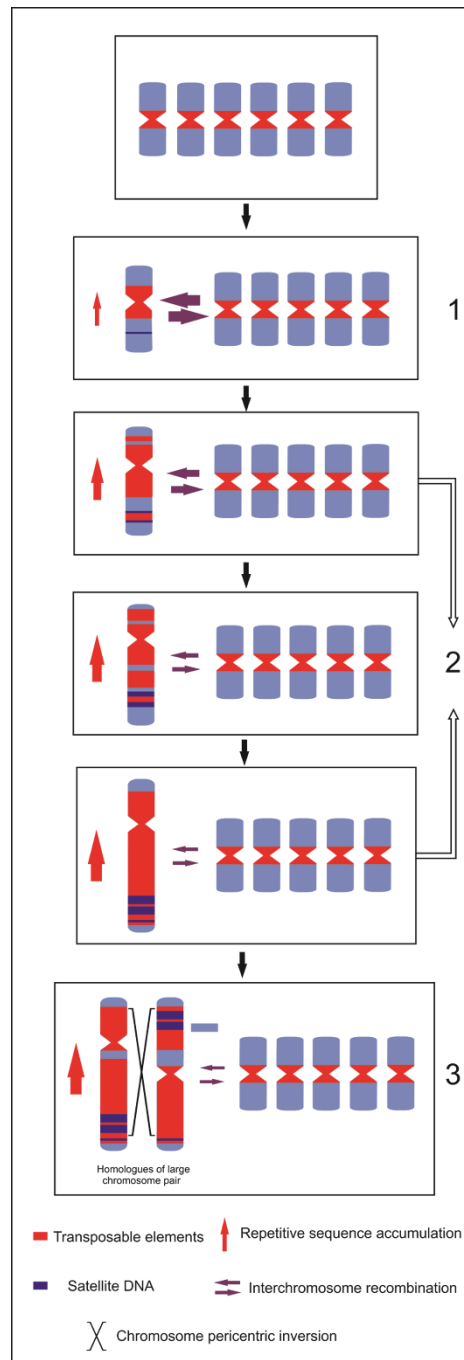


Figure 6 Bimodal karyotype evolution model for *E. bulbosa*. From a symmetric to a bimodal karyotype: 1) random increase of retroelements and satellite DNA; 2) preferential maintenance of retroelements accumulating in the large-chromosome subgenome with increasing chromosome size and decreasing gene density; 3) pericentric inversions lead to the heteromorphic homologues on the large chromosome pair, maintained by recessive lethality with selection advantages in the heterozygote status.

Table S1 List of primers used to amplify *E. bulbosa* repetitive elements.

Repetitive element	Primer name	Primer sequence
Ebusat1	Ebusat1F	TTGCCAATTGACATATTGCCAT
Ebusat1	Ebusat1R	TCAATATGTCAATTGGCAATATGTC
Ebusat2	Ebusat2F	TTTAGAACTTTGGTAAACTCTGCG
Ebusat2	Ebusat2R	AGTTTTACGTAGTTTCAAACAACCT
Ebusat3*	Ebusat3F	CCGATTTGAGTATGTGCTTGAAAAA
Ebusat3*	Ebusat3R	GCAAATTCGCATACTATAACTCACAG
Ebusat4	Ebusat4-F	AACAAACACAAACATGCACACA
Ebusat4	Ebusat4-R	TGCATCTCCGTATGTATGGCT
LTR Ty3/Gypsy CL126ctg87F Chromovirus		GGCACCGTACGAAGCCTTAT
LTR Ty3/Gypsy CL126ctg87R Chromovirus		TTCCAAGTGGCTTCCTCGTG
LTR Ty1/Copia Maximus CL19ctg162F		TTTATCATGGTACGGGCCGC
LTR Ty1/Copia Maximus CL19ctg162R		GTGATGCTTGCAACTTGGGAA
LTR Ty3/Gypsy Tat CL29ctg117F		GCCGATCGATCCACCTTCAT
LTR Ty3/Gypsy Tat CL29ctg117R		TTTTCCTCCAGCATGGCGAT
LTR Ty1/Copia Tork CL89ctg39F		ACATTGTGGCATCGTCGACT
LTR Ty1/Copia Tork CL89ctg39R		CGTACGGTTCATCCTCTCCG

*No signal at chromosomes.

Table S2 Repetitive element proportions (%) in the total genome of *E. bulbosa*. Estimated proportion of each repeat in the metacentric and acrocentric homologue of pair I, with a preferential vies amplification of the most abundant repetitive families during the whole genome amplification after microdissection.

Repetitive element		Total genome	Metacentric	Acrocentric
Satellite	Ebusat 1	4.99	14.8	5.29
	Ebusat 2	2.16	0.94	1.28
	Ebusat 3	1.4	14.88	14.0
	Ebusat 4	-	0.02	6.22
LTR-Ty3/Gypsy	Chromovirus	1.76	0.30	1.71
	Tat	14.26	0.07	0.06
LTR-Ty1/Copia	Maximus	1.78	0.30	0.68
	Tork	1.46	0.88	0.01
	TAR	1.23	2.48	0.35
	AleII	1.19	0.25	0.07
	Angela	0.98	0.18	0.002
	Ivana/Oryco	0.31	-	-
	Bianca	0.48	-	-
	AleI.Retrofit	0.031	-	-
Unclassified LTR		6.39	9.75	21.9
LINE		0.19	0.63	0.45
DNA Transposons		2.08	2.64	5.36
rDNA		0.53	1.11	-
Microsatellite		0.33	-	-
Unclassified		1.88	21.5	1.3
Total		43.04	68.8	58.8