

WANESSA ALVES LIMA

**CITOGENÔMICA COMPARATIVA DOS ELEMENTOS RT-LTR SIRE, OGRE,
TEKAY E IKEROS EM TRÊS ESPÉCIES DE *Eucalyptus* L'Her**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

Orientador: Wellington Ronildo Clarindo

**VIÇOSA - MG
2024**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

L732c
2024
Lima, Wanessa Alves, 1999-
Citogenômica comparativa dos elementos RT-LTR Sire
Ogre Tekay e Ikeros em três espécies de *Eucalyptus* L'Her /
Wanessa Alves Lima. – Viçosa, MG, 2024.
1 dissertação eletrônica (38 f.): il. (algumas color.).

Inclui apêndice.

Orientador: Wellington Ronildo Clarindo.

Dissertação (mestrado) - Universidade Federal de Viçosa,
Departamento de Biologia Geral, 2024.

Referências bibliográficas: f. 29-31.

DOI: <https://doi.org/10.47328/ufvbbt.2024.314>

Modo de acesso: World Wide Web.

1. Eucalipto - Genética. 2. Mapeamento cromossômico.
3. Citogenética. I. Clarindo, Wellington Ronildo, 1979-.
II. Universidade Federal de Viçosa. Departamento de Biologia
Geral. Programa de Pós-Graduação em Genética e
Melhoramento. III. Título.

CDD 22. ed. 576.53


WANESSA ALVES LIMA

**CITOGENÔMICA COMPARATIVA DOS ELEMENTOS RT-LTR SIRE, OGRE,
TEKAY E IKEROS EM TRÊS ESPÉCIES DE *Eucalyptus* L'Her**


Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

APROVADA: 29/02/2024

Assentimento:

Documento assinado digitalmente
 **WANESSA ALVES LIMA**
Data: 05/07/2024 11:18:32-0300
Verifique em <https://validar.iti.gov.br>

Wanessa Alves Lima
Autor

Documento assinado digitalmente
 **WELLINGTON RONILDO CLARINDO**
Data: 08/07/2024 09:45:28-0300
Verifique em <https://validar.iti.gov.br>

Wellington Ronildo Clarindo
Orientador

AGRADECIMENTOS

À Deus pela capacidade, saúde e força, para a realização da minha pós-graduação.

À Universidade Federal de Viçosa e ao Programa de Pós-Graduação em Genética e Melhoramento pela oportunidade de realizar a pós-graduação.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Código de Financiamento 001.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da minha bolsa de estudos.

Ao meu orientador Prof. Dr. Wellington Ronildo Clarindo por todo o ensino e auxílio, e a todos os meus colegas do Laboratório de Citogenética e Citometria por toda a assistência.

À minha orientadora da graduação Francielle Martins pelo encorajamento à realização desta pós graduação e por todos os conselhos.

Aos meus amigos Melissa, Geicy, Isabelle, Giovanna, Brizza, Arlam, e ao meu namorado Gabriel, por todo o incentivo.

À toda minha família, em especial aos meus pais Maria de Lourdes e Raimundo, à minha avó Maria do Carmo, aos meus irmãos Nalva, Marilene, Rosa e César, e aos meus sobrinhos pelo apoio durante minha jornada acadêmica.

RESUMO

LIMA, Wanessa Alves. Universidade Federal de Viçosa, fevereiro de 2024. **Citogenômica comparativa dos elementos RT-LTR Sire, Ogre, Tekay e Ikeros em três espécies de *Eucalyptus* L'Her.** Orientador: Wellington Ronildo Clarindo.

O genoma estrutural de espécies de *Eucalyptus* tem sido elucidado por meio do mapeamento genético e do sequenciamento genômico. Os dados de sequenciamento são base para a citogenômica, os quais evidenciam a composição e a organização genômica das diferentes porções cromossômicas. Assim, nós identificamos e classificamos o repeatoma de *Eucalyptus grandis*, *Eucalyptus urophylla* e *Eucalyptus globulus*, e, a partir desses dados, realizamos a análise filogenética e mapeamos fisicamente os retroelementos com longas sequências terminais (RT-LTR) Sire, Ogre, Tekay e Ikeros. A fração repetitiva identificada nos genomas de *E. urophylla*, *E. grandis* e *E. globulus* representou, respectivamente, 24,1726%, 26,0248% e 24,7527%. As sequências RT-LTR foram as mais prevalentes, correspondendo a 15,60% em *E. urophylla*, 16,16% em *E. grandis* e 17,01% em *E. globulus*. A superfamília Ty1_Copia foi a mais abundante e diversa em todas as espécies. Notavelmente, os retroelementos Sire e Ikeros (Ty1_Copia) e Tekay e Ogre (Ty3_Gypsy) predominaram em todas as três espécies. Portanto, essas sequências foram escolhidas para a filogenia e o mapeamento físico. Os produtos de amplificação de Sire, Ogre, Tekay e Ikeros apresentam similaridade >85% com sequências encontradas nos $x = 11$ cromossomos das três espécies, com exceção de Ogre no cromossomo 4 de *E. grandis*. A filogenia revelou que Sire e Ikeros divergiram após a especiação, enquanto Ogre e Tekay divergiram antes da especiação das três espécies. Corroborando com a filogenia, a presença dispersa de Sire e Tekay foi confirmada por meio do mapeamento físico para todos os cromossomos de *E. grandis*, *E. urophylla* e *E. globulus*, e alguns sinais intensos foram mapeados em diferentes loci cromossômicos. Por outro lado, Ogre e Ikeros foram mapeados majoritariamente em regiões pericentroméricas de poucos cromossomos nas três espécies. Este estudo pioneiro classificou os elementos RT-LTR de *E. urophylla* e *E. globulus* e mapeou os elementos móveis mais abundantes nos cromossomos metafásicos de espécies de *Eucalyptus*. Portanto, nós reportamos o papel de sequências do repeatoma na composição, organização e estrutura do genoma nuclear e cromossômico de *Eucalyptus*. Acreditamos que esses resultados representam uma base fundamental para avanços nas pesquisas sobre elementos móveis em *Eucalyptus* e em toda a família Myrtaceae.

PALAVRAS-CHAVE: Elementos Móveis; Genômica Estrutural; Mapeamento Físico; Repeatoma; Filogenia.

ABSTRACT

LIMA, Wanessa Alves. Universidade Federal de Viçosa, February, 2024. **Comparative cytogenomics of the RT-LTR elements Sire, Ogre, Tekay and Ikeros in three *Eucalyptus* L'Her species.** Adviser: Wellington Ronildo Clarindo.

Eucalyptus structural genome has been unveiled through genetic mapping and genomic sequencing. Sequencing data is a basis for cytogenomics, which evidence the genomic composition and organization of different chromosome portions. So, we identified and classified the repeatome of *Eucalyptus grandis*, *Eucalyptus urophylla* and *Eucalyptus globulus*. Based on this, we also proposed a phylogeny and mapped the long terminal repeat retroelements (RT-LTR) Sire, Ogre, Tekay, and Ikeros. The repeatome was identified, representing, respectively, 24.1726%, 26.0248% and 24.7527% of the *E. urophylla*, *E. grandis* and *E. globulus* genomes. RT-LTR sequences were the most prevalent, corresponding to 15.60% in *E. urophylla*, 16.16% in *E. grandis*, and 17.01% in *E. globulus*. Ty1_Copia superfamily was the most abundant and diverse in all species. Notably, Sire and Ikeros (Ty1_Copia) and Tekay and Ogre (Ty3_Gypsy) retroelements predominated in the three species. Therefore, these sequences were chosen for phylogeny and physical mapping. Sire, Ogre, Tekay and Ikeros amplicons show >85% similarity in relation to sequences found in x = 11 chromosomes of the three species, except Ogre on *E. grandis* chromosome 4. The phylogeny revealed that Sire and Ikeros diverged after speciation, while Ogre and Tekay diverged before speciation of the species. Corroborating to the phylogeny, Sire and Tekay were disperselly mapped across all chromosomes of *E. grandis*, *E. urophylla* and *E. globulus*, with some intense fluorescent signals in different chromosome loci. Conversely, Ogre and Ikeros were mainly mapped in pericentromeric regions of few chromosomes in the three species. For the first time, RT-LTR elements of *E. urophylla* and *E. globulus* were classified, and the most abundant mobile elements were mapped on the metaphase chromosomes of *Eucalyptus* species. Therefore, we report the repeatome role in the composition, organization and structure of the nuclear and chromosome genomes of *Eucalyptus*. We believe that these results represent a fundamental basis for further researches about mobile elements in *Eucalyptus* and other Myrtaceae species.

KEYWORDS: Mobile Elements; Structural Genomics; Physical Mapping; Repeatome; Phylogeny.

SUMÁRIO

| | |
|---|----|
| 1 INTRODUÇÃO..... | 7 |
| 2 OBJETIVOS..... | 10 |
| 2.1 Objetivo geral | 10 |
| 2.2 Objetivos específicos..... | 10 |
| 3 MATERIAIS E MÉTODOS..... | 11 |
| 3.1 Identificação e classificação das sequências repetitivas nos genomas de <i>E. grandis</i> , <i>E. urophylla</i> e <i>E. globulus</i> | 11 |
| 3.2 Filogenia dos RT-LTR mais abundantes (Sire, Ogre, Tekay e Ikeros) em <i>E. grandis</i> , <i>E. urophylla</i> e <i>E. globulus</i> | 11 |
| 3.3 Desenho dos primers, amplificação das sequências dos quatro RT-LTR mais abundantes, sequenciamento e construção das sondas | 11 |
| 3.4 Mapeamento físico das sequências de RT-LTR Sire, Ogre, Tekay e Ikeros..... | 13 |
| 4 RESULTADOS | 15 |
| 4.1 Identificação e classificação do repeatoma de <i>E. grandis</i> , <i>E. urophylla</i> e <i>E. globulus</i> | 15 |
| 4.2 Similaridade dos RT-LTR Sire, Ogre, Tekay e Ikeros | 17 |
| 4.3 Filogenia dos RT-LTR Sire, Ogre, Tekay e Ikeros nos cromossomos de <i>E. grandis</i> , <i>E. urophylla</i> e <i>E. globulus</i> | 17 |
| 4.4 Mapeamento cromossômico dos elementos LTR-RT Sire, Ogre, Tekay e Ikeros em <i>E. grandis</i> , <i>E. urophylla</i> e <i>E. globulus</i> | 19 |
| 5 DISCUSSÃO..... | 25 |
| 6 CONCLUSÃO..... | 28 |
| 7 REFERÊNCIAS | 29 |
| 8 MATERIAL SUPLEMENTAR | 32 |

1 INTRODUÇÃO

O gênero *Eucalyptus* pertence à família Myrtaceae e possui cerca de 730 espécies (Nicolle, 2015). *Eucalyptus* são árvores nativas da Austrália, amplamente cultivadas em todo o mundo, com grandes plantações principalmente no Brasil, África, sudeste da Ásia e sul da Europa (Rozefelds, 1996). O gênero apresenta abundante variabilidade intra- e interespecífica decorrente, principalmente, da alogamia (Pryor e Johnson, 1971; Oudjehih e Abdellah, 2006). Logo, informações genômicas sobre *Eucalyptus* são importantes para compreensão da diversificação, evolução e adaptação do gênero aos diferentes habitats, assim como para programas de melhoramento e conservação de recursos genéticos.

Os primeiros trabalhos de citogenética em *Eucalyptus* foram realizados para a determinação do número cromossômico $2n$. O trabalho de Oudjehih e Abdellah (2006) foi decisivo para estabelecer que, apesar de sua diversidade ecológica e morfológica, o gênero *Eucalyptus* continua sendo um táxon natural com forte constância intra- e interespecífica do número cromossômico $2n$, mantendo um único número básico $x = 11$. Das 59 espécies investigadas, todas apresentaram invariavelmente $2n = 2x = 22$ cromossomos, evidenciando o caráter diploide do cariótipo.

Dentre as espécies de *Eucalyptus*, as espécies do subgênero *Symphyomyrtus* destacam-se pela qualidade da madeira, especialmente *E. grandis* e *E. urophylla*, predominantemente cultivadas em regiões tropicais e subtropicais, e *E. globulus*, preferencialmente cultivada em regiões temperadas (Potts, 2004). Uma alta sintonia e colinearidade foi revelada por mapas genéticos de alta densidade entre os genomas dessas três espécies, sugerindo que os genomas são altamente conservados (Hudson et al., 2012). Aliado a isso, os tamanhos dos genomas e %AT são similares, tendo o genoma de *E. urophylla* $1C = 0,650$ pg (635,70 Mbp) e 61,30% de AT, *E. globulus* $1C = 0,655$ pg (640,59 Mbp) e 61,57% de AT, e *E. grandis* $1C = 0,665$ pg (650,37 Mbp) e 61,89% de AT (Carvalho et al., 2017).

E. grandis é a espécie mais estudada do gênero e teve seu genoma sequenciado com uma cobertura de 6,7x o genoma completo (Myburg et al., 2014), tornando-se a primeira referência para a ordem Myrtales. O sequenciamento revelou que o genoma de *E. grandis* possui 36.376 loci que codificam proteínas e várias outras classes de genes não codificadores. O genoma da espécie também compreende 50,1% de sequências repetitivas. Além do genoma nuclear, os genomas plastidiais e mitocondriais da espécie foram montados e anotados (Pinard et al., 2019). Em contrapartida, os genomas nucleares de *E. globulus* e de *E. urophylla* apesar de montados, ainda não foram anotados.

Com base nos dados de sequenciamento genômico, a Citogenômica tem avançado consideravelmente nos últimos anos. O mapeamento físico de sequências por meio da Hibridização Fluorescente In Situ (FISH) determina o número e o locus da sequência-alvo e, conseqüentemente, caracteriza porções cromossômicas, assim como evidencia possíveis variações genômicas entre cromossomos de um cariótipo e entre cariótipos de diferentes indivíduos e/ou espécies. Assim, os dados de Citogenômica representam a base de muitos estudos taxonômicos, evolutivos e filogenéticos (Talukdar and Sinjushin, 2015).

Até o momento, a Citogenômica em *Eucalyptus* resultou no mapeamento de genes específicos e das sequências do repeatoma relativas ao rDNA (Ribeiro et al., 2016; Sattler et al., 2019). A similaridade entre os genomas de *E. baileyana* e *Corymbia citriodora* foi constatada por meio da Hibridização Genômica In Situ (GISH) (Carvalho et al., 2017). Esses resultados contribuíram para o progresso do mapeamento físico neste gênero. Portanto, outras sequências necessitam ser mapeadas para se conhecer a estrutura genômica dos cromossomos de *Eucalyptus*.

O repeatoma, durante muito tempo, foi negligenciado e reportado como porções genômicas sem função conhecida. Com base em várias evidências “ômicas”, as sequências repetitivas passaram a ser consideradas componentes genômicos relevantes. Contribuindo para mudanças de percepção, a abordagem citogenômica comparativa tem se mostrado vital por evidenciar a ocorrência, a distribuição e a evolução das sequências repetitivas identificadas a partir do sequenciamento genômico, revelando informações sobre sua origem e evolução (Biscotti et al., 2015).

As sequências repetitivas dispersas, conhecidas também como elementos móveis (TEs), são constituídas por transposons e retrotransposons (RT). Esses elementos móveis são atestados como um dos mecanismos evolutivos que esculpe o arcabouço do genoma, uma vez que são capazes de alterar a composição e a função gênica, como resultado da sua inserção dentro dos genes ou em suas proximidades, além de terem um papel importante na organização genômica telomérica e centromérica (Lisch, 2013; Wang et al., 2013; Vitte et al., 2014; Bennetzen e Wang, 2014).

Os retrotransposons com longas sequências terminais (RT-LTR) são os TEs mais abundantes em genomas vegetais (Tollis and Boissinot, 2012). Os RT-LTR possuem dois genes, o *gag* e o *pol*. O gene *gag* é responsável pela codificação de proteínas relacionadas à geração de partículas semelhantes aos vírus, enquanto *pol* codifica domínios proteicos essenciais no processo de transposição, incluindo protease, integrase, transcriptase reversa e RnaseH (Wicker et al., 2007; Tollis e Boissinot, 2012). Ty1_Copia e Ty3_Gypsy são as duas

superfamílias de RT-LTR em plantas (Bennetzen e Wang, 2014), as quais são distinguidas pela ordem de seus domínios internos (Wicker et al., 2007) (Figura 1). A superfamília Ty1_Copia é composta por elementos geralmente com menos de 6 Kb e LTRs menores que 500 bp. Os membros da superfamília Ty3_Gypsy têm mais de 6 Kb com LTRs maiores que 500 bp (Wang e Liu, 2008).

Baseados na análise filogenética dos três domínios de poliproteínas mais conservados em RT-LTR (RT, RH e IN, Figura 1), Neumann et al., (2019) classificaram a superfamília Ty1_Copia nos clados Ale, Alesia, Angela, Bianca, Bryco, Lyco, Gymco I–IV, Ikeros, Ivana, Osser, Sire, Tar e Tork. A superfamília Ty3_Gypsy foi classificada nos clados CRM, Chlamyvir, Galadriel, Tcn1, Reina, Tekay, Athila, Tat I–III, Ogre, Retand, Phygy e Selgy.

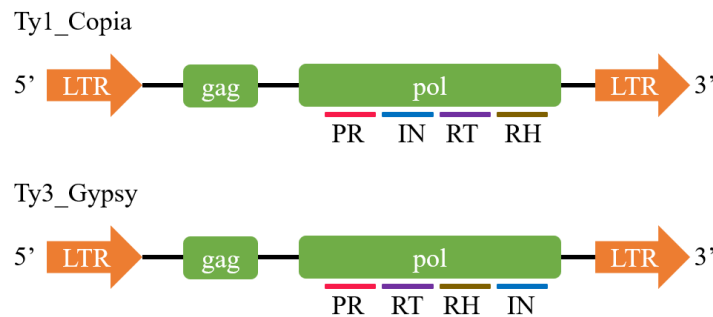


Figura 1: Estrutura das superfamílias de RT-LTR em plantas, destacando as longas repetições terminais flanqueando os genes *gag* e *pol*. No gene *pol*, as posições relativas dos domínios são detalhadas para Protease (PR), Integrase (IN), Transcriptase Reversa (RT) e RnaseH (RH) nas superfamílias Ty1_Copia e Ty3_Gypsy. Fonte: adaptado de Galindo-González et al., (2017) e de Thomas-Bulle et al., (2018).

Em *E. grandis*, os RT correspondem a maior parte do genoma (44,5%), sendo os RT-LTR a classe mais difundida (21,9%). Os transposons de DNA abrangem apenas 5,6% do genoma (Myburg et al., 2014). Recentemente, os RT-LTR de *E. grandis* foram classificados ao nível de clado dentro das superfamílias e tiveram suas idades de inserção estimadas por meio da ferramenta PlantLTRdb (Mokhtar et al., 2023). Portanto, o conhecimento acerca da estrutura e organização do genoma deve ser expandido para outras espécies do gênero, como *E. urophylla* e *E. globulus*.

2 OBJETIVOS

2.1 Objetivo geral

- Identificar, classificar e mapear sequências do repeatoma de *E. grandis*, *E. urophylla* e *E. globulus*.

2.2 Objetivos específicos

- Identificar clusters de sequências do repeatoma nos genomas de *E. grandis*, *E. urophylla* e *E. globulus*.
- Selecionar os quatro RT-LTR mais abundantes em *E. grandis*, *E. urophylla* e *E. globulus*.
- Propor uma filogenia por cromossomo dos quatro RT-LTR mais abundantes em *E. grandis*, *E. urophylla* e *E. globulus*.
- Caracterizar o cariótipo de *E. grandis*, *E. urophylla* e *E. globulus*.
- Mapear fisicamente os quatro RT-LTR mais abundantes em *E. grandis*, *E. urophylla* e *E. globulus*.

3 MATERIAIS E MÉTODOS

3.1 Identificação e classificação das sequências repetitivas nos genomas de *E. grandis*, *E. urophylla* e *E. globulus*

Reads paired-end Illumina foram obtidas do Sequence Read Archive (SRA) do banco de dados do NCBI de *E. grandis* (SRR21235325), *E. urophylla* (SRR21235326) e *E. globulus* (ERR9601529). Essas reads tiveram sua qualidade avaliada pela ferramenta FastQC (Brown et al., 2017). A análise comparativa das sequências repetitivas entre as espécies de *Eucalyptus* foi feita usando a ferramenta RepeatExplorer2 clustering, hospedada na plataforma Galaxy (<https://repeatexplorerelixir.cerit-sc.cz/galaxy/>, Novák et al., 2013). As reads foram padronizadas no tamanho de 100 bp e o número de reads analisado foi de 643.931 para *E. grandis*, de 629.406 para *E. urophylla* e de 634.248 para *E. globulus*, correspondendo a aproximadamente 0,1x do genoma de cada espécie. A partir da análise de agrupamento, os elementos repetitivos foram anotados automaticamente com base no banco de dados Viridiplantae v3.0 do REXdb (Neumann et al., 2019).

3.2 Filogenia dos RT-LTR mais abundantes (Sire, Ogre, Tekay e Ikeros) em *E. grandis*, *E. urophylla* e *E. globulus*

Com base nas sequências consenso dos RT-LTR Sire, Ogre, Tekay e Ikeros obtidas no sequenciamento, a filogenia dessas quatro sequências foi proposta para cada cromossomo ($x = 11$) de *E. grandis*, *E. urophylla* e *E. globulus*. Primeiramente, cada sequência consenso foi submetida a uma análise no BLAST contra o genoma de referência de *E. grandis* (GCA_016545825.1), *E. urophylla* (GCA_026873135.1) e *E. globulus* (GCA_014182545.1). As sequências correspondentes encontradas em cada cromossomo foram exportadas e agrupadas por RT-LTR. Em seguida, no programa MEGA 11 (Tamura et al., 2021), as sequências identificadas para cada RT-LTR foram alinhadas usando o algoritmo ClustalW e as árvores filogenéticas foram geradas pelos métodos Neighbor-Joining e Maximum Likelihood (Hall, 2013). Por fim, as árvores filogenéticas foram customizadas na ferramenta iTOL (Interactive Tree Of Life) (Letunic and Bork, 2024).

3.3 Desenho dos primers, amplificação das sequências dos quatro RT-LTR mais abundantes, sequenciamento e construção das sondas

Nós selecionamos os maiores contigs de clusters dos elementos RT-LTR Sire, Ogre, Tekay e Ikeros. Em seguida, nós analisamos esses contigs no NCBI Conserved Domain Search (Marchler-Bauer et al., 2015), a fim de identificar regiões onde há os domínios do gene que codifica a integrase, um dos domínios poliproteicos mais conservados em RT-LTR (Malik e Eickbush, 1999; Neumann et al., 2019). Uma vez identificados os domínios de integrase, os primers foram desenhados usando o software Geneious Prime para abranger essa sequência (Tabela 1).

Tabela 1: Primers *forward* (*F*) e *reverse* (*R*) utilizados para amplificar as sequências dos elementos RT-LTR Sire, Ogre, Tekay e Ikeros a partir do DNA genômico de *E. grandis*.

| RT-LTR | Primers | T _m (°C) |
|--------|--|---------------------|
| Sire | <i>F</i> - 5' ACTCAGTCTTTCATGGGTCGA 3' | 55,7 |
| | <i>R</i> - 5' AGGCAATACTTCTTACCCCA 3' | |
| Ogre | <i>F</i> - 5' TCCCCACATTGAGAAAGGC 3' | 57,3 |
| | <i>R</i> - 5' TGGCCGTGAATGAAGAGCTT 3' | |
| Tekay | <i>F</i> - 5' CGGGACACTAAGATTCCAAGGA 3' | 57 |
| | <i>R</i> - 5' GAGGTTGCAGAAGTCCTCCC 3' | |
| Ikeros | <i>F</i> - 5' TGTTGACTGTGGCGATGTCA 3' | 55,7 |
| | <i>R</i> - 5' ACTTCAGTCTTTTACCTTTTCCA 3' | |

Visto que os primers amplificam sequências comum às três espécies de *Eucalyptus*, as sequências de integrase foram amplificadas utilizando o DNA genômico de *E. grandis*. A extração do DNA genômico de *E. grandis* foi realizada utilizando o kit PureLink™ Plant Total DNA Purification (Invitrogen®). O DNA teve sua concentração e a pureza mensuradas por espectrofotometria utilizando o NanoDrop 1000 (Thermo Fisher Scientific®), enquanto sua integridade foi avaliada por meio de eletroforese em gel de agarose a 1,5%. A amplificação foi realizada em uma reação composta por: 0,5 µM de cada primer, 50 ng de DNA genômico, 200 µM de cada dNTP (Promega®), tampão de reação 1X (Promega®), 1,8 mM de MgCl₂ (Promega®) e 1,25 U de DNA polimerase GoTaq (Promega®). A amplificação foi realizada em um termociclador nas seguintes condições: desnaturação inicial a 95°C por 3 min; 30 ciclos de desnaturação a 95°C por 1 min; anelamento a 54°C por 1 min; extensão a 72°C por 1 min; e extensão final a 72°C por 2 min. Os produtos da amplificação foram

quantificados por NanoDrop 1000 (Thermo Fisher Scientific®) e analisados por eletroforese em gel de agarose 1,5% (Figura Suplementar 1).

Os produtos de PCR foram sequenciados na empresa GenomaA Biotech pelo método MinION (Oxford Nanopore Technology). Utilizando a sequência consenso de Sire, Ogre, Tekay e Ikeros (Tabela Suplementar 1), a similaridade foi verificada por meio da comparação do sequenciamento com os dados do NCBI via ferramenta Basic Local Alignment Search Tool (BLAST).

As sondas foram geradas em uma nova reação de PCR contendo: 50 ng do DNA relativo à sequência amplificada, 200 µM de dATP, dCTP e dGTP, 150 µM de dTTP, 0,05 µM dos primers, tampão de reação 1X (Promega®), 1,8 mM de MgCl₂ (Promega®), 1,25 U de DNA polimerase GoTaq (Promega®), e 0,05 mM do fluorocromo ChromaTide Alexa Fluor 488-5-dUTP (Life Technologies®) e Tetrametilrodamina 5-dUTP (Roche®) para as quatro sequências amplificadas. Os produtos de amplificação foram verificados com eletroforese em gel de agarose a 1,5%.

3.4 Mapeamento físico das sequências de RT-LTR Sire, Ogre, Tekay e Ikeros

Sementes obtidas comercialmente de *E. grandis*, *E. urophylla* e *E. globulus* foram germinadas em placa de Petri a 30°C. Os meristemas radiculares de *E. grandis* e *E. globulus* foram incubados por 3 h a 30°C em 4 µM de amiprofos-metil. Os meristemas radiculares de *E. urophylla* foram incubados por 3 h a 30°C em 60 µL de dimetilsulfóxido puro + 4 µM de amiprofos-metil. Posteriormente, os meristemas radiculares foram lavados em dH₂O por três vezes durante 5 min, e fixados em 3:1 (v/v) de metanol-ácido acético (Merck®) por 24 h. Os meristemas radiculares foram lavados por três vezes em dH₂O e macerados em pool enzimático (4% celulase + 0,4% hemicelulase + 1% pectoliase diluídas em solução de pectinase, Sigma®) diluído em dH₂O na proporção de 1:30 (pool enzimático : água) por 2 h a 36°C. Os meristemas foram lavados em dH₂O e fixados em metanol e ácido acético 3:1 (v/v) (Merck®) por 24 h. Por fim, as lâminas foram preparadas pelas técnicas de dissociação celular e secagem ao ar (Adaptado de Ribeiro et al., 2016; Silva et al., 2018).

As lâminas com cromossomos metafásicos foram envelhecidas por pelo menos 3 dias a 30°C em estufa. Em seguida, as lâminas foram lavadas em tampão de solução salina tamponada com fosfato 1x (PBS) (NaH₂PO₄ + Na₂HPO₄ + NaCl + H₂O) por 5 min e fixadas em 4% de formalina por 15 min. Após a fixação, as lâminas foram lavadas novamente em PBS 1X por 5 min. A desnaturação dos cromossomos foi realizada em solução de formamida

70% por 3 min a 72°C em banho-Maria e posteriormente desidratados em série alcoólica fria (70%, 85% e 100%) por 5 min cada. A mistura de hibridização foi composta por 3,5 µL de solução salina de citrato de sódio 2x SSC (NaCl + C₆H₅Na₃O₇ + dH₂O, Sigma[®]), 9 µL de água ultrapura, 17,5 µL de formamida (Sigma[®]), 1 µL de Herring Sperm DNA (Promega[®]) e 2 µL da sonda marcada. As sondas foram desnaturadas a 85°C por 5 min no Termociclador e imediatamente transferidas para o gelo por 5 min. Posteriormente, a mistura de hibridização foi adicionada às lâminas pré-tratadas. Logo, as lâminas foram cobertas com lamínula plástica HybriSlip (Sigma[®]) e seladas com cola Rubber Cement (Elmer's[®]). A hibridização foi conduzida no Sistema Thermobrite (Thermo Fisher Scientific[®]) a 37°C por 24 h. Após a hibridização, as lamínulas foram removidas e as lâminas foram lavadas em solução de SSC 2x a 40°C por 5 min com nível de estringência de 51% para remover o excesso de sonda. Por fim, as lâminas foram contracoradas com 50 µL de 40% glicerol/PBS + ácido 6-diamidino-2-fenilacético (adaptado de Ribeiro et al., 2016; Sattler et al., 2019; Silva et al. 2020). Para garantir resultados confiáveis, múltiplas hibridizações foram conduzidas. Além disso, as quatro sondas foram marcadas tanto com a ChromaTide Alexa Fluor 488-5-dUTP quanto com a Tetrametilrodamina 5-dUTP. As imagens foram capturadas com uma câmera digital CCD (OlympusTM DP71) acoplada a um fotomicroscópio de fluorescência Olympus BX-60 equipado com objetiva de imersão de 100x. Em seguida, as melhores metáfases foram selecionadas para sobreposição e ajuste de brilho e contraste no Adobe Photoshop 2022. Por fim, os cariótipos e o idiograma foram montados utilizando o mesmo software.

4 RESULTADOS

4.1 Identificação e classificação do repeatoma de *E. grandis*, *E. urophylla* e *E. globulus*

Usando o pipeline RepeatExplorer2 clustering, nós caracterizamos a fração repetitiva do genoma de *E. grandis*, *E. urophylla* e *E. globulus*. 3.815.170 reads foram analisadas, e destas, 33% foram incluídas nos principais clusters (Figura Suplementar 2). A análise comparativa gerou 323 clusters com abundância maior que 0,01% dos genomas (Figura Suplementar 3).

Esses clusters são formados pelas reads identificadas como sequências repetitivas e representam 24,1726%, 26,0248% e 24,7527% do genoma de *E. urophylla*, *E. grandis* e *E. globulus*, respectivamente (Tabela 2). Os RT-LTR foram as repetições mais prevalentes, correspondendo a 15,60% do genoma de *E. urophylla*, 16,16% de *E. grandis* e 17,01% de *E. globulus*. A superfamília Ty1_Copia apresentou uma notável diversidade, abrangendo nove linhagens distintas. Além disso, Ty1_Copia destacou-se como a mais prevalente, compreendendo 8,25%, 8,25% e 10,18% do genoma de *E. urophylla*, *E. grandis* e *E. globulus*, respectivamente. Já a superfamília Ty3_Gypsy compreendeu 4,41% de *E. urophylla*, 4,70% de *E. grandis* e 4,13% de *E. globulus*. Nesse contexto, os elementos Sire e Ikeros, pertencentes à superfamília Ty1_Copia, e os elementos Tekay e Ogre, pertencentes à superfamília Ty3_Gypsy, foram os mais prevalentes em todas as três espécies. Importante salientar que não há diferença significativa na abundância das superfamílias e de seus respectivos elementos entre as três espécies de *Eucalyptus*.

Os transposons de DNA e os retrotransposons não-LTR compõem uma fração menor dos genomas, representando coletivamente 1,09% do genoma de *E. urophylla*, 1,08% de *E. grandis* e 1,28% de *E. globulus*. Do mesmo modo, as sequências repetitivas em tandem, satDNA e rDNA representaram uma pequena parte do genoma das espécies, sendo 0,28% e 1,28% respectivamente em *E. urophylla*, 0,69% e 1,95% em *E. grandis*, e 0,34% e 0,58% em *E. globulus*.

Tabela 2. Composição da fração repetitiva identificada pelo RepeatExplorer2 clustering nos genomas de *E. urophylla*, *E. grandis* e *E. globulus*.

| | | | <i>E. urophylla</i> | <i>E. grandis</i> | <i>E. globulus</i> | |
|--|----------------------|-----------|---------------------|-------------------|--------------------|--------|
| Conteúdo repetitivo total (%) | | | 24,1726 | 26,0248 | 24,7527 | |
| Ordem | Superfamília | Família | Abundância | | | |
| | Não classificado | | 5,9100 | 6,1310 | 5,5273 | |
| Sequências de DNA repetitivo dispersas | | | | | | |
| Classe I – Retrotransposons | | | | | | |
| LTR | LTR não classificado | | 2,9400 | 3,1989 | 2,6967 | |
| | | Ale | 0,4555 | 0,48557 | 0,5662 | |
| | | Alesia | 0,0422 | 0,0275 | 0,0213 | |
| | | Angela | 0,2029 | 0,1971 | 0,2348 | |
| | | Bianca | 0,1359 | 0,1000 | 0,1462 | |
| | | Ty1_Copia | Ikeros | 2,3692 | 2,2293 | 2,5319 |
| | | | Ivana | 0,3224 | 0,3110 | 0,4029 |
| | | | Sire | 3,6169 | 3,8479 | 4,9720 |
| | | | TAR | 0,3833 | 0,3668 | 0,5669 |
| | | | Tork | 0,7218 | 0,6881 | 0,7460 |
| | | | Chromo/CRM | 0,2615 | 0,2655 | 0,2699 |
| | | | Chromo/Galadriel | 0,1953 | 0,1844 | 0,0980 |
| | | Ty3_Gypsy | Chromo/Tekay | 2,0751 | 2,2783 | 1,8510 |
| | | | Athila | 0,1949 | 0,1718 | 0,2548 |
| | | | Ogre | 1,6895 | 1,8094 | 1,6576 |
| Não-LTR | LINE | | 0,1060 | 0,1122 | 0,1089 | |
| | Pararetrovírus | | 0,3570 | 0,3918 | 0,4643 | |
| Total Retrotransposons | | | 16,0694 | 16,6656 | 17,5894 | |
| Classe II – Transposons | | | | | | |
| Subclasse I - TIR | hAT | | 0,1357 | 0,1233 | 0,1690 | |
| | PIF_Harbinger | | 0,2010 | 0,1945 | 0,2293 | |
| | MuDR_Mutator | | 0,2905 | 0,2619 | 0,3107 | |
| Total Transposons | | | 0,6272 | 0,5797 | 0,7090 | |
| Total repetitivo disperso | | | 16,6966 | 17,2453 | 18,2984 | |
| Sequências de DNA repetitivo em tandem | | | | | | |
| satDNA | | | 0,2860 | 0,6909 | 0,3416 | |
| rDNA | | | 1,2800 | 1,9576 | 0,5854 | |
| Total repetitivo em tandem | | | 1,5660 | 2,6485 | 0,9270 | |

4.2 Similaridade dos RT-LTR Sire, Ogre, Tekay e Ikeros

As sequências consenso de Sire, Ogre, Ikeros e Tekay apresentaram similaridade com sequências presentes em todos os cromossomos de *E. grandis*, *E. urophylla* e *E. globulus*, depositadas no NCBI, exceto Ogre no cromossomo 4 de *E. grandis* (Tabela Suplementar 2). Em *E. grandis*, Sire, Ogre, Tekay e Ikeros tiveram uma correspondência superior a 96%, 94%, 92% e 92%, respectivamente, em praticamente todos os cromossomos e RT-LTR, exceto Ogre no cromossomo 4. Por outro lado, em *E. urophylla*, esses valores foram, respectivamente, de 98%, 95%, 88% e 93%, enquanto em *E. globulus* foram, respectivamente, de 97%, 95%, 85% e 92%.

4.3 Filogenia dos RT-LTR Sire, Ogre, Tekay e Ikeros nos cromossomos de *E. grandis*, *E. urophylla* e *E. globulus*

Árvores filogenéticas dos RT-LTR Sire, Ogre, Tekay e Ikeros nos cromossomos de *E. grandis*, *E. urophylla* e *E. globulus* foram geradas pelos métodos Neighbor-Joining e Maximum Likelihood (Figura 2). As árvores filogenéticas de Sire e Ikeros (superfamília Ty1_Copia) sugerem que estas sequências divergiram-se após a especiação das três espécies de *Eucalyptus*, pois os clados nas árvores são predominantemente compostos por cromossomos da mesma espécie em ambos os métodos de estimativa. Por outro lado, as árvores filogenéticas de Ogre e Tekay (superfamília Ty3_Gypsy) indicam que estes divergiram antes da especiação das espécies, já que a maioria dos clados nas árvores de ambos os métodos de estimativa contém cromossomos das três espécies.

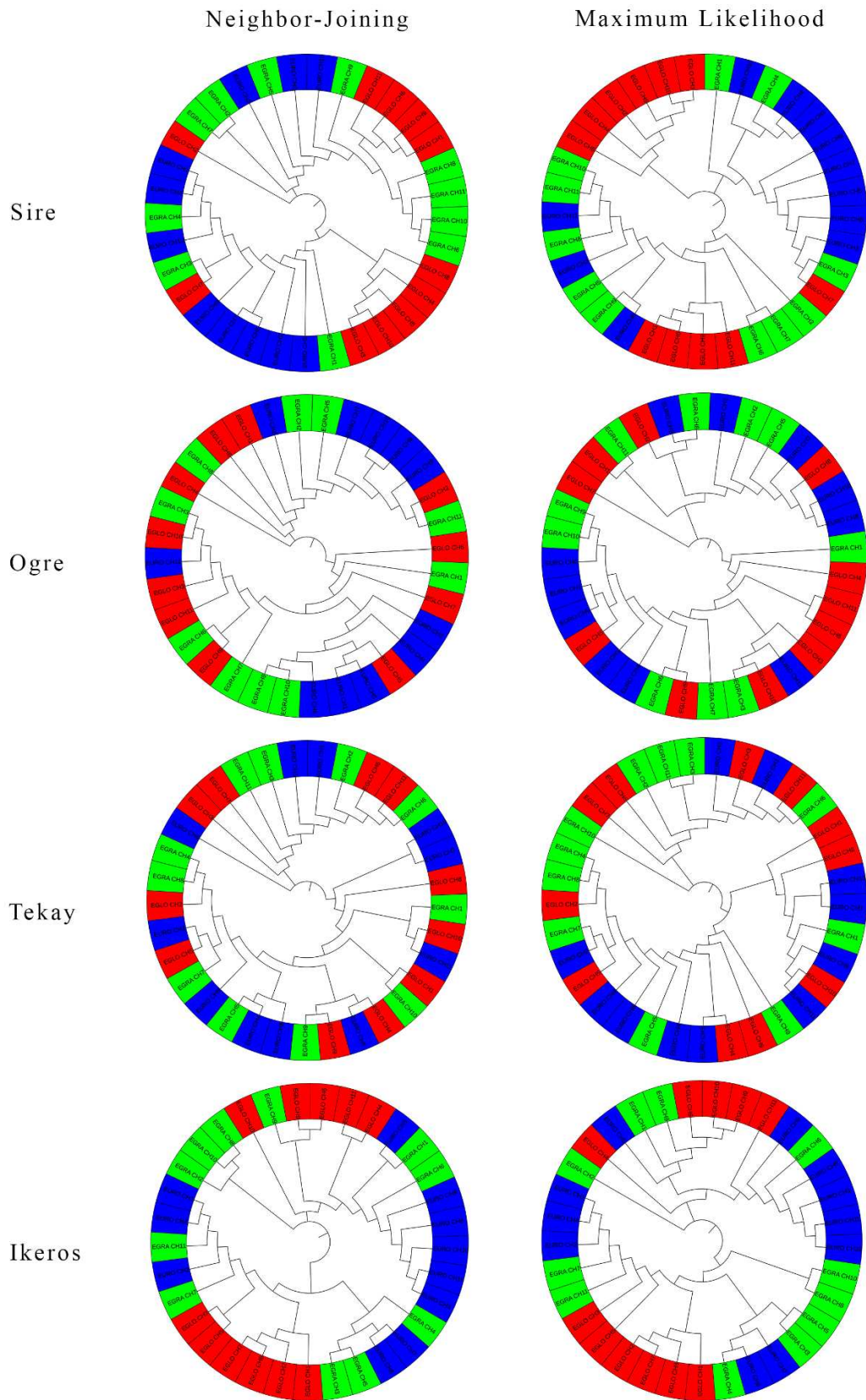


Figura 2. Árvores filogenéticas dos RT-LTR Sire, Ogre, Tekay e Ikeros por cromossomo entre as espécies *E. grandis* (EGRA) (cor verde), *E. urophylla* (EURO) (cor azul) e *E. globulus* (EGLO) (cor vermelha) pelos métodos Neighbor-Joining e Maximum Likelihood. Os números depois da sigla CH referem-se ao cromossomo.

4.4 Mapeamento cromossômico dos elementos LTR-RT Sire, Ogre, Tekay e Ikeros em *E. grandis*, *E. urophylla* e *E. globulus*

Sinais de fluorescência das sequências Sire, Ogre, Tekay e Ikeros foram detectados e mapeados nos cromossomos metafásicos de *E. grandis*, *E. urophylla* e *E. globulus*. Sire exibiu sinais de hibridização dispersos em todos os cromossomos das três espécies (Figura 3). Em *E. grandis*, sinais intensos foram mapeados nas regiões intersticiais do braço curto dos cromossomos 1 e 3. Em *E. globulus*, sinais fortes foram mapeados na região pericentromérica do cromossomo 4, assim como na região intersticial do braço curto do cromossomos 3, e na porção terminal do braço curto do cromossomo 9. Em *E. urophylla*, sinais intensos foram mapeados na região terminal do braço curto do cromossomo 10, no braço longo dos cromossomos 6 e 11, na região intersticial do braço curto dos cromossomos 1 e 7, e na região pericentromérica do cromossomo 3.

Ogre foi mapeado (Figura 4) em *E. grandis* nas regiões pericentroméricas dos cromossomos 1, 2, 3, 7 e 10. Em *E. globulus*, Ogre foi mapeado nas regiões pericentroméricas dos cromossomos 1 e 4 e na região intersticial do braço curto dos cromossomos 2 e 11. Em *E. urophylla*, Ogre foi mapeado nas regiões pericentroméricas dos cromossomos 2, 3 e 8, e na região intersticial do braço longo do cromossomo 6.

Tekay apresentou sinais fracos de hibridização dispersos em todos os cromossomos das três espécies (Figura 5). Em *E. grandis*, sinais intensos foram mapeados nas regiões intersticiais do cromossomo 8, na região terminal do braço longo do cromossomo 4 e na região intersticial do braço curto do cromossomo 10. Em *E. globulus*, Tekay exibiu sinais intensos mapeados na região intersticial dos braços curto e longo do cromossomo 2, na região intersticial do braço curto dos cromossomos 5, 6 e 10, na porção terminal do braço curto dos cromossomos 7 e 8, e no braço longo dos cromossomos 4 e 11. Em *E. urophylla*, sinais intensos foram mapeados na região terminal do braço curto dos cromossomos 2 e 8, e nas regiões pericentroméricas dos cromossomos 3 e 5.

Ikeros foi mapeado (Figura 6) em *E. grandis* nas regiões pericentroméricas dos cromossomos 2, 4, 5, 7 e 9. Em *E. globulus*, Ikeros foi mapeado nas regiões pericentroméricas dos cromossomos 1, 3 e 6 e na região intersticial do braço curto do cromossomo 7 e dos braços curto e longo do cromossomo 5. Em *E. urophylla*, Ikeros foi mapeado nas regiões pericentroméricas dos cromossomos 2, 4 e 6.

Em resumo, o mapeamento físico mostrou a abundante presença de Sire e Tekay ao longo de todos os cromossomos de *E. grandis*, *E. globulus* e *E. urophylla*, e a presença de

Ikeros e Ogre predominantemente em regiões pericentroméricas de apenas alguns cromossomos das três espécies (Figura 7).

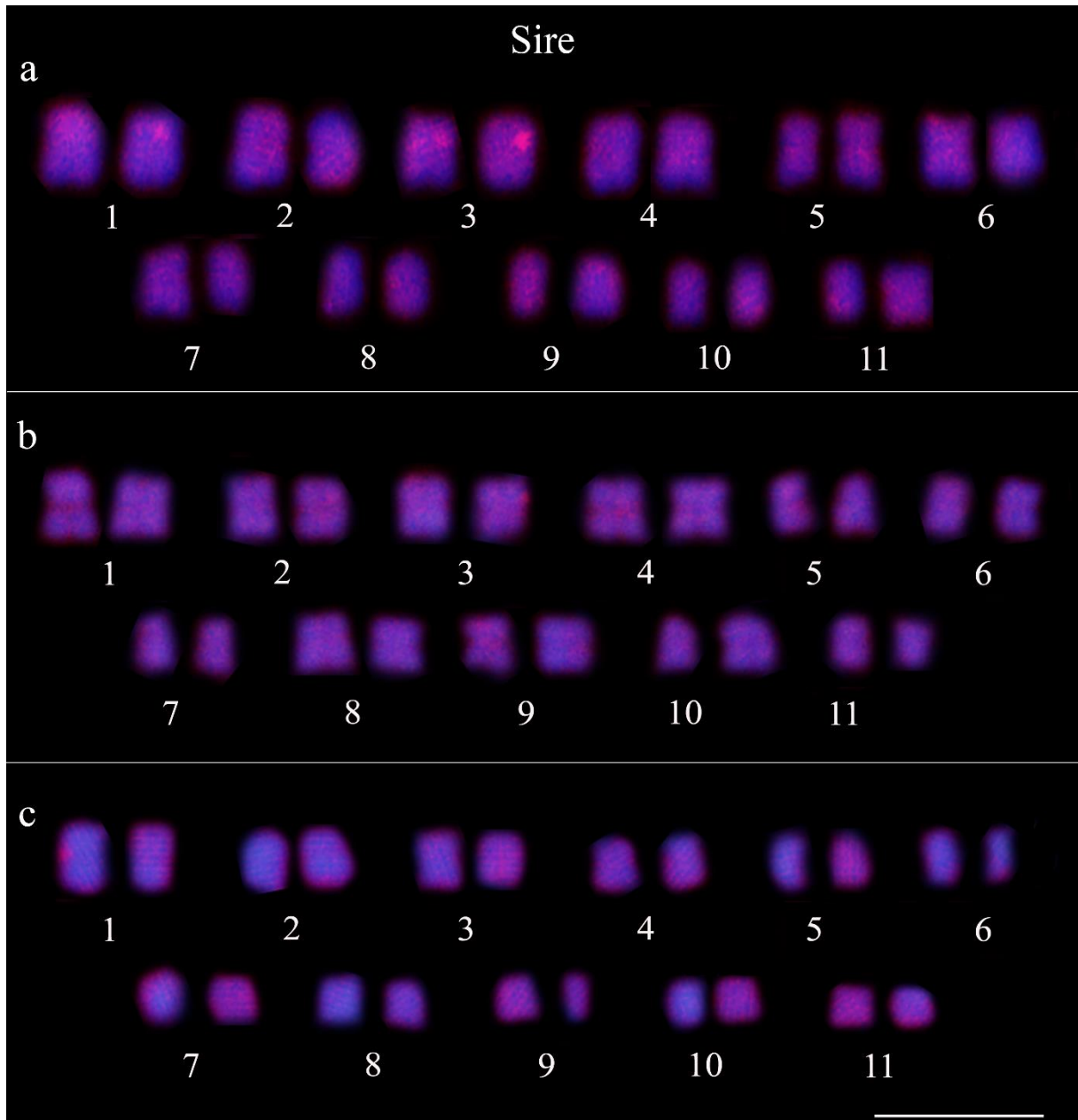


Figura 3. Mapeamento físico do elemento RT-LTR Sire da superfamília Ty1_Copia em cromossomos metafásicos de *E. grandis* (a), *E. globulus* (b) e *E. urophylla* (c). Sinais fracos e dispersos de hibridização da sonda marcada com Tetrametilrodamina 5-dUT foram mapeados ao longo de todos os cromossomos das três espécies. a) sinais fortes nas regiões intersticiais do braço curto dos cromossomos 1 e 3. b) sinais fortes na região pericentromérica do cromossomo 4, assim como na região intersticial do braço curto do cromossomos 3, e na porção terminal do braço curto do cromossomo 9. c) sinais fortes na região terminal do braço curto do cromossomo 10, no braço longo dos cromossomos 6 e 11, na região intersticial do braço curto do cromossomo 1 e 7, e na região pericentromérica do cromossomo 3. Barra = 5 μm .

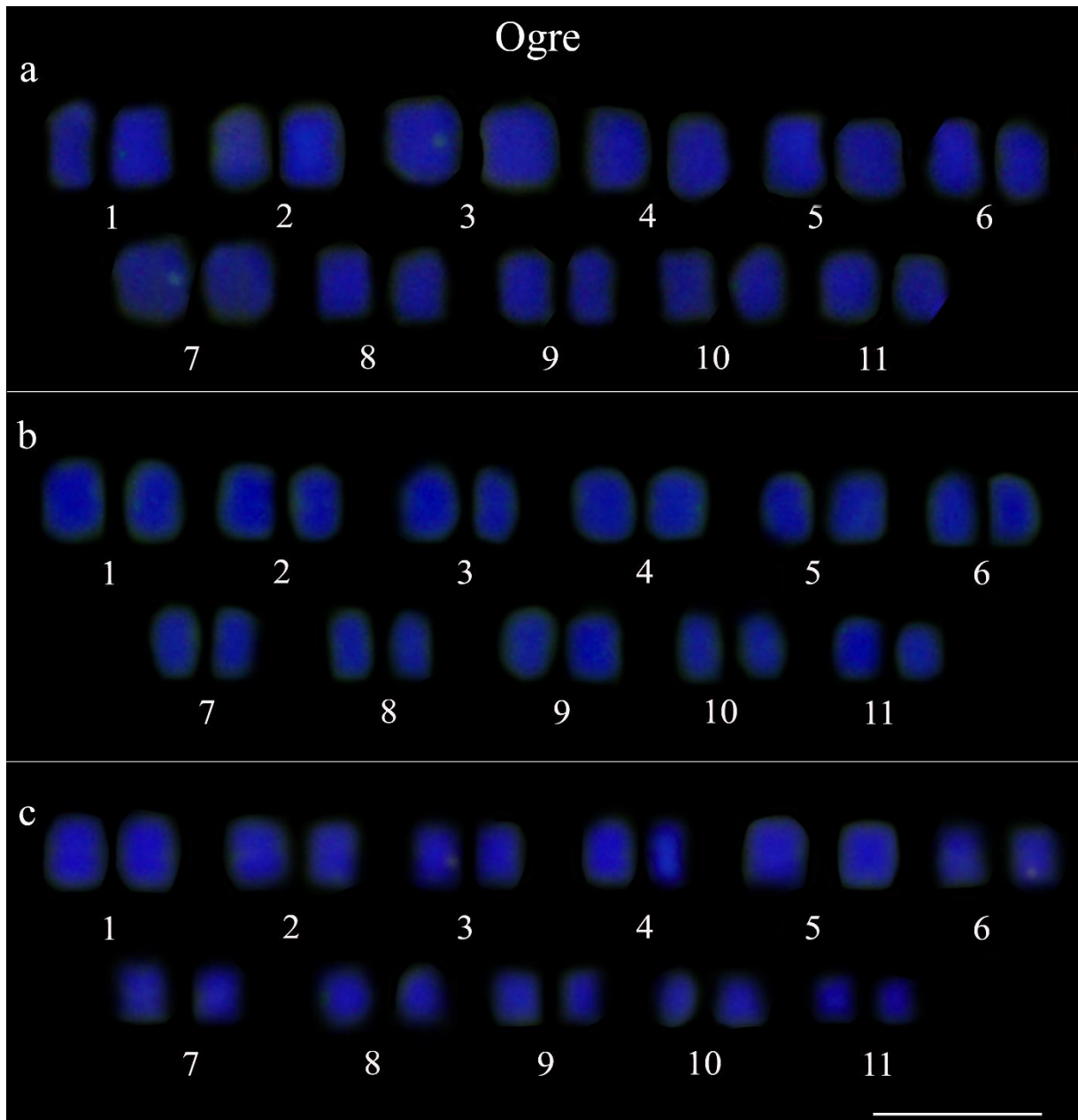


Figura 4. Mapeamento físico do elemento RT-LTR Ogre da superfamília Ty3_Gypsy em cromossomos metafásicos de *E. grandis* (a), *E. globulus* (b) e *E. urophylla* (c). a) Sinais de fluorescência da sonda marcada com ChromaTide Alexa Fluor 488-5-dUTP nas regiões pericentroméricas dos cromossomos 1, 2, 3, 7 e 10. b) Sinais nas regiões pericentroméricas dos cromossomos 1 e 4 e intersticial do braço curto dos cromossomos 2 e 11. c) Sinais nas regiões pericentroméricas dos cromossomos 2, 3 e 8, e intersticial do braço longo do cromossomo 6. Barra = 5 μ m.

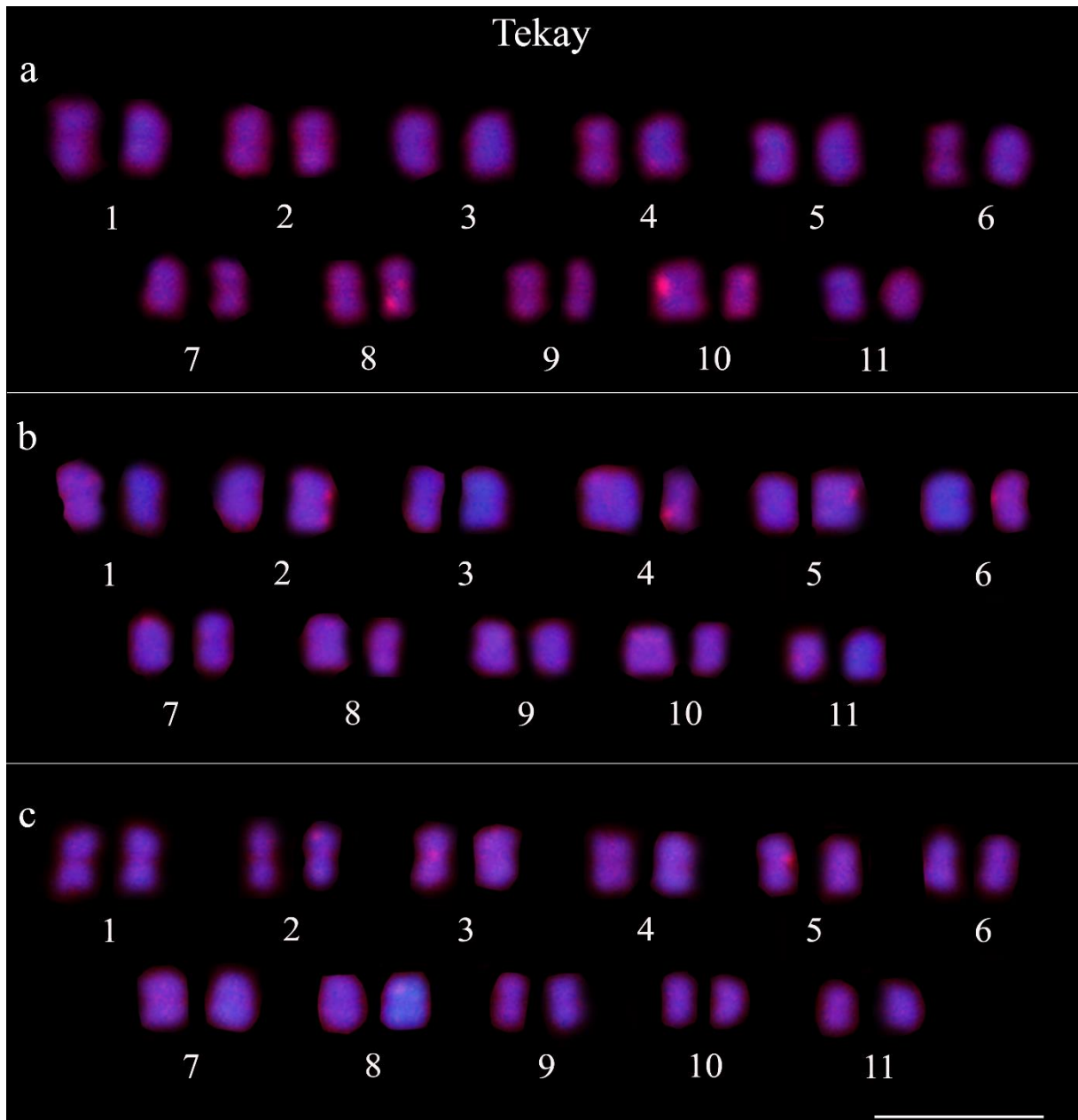


Figura 5. Mapeamento físico do elemento RT-LTR Tekay da superfamília Ty3_Gypsy em cromossomos metafásicos de *E. grandis* (a), *E. globulus* (b) e *E. urophylla* (c). Sinais de hibridização fracos e dispersos da sonda marcada com Tetrametilrodamina 5-dUT ao longo de todos os cromossomos das três espécies. a) Sinais intensos nas regiões intersticiais do cromossomo 8, na região terminal do braço longo do cromossomo 4 e na região intersticial do braço curto do cromossomo 10. b) Sinais intensos na região intersticial dos braços curto e longo do cromossomo 2, na região intersticial do braço curto dos cromossomos 5, 6 e 10, no braço longo dos cromossomos 4 e 11, e na porção terminal do braço curto dos cromossomos 7 e 8. c) Sinais intensos na região terminal do braço curto do cromossomo 2, nas regiões pericentroméricas dos cromossomos 3 e 5, e na porção terminal do braço curto do cromossomo 8. Barra = 5 μ m.

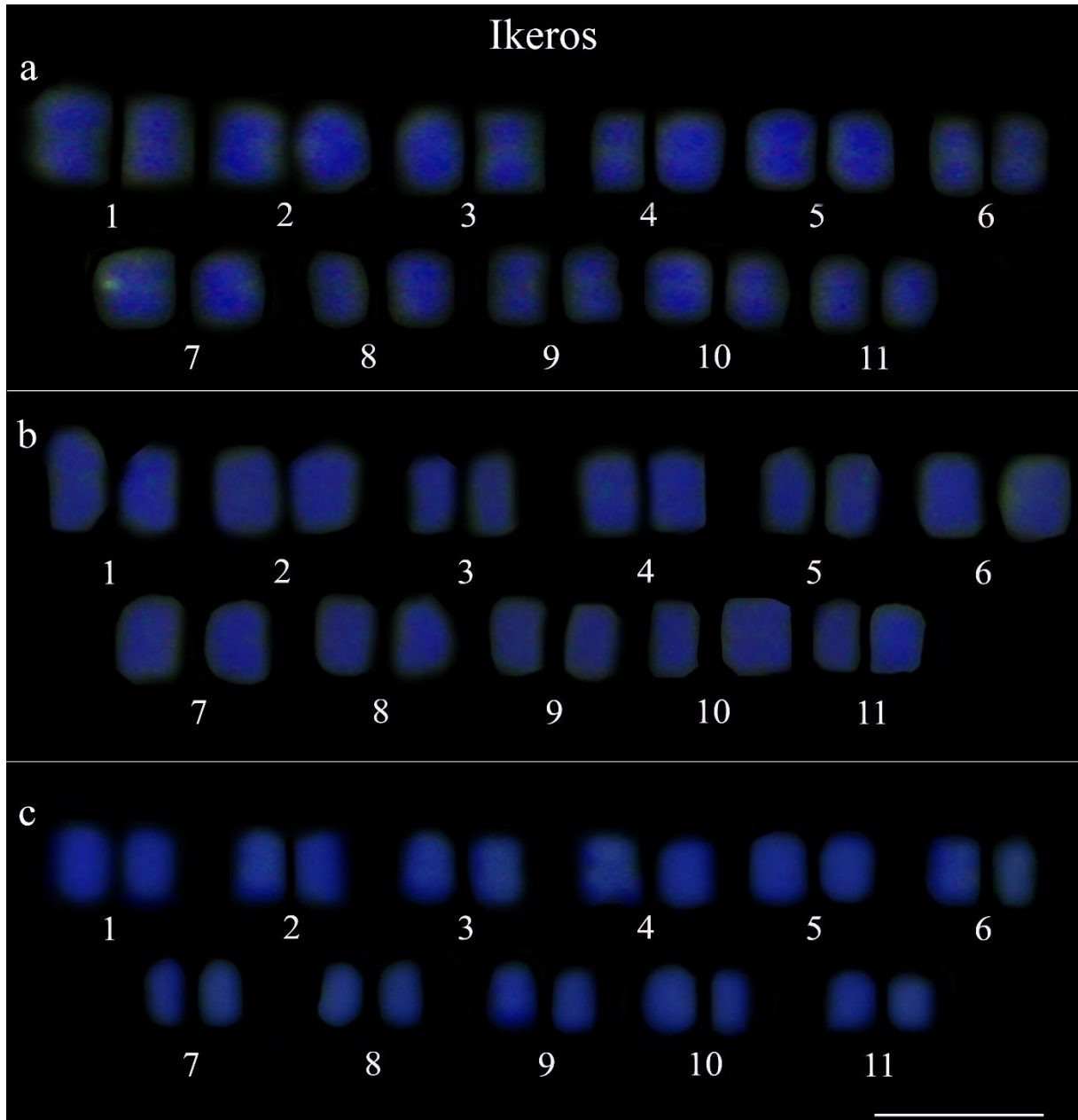


Figura 6. Mapeamento físico do elemento RT-LTR Ikeros da superfamília Ty1_Copia em cromossomos metafásicos de *E. grandis* (a), *E. globulus* (b) e *E. urophylla* (c). a) Sinais intensos de fluorescência da sonda marcada com ChromaTide Alexa Fluor 488-5-dUTP nas regiões pericentroméricas dos cromossomos 2, 4, 5, 7 e 9. b) Sinais fracos nas regiões pericentroméricas dos cromossomos 1, 3 e 6, intersticial do braço curto do cromossomos 7, e dos braços curto e longo do cromossomo 5. c) Sinais fracos nas regiões pericentroméricas dos cromossomos 2, 4 e 6. Barra = 5 μm .

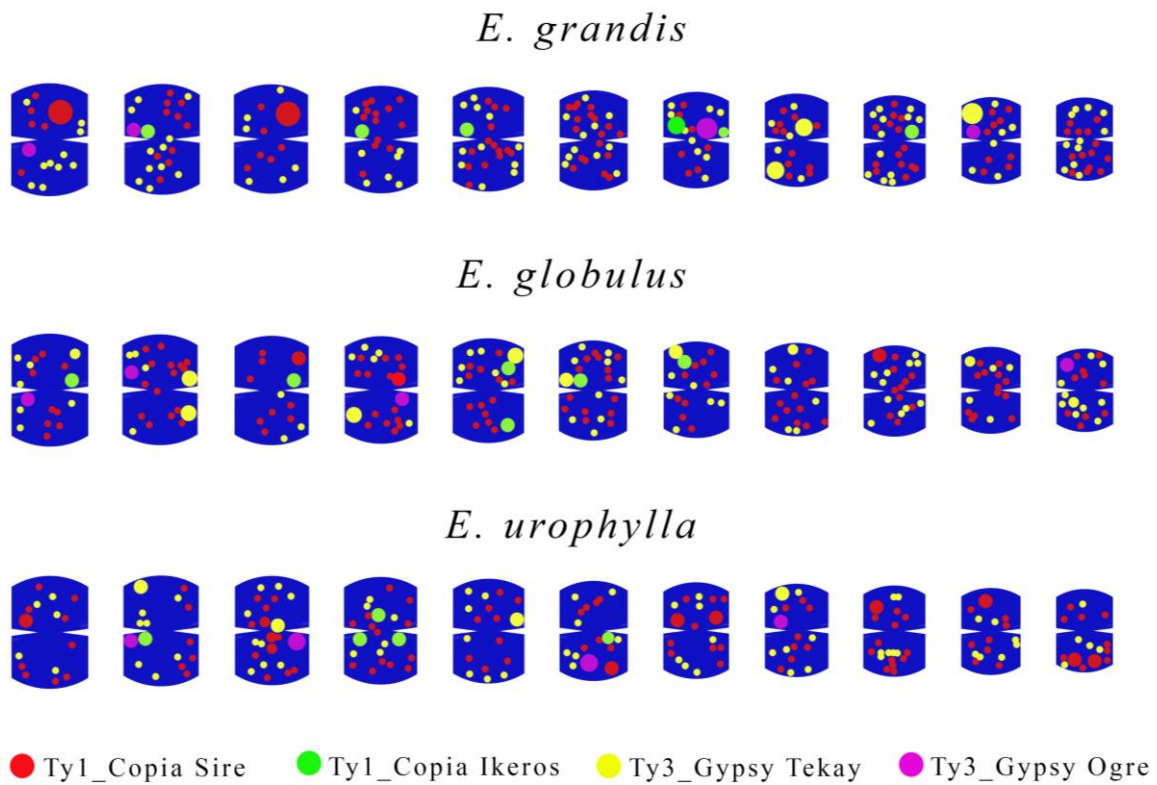


Figura 7. Idiograma representando os cromossomos e o mapeamento físico dos RT-LTR Sire, Ogre, Tekay e Ikeros nos cariótipos de *E. grandis*, *E. urophylla* e *E. globulus*. O tamanho e o locus dos círculos representam os sinais de hibridização observados. Sire (círculos vermelhos) e Tekay (círculos amarelos) foram mapeados em todos os cromossomos das três espécies de *Eucalyptus*, enquanto Ikeros (círculos verdes) e Ogre (círculos roxos) foram mapeados apenas em alguns cromossomos das três espécies.

5 DISCUSSÃO

Nós, pela primeira vez, classificamos os RT-LTR de *E. urophylla* e *E. globulus* ao nível de elemento utilizando dados genômicos, e mapeamos os quatro RT-LTR mais abundantes em cromossomos metafásicos de três espécies de *Eucalyptus*. Embora a proporção genômica do repeatoma e os valores 1C nuclear e AT% (Carvalho et al., 2017) sejam semelhantes, a distribuição dos RT-LTR mapeados evidenciou outras similaridades e diferenças cariotípicas entre as espécies.

Como esperado, os RT-LTR predominam no repeatoma de *E. grandis*, *E. urophylla* e *E. globulus*. Os RT-LTR apresentaram uma notável variedade de elementos, enquanto os RT não LTR e os transposons possuem uma baixa diversidade nos genomas das três espécies. Possivelmente, RT não LTR e os transposons não estejam amplamente representados no genoma das espécies de *Eucalyptus*, existindo sequências repetitivas com uma frequência inferior a 0,01% do genoma, o que os excluiu da anotação.

A proporção do repeatoma é relativamente semelhante entre os genomas de *E. grandis*, *E. urophylla* e *E. globulus*, variando entre 24,1726 – 26,0248 (24,98336667 ±0,773548767) Myburg et al. (2014) reportaram que a atividade recente de TEs é responsável uma diferença equivalente a 2 Mb no tamanho do genoma de *E. grandis* e *E. globulus*. A proporção do repeatoma das três espécies de *Eucalyptus* também evidencia que os TEs possuem papel fundamental na composição e evolução do genoma dessas espécies s estreitamente relacionadas, assim como relatado por Ågren e Wright (2011).

A anotação dos RT-LTR conferida pelo RepeatExplor2 clustering para *E. grandis* (presente no estudo) corrobora com os dados prévios gerados pelo PlantLTRdb (Mokhtar et al., 2023). Entretanto, segundo o PlantLTRdb, os quatro RT-LTR mais abundantes em *E. grandis* são Sire, Ikeros, Ale e Tork, diferindo, portanto, em relação a Tekay e a Ogre. Essa variação se deve à ferramenta utilizada na análise, já que o PlantLTRdb empregou diferentes pipelines, como LTRdigest e TESorter. Entretanto, o RepeatExplor2 clustering é mais adequado para anotação das sequências do repeatoma, considerando o banco de dados disponíveis para as espécies de *Eucalyptus*.

Os dados acerca do repeatoma em espécies da família Myrtaceae com genomas montados limitam-se à sua proporção no genoma. *Psidium guajava* (Feng et al., 2021) tem 43,55% do seu genoma ocupado pelo repeatoma, enquanto *Leptospermum scoparium* (Thrimawithana et al., 2019) tem 32,4%. Portanto, a classificação do repeatoma ao nível de superfamília e de elemento, anteriormente exclusiva para *E. grandis*, foi ampliada por este

trabalho, o qual apresenta dados para *E. urophylla* e *E. globulus*. Entre os RT-LTR de *Eucalyptus*, a superfamília Ty1_Copia é mais abundante que a superfamília Ty3_Gypsy, assim como em *Musa* (Novák et al., 2014), *Juncus* (Mata-Sucre et al., 2023) e em *Calendula* (Samatadze et al., 2023). A variação na distribuição entre essas duas superfamílias nas espécies vegetais parece ser amplamente influenciada por especificidades de inserção de diferentes retrotransposases em diferentes tipos de cromatina (Bennetzen, 2000; Bennetzen e Wang, 2014).

Com base na identificação e classificação do repeatoma, as quatro sequências mais predominantes foram selecionadas e filogeneticamente comparadas. Sire e Ikeros (superfamília Ty1_Copia) divergiram após a especiação, enquanto Ogre e Tekay (superfamília Ty3_Gypsy) divergiram antes da especiação. Corroborando, Ty1_Copia apresenta maior diversidade e mais inserções recentes do que Ty3_Gypsy (González and Deyholos, 2012; Galindo-González et al., 2017). Portanto, em *E. grandis*, *E. urophylla* e *E. globulus*, há mais sequências conservadas de Tekay e Ogre do que de Sire e Ikeros.

Os cromossomos pequenos e homomórficos de *E. grandis*, *E. urophylla* e *E. globulus* foram observados com maior resolução na citogenômica. O mapeamento físico revelou similaridades interespecíficas para a região cromossômica onde majoritariamente se encontraram Ogre e Ikeros, e para o padrão disperso de Sire e Tekay. Entretanto, foram observadas diferenças interespecíficas em termos de número de sinais, intensidade de fluorescência do sinal e locus cromossômico de Sire, Ikeros, Ogre e Tekay entre as três espécies de *Eucalyptus*.

Conforme mostrado na filogenia, a ocorrência de Sire em todos os cromossomos das três espécies de *Eucalyptus* foi confirmada pelo mapeamento físico. Sire foi mapeado de forma dispersa ao longo dos cromossomos, assim como em *Passiflora edulis* (Pamponét et al., 2019) e em *Humulus scandens* (Zhang et al., 2023). Sinais intensos de Sire foram mapeados na região intersticial de *E. grandis*, *E. globulus* e *E. urophylla*, na região terminal de *E. globulus* e *E. urophylla*, e na região pericentromérica de *E. urophylla*. Da mesma maneira, diferenças foram observadas no locus cromossômico com sinais intensos entre as diferentes espécies. A distribuição dispersa de Sire era esperada visto o padrão de distribuição de elementos Ty1_Copia (Galindo-González et al., 2017). A presença de elementos Ty1_Copia em regiões eucromáticas está envolvida em rearranjos cromossômicos e marcas epigenéticas (Vicient, 2010; De Setta et al., 2014; Tang et al., 2014).

Embora os loci cromossômicos tenham variado, Ikeros e Ogre foram mapeados majoritariamente na região pericentromérica de alguns dos cromossomos das três espécies de

Eucalyptus. A ocorrência de TEs nas regiões pericentroméricas tem sido associada com a resistência dos centrômeros à tensão dos microtúbulos durante segregação cromossômica na anáfase mitótica e meiótica. Além disso, essas sequências contribuem com as origens de replicação, compensando a escassez de genes na região (Freeling et al., 2015; Vergara et al., 2017). Ademais, como os elementos OGRE contêm uma região de cromodomínio. Provavelmente, eles possam interagir com proteínas CENH3, e participar da função do centrômero (Luo et al., 2012; De Castro Nunes et al., 2018).

A presença de Tekay em todos os cromossomos das três espécies *Eucalyptus* mostrada na filogenia foi confirmada pelo mapeamento físico. Além de sinais fracos dispersos em todos os cromossomos das três espécies, sinais fortes foram mapeados nas regiões pericentroméricas de *E. grandis* e *E. urophylla*, intersticiais de *E. grandis* e *E. globulus*, e terminais de *E. globulus* e *E. urophylla*. Sendo um cromovírus, Tekay apresenta um cromodomínio em seu terminal C da integrase, que o direciona a inserir-se em regiões de heterocromatina (Gao et al., 2008), tornando suas inserções menos prováveis de interferir na função gênica. Logo, a ocorrência de Tekay nas regiões heterocromáticas ao longo dos cromossomos promove o silenciamento heterocromático e mantém a estabilidade cromossômica (Gao et al., 2014, 2015).

As diferenças observadas quanto aos loci cromossômicos de Sire, OGRE, Tekay e Ikeros podem estar associadas às mutações e rearranjos cromossômicos estruturais ocorridos após a divergência de *Eucalyptus* (Ferguson et al., 2023). Além disso, a distribuição dos RT-LTR nos genomas vegetais sofre influência de características genômicas locais, como recombinação genética e densidades genéticas (Zhao e Ma, 2013).

6 CONCLUSÃO

Este estudo pioneiro reporta a classificação de RT-LTR ao nível de elemento móvel, utilizando dados genômicos de *E. urophylla* e *E. globulus*. Além disso, nós, pela primeira vez, mapeamos elementos móveis nos cromossomos metafásicos das espécies de *Eucalyptus*, representando um avanço significativo na compreensão da organização genômica dessas espécies.

A identificação e classificação das reads de *E. grandis*, *E. urophylla* e *E. globulus* revelaram uma proporção similar do repeatoma dessas espécies, com os RT-LTR destacando-se como as repetições predominantes. A superfamília Ty1_Copia mostrou-se como a mais prevalente e diversa. Os elementos Sire e Ikeros (Ty1_Copia) e Tekay e Ogre (Ty3_Gypsy) predominam em todas as espécies. A análise filogenética sugeriu que Sire e Ikeros divergiram após a especiação, enquanto Ogre e Tekay divergiram antes, indicando que em *E. grandis*, *E. urophylla* e *E. globulus*, há mais sequências conservadas de Tekay e Ogre do que de Sire e Ikeros. Além disso, a ocorrência de Sire e Tekay em todos os cromossomos das três espécies de *Eucalyptus*, mostrada na filogenia, foi confirmada pelo mapeamento físico, enquanto Ogre e Ikeros foram mapeados em poucos cromossomos nas três espécies. Além disso, Sire e Tekay apresentaram uma distribuição dispersa ao longo dos cromossomos, enquanto Ogre e Ikeros foram mapeados predominantemente em regiões pericentroméricas.

Este estudo apresenta uma análise abrangente da diversidade e distribuição cromossômica do repeatoma em *E. grandis*, *E. urophylla* e *E. globulus*. Dessa forma, acreditamos que os resultados representam uma fundamentação valiosa para o progresso das pesquisas de TEs em *Eucalyptus* e em Myrtaceae.

7 REFERÊNCIAS

- Ågren, J.A., Wright, S.I., 2011. Co-evolution between transposable elements and their hosts: a major factor in genome size evolution? *Chromosome Res.* 19, 777–786. <https://doi.org/10.1007/s10577-011-9229-0>
- Bennetzen, J.L., 2000. The many hues of plant heterochromatin. *Genome Biol.*
- Bennetzen, J.L., Wang, H., 2014. The Contributions of Transposable Elements to the Structure, Function, and Evolution of Plant Genomes. *Annu. Rev. Plant Biol.* 65, 505–530. <https://doi.org/10.1146/annurev-arplant-050213-035811>
- Biscotti, M.A., Olmo, E., Heslop-Harrison, J.S., 2015. Repetitive DNA in eukaryotic genomes. *Chromosome Res.* 23, 415–420. <https://doi.org/10.1007/s10577-015-9499-z>
- Carvalho, G.M.A., Carvalho, C.R., Soares, F.A.F., 2017. Flow cytometry and cytogenetic tools in eucalypts: genome size variation × karyotype stability. *Tree Genet. Genomes* 13, 106. <https://doi.org/10.1007/s11295-017-1186-0>
- De Assis, R., Baba, V.Y., Cintra, L.A., Gonçalves, L.S.A., Rodrigues, R., Vanzela, A.L.L., 2020. Genome relationships and LTR-retrotransposon diversity in three cultivated *Capsicum* L. (Solanaceae) species. *BMC Genomics* 21, 237. <https://doi.org/10.1186/s12864-020-6618-9>
- De Castro Nunes, R., Orozco-Arias, S., Cruzillat, D., Mueller, L.A., Strickler, S.R., Descombes, P., Fournier, C., Moine, D., De Kochko, A., Yuyama, P.M., Vanzela, A.L.L., Guyot, R., 2018. Structure and Distribution of Centromeric Retrotransposons at Diploid and Allotetraploid *Coffea* Centromeric and Pericentromeric Regions. *Front. Plant Sci.* 9, 175. <https://doi.org/10.3389/fpls.2018.00175>
- De Setta, N., Monteiro-Vitorello, C.B., Metcalfe, C.J., Cruz, G.M.Q., Del Bem, L.E., Vicentini, R., Nogueira, F.T.S., Campos, R.A., Nunes, S.L., Turrini, P.C.G., Vieira, A.P., Ochoa Cruz, E.A., Corrêa, T.C.S., Hotta, C.T., De Mello Varani, A., Vautrin, S., Da Trindade, A.S., De Mendonça Vilela, M., Lembke, C.G., Sato, P.M., De Andrade, R.F., Nishiyama, M.Y., Cardoso-Silva, C.B., Scortecci, K.C., Garcia, A.A.F., Carneiro, M.S., Kim, C., Paterson, A.H., Bergès, H., D’Hont, A., De Souza, A.P., Souza, G.M., Vincentz, M., Kitajima, J.P., Van Sluys, M.-A., 2014. Building the sugarcane genome for biotechnology and identifying evolutionary trends. *BMC Genomics* 15, 540. <https://doi.org/10.1186/1471-2164-15-540>
- Feng, Chen, Feng, Chao, Lin, X., Liu, S., Li, Y., Kang, M., 2021. A chromosome-level genome assembly provides insights into ascorbic acid accumulation and fruit softening in guava (*Psidium guajava*). *Plant Biotechnol. J.* 19, 717–730. <https://doi.org/10.1111/pbi.13498>
- Feng, R., Wang, X., Tao, M., Du, G., Wang, Q., 2017. Genome size and identification of abundant repetitive sequences in *Vallisneria spirulosa*. *PeerJ* 5, e3982. <https://doi.org/10.7717/peerj.3982>
- Ferguson, S., Jones, A., Murray, K., Andrew, R., Schwessinger, B., Borevitz, J., 2023. Plant genome evolution in the genus *Eucalyptus* driven by structural rearrangements that promote sequence divergence (preprint). *Genomics*. <https://doi.org/10.1101/2023.04.19.537464>
- Freeling, M., Xu, J., Woodhouse, M., Lisch, D., 2015. A Solution to the C-Value Paradox and the Function of Junk DNA: The Genome Balance Hypothesis. *Mol. Plant* 8, 899–910. <https://doi.org/10.1016/j.molp.2015.02.009>
- Galindo-González, L., Mhiri, C., Deyholos, M.K., Grandbastien, M.-A., 2017. LTR-retrotransposons in plants: Engines of evolution. *Gene* 626, 14–25. <https://doi.org/10.1016/j.gene.2017.04.051>
- Gao, D., Abernathy, B., Rohksar, D., Schmutz, J., Jackson, S.A., 2014. Annotation and sequence diversity of transposable elements in common bean (*Phaseolus vulgaris*). *Front. Plant Sci.* 5. <https://doi.org/10.3389/fpls.2014.00339>
- Gao, D., Jiang, N., Wing, R.A., Jiang, J., Jackson, S.A., 2015. Transposons play an important role in the evolution and diversification of centromeres among closely related species. *Front. Plant Sci.* 6. <https://doi.org/10.3389/fpls.2015.00216>
- Gao, X., Hou, Y., Ebina, H., Levin, H.L., Voytas, D.F., 2008. Chromodomains direct integration of retrotransposons to heterochromatin. *Genome Res.* 18, 359–369. <https://doi.org/10.1101/gr.7146408>
- González, L.G., Deyholos, M.K., 2012. Identification, characterization and distribution of transposable elements in the flax (*Linum usitatissimum* L.) genome. *BMC Genomics* 13, 644. <https://doi.org/10.1186/1471-2164-13-644>
- Hall, B.G., 2013. Building Phylogenetic Trees from Molecular Data with MEGA. *Mol. Biol. Evol.* 30, 1229–1235. <https://doi.org/10.1093/molbev/mst012>
- Hudson, C.J., Kullán, A.R.K., Freeman, J.S., Faria, D.A., Grattapaglia, D., Kilian, A., Myburg, A.A., Potts, B.M., Vaillancourt, R.E., 2012. High synteny and colinearity among *Eucalyptus* genomes revealed by high-density comparative genetic mapping. *Tree Genet. Genomes* 8, 339–352. <https://doi.org/10.1007/s11295-011-0444-9>
- Letunic, I., Bork, P., 2024. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* gkae268. <https://doi.org/10.1093/nar/gkae268>

- Lisch, D., 2013. How important are transposons for plant evolution? *Nat. Rev. Genet.* 14, 49–61. <https://doi.org/10.1038/nrg3374>
- Luo, S., Mach, J., Abramson, B., Ramirez, R., Schurr, R., Barone, P., Copenhaver, G., Folkerts, O., 2012. The Cotton Centromere Contains a Ty3-gypsy-like LTR Retroelement. *PLoS ONE* 7, e35261. <https://doi.org/10.1371/journal.pone.0035261>
- Malik, H.S., Eickbush, T.H., 1999. Modular Evolution of the Integrase Domain in the Ty3/Gypsy Class of LTR Retrotransposons. *J. Virol.* 73, 5186–5190. <https://doi.org/10.1128/JVI.73.6.5186-5190.1999>
- Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I., Lanczycki, C.J., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Bryant, S.H., 2015. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 43, D222–D226. <https://doi.org/10.1093/nar/gku1221>
- Mata-Sucre, Y., Matzenauer, W., Castro, N., Huettel, B., Pedrosa-Harand, A., Marques, A., Souza, G., 2023. Repeat-based phylogenomics shed light on unclear relationships in the monocentric genus *Juncus* L. (Juncaceae). *Mol. Phylogenet. Evol.* 189, 107930. <https://doi.org/10.1016/j.ympev.2023.107930>
- Mokhtar, M.M., Alsamman, A.M., El Allali, A., 2023. PlantLTRdb: An interactive database for 195 plant species LTR-retrotransposons. *Front. Plant Sci.* 14, 1134627. <https://doi.org/10.3389/fpls.2023.1134627>
- Myburg, A.A., Grattapaglia, D., Tuskan, G.A., Hellsten, U., Hayes, R.D., Grimwood, J., Jenkins, J., Lindquist, E., Tice, H., Bauer, D., Goodstein, D.M., Dubchak, I., Poliakov, A., Mizrachi, E., Kullán, A.R.K., Hussey, S.G., Pinard, D., Van Der Merwe, K., Singh, P., Van Jaarsveld, I., Silva-Junior, O.B., Togawa, R.C., Pappas, M.R., Faria, D.A., Sansaloni, C.P., Petrolí, C.D., Yang, X., Ranjan, P., Tschaplinski, T.J., Ye, C.-Y., Li, T., Sterck, L., Vanneste, K., Murat, F., Soler, M., Clemente, H.S., Saidi, N., Cassan-Wang, H., Dunand, C., Hefer, C.A., Bornberg-Bauer, E., Kersting, A.R., Vining, K., Amarasinghe, V., Ranik, M., Naithani, S., Elser, J., Boyd, A.E., Liston, A., Spatafora, J.W., Dharmwardhana, P., Raja, R., Sullivan, C., Romanel, E., Alves-Ferreira, M., Külheim, C., Foley, W., Carocha, V., Paiva, J., Kudrna, D., Brommonschenkel, S.H., Pasquali, G., Byrne, M., Rigault, P., Tibbits, J., Spokevicius, A., Jones, R.C., Steane, D.A., Vaillancourt, R.E., Potts, B.M., Joubert, F., Barry, K., Pappas, G.J., Strauss, S.H., Jaiswal, P., Grima-Pettenati, J., Salse, J., Van De Peer, Y., Rokhsar, D.S., Schmutz, J., 2014. The genome of *Eucalyptus grandis*. *Nature* 510, 356–362. <https://doi.org/10.1038/nature13308>
- Neumann, P., Novák, P., Hošťáková, N., Macas, J., 2019. Systematic survey of plant LTR-retrotransposons elucidates phylogenetic relationships of their polyprotein domains and provides a reference for element classification. *Mob. DNA* 10, 1. <https://doi.org/10.1186/s13100-018-0144-1>
- Nicolle, D., 2015. Classification of the eucalypts (*Angophora*, *Corymbia* and *Eucalyptus*).
- Novák, P., Hřibová, E., Neumann, P., Koblížková, A., Doležel, J., Macas, J., 2014. Genome-Wide Analysis of Repeat Diversity across the Family Musaceae. *PLoS ONE* 9, e98918. <https://doi.org/10.1371/journal.pone.0098918>
- Oudjehih, B., Abdallah, B., 2006. Chromosome numbers of the 59 species of *Eucalyptus* L'Herit. (Myrtaceae). *Caryologia* 59, 207–212. <https://doi.org/10.1080/00087114.2006.10797916>
- Pamponét, V.C.C., Souza, M.M., Silva, G.S., Micheli, F., De Melo, C.A.F., De Oliveira, S.G., Costa, E.A., Corrêa, R.X., 2019. Low coverage sequencing for repetitive DNA analysis in *Passiflora edulis* Sims: cytogenomic characterization of transposable elements and satellite DNA. *BMC Genomics* 20, 262. <https://doi.org/10.1186/s12864-019-5576-6>
- Pinard, D., Myburg, A.A., Mizrachi, E., 2019. The plastid and mitochondrial genomes of *Eucalyptus grandis*. *BMC Genomics* 20, 132. <https://doi.org/10.1186/s12864-019-5444-4>
- Potts, B.M., 2004. TREE BREEDING, PRACTICES | Genetic Improvement of Eucalypts, in: *Encyclopedia of Forest Sciences*. Elsevier, pp. 1480–1490. <https://doi.org/10.1016/B0-12-145160-7/00084-3>
- Pryor, L.D., Johnson, L.A., 1971. A classification of the Eucalyptus. Canberra: Australian National University.
- Ribeiro, T., Barrela, R.M., Bergès, H., Marques, C., Loureiro, J., Morais-Cecílio, L., Paiva, J.A.P., 2016. Advancing *Eucalyptus* Genomics: Cytogenomics Reveals Conservation of *Eucalyptus* Genomes. *Front. Plant Sci.* 7. <https://doi.org/10.3389/fpls.2016.00510>
- Rozefelds, A.C., 1996. *Eucalyptus* phylogeny and history: a brief summary 8.
- Samatadze, T.E., Yurkevich, O.Yu., Khazieva, F.M., Basalaeva, I.V., Savchenko, O.M., Zoshchuk, S.A., Morozov, A.I., Amosova, A.V., Muravenko, O.V., 2023. Genome Studies in Four Species of *Calendula* L. (Asteraceae) Using Satellite DNAs as Chromosome Markers. *Plants* 12, 4056. <https://doi.org/10.3390/plants12234056>
- Sattler, M.C., Soares, F.A.F., Silva, J.C., Carvalho, C.R., Clarindo, W.R., 2019. Physical Mapping of 5S rDNA in *Eucalyptus dunnii* Maiden and *Zea mays* L. by PRINS. *Cytologia (Tokyo)* 84, 77–83. <https://doi.org/10.1508/cytologia.84.77>
- Silva, J.C., Carvalho, C.R., Clarindo, W.R., 2018. Updating the maize karyotype by chromosome DNA sizing. *PLOS ONE* 13, e0190428. <https://doi.org/10.1371/journal.pone.0190428>

- Tamura, K., Stecher, G., Kumar, S., 2021. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* 38, 3022–3027. <https://doi.org/10.1093/molbev/msab120>
- Tang, X., Datema, E., Guzman, M.O., De Boer, J.M., Van Eck, H.J., Bachem, C.W.B., Visser, R.G.F., De Jong, H., 2014. Chromosomal organizations of major repeat families on potato (*Solanum tuberosum*) and further exploring in its sequenced genome. *Mol. Genet. Genomics* 289, 1307–1319. <https://doi.org/10.1007/s00438-014-0891-8>
- Thomas-Bulle, C., Piednoël, M., Donnart, T., Filée, J., Jollivet, D., Bonnivard, É., 2018. Mollusc genomes reveal variability in patterns of LTR-retrotransposons dynamics. *BMC Genomics* 19, 821. <https://doi.org/10.1186/s12864-018-5200-1>
- Thrimawithana, A.H., Jones, D., Hilario, E., Grierson, E., Ngo, H.M., Liachko, I., Sullivan, S., Bilton, T.P., Jacobs, J.M.E., Bicknell, R., David, C., Deng, C., Nieuwenhuizen, N., Lopez-Girona, E., Tobias, P.A., Morgan, E., Perry, N.B., Lewis, D.H., Crowhurst, R., Davies, K.M., Chagné, D., Schwinn, K.E., 2019. A whole genome assembly of *Leptospermum scoparium* (Myrtaceae) for mānuka research. *N. Z. J. Crop Hortic. Sci.* 47, 233–260. <https://doi.org/10.1080/01140671.2019.1657911>
- Tollis, M., Boissinot, S., 2012. The Evolutionary Dynamics of Transposable Elements in Eukaryote Genomes, in: Garrido-Ramos, M.A. (Ed.), *Genome Dynamics*. S. Karger AG, pp. 68–91. <https://doi.org/10.1159/000337126>
- Vergara, Z., Sequeira-Mendes, J., Morata, J., Peiró, R., Hénaff, E., Costas, C., Casacuberta, J.M., Gutierrez, C., 2017. Retrotransposons are specified as DNA replication origins in the gene-poor regions of *Arabidopsis* heterochromatin. *Nucleic Acids Res.* 45, 8358–8368. <https://doi.org/10.1093/nar/gkx524>
- Vicient, C.M., 2010. Transcriptional activity of transposable elements in maize. *BMC Genomics* 11, 601. <https://doi.org/10.1186/1471-2164-11-601>
- Vitte, C., Fustier, M.-A., Alix, K., Tenaillon, M.I., 2014. The bright side of transposons in crop evolution. *Brief. Funct. Genomics* 13, 276–295. <https://doi.org/10.1093/bfpg/elu002>
- Wang, H., Liu, J.-S., 2008. LTR retrotransposon landscape in *Medicago truncatula*: more rapid removal than in rice. *BMC Genomics* 9, 382. <https://doi.org/10.1186/1471-2164-9-382>
- Wang, X., Weigel, D., Smith, L.M., 2013. Transposon Variants and Their Effects on Gene Expression in *Arabidopsis*. *PLoS Genet.* 9, e1003255. <https://doi.org/10.1371/journal.pgen.1003255>
- Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P., Schulman, A.H., 2007. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* 8, 973–982. <https://doi.org/10.1038/nrg2165>
- Zhang, G.-J., Jia, K.-L., Wang, J., Gao, W.-J., Li, S.-F., 2023. Genome-wide analysis of transposable elements and satellite DNA in *Humulus scandens*, a dioecious plant with XX/XY1Y2 chromosomes. *Front. Plant Sci.* 14, 1230250. <https://doi.org/10.3389/fpls.2023.1230250>
- Zhao, M., Ma, J., 2013. Co-evolution of plant LTR-retrotransposons and their host genomes. *Protein Cell* 4, 493–501. <https://doi.org/10.1007/s13238-013-3037-6>

8 MATERIAL SUPLEMENTAR

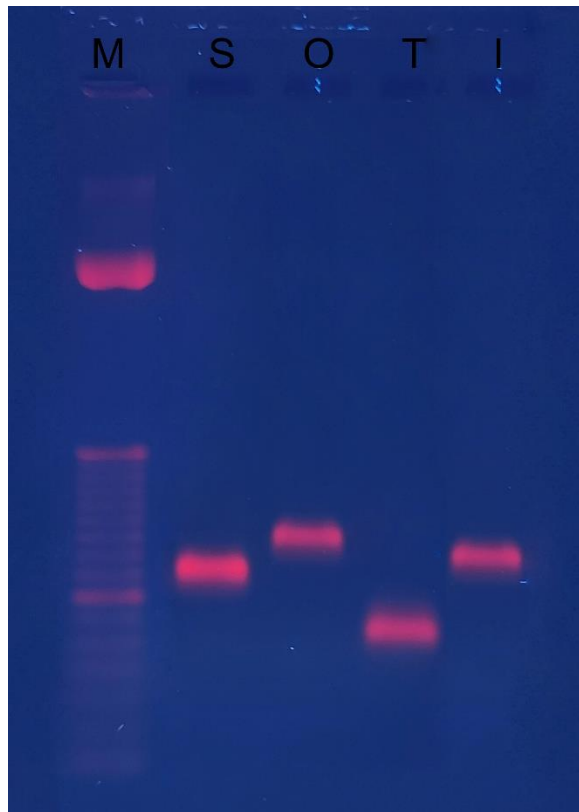


Figura Suplementar 1: Gel de agarose a 1,5% mostrando os produtos de amplificação gerados por PCR usando os primers da Tabela 1. Sire (S), Ogre (O), Tekay (T), Ikeros (I) e DNA Ladder 100 pb (M).

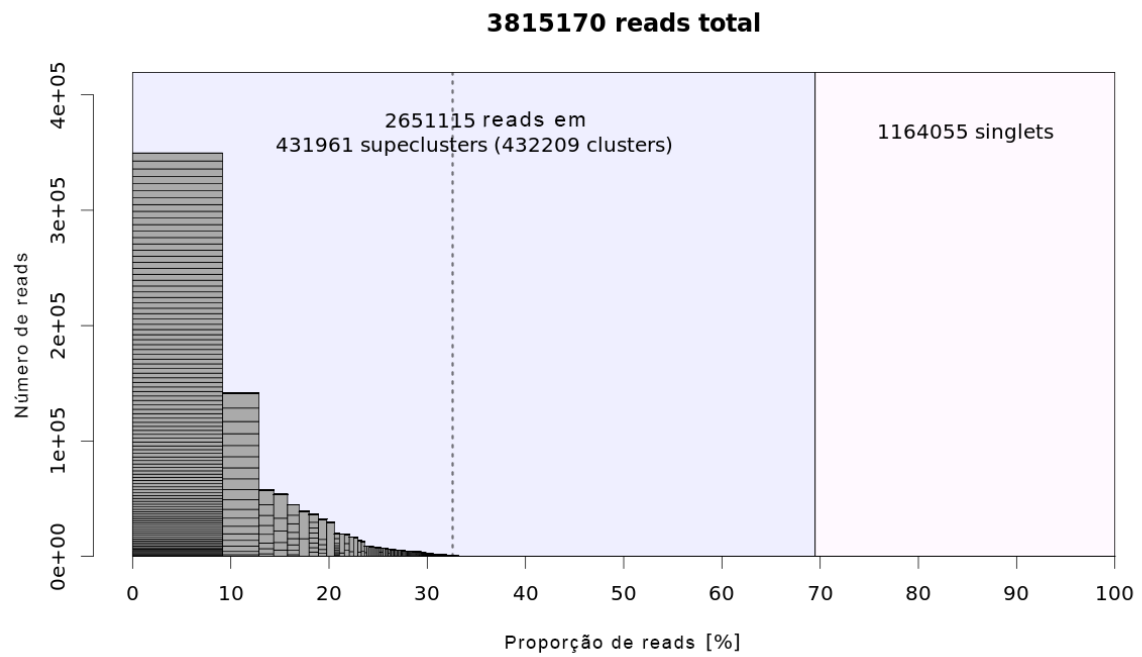


Figura Suplementar 2: Resumo gráfico dos resultados de agrupamento de sequências repetitivas para os genomas de *E. grandis*, *E. urophylla* e *E. globulus* geradas pelo RepeatExplorer2 clustering. As barras representam superaglomerados, com suas alturas e larguras correspondentes ao número de leituras nos superaglomerados (eixo y) e às suas proporções em todas as leituras analisadas (eixo x), respectivamente. Os retângulos dentro das barras do superaglomerado representam clusters individuais. Os clusters superiores estão à esquerda da linha pontilhada.

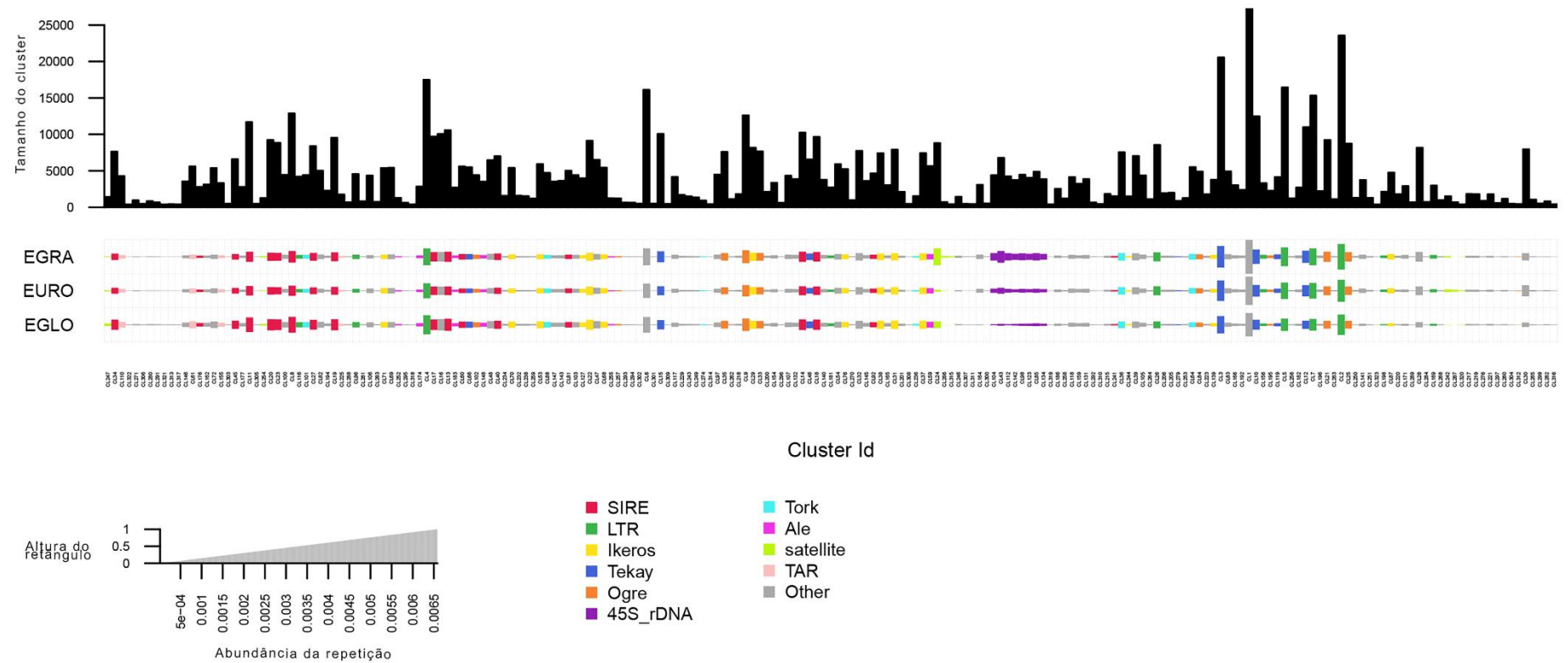


Figura Suplementar 3: Clusters de sequências repetitivas com abundância $> 0,01\%$ nos genomas de *E. grandis*, *E.s urophylla* e *E. globulus*, obtidos a partir do pipeline RepeatExplorer2 clustering. O gráfico de barras na parte superior mostra o tamanho dos clusters individuais. O tamanho dos retângulos no painel inferior é proporcional ao número de leituras em um cluster para cada espécie.

Tabela Suplementar 1: Sequências consenso obtidas a partir do sequenciamento de amplicons dos RT-LTR Sire, Ogre, Tekay e Ikeros, os quais foram amplificados a partir do DNA genômico de *E. grandis* utilizando primers desenhados a partir da análise realizada pelo RepeatExplorer2 clustering.

| RT-LTR | Sequência consenso | Tamanho (bp) |
|--------|---|--------------|
| Sire | CAATACTTCTTACCCCCAATACTTTGAGTTCTGGTTGGTCCGA AGAGATCCATATGCAGCAACTGTAGTACATGATTAGTAGAGA CATAATTAATTGGCTTAAAAGAATTTCTTACCTGCTTTCCAG AATACATGGAGTGCATGAATCAGTCTTTTGGTATGGCAATTTG GGTAGTCCTCGAACAAGCTGTTTTGAAGAGATTTTGGCTAATT GCTTCATGTTGACATGACCAAGCTTTTTGTGCCATAAGCTTGC TTCGTCTTGAATTGAGATAAGGCATTGCGATTCAATTTGGTTTC ACATCCAGAAGATAGATATTTCCATGTCTTCGACCCATGAAAG ACTGAG | 349 |
| Ogre | CCCCACATTGAGAAAGGCCATGGCTCAGATAGCTGGTGCAACT CCGTTGGAGGAACATTTATTTTATCTCCATGAATTTGGCAACGA TGGCAACTTCTAACATGCTGTATGCAGTCACTTTCCAGCGCAAG CCAATAGTAACCTAGTCTCATGATCTTCTTGGCTAACATGTGCC CATTTCATGTGTGGGCCACAACTCCTTCATGCACTTCTTGCATG ATTTGGGTGGCTTCTGCTTCATCTACGCATCTCAATAAAACTGA ATCATAAGATCTCTTGTATAAGTTCTTACCACTGACGAAGA TTGAGGCCATCTTCATTATGTATTTTCTATCGGCTGTAGTGCTCC CTTCAGGAAATTCCTGCTTTTGGATATAGTTCATGATGTCGCAA TACCATGGCTTCTCATCAAGCTCTTCATTAGC | 428 |
| Tekay | GGAAACGTAGTGCCTTATATCGCTATTATCTTCGGCGTCTGCTT GGGTGTTTAACCTGAGGTTGCAGAATCCTGCGGCTTGCAATGCT GAGCTTTTACTTGTACATGTCAAACACTGGGCCACATGTTTGGC GATATCCGCCTTCATACCCGACCACCAGTAGTGTGACGTAGGT TCTGATACATCTTTTCGTGCTTCCAGGATGAACGCTGTA ACTGCTACGATGTGCTTACAGATAAAATTTCTTCTCTAAGCTCTACATCGT CAGGTACAACCAATCGTCCTTGGAAATCTTAGTGTCG | 300 |
| Ikeros | GTTGACTGTGGCGATGTCAATATGTTATCCCCTTGTTTACAGTCT TCTATCTTATATGATAATGATGTGAATATGATGGCATGCTAGAC TTGGTCATATGGGCCAACACATATGGAAGAAGACTAGCCAAA GAGAGCTTGTGGGCAACATTGAAAAAGTCGATTTGCCCATATG TAAGCATTGCCTAGTAGGGAAAATGACAATGAGACCATGGAAA AGGTAAAAGAGCTGAAGTATGTTGACTGTGGCGATGTCGATAT ATGTTATTCCTTGTTTAAGCATCTACCTACTTTATATGATAATG AT | 309 |

Tabela Suplementar 2: Similaridade dos RT-LTR Sire, Ogre, Tekay e Ikeros sequenciados com dados do NCBI. As sequências consenso dos quatro RT-LTR apresentaram similaridade com sequências presentes em todos os cromossomos das três espécies, exceto Ogre no cromossomo 4 de *E. grandis*. Em *E. grandis*, Sire, Ogre, Tekay e Ikeros exibiram uma correspondência de mais de 96%, 94%, 92% e 92%, respectivamente, em todos os cromossomos (exceto Ogre no cromossomo 4). Por outro lado, em *E. urophylla*, as taxas de correspondência foram de acima de 98%, 95%, 88% e 93%, respectivamente, enquanto em *E. globulus*, foram maior que 97%, 95%, 85% e 92%, respectivamente.

| Description | Max score | Total Score | Query Cover | E value | Per Ident | Acc. Len | Acession |
|---|-----------|-------------|-------------|---------|-----------|----------|-------------|
| <i>Eucalyptus grandis</i> - Sire | | | | | | | |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 1, ASM1654582v1 | 606 | 38927 | 100% | 4e-172 | 98.00% | 53951317 | NC_052612.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 2, ASM1654582v1 | 617 | 41950 | 100% | 2e-175 | 98.57% | 58486256 | NC_052613.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 3, ASM1654582v1 | 612 | 63013 | 100% | 9e-174 | 98.28% | 72470685 | NC_052614.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 4, ASM1654582v1 | 612 | 34787 | 100% | 9e-174 | 98.83% | 39120311 | NC_052615.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 5, ASM1654582v1 | 616 | 54719 | 100% | 7e-175 | 98.57% | 62829834 | NC_052616.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 6, ASM1654582v1 | 606 | 37517 | 100% | 4e-172 | 97.99% | 57658180 | NC_052617.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 7, ASM1654582v1 | 603 | 57456 | 100% | 5e-171 | 97.99% | 60484081 | NC_052618.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 8, ASM1654582v1 | 606 | 59847 | 100% | 4e-172 | 98.55% | 74572912 | NC_052619.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 9, ASM1654582v1 | 617 | 43075 | 100% | 2e-175 | 98.57% | 44060374 | NC_052620.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 10, ASM1654582v1 | 575 | 30798 | 100% | 1e-162 | 96.31% | 38819584 | NC_052621.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 11, ASM1654582v1 | 612 | 30317 | 100% | 9e-174 | 98.84% | 50013132 | NC_052622.1 |
| <i>Eucalyptus grandis</i> - Ogre | | | | | | | |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 1, ASM1654582v1 | 719 | 2521 | 99% | 0.0 | 97.18% | 53951317 | NC_052612.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 2, ASM1654582v1 | 747 | 2836 | 99% | 0.0 | 98.35% | 58486256 | NC_052613.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 3, ASM1654582v1 | 725 | 4306 | 99% | 0.0 | 97.41% | 72470685 | NC_052614.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 5, ASM1654582v1 | 756 | 3977 | 99% | 0.0 | 98.82% | 62829834 | NC_052616.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 6, ASM1654582v1 | 725 | 4218 | 99% | 0.0 | 97.41% | 57658180 | NC_052617.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 7, ASM1654582v1 | 752 | 2148 | 99% | 0.0 | 98.59% | 60484081 | NC_052618.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 8, ASM1654582v1 | 712 | 2804 | 99% | 0.0 | 96.94% | 74572912 | NC_052619.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 9, ASM1654582v1 | 741 | 2881 | 99% | 0.0 | 98.13% | 44060374 | NC_052620.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 10, ASM1654582v1 | 660 | 660 | 99% | 0.0 | 94.87% | 38819584 | NC_052621.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 11, ASM1654582v1 | 680 | 1049 | 99% | 0.0 | 95.53% | 50013132 | NC_052622.1 |
| <i>Eucalyptus grandis</i> - Tekay | | | | | | | |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 1, ASM1654582v1 | 351 | 9716 | 76% | 2e-95 | 93.67% | 53951317 | NC_052612.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 2, ASM1654582v1 | 379 | 11390 | 76% | 9e-104 | 95.78% | 58486256 | NC_052613.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 3, ASM1654582v1 | 357 | 14725 | 76% | 4e-97 | 94.09% | 72470685 | NC_052614.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 4, ASM1654582v1 | 363 | 7633 | 76% | 9e-99 | 94.51% | 39120311 | NC_052615.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 5, ASM1654582v1 | 366 | 14800 | 76% | 7e-100 | 94.94% | 62829834 | NC_052616.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 6, ASM1654582v1 | 374 | 12376 | 76% | 4e-102 | 95.36% | 57658180 | NC_052617.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 7, ASM1654582v1 | 363 | 14747 | 76% | 9e-99 | 94.51% | 60484081 | NC_052618.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 8, ASM1654582v1 | 364 | 13872 | 76% | 2e-99 | 94.56% | 74572912 | NC_052619.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 9, ASM1654582v1 | 340 | 8918 | 76% | 4e-92 | 92.83% | 44060374 | NC_052620.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 10, ASM1654582v1 | 379 | 8097 | 76% | 9e-104 | 95.78% | 38819584 | NC_052621.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 11, ASM1654582v1 | 339 | 9366 | 76% | 1e-91 | 93.16% | 50013132 | NC_052622.1 |
| <i>Eucalyptus grandis</i> - Ikeros | | | | | | | |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 1, ASM1654582v1 | 351 | 9716 | 76% | 2e-95 | 93.67% | 53951317 | NC_052612.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 2, ASM1654582v1 | 379 | 11390 | 76% | 9e-104 | 95.78% | 58486256 | NC_052613.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 3, ASM1654582v1 | 357 | 14725 | 76% | 4e-97 | 94.09% | 72470685 | NC_052614.1 |

| | | | | | | | |
|---|-----|-------|------|--------|--------|----------|-------------------|
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 4, ASM1654582v1 | 363 | 7633 | 76% | 9e-99 | 94.51% | 39120311 | NC_052615.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 5, ASM1654582v1 | 366 | 14800 | 76% | 7e-100 | 94.94% | 62829834 | NC_052616.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 6, ASM1654582v1 | 374 | 12376 | 76% | 4e-102 | 95.36% | 57658180 | NC_052617.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 7, ASM1654582v1 | 363 | 14747 | 76% | 9e-99 | 94.51% | 60484081 | NC_052618.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 8, ASM1654582v1 | 364 | 13872 | 76% | 2e-99 | 94.56% | 74572912 | NC_052619.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 9, ASM1654582v1 | 340 | 8918 | 76% | 4e-92 | 92.83% | 44060374 | NC_052620.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 10, ASM1654582v1 | 379 | 8097 | 76% | 9e-104 | 95.78% | 38819584 | NC_052621.1 |
| <i>Eucalyptus grandis</i> isolate ANBG69807.140 chromosome 11, ASM1654582v1 | 339 | 9366 | 76% | 1e-91 | 93.16% | 50013132 | NC_052622.1 |
| <i>Eucalyptus urophylla</i> - Sire | | | | | | | |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 1, whole genome shotgun sequence | 628 | 29486 | 100% | 8e-179 | 99.14% | 45562418 | JAOPUP010000001.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 2, whole genome shotgun sequence | 628 | 28132 | 100% | 8e-179 | 99.14% | 46294847 | JAOPUP010000002.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 3, whole genome shotgun sequence | 640 | 38956 | 100% | 0.0 | 99.71% | 49881599 | JAOPUP010000003.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 4, whole genome shotgun sequence | 634 | 31414 | 100% | 2e-180 | 99.43% | 35561653 | JAOPUP010000004.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 5, whole genome shotgun sequence | 634 | 26977 | 100% | 2e-180 | 99.43% | 40242915 | JAOPUP010000005.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 6, whole genome shotgun sequence | 628 | 26571 | 100% | 8e-179 | 99.14% | 47639447 | JAOPUP010000006.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 7, whole genome shotgun sequence | 623 | 26458 | 100% | 4e-177 | 98.85% | 42454922 | JAOPUP010000007.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 8, whole genome shotgun sequence | 640 | 38001 | 100% | 0.0 | 99.71% | 60186531 | JAOPUP010000008.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 9, whole genome shotgun sequence | 628 | 21121 | 100% | 8e-179 | 99.14% | 35568223 | JAOPUP010000009.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 10, whole genome shotgun sequence | 628 | 25028 | 100% | 8e-179 | 99.14% | 40619429 | JAOPUP010000010.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 11, whole genome shotgun sequence | 621 | 13901 | 100% | 1e-176 | 98.85% | 37154267 | JAOPUP010000011.1 |
| <i>Eucalyptus urophylla</i> - OGRE | | | | | | | |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 1, whole genome shotgun sequence | 741 | 3580 | 99% | 0.0 | 98.12% | 45562418 | JAOPUP010000001.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 2, whole genome shotgun sequence | 747 | 2927 | 99% | 0.0 | 98.35% | 46294847 | JAOPUP010000002.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 3, whole genome shotgun sequence | 752 | 3287 | 99% | 0.0 | 98.59% | 49881599 | JAOPUP010000003.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 4, whole genome shotgun sequence | 741 | 1455 | 99% | 0.0 | 98.12% | 35561653 | JAOPUP010000004.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 5, whole genome shotgun sequence | 741 | 2204 | 99% | 0.0 | 98.12% | 40242915 | JAOPUP010000005.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 6, whole genome shotgun sequence | 758 | 2659 | 99% | 0.0 | 98.82% | 47639447 | JAOPUP010000006.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 7, whole genome shotgun sequence | 691 | 1497 | 99% | 0.0 | 96.00% | 42454922 | JAOPUP010000007.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 8, whole genome shotgun sequence | 736 | 4197 | 99% | 0.0 | 97.88% | 60186531 | JAOPUP010000008.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 9, whole genome shotgun sequence | 686 | 1956 | 99% | 0.0 | 95.76% | 35568223 | JAOPUP010000009.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 10, whole genome shotgun sequence | 730 | 2165 | 99% | 0.0 | 97.65% | 40619429 | JAOPUP010000010.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 11, whole genome shotgun sequence | 680 | 680 | 99% | 0.0 | 95.53% | 37154267 | JAOPUP010000011.1 |
| <i>Eucalyptus urophylla</i> - Tekay | | | | | | | |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 1, whole genome shotgun sequence | 339 | 5065 | 80% | 1e-91 | 91.87% | 45562418 | JAOPUP010000001.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 2, whole genome shotgun sequence | 361 | 3210 | 80% | 3e-98 | 93.50% | 46294847 | JAOPUP010000002.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 3, whole genome shotgun sequence | 366 | 3971 | 80% | 6e-100 | 93.90% | 49881599 | JAOPUP010000003.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 4, whole genome shotgun sequence | 359 | 3157 | 80% | 1e-97 | 93.47% | 35561653 | JAOPUP010000004.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 5, whole genome shotgun sequence | 327 | 2530 | 80% | 3e-88 | 91.06% | 40242915 | JAOPUP010000005.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 6, whole genome shotgun sequence | 309 | 1637 | 80% | 1e-82 | 89.80% | 47639447 | JAOPUP010000006.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 7, whole genome shotgun sequence | 294 | 4222 | 80% | 3e-78 | 88.62% | 42454922 | JAOPUP010000007.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 8, whole genome shotgun sequence | 361 | 3971 | 80% | 3e-98 | 93.50% | 60186531 | JAOPUP010000008.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 9, whole genome shotgun sequence | 350 | 3674 | 80% | 6e-95 | 92.68% | 35568223 | JAOPUP010000009.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 10, whole genome shotgun sequence | 311 | 2112 | 80% | 3e-83 | 89.84% | 40619429 | JAOPUP010000010.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 11, whole genome shotgun sequence | 320 | 1471 | 80% | 5e-86 | 92.17% | 37154267 | JAOPUP010000011.1 |
| <i>Eucalyptus urophylla</i> - Ikeros | | | | | | | |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 1, whole genome shotgun sequence | 357 | 10872 | 76% | 4e-97 | 94.09% | 45562418 | JAOPUP010000001.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 2, whole genome shotgun sequence | 344 | 9858 | 76% | 3e-93 | 93.25% | 46294847 | JAOPUP010000002.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 3, whole genome shotgun sequence | 346 | 9095 | 76% | 8e-94 | 93.28% | 49881599 | JAOPUP010000003.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 4, whole genome shotgun sequence | 357 | 10441 | 76% | 4e-97 | 94.09% | 35561653 | JAOPUP010000004.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 5, whole genome shotgun sequence | 374 | 12339 | 76% | 4e-102 | 95.36% | 40242915 | JAOPUP010000005.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 6, whole genome shotgun sequence | 351 | 11431 | 76% | 2e-95 | 93.67% | 47639447 | JAOPUP010000006.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 7, whole genome shotgun sequence | 368 | 12830 | 76% | 2e-100 | 94.94% | 42454922 | JAOPUP010000007.1 |

| | | | | | | | |
|---|-----|-------|------|--------|--------|----------|------------------|
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 8, whole genome shotgun sequence | 357 | 14461 | 76% | 4e-97 | 94.09% | 60186531 | JAOPUP01000008.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 9, whole genome shotgun sequence | 346 | 9166 | 76% | 8e-94 | 93.25% | 35568223 | JAOPUP01000009.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 10, whole genome shotgun sequence | 346 | 7013 | 76% | 8e-94 | 93.59% | 40619429 | JAOPUP01000010.1 |
| <i>Eucalyptus urophylla</i> isolate SAP_URO_FK1756 chromosome 11, whole genome shotgun sequence | 351 | 8176 | 76% | 2e-95 | 94.02% | 37154267 | JAOPUP01000011.1 |
| <i>Eucalyptus globulus</i> - Sire | | | | | | | |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 1, whole genome shotgun sequence | 590 | 43067 | 100% | 4e-167 | 97.67% | 43159921 | JABKBK01000001.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 2, whole genome shotgun sequence | 601 | 45864 | 100% | 2e-170 | 97.71% | 51386561 | JABKBK01000002.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 3, whole genome shotgun sequence | 612 | 67298 | 100% | 8e-174 | 98.28% | 67249429 | JABKBK01000003.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 4, whole genome shotgun sequence | 586 | 46108 | 100% | 5e-166 | 97.38% | 36876699 | JABKBK01000004.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 5, whole genome shotgun sequence | 601 | 66471 | 100% | 2e-170 | 97.71% | 57417439 | JABKBK01000005.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 6, whole genome shotgun sequence | 599 | 51855 | 100% | 6e-170 | 97.71% | 54598197 | JABKBK01000006.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 7, whole genome shotgun sequence | 616 | 46836 | 100% | 6e-175 | 99.12% | 46883028 | JABKBK01000007.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 8, whole genome shotgun sequence | 612 | 75516 | 100% | 8e-174 | 98.28% | 64768990 | JABKBK01000008.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 9, whole genome shotgun sequence | 593 | 47259 | 100% | 3e-168 | 97.42% | 39375553 | JABKBK01000009.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 10, whole genome shotgun sequence | 593 | 32641 | 100% | 3e-168 | 97.42% | 37343655 | JABKBK01000010.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 11, whole genome shotgun sequence | 606 | 39845 | 100% | 4e-172 | 98.00% | 42041974 | JABKBK01000011.1 |
| <i>Eucalyptus globulus</i> - Ogre | | | | | | | |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 1, whole genome shotgun sequence | 741 | 2117 | 99% | 0.0 | 98.12% | 43159921 | JABKBK01000001.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 2, whole genome shotgun sequence | 712 | 1413 | 99% | 0.0 | 96.94% | 51386561 | JABKBK01000002.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 3, whole genome shotgun sequence | 697 | 5809 | 99% | 0.0 | 96.24% | 67249429 | JABKBK01000003.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 4, whole genome shotgun sequence | 706 | 706 | 99% | 0.0 | 96.70% | 36876699 | JABKBK01000004.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 5, whole genome shotgun sequence | 725 | 3486 | 99% | 0.0 | 97.41% | 57417439 | JABKBK01000005.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 6, whole genome shotgun sequence | 717 | 1414 | 99% | 0.0 | 97.18% | 54598197 | JABKBK01000006.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 7, whole genome shotgun sequence | 678 | 1475 | 99% | 0.0 | 95.53% | 46883028 | JABKBK01000007.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 8, whole genome shotgun sequence | 730 | 4280 | 99% | 0.0 | 97.65% | 64768990 | JABKBK01000008.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 9, whole genome shotgun sequence | 706 | 2816 | 99% | 0.0 | 96.71% | 39375553 | JABKBK01000009.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 10, whole genome shotgun sequence | 689 | 689 | 99% | 0.0 | 96.00% | 37343655 | JABKBK01000010.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 11, whole genome shotgun sequence | 708 | 1402 | 99% | 0.0 | 96.71% | 42041974 | JABKBK01000011.1 |
| <i>Eucalyptus globulus</i> - Tekay | | | | | | | |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 1, whole genome shotgun sequence | 244 | 682 | 80% | 3e-63 | 85.25% | 43159921 | JABKBK01000001.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 2, whole genome shotgun sequence | 344 | 1519 | 80% | 3e-93 | 92.28% | 51386561 | JABKBK01000002.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 3, whole genome shotgun sequence | 324 | 2755 | 80% | 4e-87 | 90.73% | 67249429 | JABKBK01000003.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 4, whole genome shotgun sequence | 344 | 595 | 80% | 3e-93 | 92.28% | 36876699 | JABKBK01000004.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 5, whole genome shotgun sequence | 355 | 3875 | 80% | 1e-96 | 93.09% | 57417439 | JABKBK01000005.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 6, whole genome shotgun sequence | 272 | 751 | 80% | 1e-71 | 86.99% | 54598197 | JABKBK01000006.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 7, whole genome shotgun sequence | 339 | 2064 | 80% | 1e-91 | 91.87% | 46883028 | JABKBK01000007.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 8, whole genome shotgun sequence | 311 | 1481 | 80% | 3e-83 | 89.84% | 64768990 | JABKBK01000008.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 9, whole genome shotgun sequence | 305 | 4014 | 80% | 1e-81 | 89.43% | 39375553 | JABKBK01000009.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 10, whole genome shotgun sequence | 302 | 948 | 79% | 2e-80 | 89.63% | 37343655 | JABKBK01000010.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 11, whole genome shotgun sequence | 322 | 600 | 80% | 1e-86 | 90.65% | 42041974 | JABKBK01000011.1 |
| <i>Eucalyptus globulus</i> - Ikeros | | | | | | | |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 1, whole genome shotgun sequence | 335 | 7377 | 76% | 2e-90 | 92.41% | 43159921 | JABKBK01000001.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 2, whole genome shotgun sequence | 340 | 8536 | 76% | 4e-92 | 92.83% | 51386561 | JABKBK01000002.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 3, whole genome shotgun sequence | 350 | 11516 | 76% | 6e-95 | 93.64% | 67249429 | JABKBK01000003.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 4, whole genome shotgun sequence | 340 | 7570 | 76% | 4e-92 | 92.83% | 36876699 | JABKBK01000004.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 5, whole genome shotgun sequence | 346 | 14615 | 76% | 8e-94 | 93.25% | 57417439 | JABKBK01000005.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 6, whole genome shotgun sequence | 335 | 8459 | 76% | 2e-90 | 92.44% | 54598197 | JABKBK01000006.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 7, whole genome shotgun sequence | 335 | 9282 | 76% | 2e-90 | 92.44% | 46883028 | JABKBK01000007.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 8, whole genome shotgun sequence | 335 | 14441 | 76% | 2e-90 | 92.41% | 64768990 | JABKBK01000008.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 9, whole genome shotgun sequence | 344 | 8889 | 76% | 3e-93 | 93.25% | 39375553 | JABKBK01000009.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 10, whole genome shotgun sequence | 346 | 6357 | 76% | 8e-94 | 93.25% | 37343655 | JABKBK01000010.1 |
| <i>Eucalyptus globulus</i> isolate CBG9404798.5 chromosome 11, whole genome shotgun sequence | 346 | 6706 | 76% | 8e-94 | 93.25% | 42041974 | JABKBK01000011.1 |