



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE BIOCIÊNCIAS
CURSO DE GRADUAÇÃO
BACHARELADO EM CIÊNCIAS BIOLÓGICAS COM ÊNFASE EM CIÊNCIAS
AMBIENTAIS

AUGUSTO CÉSAR MUNIZ BEZERRA

ANÁLISE GENÔMICA COMPARATIVA DA GIMNOSPERMA RELICTUAL *Podocarpus*
***sellowii* KLOTZSCH EX ENDL. (PODOCARPACEAE) EM POPULAÇÕES DISJUNTAS**
AO LONGO DO BRASIL

Recife
2024

AUGUSTO CÉSAR MUNIZ BEZERRA

**ANÁLISE GENÔMICA COMPARATIVA DA GIMNOSPERMA RELICTUAL *Podocarpus
sellowii* KLOTZSCH EX ENDL. (PODOCARPACEAE) EM POPULAÇÕES DISJUNTAS
AO LONGO DO BRASIL**

Trabalho de Conclusão de Curso de Graduação
apresentado ao Bacharelado em Ciências
Biológicas com ênfase em Ciências Ambientais
da Universidade Federal de Pernambuco, como
requisito parcial para obtenção do título de
bacharel.

Orientador (a): Luiz Gustavo Rodrigues Souza

Recife

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Bezerra, Augusto César Muniz.

Análise genômica comparativa da gimnosperma relictual *Podocarpus sellowii* Klotzsch ex Endl. (Podocarpaceae) em populações disjuntas ao longo do Brasil / Augusto César Muniz Bezerra. - Recife, 2024.

69 : il., tab.

Orientador(a): Luiz Gustavo Rodrigues Souza

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Biociências, Ciências Biológicas /Ciências Ambientais - Bacharelado, 2024.

Inclui referências.

1. Diversidade Genética. 2. Análise Genômica. 3. Repeatoma. 4. Filogenômica. 5. Estrutura Populacional. I. Souza, Luiz Gustavo Rodrigues. (Orientação). II. Título.

580 CDD (22.ed.)

AUGUSTO CÉSAR MUNIZ BEZERRA

**ANÁLISE GENÔMICA COMPARATIVA DA GIMNOSPERMA RELICTUAL *Podocarpus
sellowii* KLOTZSCH EX ENDL. (PODOCARPACEAE) EM POPULAÇÕES DISJUNTAS
AO LONGO DO BRASIL**

Trabalho de Conclusão de Curso de Graduação
apresentado ao Bacharelado em Ciências
Biológicas com ênfase em Ciências Ambientais
da Universidade Federal de Pernambuco, como
requisito parcial para obtenção do título de
bacharel.

Aprovado em: 14/10/2024

COMISSÃO EXAMINADORA

Documento assinado digitalmente
 **LUIZ GUSTAVO RODRIGUES SOUZA**
Data: 22/10/2024 13:02:56-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Luiz Gustavo Rodrigues Souza (Orientador)
Universidade Federal de Pernambuco – UFPE

Documento assinado digitalmente
 **LUCAS ALEXANDRE DE SOUZA COSTA**
Data: 21/10/2024 21:23:01-0300
Verifique em <https://validar.iti.gov.br>

Dr. Lucas Alexandre de Souza Costa
Universidade Federal de Pernambuco – UFPE

Documento assinado digitalmente
 **BRENA VAN LUME DO NASCIMENTO**
Data: 21/10/2024 21:04:12-0300
Verifique em <https://validar.iti.gov.br>

Dra. Brena Van-Lume do Nascimento
Universidade Federal de Pernambuco – UFPE

Recife
2024

À minha mãe, Josineide, e à minha avó, Maria Eneide, por todo o amor incondicional.

AGRADECIMENTOS

Quero começar deixando claro que, para mim, agradecer não é difícil, pois a gratidão é o que me move na vida. Apesar disso, não imaginei que seria tão difícil expressar, em palavras, uma fração desse sentimento a todas e todos que me acompanharam e estiveram comigo durante minha trajetória acadêmica, até chegarmos a esta monografia. Primeiramente, gostaria de agradecer infinitamente à **minha mãe, Josineide**, e à **minha avó, Maria Eneide**, pela mais genuína forma de amor incondicional que tive a sorte de receber em minha vida, pela força e pelos exemplos de dedicação e resiliência em todos os momentos. Elas são a base de tudo que sou e conquistei, os grandes amores da minha vida. Agradeço também ao **meu avô, José Muniz**, por ter sido meu exemplo de pai, por sempre ter me incentivado e investido para que eu fosse "alguém na vida". Te amo, vô, saudades eternas. Gostaria de agradecer também aos **meus irmãos, Arthur e Ari**, os quais tanto amo e que são exemplos para mim. Tenho muita sorte da vida ter me dado vocês dois.

Gostaria de expressar também minha imensa gratidão a todas as pessoas incríveis que conheci durante a graduação, em especial às minhas amigas, **Elisa Maia, Ana Júlia Monteiro, Bárbara Coelho, Flora Tomás, Deborah Lacerda e Cauane Assunção**, as quais tanto amo. Obrigado por compartilharem comigo tantos momentos bons e inesquecíveis, e por estarem comigo até mesmo nos momentos ruins. Vocês são o maior presente que a UFPE poderia ter me dado!

Também não poderia deixar de agradecer a todas e todos do **Laboratório de Citogenética e Evolução Vegetal** da UFPE, cuja colaboração e apoio foram fundamentais para a minha formação acadêmica e científica. Agradeço especialmente a **Natália Castro, Lucas Costa e Brena Van-Lume**, por toda a paciência e dedicação para compartilhar e discutir ciência comigo. Também agradeço ao meu orientador, Prof. **Gustavo Souza**, por sua orientação, seus conselhos e por sempre acreditar no meu potencial, mesmo nos momentos mais desafiadores. À Profa. **Andrea Pedrosa**, sou imensamente grato pelo acolhimento, pela generosidade em compartilhar seu conhecimento e por ser tão incrível, fonte eterna de inspiração.

Por fim e não menos importante, quero agradecer a mim mesmo por não ter desistido de mim e da biologia em todos os momentos em que esse pareceu ser o caminho mais fácil. A vida e os sonhos valem a pena!

RESUMO

A fração repetitiva dos genomas (repeatoma) é a porção mais abundante nos genomas eucarióticos, principalmente devido à sua capacidade de gerar cópias dos seus elementos, contribuindo significativamente para a variação genômica entre populações. A rápida evolução do repeatoma pode ser influenciada por pressões ambientais, sugerindo que sua composição e abundância refletem adaptações a condições ecológicas específicas. Isso é especialmente importante para espécies com distribuição fragmentada e habitats especializados, como *Podocarpus sellowii* Klotzsch ex Endl, uma das poucas gimnospermas nativas do Brasil. Sua distribuição em pequenas populações isoladas na Mata Atlântica e Cerrado reflete um padrão relictual causado por vicariância nos últimos 20 milhões de anos. Neste estudo, analisamos comparativamente o genoma de 10 indivíduos de populações disjuntas de *P. sellowii* ao longo de um gradiente latitudinal, do Ceará a Santa Catarina. Para isso, caracterizamos os *repeats* mais abundantes das populações no *pipeline RepeatExplorer2*, usando *reads* de baixa cobertura (Illumina HiSeq2500). Coordenadas geográficas foram usadas para extrair variáveis climáticas para cada população através da plataforma WorldClim. Os perfis de repeatoma foram interpretados considerando relações filogenéticas baseadas em plastomas inteiros montados para este estudo. Utilizando PGLS (do inglês, *Phylogenetic Generalized Least Squares*), foi possível calcular a correlação entre *repeats* e variáveis ecológicas. Em geral, os repeatomas de *P. sellowii* apresentaram alta abundância de retrotransposons LTR (do inglês, *Long Terminal Repeats*), especialmente Ty1/copia-Ivana (29,65% a 35,13%), Ty3/gypsy-Ogre (10,88% a 11,29%) e Ty1/copia-SIRE (8,06% a 10,46%). Como as abundâncias de cada *repeat* foram polimórficas entre nossas amostras de cada população, testamos a influência da ecologia e/ou filogenia sobre isso. Identificamos uma correlação significativa entre vários LTRs e transposons de DNA com características de temperatura e precipitação. A árvore filogenética baseada em *repeats* revelou grupos ecológicos, separando indivíduos da Mata Atlântica dos do Cerrado, em vez de relações genealógicas. Nossos dados mostram que cerca de 40% da fração genômica repetitiva é diretamente influenciada por variáveis ecológicas. Isso destaca interações significativas entre variáveis ambientais e a abundância de *repeats*, sugerindo que a especificidade ecológica dessas plantas relictuais impactou a evolução de seus genomas.

Palavras-chave: Diversidade Genética; Análise Genômica; Repeatoma; Filogenômica; Estrutura Populacional.

ABSTRACT

The repetitive fraction of genomes (the repeatome) is the most abundant portion in eukaryotic genomes, primarily due to its ability to generate copies of its elements, significantly contributing to genomic variation among populations. The rapid evolution of the repeatome can be influenced by environmental pressures, suggesting that its composition and abundance reflect adaptations to specific ecological conditions. This is particularly important for species with fragmented distributions and specialized habitats, such as *Podocarpus sellowii* Klotzsch ex Endl, one of the few native gymnosperms of Brazil. Its distribution in small, isolated populations in the Atlantic Forest and Cerrado reflects a relictual pattern caused by vicariance over the past 20 million years. In this study, we conducted a comparative genome analysis of 10 individuals from disjunct populations of *P. sellowii* along a latitudinal gradient, from Ceará to Santa Catarina. To achieve this, we characterized the most abundant repeats in the populations using the RepeatExplorer2 pipeline, based on low-coverage reads (Illumina HiSeq2500). Geographic coordinates were used to extract climate variables for each population via the WorldClim platform. The repeatome profiles were interpreted considering phylogenetic relationships based on whole plastomes assembled for this study. Using Phylogenetic Generalized Least Squares (PGLS), we were able to calculate the correlation between repeats and ecological variables. Overall, the repeatomes of *P. sellowii* showed a high abundance of LTR retrotransposons (Long Terminal Repeats), particularly Ty1/copia-Ivana (29.65% to 35.13%), Ty3/gypsy-Ogre (10.88% to 11.29%), and Ty1/copia-SIRE (8.06% to 10.46%). Since the abundance of each repeat was polymorphic across our samples from each population, we tested the influence of ecology and/or phylogeny on this variation. We identified a significant correlation between several LTRs and DNA transposons with temperature and precipitation characteristics. The phylogenetic tree based on repeats revealed ecological groups, separating individuals from the Atlantic Forest from those of the Cerrado, rather than genealogical relationships. Our data show that approximately 40% of the repetitive genomic fraction is directly influenced by ecological variables. This highlights significant interactions between environmental variables and repeat abundance, suggesting that the ecological specificity of these relict plants has impacted the evolution of their genomes.

Keywords: Genetic Diversity; Genomic Analysis; Repeatome; Phylogenomics; Population Structure.

SUMÁRIO

1 INTRODUÇÃO	9
1.1 OBJETIVOS	11
1.1.1 Geral	11
1.1.2 Específicos	11
2 REFERENCIAL TEÓRICO	12
2.1 DNA repetitivo	12
2.2 FERRAMENTAS para A caracterização do DNA repetitivo	13
2.3 Métodos Filogenéticos Comparativos	14
2.4 Estudo da Interação Genoma-Ambiente	16
2.5 A Família Podocarpaceae e a espécie <i>Podocarpus sellowii</i>	17
3 MATERIAIS E MÉTODOS	20
3.1 Caracterização e Análise <i>in silico</i> da Fração Repetitiva	20
3.2 Filogenia Molecular	20
3.3 Filogenia de <i>Repeats</i> e Sinal Filogenético	21
3.4 Distribuição Geográfica e Seleção de Variáveis Ecológicas	21
3.5 Interação Genoma-Ambiente	22
REFERÊNCIAS	23
4 RESULTADOS	30
5. CONCLUSÕES	55
REFERÊNCIAS	56

1 INTRODUÇÃO

Sequências repetitivas de DNA (*repeats*) constituem a maior parte do genoma em eucariotos e têm um papel central na estruturação e evolução genômica. Sua classificação básica distingue entre (1) elementos transponíveis (TEs, do inglês *transposable elements*) e (2) sequências em tandem, como DNAsat, teloméricos e centroméricos (Biscotti, Olmo e Heslop-Harrison, 2015). Os TEs, por exemplo, podem se mover ao longo do genoma, provocando mutações e rearranjos significativos, com efeitos potenciais na variabilidade genética e na estrutura cromossômica (Bourque *et al.*, 2018; Hénault *et al.*, 2024). Eles se dividem em retrotransposons (Classe I), que utilizam um intermediário de RNA para sua replicação, e transposons de DNA (Classe II), que se movem por um mecanismo de "cortar e colar" via intermediários de DNA (Bourque *et al.*, 2018). Já as sequências em tandem, organizadas em blocos adjacentes, desempenham funções estruturais essenciais, como nos centrômeros e telômeros, além de influenciarem a formação de heterocromatina (Hartley & O'Neill, 2019).

Em algumas plantas, esses elementos repetitivos podem compor até 80% do genoma total, como é o caso de *Zea mays* L., sendo os principais responsáveis pela variação no tamanho do genoma (GS, do inglês *Genome Size*) entre diferentes espécies (Haberer *et al.*, 2020; González; Poggio, 2021). Os TEs desempenham papéis importantes nos processos adaptativos das plantas, principalmente em resposta a estresses ambientais (Horváth; Merenciano; González, 2017; Chumová *et al.*, 2022). Embora os TEs evoluam majoritariamente por deriva genética (Bourque *et al.*, 2018), sua atividade pode ser modulada por fatores ambientais, aumentando assim sua influência na adaptação (Horváth; Merenciano; González, 2017). A relação entre a abundância de TEs, a estrutura heterocromática e o GS, e como essas características variam em resposta a fatores ecológicos, como temperatura, precipitação e tipo de solo, é um campo emergente dentro da ecologia evolutiva (Chenais *et al.*, 2012; Castro *et al.*, 2024).

Métodos filogenéticos comparativos têm se mostrado ferramentas essenciais para testar a correlação entre características genômicas e variáveis ecológicas (Uyeda *et al.*, 2018). Tais abordagens permitem incorporar informações filogenéticas e compará-las com variáveis ambientais, facilitando a compreensão dos processos adaptativos em uma perspectiva evolutiva mais ampla (Garamszegi, 2014). O avanço das tecnologias de Sequenciamento de Nova Geração (NGS, do inglês *Next-Generation Sequencing*) tem facilitado a geração de grandes volumes de dados a baixos custos, permitindo a análise detalhada da evolução genômica em grupos taxonômicos antes pouco explorados

(Dodsworth *et al.*, 2019). Essa abordagem tem sido particularmente útil para estudar grupos com histórias evolutivas complexas, como as gimnospermas, com linhagens relictuais que datam de aproximadamente 300 milhões de anos (Murray, 2013).

A família Podocarpaceae Endl., por exemplo, é uma das poucas gimnospermas que persistem nas florestas tropicais, com destaque para o gênero *Podocarpus* L'Hér. ex Pers., que possui uma diversidade significativa em ambientes úmidos do Hemisfério Sul (Farjon, 2017). Este gênero, que se originou no clima frio e úmido de Gondwana, representa um grupo-chave para estudos filogenéticos e biogeográficos devido à sua distribuição ampla e à sua persistência em habitats tropicais e subtropicais (Page, 1990; Mill, 2003). A espécie *Podocarpus sellowii* Klotzsch ex Endl. é particularmente interessante, pois apresenta uma distribuição fragmentada ao longo de diferentes regiões do Brasil, com populações isoladas no Sul, Sudeste, Centro-Oeste e Nordeste (Lima; Barbosa, 1998; CNCFlora, 2024). A adaptação de *P. sellowii* a condições ecológicas variadas faz dela um modelo ideal para investigar como fatores ambientais moldam a evolução de seu genoma.

Estudos filogeográficos anteriores com *P. sellowii* baseados em marcadores nucleares e plastidiais revelaram padrões de divergência populacional entre diferentes regiões brasileiras (Melo *et al.*, 2019). No entanto, essas análises se concentraram em marcadores moleculares específicos e amostragens limitadas, o que não permitiu explorar completamente a influência dos *repeats* na evolução da espécie. Com o avanço das tecnologias de sequenciamento, tornou-se possível acessar com maior resolução a composição genômica, abrindo novas possibilidades para investigar como as variações na abundância e na distribuição dos TEs estão associadas a fatores ambientais e à história evolutiva de *P. sellowii*.

Neste contexto, o presente estudo busca investigar como os *repeats* do genoma de *P. sellowii* estão relacionados à sua história demográfica e à variação ambiental ao longo de sua distribuição fragmentada no Brasil. A análise das interações genoma-ambiente em espécies relictuais como *P. sellowii* contribui para o entendimento dos mecanismos de adaptação e diversificação genômica em plantas tropicais. Além disso, esse tipo de abordagem é relevante para fundamentar a conservação de populações geneticamente isoladas e adaptadas a diferentes condições ecológicas. Assim, este trabalho visa aprofundar o conhecimento da história demográfica de *P. sellowii* na região Neotropical, avaliando suas relações filogenéticas e o impacto de eventos climáticos na evolução de sua fração repetitiva. A espécie também é utilizada como modelo para entender a interação genoma-ambiente por meio de métodos filogenéticos comparativos, fornecendo uma compreensão mais abrangente da dinâmica evolutiva de populações fragmentadas.

1.1 OBJETIVOS

1.1.1 Geral

- Caracterizar a fração repetitiva do genoma de 10 indivíduos de *Podocarpus sellowii* Klotzsch ex Endl. (Podocarpaceae) em sua distribuição ao longo de um gradiente latitudinal no Brasil, usando os polimorfismos nos *repeats* como fontes de dados para entender a história demográfica da espécie.

1.1.2 Específicos

- Realizar análises comparativas das frações repetitivas caracterizadas, estabelecendo dados de abundância e similaridade de cada *cluster*;
- Gerar topologias baseadas em abundância dos *repeats* para testar as hipóteses filogeográficas estabelecidas por Melo *et al.*, 2019;
- Comparar topologias baseadas em abundância e similaridade dos *repeats* com as redes de plastomas e DNAr;
- Testar o impacto do ambiente (variáveis ecológicas) e das principais classes de *repeats* por métodos filogenéticos comparativos.

2 REFERENCIAL TEÓRICO

2.1 DNA REPETITIVO

Sequências repetitivas, também conhecidas como elementos repetitivos, ou *repeats*, são padrões curtos ou longos de ácidos nucleicos (DNA ou RNA) que ocorrem em múltiplas cópias ao longo do genoma (Schley *et al.*, 2022). As sequências repetitivas são categorizadas em diferentes classes, dependendo de características como estrutura, comprimento, localização, origem e modo de multiplicação (Mitchell, 2018). O DNA repetitivo constitui uma parte significativa do genoma dos eucariotos (Wells; Feschotte, 2020), sendo esse tipo de sequência dividido em duas principais categorias: DNA repetitivo em tandem e DNA repetitivo disperso (Bourque *et al.*, 2018; Paço *et al.*, 2019). Essas categorias são definidas com base na organização e no padrão de distribuição dos *repeats*: as sequências em tandem são organizadas em blocos contínuos, enquanto as dispersas estão distribuídas aleatoriamente ao longo do genoma (Bourque *et al.*, 2018).

Dentro da fração repetida em tandem, encontram-se, por exemplo, o DNA ribossomal (DNAr) e os DNA satélites (DNAsat) (Bourque *et al.*, 2018). As sequências dispersas são geralmente associadas aos elementos transponíveis (TEs), que possuem a capacidade de se deslocar ao longo do genoma. Estes elementos são classificados em duas classes conforme seu mecanismo de transposição: os Elementos de Classe I (retrotransposons), que utilizam um mecanismo de “copiar e colar” através de uma molécula de RNA intermediário, e Elementos de Classe II (transposons), que utilizam um mecanismo de “cortar e colar” com um intermediário de DNA, embora existam variações de acordo com a família do retrotransposon (Bourque *et al.*, 2018; Cain *et al.*, 2020). Os elementos de Classe I são mais comuns em plantas e podem ser subdivididos em LTR (do inglês, *Long Terminal Repeats*) e não-LTR, que recebem esse nome devido às longas repetições terminais que flanqueiam a sequência interna (Miousse *et al.*, 2015; Macas *et al.*, 2015). Eles são classificados em cinco superfamílias: Ty1/copia, Ty3/gypsy, Bel-Pao, Retrovírus e Retrovírus endógenos (Bennetzen, 2014). Contudo, apenas as superfamílias Ty1/copia e Ty3/gypsy são encontradas em plantas (Neumann *et al.*, 2019), influenciando a evolução e variação no tamanho do genoma (Bennetzen, 2014).

Além de variações no GS, devido à sua capacidade de alterar sua posição dentro do genoma, os TEs podem modificar a sequência cromossômica original, resultando em diversos efeitos e rearranjos que afetam, por exemplo, a evolução da regulação gênica (Kazazian, 2004) e o tamanho do genoma (Sader, 2020). Além disso, a deriva genética

também exerce um papel crucial na evolução dos elementos transponíveis. Devido ao fato de que esses elementos podem se mover e se replicar de forma autônoma dentro do genoma, muitas vezes sem exercer uma função benéfica para o organismo, sua presença pode ser considerada neutra ou até mesmo deletéria em algumas circunstâncias (Lynch, 2007). Em populações pequenas, a deriva genética pode permitir a fixação de elementos transponíveis que, em condições normais, seriam eliminados por seleção natural (Szitenberg *et al.*, 2016). Esse processo de fixação aleatória pode resultar em uma acumulação desses elementos, contribuindo para a expansão do genoma e o aumento da variabilidade genômica, como observado em várias espécies de plantas e animais (Bourque *et al.*, 2018). Assim, a interação entre a deriva genética e os TEs exemplifica como processos estocásticos influenciam a evolução do genoma, destacando o papel da evolução não adaptativa na dinâmica desses elementos (Szitenberg *et al.*, 2016).

A alta variabilidade na fração repetitiva dos genomas pode, também, estar relacionada a condições ecológicas. Retrotransposons podem ser amplificados ou desamplificados sob a influência de fatores ambientais, especialmente durante situações de estresse térmico, salinidade, precipitação e outras características presentes no nicho ecológico dos organismos. (Chénais *et al.*, 2012; Negi; Rai; Suprasanna, 2016; Lyu *et al.*, 2018). Esses fatores possuem a capacidade de produzir rapidamente novos genótipos e fenótipos em resposta a perturbações ambientais (Pimpinelli; Piacentini, 2020).

2.2 FERRAMENTAS PARA A CARACTERIZAÇÃO DO DNA REPETITIVO

A tecnologia de sequenciamento de próxima geração (do inglês, *Next Generation Sequencing* – NGS) tem revolucionado os campos da genômica, filogeografia, genética de populações, ecologia molecular e biologia da conservação. Nesse contexto, o NGS tem facilitado o acesso aos genomas de diversos organismos, tornando possível a ampla caracterização da fração repetitiva desses genomas (Weiss-Schneeweiss *et al.*, 2015). Além disso, o desenvolvimento de tecnologias NGS não só viabilizaram a montagem de genomas plastidiais, mitocondriais e nucleares em um tempo menor, como também facilitaram a caracterização de sequências repetitivas, como elementos transponíveis, DNAs satélites, dentre outros (Buermans; Den Dunnen, 2014).

Desse modo, atualmente, é possível estabelecer abordagens filogenéticas clássicas mais robustas, como por exemplo, baseadas em plastomas e/ou DNAr inteiros, e comparar com tendências evolutivas de linhagens específicas de *repeats*, inferindo sobre a sua evolução de poder informativo para o estabelecimento de relações de ancestralidade (Wang

et al., 2024). Diferentes abordagens moleculares foram desenvolvidas, como o *Genome Skimming* (sequenciamento de baixa cobertura), que se destaca por sequenciar pequenos fragmentos aleatórios do genoma, permitindo o desenvolvimento de várias ferramentas de bioinformática para análise genômica (Novak *et al.*, 2013). Essa abordagem apresenta a vantagem de poder ser aplicada em organismos não-modelo e vem se popularizando com uma gradativa redução de custos para a obtenção dos dados NGS (Ekblom & Galindo, 2011).

Utilizando dados de *genome skimming* e a caracterização quantitativa e qualitativa dos *repeats* é possível gerar dados de abundância e/ou similaridade dos *repeats* para estabelecer hipóteses filogenéticas (Dodsworth, 2015a). De modo que essa técnica é capaz de eleger como “estados de caráter” cada classe de *repeat (cluster)*, tal como sua abundância nos diferentes terminais (Dodsworth *et al.*, 2015a).

A plataforma *online RepeatExplorer* (<http://www.repeatexplorer.org/>), utiliza um princípio de “clusterização” para caracterização dos *repeats* por meio de uma comparação par-a-par de todos os *reads* gerados no sequenciamento (*all-to-all pairwise*). Essa análise pode ser feita de forma comparativa entre diferentes indivíduos ou espécies (Novák *et al.*, 2010; Novák *et al.*, 2013). A comparação da abundância e/ ou similaridade de *clusters* de *repeats* parecem conter uma linha filogenética em nível de gênero e família, podendo ser, inclusive, mais informativa que sequências tradicionalmente utilizadas por sequenciamento tipo Sanger (Dodsworth *et al.*, 2015a; 2015b; 2016; 2017; 2018).

Porém, ainda não há uma grande quantidade de estudos desse tipo de abordagem para estabelecer relações evolutivas entre linhagens dentro de uma mesma espécie, e essa pode ser uma abordagem inovadora. Como parte desses *repeats* (especialmente TEs) são sensíveis às variações ambientais, é possível utilizar esse tipo de abordagem para testar correlações entre variáveis genômicas e ecológicas (Chumová *et al.*, 2022).

2.3 MÉTODOS FILOGENÉTICOS COMPARATIVOS

Os métodos filogenéticos comparativos, ou PCMs (do inglês, *Phylogenetic Comparative Methods*), são amplamente utilizados em estudos macroevolutivos e são considerados um dos avanços metodológicos mais significativos da biologia evolutiva nas últimas décadas (Paradis, 2014). Os PCMs surgiram como um conjunto de metodologias para testar a “não independência” das observações (Paradis, 2014). Eles combinam vários modelos estatísticos para o estudo analítico de espécies, populações e indivíduos, com o

objetivo de inferir hipóteses sobre o histórico evolutivo dos grupos biológicos em relação a diversas variáveis (Garamszegi, 2014; Paradis, 2014).

A PGLS (do inglês, *Phylogenetic Generalised Least Squares*) é uma técnica que identifica a quantidade de correlação esperada entre espécies com base em sua história evolutiva compartilhada, utilizando a filogenia como base (Housworth; Martins; Lynch, 2004). Trata-se provavelmente do método de modelos comparativos filogenéticos (PCM) mais comumente utilizado (Kamilar; Cooper, 2013; Losos, 2011). A PGLS é amplamente empregada para estabelecer associações evolutivas entre dois ou mais traços biológicos (Symonds; Blomberg, 2014), como a relação entre o tamanho do genoma (GS) e a quantidade de DNA repetitivo (Reinar et al., 2023) ou a relação entre variáveis ecológicas e o aumento da abundância de elementos repetitivos (Schley et al., 2022).

A não independência dos traços pode ser medida também pelo sinal filogenético (Borges et al., 2019). O sinal filogenético pode ser mensurado e testado por diferentes métodos estatísticos, como o I de Moran (Morales-Castilla et al., 2015; Borges et al., 2019), que mede a autocorrelação entre as espécies de acordo com suas distâncias filogenéticas; os testes de média Abouheif C (Pavoine; Ricotta, 2020), que se baseiam na soma das diferenças quadradas sucessivas entre valores de características de espécies vizinhas; o K de Blomberg (Münkemüller et al., 2012; Mitteroecker et al., 2024), estimado pela razão entre o erro quadrado médio dos terminais e o erro corrigido pela covariância filogenética; e o λ de Pagel (Smith; O'Meara, 2019), onde o valor de λ é estimado por máxima verossimilhança e comparado à verossimilhança de um valor hipotético.

Considera-se um sinal filogenético forte como indicativo de que o traço de interesse tende a ser compartilhado por espécies filogeneticamente próximas, enquanto em espécies filogeneticamente distantes o traço seria menos similar, indicando um sinal filogenético fraco (Revell, 2013; Münkemüller et al., 2012; Smith; O'Meara, 2019). O sinal filogenético é uma ferramenta importante em diversas áreas, permitindo, por exemplo, analisar traços como a evolução do padrão de bandas CMA/DAPI (Van-Lume et al., 2017) e o GS (Souza et al., 2019). Os métodos filogenéticos comparativos surgiram então como uma solução estatística multivariada para investigar a evolução dos diferentes tipos de caracteres, sem desconsiderar as relações de parentesco (Uyeda et al., 2018). Essas abordagens permitem a integração de dados genéticos e ecológicos, possibilitando análises evolutivas multivariadas que consideram as relações filogenéticas entre os organismos. Avanços recentes, como a aplicação de estatísticas de sinal filogenético e a incorporação de variáveis ambientais, têm contribuído para a identificação de padrões adaptativos e de estruturação genômica em várias espécies. (Castro et al., 2024)

2.4 ESTUDO DA INTERAÇÃO GENOMA-AMBIENTE

A interação genoma-ambiente atualmente pode ser melhor estudada através dos Métodos Filogenéticos Comparativos, sendo possível realizar um conjunto complexo de análises multivariadas que estimam a relação entre variáveis incluindo nos algoritmos as relações evolutivas (Garamszegi, 2014). A complexa interação entre componentes do genoma (como por exemplo: GS, atividade gênica e elementos transponíveis) e variáveis ecológicas (ex: solo, temperatura, precipitação etc.) são um desafio para a biologia evolutiva (Bourgeois *et al.*, 2020). Notavelmente, nenhum padrão geral absoluto na interação genoma-ambiente foi descrito até o presente momento, assim os resultados muitas vezes parecem mostrar tendências contraditórias, como o caso da relação entre GS e latitude, embora algumas tendências tenham sido observadas, como o aumento do GS em espécies de regiões temperadas em comparação com aquelas de regiões tropicais, essas correlações não são consistentes entre diferentes grupos taxonômicos (Bureš *et al.*, 2024).

Recentemente, alguns estudos demonstraram que o DNA repetitivo (*repeats*) pode ser afetado pelo ambiente, resultando em ativação/desativação de genes e mudanças na abundância/composição do DNA repetitivo (Canapa *et al.*, 2020). Nessa perspectiva, fica evidente que o ambiente tem um impacto significativo na evolução do genoma (Bogan & Yi, 2024), influenciando, por sua vez, as características estruturais e funcionais das células, mediando processos evolutivos e facilitando a adaptação a novos ambientes (Lyu *et al.*, 2018). Além disso, as interações entre fatores ambientais e componentes genômicos geram variabilidade sobre a qual a seleção natural pode atuar, integrando processos típicos de especiação (Johnson, 2022). Dessa forma, as variações ambientais, como mudanças na temperatura, podem modular a abundância de *repeats* no genoma, influenciando diretamente a dinâmica evolutiva de espécies. Essas alterações na frequência e mobilidade dos *repeats* podem, em certos contextos, atuar como mecanismos de pré-adaptação, facilitando respostas evolutivas rápidas e eficazes frente às novas pressões ambientais (Canapa *et al.*, 2020).

A diversidade na composição dos *repeats* dos genomas das plantas resulta de uma interação complexa entre os *repeats*, os genes e o ambiente (Schley *et al.*, 2022). Embora a temperatura seja um dos principais parâmetros que afetam diretamente os elementos móveis, outros fatores ambientais, como precipitação, salinidade, altitude e radiação, também desempenham papéis cruciais na modulação dessas frações repetitivas. A salinidade, por exemplo, pode influenciar a atividade de elementos transponíveis em

ambientes áridos, afetando a estabilidade e a evolução genômica (Schley *et al.*, 2022). Altitudes elevadas, por sua vez, aumentam a exposição à radiação UV, o que pode induzir mutações e alterar a atividade de transposons, contribuindo para a plasticidade genômica e para a adaptação das espécies a condições ambientais adversas (Canapa *et al.*, 2020; Schley *et al.*, 2022). Esses fatores ambientais, em conjunto com a temperatura, agem de forma integrada para moldar a evolução dos genomas, influenciando tanto a composição quanto a abundância dos *repeats*. Análises bioquímicas sugerem que a atividade dos genes de TEs pode ser predominante em resposta ao estresse. Estudos experimentais mostraram que a sensibilidade à temperatura da protease dos elementos Ty1/copia pode aumentar sua eficiência de transposição (Gazda *et al.*, 2020). Essas análises indicam que os genes de TEs são inerentemente sensíveis à temperatura, podendo maximizar os eventos de transposição em situações de elevação de temperatura (Gazda *et al.*, 2020). Embora as interações genoma-temperatura sejam um fenômeno relevante, ainda existem poucos estudos que correlacionam a composição e a abundância de DNAs repetitivos com variáveis ambientais (Lyu *et al.*, 2018; Schley *et al.*, 2022).

2.5 A FAMÍLIA PODOCARPACEAE E A ESPÉCIE *Podocarpus sellowii*

As gimnospermas atuais representam linhagens de plantas relictuais com sementes que datam de aproximadamente 350 milhões de anos atrás (Murray, 2013; Biswas & Johri, 2013). Embora as gimnospermas formem um grupo com uma boa variação morfológica, ainda são consideradas estruturalmente conservadas ao longo do tempo em relação às angiospermas (Melo *et al.*, 2019). No período que compreende o Cretáceo médio e tardio, as gimnospermas foram quase eliminadas dos trópicos, sendo forçadas a ambientes marginais devido à competição com as angiospermas, que por sua vez, eram fisiologicamente mais inovadoras (Midgley & Bond, 1991; Feild & Arens, 2005; Berendse & Scheffer, 2009; Brodribb & Feild, 2010). Contudo, apesar da forte pressão competitiva exercida pelas angiospermas, a família Podocarpaceae conseguiu ser uma das poucas que persistem nas florestas tropicais em todo o mundo. A família Podocarpaceae é a segunda mais diversa entre as coníferas, abrangendo aproximadamente 200 espécies distribuídas em 17 a 19 gêneros (Migliore *et al.*, 2020; Quiroga *et al.*, 2016). Essas espécies estão presentes em alguns dos ecossistemas mais diversos e ameaçados do mundo e estão distribuídas principalmente no Hemisfério Sul, sendo quase que inteiramente restritas a florestas ou ambientes de altitude. O gênero mais representativo, *Podocarpus* L'Hér. ex Pers., originou-se no clima frio e úmido do sul de Gondwana, o que sugere que tenha vindo

da flora da Antártida (Migliore *et al.*, 2020). Nesse gênero, *Podocarpus sellowii* se destaca por possuir uma distribuição geográfica fragmentada em pequenas populações isoladas desde o Nordeste ao Sul do Brasil (Lima & Barbosa, 1998). Essa espécie já teve suas populações disjuntas do Nordeste brasileiro previamente analisadas por marcadores moleculares SSR e ISSR (*Simple Sequence Repeats* e *Inter Simple Sequence Repeats*, respectivamente). E apesar da baixa diversidade genética observada, houve uma elevada diferenciação entre as populações (Dantas *et al.*, 2015). Todavia, devido à limitação dos marcadores e à baixa quantidade de populações estudadas, não puderam ser realizadas interpretações macroevolutivas.

A maioria das espécies de *Podocarpus* s.s. está hoje presente em ambientes de florestas úmidas e amplamente dominados por Angiospermas (Melo *et al.*, 2019). O nome genérico *Podocarpus* deriva do grego *podos* (pé) e *karpós* (fruto), explicando-se pelo fato de a semente ser sustentada por um pedúnculo carnoso, conhecido como epimácio (Marchiori, 1995). *Podocarpus* L'Hér. ex Pers. é um dos principais gêneros da família Podocarpaceae, e, embora ainda seja considerado o maior gênero de coníferas tropicais no Hemisfério Sul, a classificação da família passou por revisões mais recentes que ampliam a compreensão sobre sua diversidade e distribuição (Khan *et al.*, 2023). Embora historicamente descrito como o único gênero tropical das Coniferales no Hemisfério Sul (Duarte, 1973), estudos mais recentes indicam que a família Podocarpaceae inclui mais de 20 gêneros e 200 espécies distribuídas também em regiões temperadas do Hemisfério Sul, com algumas ocorrências no Hemisfério Norte, como no México e no Caribe (Farjon, 2018). A diversidade de *Podocarpus* é maior em hotspots como Nova Zelândia, Nova Caledônia e Tasmânia, mas a espécie também ocorre em outras partes da Australásia, América do Sul e África do Sul (Khan *et al.*, 2023). No Brasil, existem três espécies descritas de *Podocarpus* (Guerra *et al.*, 2022): *Podocarpus sellowii*, amplamente distribuída em fragmentos da Mata Atlântica e do Cerrado; *Podocarpus lambertii* Klotzsch ex Endl., presente principalmente nas regiões Sul e Sudeste, associada à floresta de araucárias; e *Podocarpus transiens* (Pilg.) de Laub. ex Silba, que ocorre em locais específicos, mas com registros menos frequentes. Há ainda uma quarta espécie, *Podocarpus brasiliensis* de Laub., cuja distinção taxonômica é discutida, sendo possivelmente uma variação de *P. sellowii*, conforme sugerem estudos moleculares recentes (Bernardi *et al.*, 2024). *P. sellowii* apresenta grande variabilidade fenotípica quanto ao porte, tamanho das folhas e número de cones polínicos por axila foliar (Garcia, 2002). Esta espécie se assemelha a *P. lambertii*, mas distingue-se por ter folhas lanceoladas maiores e ausência de um pedúnculo comum aos grupos de estróbilos masculinos (Marchiori, 1996). O epíteto específico, *sellowii*, homenageia

Friedrich Sellow (1789 – 1831), botânico alemão que integrou a comitiva de naturalistas que veio ao Brasil acompanhando Dona Leopoldina, noiva de Dom Pedro I (Marchiori, 1995).

De maneira geral, *P. sellowii* é uma espécie perene e nativa do Brasil, apresentando uma ampla distribuição, que cresce em altitudes entre 800 e 1800 m, apesar de já ter sido documentada e coletada em locais menor altitude (Melo *et al.*, 2019) em locais relativamente secos com solos leves, onde o dossel da floresta é mais aberto e geralmente mais baixo que 20 m, muitas vezes não ultrapassando 10–12 m (Neale *et al.*, 2019). No entanto, suas populações ocorrem de forma descontínua (Lima e Barbosa 1998; Dantas *et al.*, 2015) e são mais adensadas em áreas elevadas na extensão da Floresta Atlântica e raramente encontrada em Floresta de Araucária mais ao sul (Turner & Cernusak, 2011). Ocorre também no Cerrado, em matas de galeria e em zonas de transição (Cerrado – Floresta Atlântica) (Gonçalves *et al.*, 2016). Não está claro como se deu este padrão específico de distribuição, porém é sabido que outras espécies do gênero *Podocarpus* na América do Sul também apresentam esse padrão disjunto, o que sugere associação a habitats especializados e distribuição em refúgios ecológicos (Dalling *et al.*, 2011).

P. sellowii ocorre naturalmente em solos de fertilidade química variável, geralmente pobres, rasos, bem drenados e com textura que varia de franca a argilosa. É uma espécie esciófila (adaptada a ambientes sombreados) a heliófila (adaptada a ambientes com alta incidência de luz), que não tolera baixas temperaturas e apresenta disposição simpodial (crescimento em que o caule principal é substituído por ramos laterais, o que resulta em uma ramificação complexa), uma característica incomum entre as coníferas (Valladares *et al.* 2016; Yang, 2024). *P. sellowii* está na lista de espécies raras ou ameaçadas de extinção no Distrito Federal (Filgueiras; Pereira, 1990). Leite *et al.*, (1986) mencionam sua ocorrência no sul de Mato Grosso do Sul, mas Souza *et al.*, (1997) não confirmaram a presença dessa espécie na área citada. Lima (1991) na década de 90 já fazia um apelo para garantir a preservação dessa gimnosperma, que já se encontrava em vias de extinção devido aos constantes desmatamentos registrados em sua única área de ocorrência em Pernambuco, o Brejo de Altitude de Serra dos Cavalos, na cidade de Caruaru. Atualmente, a espécie está listada como Em Perigo (EN) na Lista Vermelha da IUCN devido ao grande declínio de seu habitat no passado e ao pequeno tamanho atual de suas populações em áreas restritas (Farjon, 2018). Embora esteja bem documentada em várias regiões do sul e sudeste do Brasil, sua presença em áreas protegidas é limitada, dificultando a conservação. A perda de habitat causada pela expansão agrícola e outras atividades humanas continua sendo uma ameaça significativa (Farjon, 2018).

3 MATERIAIS E MÉTODOS

3.1 CARACTERIZAÇÃO E ANÁLISE *IN SILICO* DA FRAÇÃO REPETITIVA

O DNA genômico de 10 indivíduos representativos de populações de *Podocarpus sellowii* foi sequenciado utilizando a plataforma Illumina HiSeq2500 (2 × 150 bp). A análise comparativa da fração repetitiva do genoma foi realizada no pipeline Galaxy/RepeatExplorer2. O número de leituras utilizado na análise foi ajustado para garantir uma cobertura de 0,10 vezes para cada população, excluindo-se as sequências de cloroplasto e mitocôndria, por representarem possíveis contaminantes. Para a análise comparativa entre populações, as leituras foram filtradas por qualidade, com 95% das bases apresentando qualidade igual ou superior ao valor de corte de 10. A cobertura foi ajustada para ~0,04× para cada população, com agrupamento baseado em 90% de similaridade sobre um mínimo de 55% de sobreposição de sequência. Apenas *clusters* com abundância superior a 0,01% foram considerados (Novák *et al.*, 2020). As proporções individuais e comparativas dos clados de *repeats*, referindo-se ao total de unidades repetitivas classificadas em cada categoria, foram calculadas com base no número de leituras agrupadas em relação ao total de leituras utilizadas na análise, após a exclusão das leituras de cloroplasto e mitocôndria (Neumann *et al.*, 2019). A filogenia dos *repeats* foi construída utilizando dois métodos: uma análise de *Neighbor Joining* baseada na abundância dos *repeats* (Dodsworth *et al.*, 2015a) e uma matriz de similaridade gerada a partir dos *clusters* mais abundantes, representando mais de 0,01% do genoma, utilizando o *software* Past 4.16 (Hammet; Harper, 2001). Adicionalmente, o *pipeline* TAREAN (*Tandem Repeat Analyser*), também implementado no Galaxy/RepeatExplorer2, foi empregado para identificar repetições em *tandem* (Novák *et al.*, 2017).

3.2 FILOGENIA MOLECULAR

A filogenia molecular foi conduzida com base em sequências de plastomas completos e DNAr nuclear para investigar a evolução do repeatoma no contexto filogeográfico de *Podocarpus sellowii*. Os plastomas foram previamente montados conforme descrito por Nieto-Blázquez *et al.* (2020), enquanto as sequências de DNAr seguiram as diretrizes de Melo *et al.* (2019). Ambas as sequências foram alinhadas utilizando o algoritmo MAFFT, assegurando alta qualidade nos alinhamentos. Para inferir as topologias, aplicou-se o método de *Network* por meio do *software* Splitstree4 (Huson;

Bryant, 2005). A análise de *Network* foi baseada na máxima parcimônia, aplicada tanto aos plastomas quanto ao DNAr. O método de inferência filogenética baseado na abundância de *repeats* (Dodsworth, 2015) foi utilizado para analisar divergências genéticas e delimitar clados taxonômicos em 10 populações de *P. sellowii*.

3.3 FILOGENIA DE *REPEATS* E SINAL FILOGENÉTICO

A filogenia dos *repeats* nos genomas das populações de *Podocarpus sellowii* foi realizada para complementar as topologias baseadas em plastomas e DNAr. Utilizou-se o método de Máxima Verossimilhança (ML) com 1000 replicações no *software* Geneious (v. 9.1.8) e o *plugin* FastTree (Price *et al.*, 2009) para reconstruir as relações filogenéticas. Para avaliar possíveis incongruências entre as topologias geradas a partir de diferentes marcadores, aplicou-se o pacote Cophylo no *software* R (Revell, 2012), permitindo a comparação simultânea das árvores filogenéticas. A rede filogenética foi construída com base em dados combinados de plastomas e DNAr nuclear utilizando o programa SplitsTree4 (Huson; Bryant, 2006), conforme a metodologia de Holland; Moulton (2003). Adicionalmente, foi realizada uma análise do sinal filogenético dos *repeats*, correlacionando a abundância desses elementos com preditores ecológicos extraídos de variáveis ambientais do WorldClim. Para identificar as variáveis que melhor explicam as abundâncias dos *repeats*, foram empregados modelos lineares generalizados com regressão quase-binomial, utilizando o pacote glmulti (Calcagno, 2010) no *software* R (Bartoń, 2023). A seleção dos modelos baseou-se no critério de informação de Akaike (AIC) (Akaike, 1973; Hurvich; Tsai, 1989), selecionando os modelos com menores valores de AIC como mais robustos para representar a relação entre variáveis ecológicas e a distribuição dos repetitivos.

3.4 DISTRIBUIÇÃO GEOGRÁFICA E SELEÇÃO DE VARIÁVEIS ECOLÓGICAS

Para investigar a influência de variáveis ecológicas não colineares sobre os diferentes clados repetitivos nos genomas de *Podocarpus sellowii*, foram selecionadas 10 populações naturais, representando diferentes regiões geográficas, localizadas e codificadas da seguinte maneira: Baturité - CE (BTT), Caruaru - PE (BJC), Itabaiana - SE (ITB), Comodoro - MT (COM), Alto Paraíso de Goiás - GO (CPV), Una - BA (BAH), Brasília - DF (BRA), Carrancas - MG (CAR), Paranapiacaba - SP (PAR) e Santo Amaro - SC (STA). As coordenadas geográficas dessas populações foram registradas e utilizadas para

obtenção de dados climáticos. Um único indivíduo por população foi selecionado para caracterizar a fração repetitiva do genoma. As variáveis ecológicas foram obtidas da plataforma WorldClim 2.1 (Fick; Hijmans, 2017), e os dados geoclimáticos processados no *software* QGIS (QGIS - Development Team, 2023). Foram extraídas 19 variáveis bioclimáticas (Hijmans *et al.*, 2005), e realizada uma análise de colinearidade entre elas. Variáveis com alta correlação (coeficiente de Pearson $> 0,7$) ou baixa contribuição nas primeiras cinco componentes principais de uma análise PCA foram excluídas. O conjunto final de variáveis incluiu cinco bioclimáticas: BIO2 (variação da temperatura diurna), BIO5 (temperatura máxima do mês mais quente), BIO6 (temperatura mínima do mês mais frio), BIO13 (precipitação do mês mais úmido) e BIO14 (precipitação do mês mais seco). Os dados das variáveis selecionadas foram organizados em tabelas utilizando o *software* Microsoft Excel 2021 (Microsoft Corporation, 2021) para facilitar a visualização e o tratamento dos dados. Foi gerado um gráfico bidimensional para representar a separação geoclimática das populações com base nas cinco variáveis selecionadas, visando identificar padrões de distribuição das linhagens repetitivas nos gradientes ambientais.

3.5 INTERAÇÃO GENOMA-AMBIENTE

Para investigar a relação entre as abundâncias dos principais repeats do genoma e as variáveis ambientais, foram realizadas análises comparativas e regressões. As variáveis climáticas obtidas pela plataforma WorldClim 2.1 (Fick; Hijmans, 2017) foram usadas como preditoras. A correlação entre as abundâncias de linhagens repetitivas e as variáveis climáticas foi avaliada por correlação de Pearson, excluindo variáveis com coeficiente superior a 0,7. Um modelo Species Packing com distribuição Gaussiana, implementado no *software* PAST 4.16 (Hammer *et al.*, 2001), foi utilizado para modelar a resposta das abundâncias dos repeats ao longo dos gradientes ecológicos. Além disso, uma regressão linear simples foi aplicada para avaliar a relação entre a abundância de repeats e as variáveis bioclimáticas selecionadas (BIO2, BIO5, BIO6, BIO13 e BIO14). Por fim, uma análise de Mínimos Quadrados Generalizados Filogenéticos (PGLS - *Phylogenetic Generalized Least Squares*) foi aplicada para examinar a influência das variáveis geoclimáticas sobre a diversidade genômica plastidial das populações de *P. sellowii*, utilizando o *software* R Studio e o pacote Caper (Orme *et al.*, 2013). A matriz filogenética foi derivada dos dados plastidiais, assumindo um modelo Browniano para a evolução dos caracteres, sendo o modelo com menor valor de AIC selecionado para identificar as associações mais adequadas entre o genoma plastidial e os fatores geoclimáticos.

REFERÊNCIAS

AKAIKE, Htrotugu. Maximum likelihood identification of Gaussian autoregressive moving average models. **Biometrika**, v. 60, n. 2, p. 255-265, 1973.

ÁLVAREZ-YÉPIZ, J. C.; DOVČIAK, M.; BÚRQUEZ, A. Persistence of a rare ancient cycad: Effects of environment and demography. *Biological Conservation*, v. 144, n. 1, p. 122–130, 1 jan. 2011.

BARTON, Kamil. MuMIn: multi-model inference. R package version 1.43. 17. 2020.

BENNETZEN, Jeffrey L.; WANG, Hao. 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, 2014.

BENNETZEN, Jeffrey L.; WANG, Hao. 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, 2014.

BERENDSE, Frank; SCHEFFER, Marten. The angiosperm radiation revisited, an ecological explanation for Darwin's 'abominable mystery'. **Ecology letters**, v. 12, n. 9, p. 865-872, 2009.

BERNARDI, Alison Paulo *et al.* Genetic diversity and fine-scale genetic structure of the conifer in a forest-grassland mosaic from Southern Brazil. **Silvae Genetica**, v. 73, n. 1, p. 60-69.

BISCOTTI, Maria Assunta; OLMO, Ettore; HESLOP-HARRISON, J. S. Repetitive DNA in eukaryotic genomes. **Chromosome Research**, v. 23, p. 415-420, 2015.

BISWAS, Chhaya; JOHRI, Brij Mohan. **The gymnosperms**. Springer Science & Business Media, 2013.

BOGAN, Samuel N.; YI, Soojin V. Potential role of DNA methylation as a driver of plastic responses to the environment across cells, organisms, and populations. **Genome biology and evolution**, v. 16, n. 2, p. evae022, 2024.

BORGES, Rui *et al.* Measuring phylogenetic signal between categorical traits and phylogenies. **Bioinformatics**, v. 35, n. 11, p. 1862-1869, 2019.

BOURGEOIS, Yann *et al.* Disentangling the determinants of transposable elements dynamics in vertebrate genomes using empirical evidences and simulations. **PLoS genetics**, v. 16, n. 10, p. e1009082, 2020.

BOURQUE, Guillaume *et al.* Ten things you should know about transposable elements. **Genome biology**, v. 19, p. 1-12, 2018.

BRODRIBB, Tim J.; FEILD, Taylor S.; SACK, Lawren. Viewing leaf structure and evolution from a hydraulic perspective. **Functional Plant Biology**, v. 37, n. 6, p. 488-498, 2010.

BUERMANS, H. P. J.; DEN DUNNEN, J. T. Next generation sequencing technology: advances and applications. **Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease**, v. 1842, n. 10, p. 1932-1941, 2014.

BUREŠ, Petr *et al.* The global distribution of angiosperm genome size is shaped by climate. **New Phytologist**, v. 242, n. 2, p. 744-759, 2024.

CAIN, Amy K. *et al.* A decade of advances in transposon-insertion sequencing. **Nature Reviews Genetics**, v. 21, n. 9, p. 526-540, 2020.

CALCAGNO, Vincent; DE MAZANCOURT, Claire. glmulti: an R package for easy automated model selection with (generalized) linear models. **Journal of statistical software**, v. 34, p. 1-29, 2010.

CANAPA, A. *et al.* Shedding light upon the complex net of genome size, genome composition and environment in chordates. **The European Zoological Journal**, v. 87, n. 1, p. 192-202, 2020.

CARDOSO, Thiago Kossmann *et al.* Diversity and conservation of macrofungi from the Aparados da Serra with an emphasis on forest ecosystems. 2022.

CASTRO, Natália *et al.* Repeatoma evolution across space and time: Unravelling *repeats* dynamics in the plant genus *Erythrostemon* Klotzsch (Leguminosae Juss). **Molecular Ecology**, p. e17510, 2024.

CHÉNAIS, Benoît *et al.* The impact of transposable elements on eukaryotic genomes: from genome size increase to genetic adaptation to stressful environments. **Gene**, v. 509, n. 1, p. 7-15, 2012.

CHÉNAIS, Benoît *et al.* The impact of transposable elements on eukaryotic genomes: from genome size increase to genetic adaptation to stressful environments. **Gene**, v. 509, n. 1, p. 7-15, 2012.

CHUMOVÁ, Zuzana *et al.* The relationship between transposable elements and ecological niches in the Greater Cape Floristic Region: A study on the genus *Pteronia* (Asteraceae). **Frontiers in Plant Science**, v. 13, p. 982852, 2022.

CNCFLORA. *Podocarpus sellowii*. Rio de Janeiro: Centro Nacional de Conservação da Flora, 2024. Disponível em: <https://cncflora.jbrj.gov.br/>. Acesso em: 4 set. 2024.

DANTAS, Liliane G. *et al.* Low genetic diversity and high differentiation among relict populations of the neotropical gymnosperm *Podocarpus sellowii* (Klotz.) in the Atlantic Forest. **Genetica**, v. 143, p. 21-30, 2015.

DODSWORTH, Steven *et al.* Genomic *repeat* abundances contain phylogenetic signal. **Systematic biology**, v. 64, n. 1, p. 112-126, 2015.

DODSWORTH, Steven *et al.* Hyb-Seq for flowering plant systematics. **Trends in Plant Science**, v. 24, n. 10, p. 887-891, 2019.

DODSWORTH, Steven. Genome skimming for next-generation biodiversity analysis. **Trends in Plant Science**, v. 20, n. 9, p. 525-527, 2015.

FARJON, A. 2013. *Podocarpus sellowii*. The IUCN Red List of Threatened Species 2013: e.34099A2845580. <http://dx.doi.org/10.2305/IUCN.UK.20131.RLTS.T34099A2845580.en>. Downloaded on 09 August 2024.

FARJON, Aljos. The Kew review: conifers of the world. **Kew Bulletin**, v. 73, p. 1-16, 2018.

FARJON, Aljos. **World checklist and bibliography of conifers**. 2017.

FEILD, Taylor S.; ARENS, Nan Crystal. Form, function and environments of the early angiosperms: merging extant phylogeny and ecophysiology with fossils. **New Phytologist**, v. 166, n. 2, p. 383-408, 2005.

FICK, Stephen E.; HIJMANS, Robert J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. **International journal of climatology**, v. 37, n. 12, p. 4302-4315, 2017.

GARAMSZEGI, László Zsolt (Ed.). **Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice**. Springer, 2014.

GARCIA, Lucinda Carneiro. Aspectos morfo-anatômicos e tolerância à dessecação de sementes de *Podocarpus lambertii* Klotz. e *Podocarpus sellowii* Klotz. (Podocarpaceae). 2003.

GARCIA, Lucinda Carneiro. Aspectos morfo-anatômicos e tolerância à dessecação de sementes de *Podocarpus lambertii* Klotz. e *Podocarpus sellowii* Klotz. (Podocarpaceae). 2003.

GAZDA, Małgorzata A. *et al.* A genetic mechanism for sexual dichromatism in birds. **Science**, v. 368, n. 6496, p. 1270-1274, 2020

GONÇALVES, F. R.; VIEIRA, F. A.; CARVALHO, D. Naturally fragmented but not genetically isolated populations of *Podocarpus sellowii* Klotzsch (Podocarpaceae) in southeast Brazil. **Genetics and Molecular Research**, v. 15, n. 4, p. 1-17, 2016.

GONZÁLEZ, Graciela Esther; POGGIO, Lidia. Intragenomic conflict between knob heterochromatin and B chromosomes is the key to understand genome size variation along altitudinal clines in maize. **Plants**, v. 10, n. 9, p. 1859, 2021.

HABERER, Georg *et al.* European maize genomes highlight intraspecies variation in repeat and gene content. **Nature Genetics**, v. 52, n. 9, p. 950-957, 2020.

HAMMER, Øyvind. PAST: Paleontological statistics software package for education and data analysis. **Palaeontol electron**, v. 4, p. 9, 2001.

HAMMER, Øyvind; HARPER, David AT. Past: paleontological statistics software package for education and data analysis. **Palaeontologia electronica**, v. 4, n. 1, p. 1, 2001.

HARTLEY, Gabrielle; O'NEILL, Rachel J. Centromere *repeats*: hidden gems of the genome. **Genes**, v. 10, n. 3, p. 223, 2019.

HÉNAULT, Mathieu *et al.* The genomic landscape of transposable elements in yeast hybrids is shaped by structural variation and genotype-specific modulation of transposition rate. **Elife**, v. 12, p. RP89277, 2024.

HIJMANS, Robert J. *et al.* Very high resolution interpolated climate surfaces for global land areas. **International Journal of Climatology: A Journal of the Royal Meteorological Society**, v. 25, n. 15, p. 1965-1978, 2005.

HILL, Robert S.; BRODRIBB, Tim J. Southern conifers in time and space. **Australian Journal of Botany**, v. 47, n. 5, p. 639-696, 1999.

HOLLAND, Barbara; MOULTON, Vincent. Consensus networks: a method for visualising incompatibilities in collections of trees. In: **International workshop on algorithms in bioinformatics**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003. p. 165-176.

HORVÁTH, Vivien; MERENCIANO, Miriam; GONZÁLEZ, Josefa. Revisiting the relationship between transposable elements and the eukaryotic stress response. **Trends in Genetics**, v. 33, n. 11, p. 832-841, 2017.

HOUSWORTH, Elizabeth A.; MARTINS, Emilia P.; LYNCH, Michael. The phylogenetic mixed model. **The American Naturalist**, v. 163, n. 1, p. 84-96, 2004.

HURVICH, Clifford M.; TSAI, Chih-Ling. Regression and time series model selection in small samples. **Biometrika**, v. 76, n. 2, p. 297-307, 1989.

HUSON, Daniel H.; BRYANT, David. Application of phylogenetic networks in evolutionary studies. **Molecular biology and evolution**, v. 23, n. 2, p. 254-267, 2006.

JOHNSON, Norman A. Speciation: genomic sequence data and the biogeography of speciation. *National Science Review*, v. 9, n. 12, p. nwac294, 2022.

KAMILAR, Jason M.; COOPER, Natalie. Phylogenetic signal in primate behaviour, ecology and life history. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 368, n. 1618, p. 20120341, 2013.

KAZAZIAN JR, Haig H. Mobile elements: drivers of genome evolution. **science**, v. 303, n. 5664, p. 1626-1632, 2004.

KHAN, Raees *et al.* Diversity, distribution, systematics and conservation status of Podocarpaceae. **Plants**, v. 12, n. 5, p. 1171, 2023.

KHAN, Raees *et al.* Diversity, distribution, systematics and conservation status of Podocarpaceae. **Plants**, v. 12, n. 5, p. 1171, 2023.

LIMA, T. V.; BARBOSA, D. C. A. Levantamento da distribuição espacial de plantas jovens não estabelecidas e estabelecidas de *Podocarpus sellowii* Klotzsch, em Serra dos Cavalos, Caruaru—PE. **Trabalho técnico, Imprensa Universitária, UFRPE, Recife**, v. 32, 1998.

LOSOS, Jonathan B. **Lizards in an evolutionary tree: ecology and adaptive radiation of anoles**. Univ of California Press, 2011.

LYNCH, Michael; WALSH, Bruce. **The origins of genome architecture**. Sunderland, MA: Sinauer associates, 2007.

MACAS, Jiří *et al.* In depth characterization of repetitive DNA in 23 plant genomes reveals sources of genome size variation in the legume tribe Fabeae. **PloS one**, v. 10, n. 11, p. e0143424, 2015.

MARCHIORI, José Newton Cardoso. Anatomia do xilema secundário de *Mimosa incana* (Spreng.) Benth. **Ciência Florestal**, v. 6, n. 1, p. 53-63, 1996.

MELO, Tiago Esposito Couceiro. Origem das disjunções sul-americanas de *Podocarpus sellowii* Klotzsch ex Endl (Podocarpaceae): filogeografia, modelagem de nicho e filogenômica. 2019. Tese (Doutorado em Biologia Vegetal) – Universidade Federal de Pernambuco, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; Orientador: Andrea Pedrosa Harand.

MICROSOFT CORPORATION. Microsoft Excel (Version 16.0). Redmond: Microsoft, 2021.

MIDGLEY, Jeremy J.; BOND, William J. Ecological aspects of the rise of angiosperms: a challenge to the reproductive superiority hypotheses. **Biological Journal of the Linnean Society**, v. 44, n. 2, p. 81-92, 1991.

MIGLIORE, Jérémy; LÉZINE, Anne-Marie; HARDY, Olivier J. The recent colonization history of the most widespread *Podocarpus* tree species in Afromontane forests. **Annals of Botany**, v. 126, n. 1, p. 73-83, 2020.

MILL, Robert R. Towards a biogeography of the Podocarpaceae. In: **IV International Conifer Conference 615**. 1999. p. 137-147.

MIOUSSE, Isabelle R. *et al.* Response of transposable elements to environmental stressors. **Mutation Research/Reviews in Mutation Research**, v. 765, p. 19-39, 2015.

MITCHELL, Sandra D. Perspectives, representation, and integration. **Understanding perspectivism: Scientific challenges and methodological prospects**, p. 178-193, 2020.

MITTEROECKER, Philipp; COLLYER, Michael L.; ADAMS, Dean C. Exploring Phylogenetic Signal in Multivariate Phenotypes by Maximizing Blomberg's K. *Systematic Biology*, p. syae035, 2024.

MORALES-CASTILLA, Ignacio *et al.* Inferring biotic interactions from proxies. **Trends in ecology & evolution**, v. 30, n. 6, p. 347-356, 2015.

MÜNKEMÜLLER, Tamara *et al.* How to measure and test phylogenetic signal. **Methods in Ecology and Evolution**, v. 3, n. 4, p. 743-756, 2012.

MURRAY, Brian G. Karyotype variation and evolution in gymnosperms. In: **Plant Genome Diversity Volume 2: Physical Structure, Behaviour and Evolution of Plant Genomes**. Vienna: Springer Vienna, 2012. p. 231-243.

NEALE, David B. *et al.* Noncoding and repetitive DNA. *The Conifers: Genomes, Variation And Evolution*, p. 61-74, 2019.

NEALE, David B. *et al.* **The conifers**. Springer International Publishing, 2019.

NEGI, Pooja; RAI, Archana N.; SUPRASANNA, Penna. Moving through the stressed genome: emerging regulatory roles for transposons in plant stress response. **Frontiers in plant science**, v. 7, p. 1448, 2016.

NIETO-BLÁZQUEZ, María Esther; PEÑA-CASTILLO, Lourdes; RONCAL, Julissa. Historical biogeography of Caribbean *Podocarpus* does not support the progression rule. **Journal of Biogeography**, v. 48, n. 3, p. 690-702, 2021.

NOVÁK, Petr *et al.* RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. **Bioinformatics**, v. 29, n. 6, p. 792-793, 2013.

NOVÁK, Petr *et al.* TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. **Nucleic acids research**, v. 45, n. 12, p. e111-e111, 2017.

NOVÁK, Petr; NEUMANN, Pavel; MACAS, Jiří. Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. **Nature Protocols**, v. 15, n. 11, p. 3745-3776, 2020.

ORME, David *et al.* The caper package: comparative analysis of phylogenetics and evolution in R. **R package version**, v. 5, n. 2, p. 1-36, 2013.

PAÇO, Ana; FREITAS, Renata; VIEIRA-DA-SILVA, Ana. Conversion of DNA sequences: from a transposable element to a tandem *repeat* or to a gene. **Genes**, v. 10, n. 12, p. 1014, 2019.

PAGE, C. N. New and maintained genera in the conifer families Podocarpaceae and Pinaceae. **Notes RBG Edinb**, 1988.

PAGE, Christopher N. New and maintained genera in the taxonomic alliance of Prumnopitys s. l. (Podocarpaceae), and circumscription of a new genus: Pectinopitys. **New Zealand Journal of Botany**, v. 57, n. 3, p. 137-153, 2019.

PAGE, Roderic DM. New Zealand and the new biogeography. **New Zealand Journal of Zoology**, v. 16, n. 4, p. 471-483, 1990.

PARADIS, Emmanuel. **Analysis of Phylogenetics and Evolution with R**. New York: Springer, 2012.

PAVOINE, Sandrine; RICOTTA, Carlo. Functional and phylogenetic similarity among communities. **Methods in Ecology and Evolution**, v. 5, n. 7, p. 666-675, 2014.

PIMPINELLI, Sergio; PIACENTINI, Lucia. Environmental change and the evolution of genomes: Transposable elements as translators of phenotypic plasticity into genotypic variability. **Functional Ecology**, v. 34, n. 2, p. 428-441, 2020.

PRICE, Morgan N.; DEHAL, Paramvir S.; ARKIN, Adam P. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. **Molecular biology and evolution**, v. 26, n. 7, p. 1641-1650, 2009.

QGIS DEVELOPMENT TEAM. QGIS Geographic Information System (Version 3.28). Open Source Geospatial Foundation Project, 2023. Disponível em: <https://qgis.org/>.

QUIROGA, María Paula *et al.* Molecular and fossil evidence disentangle the biogeographical history of *Podocarpus*, a key genus in plant geography. **Journal of Biogeography**, v. 43, n. 2, p. 372-383, 2016.

REINAR, William B. *et al.* Teleost genomic *repeat* landscapes in light of diversification rates and ecology. **Mobile DNA**, v. 14, n. 1, p. 14, 2023.

REVELL, Liam J. phytools: an R package for phylogenetic comparative biology (and other things). **Methods in ecology and evolution**, n. 2, p. 217-223, 2012.

REVELL, Liam J. Two new graphical methods for mapping trait evolution on phylogenies. **Methods in Ecology and Evolution**, v. 4, n. 8, p. 754-759, 2013.

SCHLEY, Rowan J. *et al.* The ecology of palm genomes: *repeat*-associated genome size expansion is constrained by aridity. **New Phytologist**, v. 236, n. 2, p. 433-446, 2022.

SMITH, Stephen A.; O'MEARA, Brian C. treePL: divergence time estimation using penalized likelihood for large phylogenies. **Bioinformatics**, v. 28, n. 20, p. 2689-2690, 2012.

SYMONDS, Matthew RE; BLOMBERG, Simon P. A primer on phylogenetic generalised least squares. **Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice**, p. 105-130, 2014.

SZITENBERG, Amir *et al.* Genetic drift, not life history or RNAi, determine long-term evolution of transposable elements. **Genome biology and evolution**, v. 8, n. 9, p. 2964-2978, 2016.

TURNER, Benjamin L.; CERNUSAK, Lucas A. Ecology of the Podocarpaceae in tropical forests. **Smithsonian Contributions to Botany**, 2011.

UYEDA, Josef C.; ZENIL-FERGUSON, Rosana; PENNELL, Matthew W. Rethinking phylogenetic comparative methods. **Systematic Biology**, v. 67, n. 6, p. 1091-1109, 2018.

VALLADARES, Fernando *et al.* Shedding light on shade: ecological perspectives of understory plant life. **Plant Ecology & Diversity**, v. 9, n. 3, p. 237-251, 2016.

VAN-LUME, Brena *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, 2017.

WANG, Jie *et al.* The pan-plastome of *Prunus mume*: insights into *Prunus* diversity, phylogeny, and domestication history. **Frontiers in Plant Science**, v. 15, p. 1404071, 2024.

WEISS-SCHNEEWEISS, Hanna *et al.* Employing next generation sequencing to explore the *repeat* landscape of the plant genome. **Next generation sequencing in plant systematics. Regnum Vegetabile**, v. 157, p. 155-179, 2015.

WELLS, Jonathan N.; FESCHOTTE, Cédric. A field guide to eukaryotic transposable elements. **Annual review of genetics**, v. 54, n. 1, p. 539-561, 2020.

YANG, Yong. Systematic classification of gymnosperms: past, present and future. **JOURNAL OF NANJING FORESTRY UNIVERSITY**, v. 48, n. 3, p. 14, 2024.

4 RESULTADOS

4.1 ARTIGO:

Análise Genômica Comparativa da Gimnosperma Relictual *Podocarpus sellowii* Klotzsch Ex Endl. (Podocarpaceae) em Populações Disjuntas ao Longo do Brasil

Augusto César Muniz, Natália Castro, Tiago Esposito Meto, Andrea Pedrosa-Harand, Gustavo Souza

Resumo

A fração repetitiva dos genomas (repeatoma) é a mais abundante nos genomas eucarióticos, principalmente devido à sua capacidade de gerar cópias dos seus elementos, contribuindo significativamente para a variação genômica entre populações. A rápida evolução do repeatoma pode ser influenciada por pressões ambientais, sugerindo que sua composição e abundância refletem adaptações a condições ecológicas específicas. Isso é especialmente importante para espécies com distribuição fragmentada e habitats especializados, como *Podocarpus sellowii* Klotzsch ex Endl, uma das poucas gimnospermas nativas do Brasil. Sua distribuição em pequenas populações isoladas na Mata Atlântica e Cerrado reflete um padrão relictual causado por vicariância nos últimos 20 milhões de anos. Neste estudo, analisamos comparativamente o genoma de 10 indivíduos de populações disjuntas de *P. sellowii* ao longo de um gradiente latitudinal, do Ceará a Santa Catarina. Para isso, caracterizamos os *repeats* mais abundantes das populações no *pipeline RepeatExplorer2*, usando *reads* de baixa cobertura (Illumina HiSeq2500). Coordenadas geográficas foram usadas para extrair variáveis climáticas para cada população através da plataforma WorldClim. Os perfis de repeatoma foram interpretados considerando relações filogenéticas baseadas em plastomas inteiros montados para este estudo. Utilizando PGLS (do inglês, *Phylogenetic Generalized Least Squares*), foi possível calcular a correlação entre *repeats* e variáveis ecológicas. Em geral, os repeatomas de *P. sellowii* apresentaram alta abundância de retrotransposons LTR (do inglês, *Long Terminal Repeats*), especialmente Ty1/copia-Ivana (29,65% a 35,13%), Ty3/gypsy-Ogre (10,88% a 11,29%) e Ty1/copia-SIRE (8,06% a 10,46%). Como as abundâncias de cada *repeat* foram polimórficas entre nossas amostras de cada população, testamos a influência da ecologia e/ou filogenia sobre isso. Identificamos uma correlação significativa entre vários LTRs e transposons de DNA com características de temperatura e precipitação. A árvore filogenética baseada em *repeats* revelou grupos ecológicos, separando indivíduos da Mata

Atlântica dos do Cerrado, em vez de relações genealógicas. Nossos dados mostram que cerca de 40% da fração genômica repetitiva é diretamente influenciada por variáveis ecológicas. Isso destaca interações significativas entre variáveis ambientais e a abundância de *repeats*, sugerindo que a especificidade ecológica dessas plantas relictuais impactou a evolução de seus genomas.

Palavras-chave: Diversidade Genética; Análise Genômica; Repeatoma; Filogenômica; Estrutura Populacional.

Abstract

The repetitive fraction of genomes (the repeatome) is the most abundant in eukaryotic genomes, primarily due to its ability to generate copies of its elements, significantly contributing to genomic variation among populations. The rapid evolution of the repeatome can be influenced by environmental pressures, suggesting that its composition and abundance reflect adaptations to specific ecological conditions. This is particularly important for species with fragmented distributions and specialized habitats, such as *Podocarpus sellowii* Klotzsch ex Endl, one of the few native gymnosperms of Brazil. Its distribution in small, isolated populations in the Atlantic Forest and Cerrado reflects a relictual pattern caused by vicariance over the past 20 million years. In this study, we conducted a comparative genome analysis of 10 individuals from disjunct populations of *P. sellowii* along a latitudinal gradient, from Ceará to Santa Catarina. To achieve this, we characterized the most abundant repeats in the populations using the RepeatExplorer2 pipeline, based on low-coverage reads (Illumina HiSeq2500). Geographic coordinates were used to extract climate variables for each population via the WorldClim platform. The repeatome profiles were interpreted considering phylogenetic relationships based on whole plastomes assembled for this study. Using Phylogenetic Generalized Least Squares (PGLS), we were able to calculate the correlation between repeats and ecological variables. Overall, the repeatomes of *P. sellowii* showed a high abundance of LTR retrotransposons (Long Terminal Repeats), particularly Ty1/copia-Ivana (29.65% to 35.13%), Ty3/gypsy-Ogre (10.88% to 11.29%), and Ty1/copia-SIRE (8.06% to 10.46%). Since the abundance of each repeat was polymorphic across our samples from each population, we tested the influence of ecology and/or phylogeny on this variation. We identified a significant correlation between several LTRs and DNA transposons with temperature and precipitation characteristics. The phylogenetic tree based on repeats revealed ecological groups, separating individuals from

the Atlantic Forest from those of the Cerrado, rather than genealogical relationships. Our data show that approximately 40% of the repetitive genomic fraction is directly influenced by ecological variables. This highlights significant interactions between environmental variables and repeat abundance, suggesting that the ecological specificity of these relict plants has impacted the evolution of their genomes.

Keywords: Genetic Diversity; Genomic Analysis; Repeatome; Phylogenomics; Population Structure.

Introdução

Sequências repetitivas de DNA, ou *repeats*, representam a maior fração do genoma de eucariotos e desempenham papéis fundamentais na evolução e estruturação genômica (Šatović-Vukšić; Plohl, 2023). Esses elementos podem ser classificados em dois grupos principais: DNA repetitivo disperso, também conhecido como elementos transponíveis (TEs), e sequências em tandem, como DNAsat, teloméricos e centroméricos (Neumann *et al.*, 2019; Šatović-Vukšić; Plohl, 2023). Os elementos transponíveis (do inglês, Transposable Elements - TEs), por exemplo, são conhecidos por sua mobilidade dentro do genoma, o que pode gerar mutações e rearranjos significativos (Li *et al.*, 2017), enquanto as sequências em tandem, organizadas em blocos adjacentes, desempenham funções estruturais essenciais (Hartley; O'Neill, 2019). A abundância e a similaridade dessas sequências repetitivas têm sido amplamente utilizadas para formular hipóteses filogenéticas e compreender padrões evolutivos (Dodsworth, 2015a).

Em determinadas espécies vegetais, esses elementos repetitivos podem constituir até 80% do genoma total, como é o caso de *Zea mays* L., sendo os principais responsáveis pela variação no tamanho do genoma (GS, em inglês *Genome Size*) entre diferentes espécies (Haberer *et al.*, 2020; González; Poggio, 2021). Esses elementos desempenham papéis relevantes em processos adaptativos, especialmente em resposta a estresses ambientais, influenciando a variabilidade genética adaptativa (Horváth; Merenciano; González, 2017; Chumová *et al.*, 2022). Assim, a relação entre variáveis ecológicas, como temperatura, precipitação e tipo de solo, e a abundância de *repeats*, em diferentes espécies, tem se tornado um foco crescente em estudos de ecologia evolutiva (Castro *et al.*, 2024; Chumová *et al.*, 2022; Moraes *et al.*, 2022).

O gênero *Podocarpus* Endl. (Podocarpaceae), é um grupo vegetal que oferece um modelo intrigante para esse tipo de estudo. Entre as coníferas, *Podocarpus* se destaca por

sua ampla distribuição no Hemisfério Sul e por ser uma das poucas que persistem em florestas tropicais (Brodrigg; Hill, 1999). A espécie *Podocarpus sellowii* Klotzsch Ex Endl. possui uma distribuição geográfica fragmentada, com populações disjuntas localizadas no Sul, Sudeste, Centro-Oeste e Nordeste do Brasil (Lima; Barbosa, 1998; CNCFlora, 2024). A adaptação dessa espécie a diferentes ambientes faz de *P. sellowii* um excelente modelo para estudos sobre a interação genoma-ambiente.

Estudos filogeográficos anteriores com *P. sellowii* revelaram padrões de subdivisão populacional, baseados em análises de regiões nucleares e plastidiais (Melo *et al.*, 2019). No entanto, esses estudos não exploraram de forma abrangente a influência dos *repeats* na evolução da espécie. Os marcadores moleculares utilizados apresentaram limitações, e o número restrito de populações amostradas comprometeu conclusões mais robustas (Melo *et al.*, 2019). Embora as novas tecnologias de sequenciamento tenham permitido um maior detalhamento das frações repetitivas do genoma, ainda é necessário avançar na compreensão da relação entre a abundância de *repeats* e as variáveis ecológicas que influenciam a evolução de populações fragmentadas.

Diante desse contexto, o presente estudo busca utilizar a fração repetitiva dos genomas (repeatoma) de indivíduos de populações naturais de *Podocarpus sellowii* como fonte de informação para entender a história demográfica e interações ecológicas da espécie ao longo de sua distribuição no Brasil. A investigação dessas interações em espécies relictuais como *P. sellowii* contribui para o entendimento dos mecanismos de adaptação e diversificação genômica em plantas de florestas tropicais. Além disso, essa compreensão é essencial para a conservação de populações geneticamente isoladas e adaptadas a diferentes condições ecológicas. Assim, o presente trabalho visa aprofundar o conhecimento da história demográfica de *P. sellowii* na região Neotropical, avaliando suas relações filogenéticas e o impacto dos eventos climáticos na evolução de sua fração repetitiva, utilizando métodos filogenéticos comparativos para entender melhor as interações genoma-ambiente.

Metodologia

Caracterização e Análise *in silico* da Fração Repetitiva

O DNA genômico de 10 indivíduos representativos de populações de *Podocarpus sellowii*, distribuídos em diferentes regiões geográficas do Brasil, localizadas e codificadas da seguinte maneira: Baturité - CE (BTT), Caruaru - PE (BJC), Itabaiana - SE (ITB),

Comodoro - MT (COM), Alto Paraíso de Goiás - GO (CPV), Una - BA (BAH), Brasília - DF (BRA), Carrancas - MG (CAR), Paranapiacaba - SP (PAR) e Santo Amaro - SC (STA), foi sequenciado utilizando a plataforma Illumina HiSeq2500 (2 × 150 bp). A análise comparativa da fração repetitiva do genoma foi realizada com a ferramenta Galaxy/RepeatExplorer2, que executa um agrupamento baseado em gráficos para identificar os *repeats* mais abundantes, agrupando-os conforme a similaridade e gerando clusters para diferentes famílias de DNA repetitivo (Novák *et al.*, 2013). O número de leituras utilizado na análise foi ajustado para garantir uma cobertura de 0,10 vezes para cada população, excluindo-se as sequências de cloroplasto e mitocôndria por representarem possíveis contaminantes. Para a análise comparativa entre populações, as leituras foram filtradas por qualidade, com 95% das bases apresentando qualidade igual ou superior ao valor de corte de 10. A cobertura foi ajustada para ~0,04× para cada população, com agrupamento baseado em 90% de similaridade sobre um mínimo de 55% de sobreposição de sequência. Apenas clusters com abundância superior a 0,01% foram considerados (Novák *et al.*, 2020). As proporções individuais e comparativas dos clados de *repeats*, referindo-se ao total de unidades repetitivas classificadas em cada categoria, foram calculadas com base no número de leituras agrupadas em relação ao total de leituras utilizadas na análise, após a exclusão das leituras de cloroplasto e mitocôndria (Neumann *et al.*, 2019). Adicionalmente, o pipeline TAREAN (Tandem Repeat Analyser), também implementado no Galaxy/RepeatExplorer2, foi empregado para identificar *repeats* em tandem, permitindo a caracterização de satDNA (Novák *et al.*, 2017).

Filogenia Molecular

A filogenia molecular foi conduzida com base em sequências de plastomas completos e DNAr nuclear para investigar a evolução do repetoma no contexto filogeográfico de *Podocarpus sellowii*. Os plastomas foram previamente montados conforme descrito por Nieto-Blázquez *et al.* (2020), enquanto as sequências de DNAr seguiram as diretrizes de Melo *et al.* (2019). Ambas as sequências foram alinhadas utilizando o algoritmo MAFFT, assegurando alta qualidade nos alinhamentos. Para inferir as topologias, aplicou-se o método de *Network* por meio do software Splitstree4 (Huson & Bryant, 2005). A análise de *Network* foi baseada na máxima parcimônia, aplicada tanto aos plastomas quanto ao DNAr, proporcionando uma visão integrada das divergências genéticas entre os diferentes clados. Além disso, foi utilizado o método de inferência filogenética baseado na abundância de *repeats* (Dodsworth, 2015), permitindo verificar a

acumulação diferencial desses elementos nas populações de *P. sellowii* analisadas, com o objetivo de delimitar clados taxonômicos.

Filogenia de *Repeats* e Sinal Filogenético

A filogenia dos *repeats* nos genomas das populações de *Podocarpus sellowii* foi realizada para complementar as topologias baseadas em plastomas e DNAr. Utilizou-se o método de Máxima Verossimilhança (ML) com 1000 replicações no software Geneious (v. 9.1.8) e o plugin FastTree (Price *et al.*, 2009) para reconstruir as relações filogenéticas. Para avaliar possíveis incongruências entre as topologias geradas a partir de diferentes marcadores, aplicou-se o pacote Cophylo no software R (Revell, 2012), permitindo a comparação simultânea das árvores filogenéticas. O sinal filogenético foi avaliado pela estatística λ , onde $\lambda = 1$ indica estrutura filogenética de acordo com o modelo de movimento browniano, e $\lambda = 0$ indica ausência de sinal filogenético. Os valores de $p < 0,05$ foram considerados significativos. A rede filogenética foi construída com base em dados combinados de plastomas e DNAr nuclear utilizando o programa SplitsTree4 (Huson; Bryant, 2006), conforme a metodologia de Holland; Moulton (2003). Essa análise ofereceu uma perspectiva detalhada sobre a distribuição das linhagens repetitivas entre as populações, contribuindo para a compreensão da influência de eventos históricos e ecológicos. Adicionalmente, foi realizada uma análise do sinal filogenético dos *repeats*, correlacionando a abundância desses elementos com preditores ecológicos extraídos de variáveis ambientais do WorldClim. Para identificar as variáveis que melhor explicam as abundâncias dos *repeats*, foram empregados modelos lineares generalizados com regressão quase-binomial, utilizando o pacote glmulti (Calcagno, 2010) no software R (Bartoń, 2023). A seleção dos modelos baseou-se no critério de informação de Akaike (AIC) (Akaike, 1973; Hurvich; Tsai, 1989), selecionando os modelos com menores valores de AIC como mais robustos para representar a relação entre variáveis ecológicas e a distribuição dos repetitivos.

Distribuição Geográfica e Seleção de Variáveis Ecológicas

Para investigar a influência de variáveis ecológicas não colineares sobre os diferentes clados repetitivos nos genomas de *Podocarpus sellowii*, foram selecionadas 10 populações naturais do Brasil, representando diferentes regiões geográficas, localizadas e codificadas da seguinte maneira: Baturité - CE (BTT), Caruaru - PE (BJC), Itabaiana - SE

(ITB), Comodoro - MT (COM), Alto Paraíso de Goiás - GO (CPV), Una - BA (BAH), Brasília - DF (BRA), Carrancas - MG (CAR), Paranapiacaba - SP (PAR) e Santo Amaro - SC (STA). As coordenadas geográficas dessas populações foram registradas e utilizadas para obtenção de dados climáticos. Um único indivíduo por população foi selecionado para caracterizar a fração repetitiva do genoma. As variáveis ecológicas foram obtidas da plataforma WorldClim 2.1 (Fick; Hijmans, 2017), e os dados geoclimáticos processados no *software* QGIS (QGIS - Development Team, 2023). Foram extraídas 19 variáveis bioclimáticas (Hijmans *et al.*, 2005), e realizada uma análise de colinearidade entre elas. Variáveis com alta correlação (coeficiente de Pearson $> 0,7$) ou baixa contribuição nas primeiras cinco componentes principais de uma análise PCA foram excluídas. O conjunto final de variáveis incluiu cinco bioclimáticas: BIO2 (variação da temperatura diurna), BIO5 (temperatura máxima do mês mais quente), BIO6 (temperatura mínima do mês mais frio), BIO13 (precipitação do mês mais úmido) e BIO14 (precipitação do mês mais seco). Os dados das variáveis selecionadas foram organizados em tabelas utilizando o *software* Microsoft Excel 2021 (Microsoft Corporation, 2021) para facilitar a visualização e o tratamento dos dados. Foi gerado um gráfico bidimensional para representar a separação geoclimática das populações com base nas cinco variáveis selecionadas, visando identificar padrões de distribuição das linhagens repetitivas nos gradientes ambientais.

Interação Genoma-Ambiente

Para explorar a relação entre as abundâncias dos principais *repeats* do genoma e as variáveis ambientais, foram realizadas análises comparativas e regressões. As variáveis climáticas previamente obtidas da plataforma WorldClim 2.1 (Fick & Hijmans, 2017) foram utilizadas como preditoras. A avaliação da correlação entre abundâncias de linhagens repetitivas e variáveis climáticas foi realizada por meio de correlação de Pearson, excluindo-se variáveis com coeficiente maior que 0,7. Um modelo *Species Packing* com distribuição Gaussiana, implementado no *software* PAST 4.16 (Hammer *et al.*, 2001), foi utilizado para modelar a resposta das abundâncias dos repetitivos ao longo dos gradientes ecológicos. Adicionalmente, uma análise de regressão linear simples foi realizada para avaliar a relação entre a abundância de *repeats* e as variáveis bioclimáticas selecionadas (BIO2, BIO5, BIO6, BIO13 e BIO14). Por fim, uma análise de Mínimos Quadrados Generalizados Filogenéticos (PGLS, do inglês *Phylogenetic Generalized Least Squares*) foi aplicada para examinar a influência das variáveis geoclimáticas sobre a diversidade genômica plastidial das populações de *P. sellowii*, utilizando o *software* R Studio e o pacote *caper* (Orme *et al.*,

2013). A matriz filogenética foi derivada dos dados plastidiais, assumindo um modelo Browniano para a evolução dos caracteres, sendo o modelo com menor valor de AIC selecionado para identificar as associações mais adequadas entre o genoma plastidial e os fatores geoclimáticos.

Resultados e Discussão

Composição da fração repetitiva de indivíduos de *P. sellowii*

Com a análise das frações repetitivas dos genomas, realizada por meio da plataforma RepeatExplorer (Novak *et al.*, 2013), foi possível obter a caracterização detalhada dos principais elementos repetitivos nos genomas dos 10 indivíduos de populações naturais de *Podocarpus sellowii*. Os resultados evidenciaram padrões distintos na abundância e diversidade de repeats entre os indivíduos analisados (Tabela 1, Figura 1). Alguns estudos indicam que variações em repeats do genoma podem estar associadas a respostas diferenciadas a pressões ambientais, afetando a capacidade adaptativa e evolutiva de populações frente a desafios ecológicos, como competição por recursos e mudanças no habitat (Avolio *et al.*, 2012). Esses padrões de variação genética e diversidade fenotípica têm sido observados em diferentes contextos ecológicos, reforçando a ideia de que características genômicas, principalmente relacionadas à dinâmica dos elementos repetitivos do genoma, podem desempenhar um papel crucial na resiliência e adaptação de populações a novos ambientes (Pasala *et al.*, 2020). No repeatoma das populações analisadas, os elementos mais abundantes foram: a linhagem/classe Ty1/Copia - Ivana, sendo o elemento mais abundante no genoma, com valores de abundância variando entre 29,65% e 35,13% nas diferentes populações amostradas. A linhagem/classe Ty3/Gypsy - Ogre, embora menos abundante que Ivana, apresenta alta abundância, variando de 10,88% a 11,29% nas populações. A linhagem/classe Ty1/Copia - SIRE também apresenta abundância significativa, com valores entre 8,06% e 10,46%. Entre os elementos não-LTR, a linhagem LINE é a mais abundante, variando de 0,68% a 0,95%. Nessa perspectiva, ficou evidente que as linhagens Copia predominam sobre as linhagens Gypsy, com destaque expressivo para o Ivana. Além disso, os elementos LTR são mais abundantes do que os não-LTR no genoma desta espécie. Esse padrão de maior abundância de LTR também foi observado em representantes de outras plantas, notadamente *Zea mays* L. (milho), *Glycine max* (L.) Merr. (soja) e *Oryza sativa* L. (arroz) (Baucom *et al.*, 2009; Tenaillon *et al.*, 2011; Munoz-Amatriain *et al.*, 2013). Em *Zea mays*,

os elementos LTR representam a maior parte do conteúdo repetitivo, contribuindo significativamente para o tamanho do genoma (Tenaillon *et al.*, 2011). De forma semelhante, em *Glycine max*, retrotransposons da família LTR são predominantes, desempenhando um papel importante na variação genômica (Du *et al.*, 2010). Já em *Oryza sativa*, os elementos LTR foram identificados como os mais abundantes entre os *repeats* (Baucom *et al.*, 2009). Além disso, variação intraespecífica na composição de *repeats* foi observada em diversas espécies de plantas. Em *Hordeum vulgare* L. (cevada), por exemplo, as diferenças no conteúdo de LTRs entre linhagens refletem variações no tamanho do genoma (Munoz-Amatriain *et al.*, 2013). No *Triticum aestivum* L. (trigo), também foram observadas diferenças significativas na abundância de elementos LTR entre variedades e ecótipos (Wicker *et al.*, 2018), sugerindo um impacto importante desses elementos na plasticidade genômica e na adaptação às condições ambientais. Levando isso em consideração, a predominância das linhagens Copia sobre as Gypsy sugere que essas sequências repetitivas podem ter desempenhado um papel significativo na evolução e na estrutura genômica de *Podocarpus sellowii*. Os elementos Copia apresentam uma distribuição mais uniforme ao longo dos cromossomos, com menor especificidade pela heterocromatina em comparação aos Gypsy. Essa característica pode influenciar sua relação com a estrutura genômica e potencialmente com aspectos adaptativos (Bourgeois *et al.*, 2021).

A topologia da árvore filogenética baseada em repeats não apresentou similaridade com as redes construídas a partir dos plastomas ou DNAr, o que sugere uma evolução independente da fração repetitiva do genoma em relação às linhagens filogenéticas tradicionais. Isso sugere que os elementos repetitivos no genoma de *Podocarpus sellowii* aparentemente possuem uma dinâmica evolutiva pouco correlacionada com as relações de parentesco genético, indicando a influência de outros fatores. Por outro lado, a árvore baseada em repeats revelou uma divisão em dois grandes clados: o primeiro inclui as populações do Nordeste do Brasil (BTT, BJC, ITB, BAH) e outras populações mais próximas ao litoral (PAR e STA, com exceção de BRA), enquanto o segundo é composto por populações do Centro-Oeste e Sudeste (COM, CPV, CAR). Essa divisão filogeográfica está fortemente associada à distribuição geográfica das populações, sugerindo que os agrupamentos da topologia dos repeats está refletindo relações ecológicas. Castro *et al.*, (2024), destacam a influência de gradientes ambientais sobre a composição do repeatoma em diferentes populações de espécies do gênero *Erythrostemon* Klotzsch (Leguminosae Juss). No entanto, essa topologia dos repeats apresentou baixos valores de suporte (Figura 2), o que pode estar relacionado ao tipo de dado (abundância dos *repeats*) usados para a

construção da árvore e/ou baixa resolução filogenética dos repeats nessa espécie (Dodsworth *et al.* 2015).

Alternativamente, esses baixos valores de suporte podem sugerir que a divergência entre as populações pode não ser tão robusta ou linear quando comparadas às filogenias obtidas por outros parâmetros, revelando que outros fatores, como evolução convergente de *repeats* ou pressões seletivas locais, além das próprias variáveis ecológicas, podem estar atuando na diversificação das populações (Ord; Summers, 2015; Bourgeois *et al.*, 2019). A discrepância observada entre esta árvore dos *repeats* e as redes baseadas em plastomas e DNAr, juntamente com o baixo sinal filogenético observado (Tabela 3), levantam o indício de uma evolução independente dos *repeats* em relação às sequências genômicas conservadas, o que é inesperado, tendo em vista o forte significado filogenético dos repeatomas apontado em estudos anteriores (Dodsworth *et al.* 2015; Du *et al.*, 2022; Jost *et al.*, 2022). Isso pode sugerir que em análises numa escala intraespecífica com relações não-dicotômicas os *repeats* podem não ser bons marcadores de relações de parentesco.

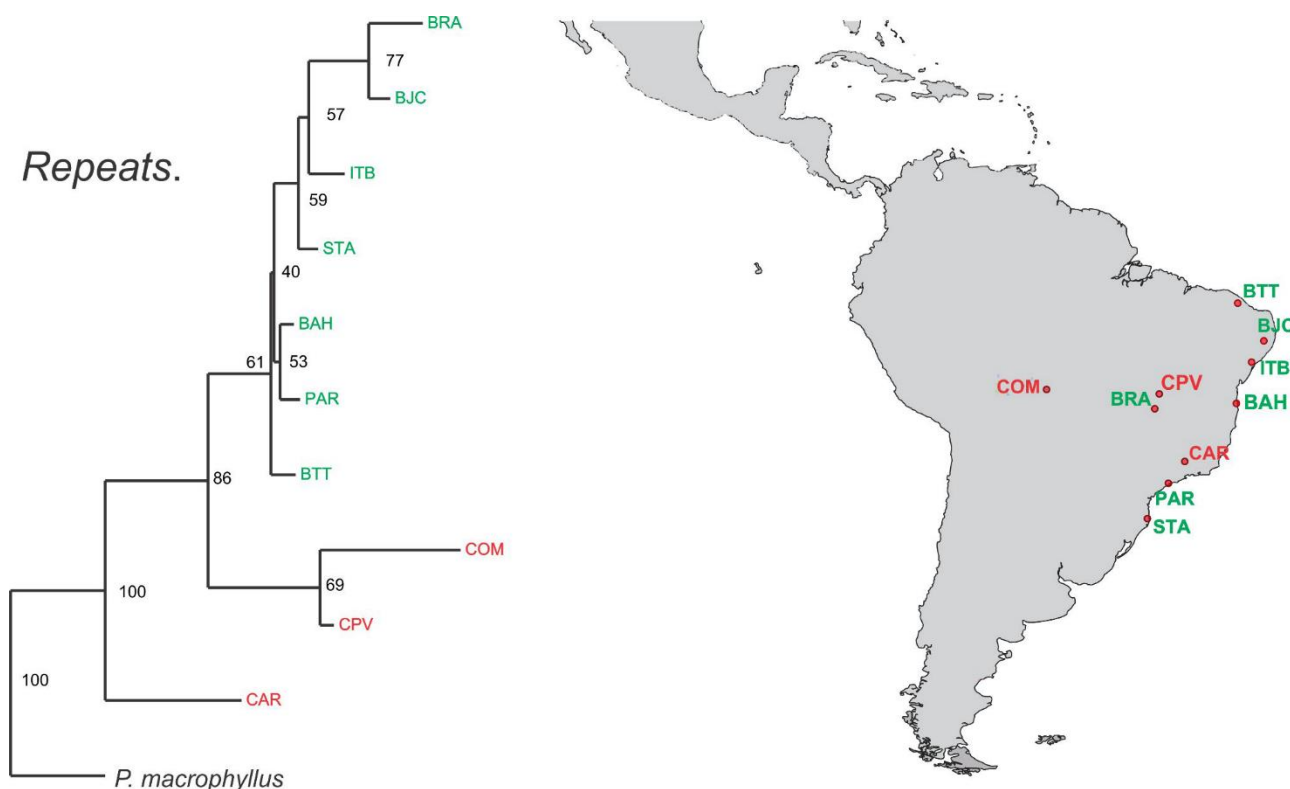


Figura 2. Árvore filogenética baseada na abundância de *repeats* nos genomas de *Podocarpus sellowii* e mapa de ocorrência das 10 populações analisadas. Fonte: O autor (2024).

Análise filogenética, Seleção de Modelo Ambiental e Sinal Filogenético

Utilizando o método de Máxima Verossimilhança (ML) com 1000 replicações e redes filogenéticas, obtivemos informações importantes sobre a relação entre a filogenia e a abundância dos *repeats* no genoma (Tabela 3). Os resultados da avaliação do sinal filogenético, utilizando a estatística λ , indicaram uma estrutura filogenética alinhada ao modelo de movimento browniano em diversas características analisadas, com λ próximo de 1. Em contrapartida, observou-se ausência de sinal filogenético em outros casos, com valores de λ tendendo a 0. Só houve significância estatística confirmada com valores de $p < 0,05$, indicando a robustez dos resultados obtidos. Para os plastomas, observou-se que a maioria dos *repeats* não apresentou sinal filogenético significativo, com exceção da família Athila ($\lambda = 1$, $p = 0,043$) e da família CACTA ($\lambda = 1$, $p = 0,069$). Esses resultados indicam que a evolução dessas famílias de repetitivos pode estar correlacionada com a estrutura filogenética das populações de *P. sellowii*. Em contrapartida, as linhagens Ale, CRM, Gymco, LINE, Ogre, Reina, SIRE, SOLA2, TatII, Tekay, Tork, PARA e satellite não apresentaram sinal filogenético significativo ($p = 1,000$), sugerindo que a evolução dessas famílias não segue o padrão de divergência filogenética esperado pelo modelo browniano. No DNAr nuclear, o sinal filogenético foi mais evidente em alguns *repeats*, como Ivana ($\lambda = 0,995$, $p = 0,034$), SIRE ($\lambda = 0,971$, $p = 0,001$), PARA ($\lambda = 1$, $p = 0,003$) e satellite ($\lambda = 0,999$, $p = 0,026$). Esses resultados sugerem uma correlação entre a evolução da fração repetitiva e a divergência filogenética nas populações de *P. sellowii*, especialmente para esses elementos. Entretanto, outros repetitivos como Athila ($\lambda = 0,794$, $p = 0,348$), CACTA ($\lambda = 0,783$, $p = 0,145$) e Ogre ($\lambda = 0,988$, $p = 0,297$) não apresentaram sinal significativo, indicando uma evolução mais independente desses elementos em relação à história filogenética das populações. A discrepância entre as análises baseadas em plastomas e DNAr nuclear sugere uma evolução distinta dos *repeats* em diferentes partes do genoma, possivelmente influenciada por fatores ecológicos e ambientais. Modelos baseados em variáveis ecológicas do WorldClim reforçam essa hipótese, mostrando que preditores climáticos desempenham um papel importante na diversificação genômica de *Podocarpus sellowii*. Assim, a evolução da fração repetitiva do genoma dessa espécie não parece estar estritamente ligada à filogenia, sendo que o sinal filogenético em alguns elementos pode refletir herança, enquanto a ausência desse sinal em outros sugere a atuação de outros fatores, como influências ambientais, ou evolução convergente. Estudos prévios corroboram a variação observada nos sinais filogenéticos em elementos repetitivos. Münkemüller *et al.* (2012) enfatizam que, enquanto algumas características genômicas podem seguir um modelo de evolução filogenética baseado no movimento browniano, outras, especialmente em grupos vegetais, apresentam evolução não filogenética devido a

pressões ambientais ou processos de convergência evolutiva. Delsuc *et al.* (2018) demonstram que a evolução de elementos repetitivos, como transposons, pode ser modulada por fatores ecológicos e climáticos, reforçando a hipótese de que a evolução de frações genômicas específicas não está estritamente relacionada à história filogenética, mas influenciada por variações ambientais. Esses fatores podem explicar a ausência de sinal filogenético significativo em diversas linhagens de *P. sellowii*, particularmente em relação aos *repeats* que seguem padrões evolutivos independentes de sua filogenia.

Tabela 3. Sinal filogenético (λ) e valores de p para as principais linhagens de *repeats* nos genomas de *Podocarpus sellowii*, com base em plastomas e DNAr nuclear. Valores de λ indicam a presença ($\lambda = 1$) ou ausência ($\lambda = 0$) de sinal filogenético, enquanto valores de $p < 0,05$ indicam significância estatística.

Plastomas				DNAr		
<i>Repeat</i>	λ	p		<i>Repeat</i>	λ	p
Ale	0,000	1,000		Ale	0,000	1,000
Athila	1,000	0,043		Athila	0,794	0,348
CRM	0,000	1,000		CRM	0,000	1,000
CACTA	1,000	0,069		CACTA	0,783	0,145
Gymco	0,000	1,000		Gymco	0,000	1,000
hAT	1,000	0,069		hAT	0,000	1,000
Ivana	0,735	0,309		Ivana	0,995	0,034
LINE	0,000	1,000		LINE	0,000	1,000
Ogre	0,000	1,000		Ogre	0,988	0,297
Reina	0,000	1,000		Reina	0,000	1,000
Retand	0,000	1,000		Retand	0,000	1,000
SIRE	0,000	1,000		SIRE	0,971	0,001
SOLA2	0,000	1,000		SOLA2	0,569	0,408
TatII	0,000	1,000		TatII	0,000	1,000
Tekay	0,000	1,000		Tekay	0,000	1,000
Tork	0,000	1,000		Tork	0,000	1,000
PARA	0,000	1,000		PARA	1,000	0,003
satellite	0,000	1,000		satellite	0,999	0,026

Fonte: O autor (2024)

Distribuição das Populações de *Podocarpus sellowii*

A análise das populações de *Podocarpus sellowii* ao longo de um gradiente latitudinal no Brasil (Figuras 2 e 3) revelou variação genômica significativa. A abordagem NGS gerou dados informativos que permitiram identificar padrões genéticos importantes. As análises de *Network* revelaram um cenário relativamente similar entre os dados plastidiais e de DNAr (Figura 4). As análises nucleares e plastidiais revelaram *P. sellowii* uma tendência de separação entre as populações do Nordeste e as do Centro-Oeste/Sudeste/Sul (Figura 2). As populações mais próximas ao litoral do Nordeste exibem ramos mais curtos nas redes, indicando maior similaridade entre si e um possível evento mais recente de diversificação, o que é corroborado por análises prévias (Melo *et al.*, 2019). Em contraste, as populações do Centro-Oeste, Sudeste e Sul mostram ramos mais longos, uma maior complexidade nas relações com maiores reticulações, o que sugere uma maior complexidade filogenética associada a um tempo de divergência mais longo (Melo *et al.*, 2019) (Figura 4). Essa diferenciação pode estar ligada a um processo de evolução mais antigo das populações do Centro/ Sul, o que reforça a hipótese de que a rota de colonização de *P. sellowii* para a Mata Atlântica se deu por conexões com a Amazônia (centro de origem do grupo) via a região central do Brasil durante períodos interglaciais (Quiroga *et al.*, 2016; Melo *et al.*, 2019). A variação genômica significativa observada nas populações de *P. sellowii* ao longo do gradiente latitudinal pode ser explicada por eventos históricos e ecológicos que moldaram a distribuição genética das populações. Pennington *et al.* (2004) discutem a importância das conexões históricas entre a Amazônia e a Mata Atlântica durante os períodos interglaciais, corroborando a hipótese de que a colonização da Mata Atlântica por *P. sellowii* ocorreu via a região central do Brasil. Além do mais, Nieto-Blázquez (2019) sugere que eventos históricos e fatores ecológicos desempenham um papel crucial na diversificação de coníferas e outros grupos vegetais, como evidenciado no estudo da biogeografia de *Podocarpus*. Esses trabalhos reforçam eventos que podem explicar a separação entre as populações do Nordeste e aquelas do Centro-Oeste, Sudeste e Sul, observada nas análises nucleares e plastidiais, e indicam que o padrão de divergência no Nordeste pode estar relacionado a eventos de diversificação das principais linhagens da espécie, que se deu há 18 milhões de anos, durante o Mioceno, como observado por Melo *et al.* (2019).

Além disso, as análises de *Network* indicaram que as populações do Nordeste de *P. sellowii* permanecem monofiléticas, mostrando maior similaridade genética entre essas populações em relação às do Centro-Oeste, Sudeste e Sul (Figura 4). Esse padrão é

observado tanto nos dados nucleares quanto nos plastidiais (Figura 4, A e B), mas com uma diferença notável: a rede filogenética baseada em DNAr apresenta menor resolução e maior reticulação em comparação com a rede plastidial. Acredita-se que essa diferença seja resultado da herança biparental do DNAr, que aumenta a recombinação e a variabilidade genética, levando a uma topologia mais complexa e reticulada nas redes nucleares. Esse fenômeno tem sido amplamente discutido em estudos de filogenia de plantas, como os de Hollingsworth *et al.* (2011) que destacam que a herança biparental do DNA nuclear complica a resolução de filogenias, aumentando o grau de reticulação em redes filogenéticas. Além disso, a pesquisa de Li *et al.* (2021) também observou que sequências plastidiais, devido à sua herança uniparental, fornecem uma resolução filogenética mais clara, corroborando os resultados encontrados nas redes plastidiais de *P. sellowii*. Mais recentemente, estudos reforçam essa distinção, mostrando que a análise de plastomas frequentemente resulta em maior eficiência na discriminação de espécies devido à menor recombinação e à herança direta da mãe, facilitando a reconstrução de topologias filogenéticas bem-resolvidas (Fu *et al.*, 2024).



Figura 3. Mapa com as populações naturais amostradas de *Podocarpus sellowii*. Siglas indicam as localidades de cada população, conforme legenda disponível na Tabela 1. Fonte: O autor (2024).

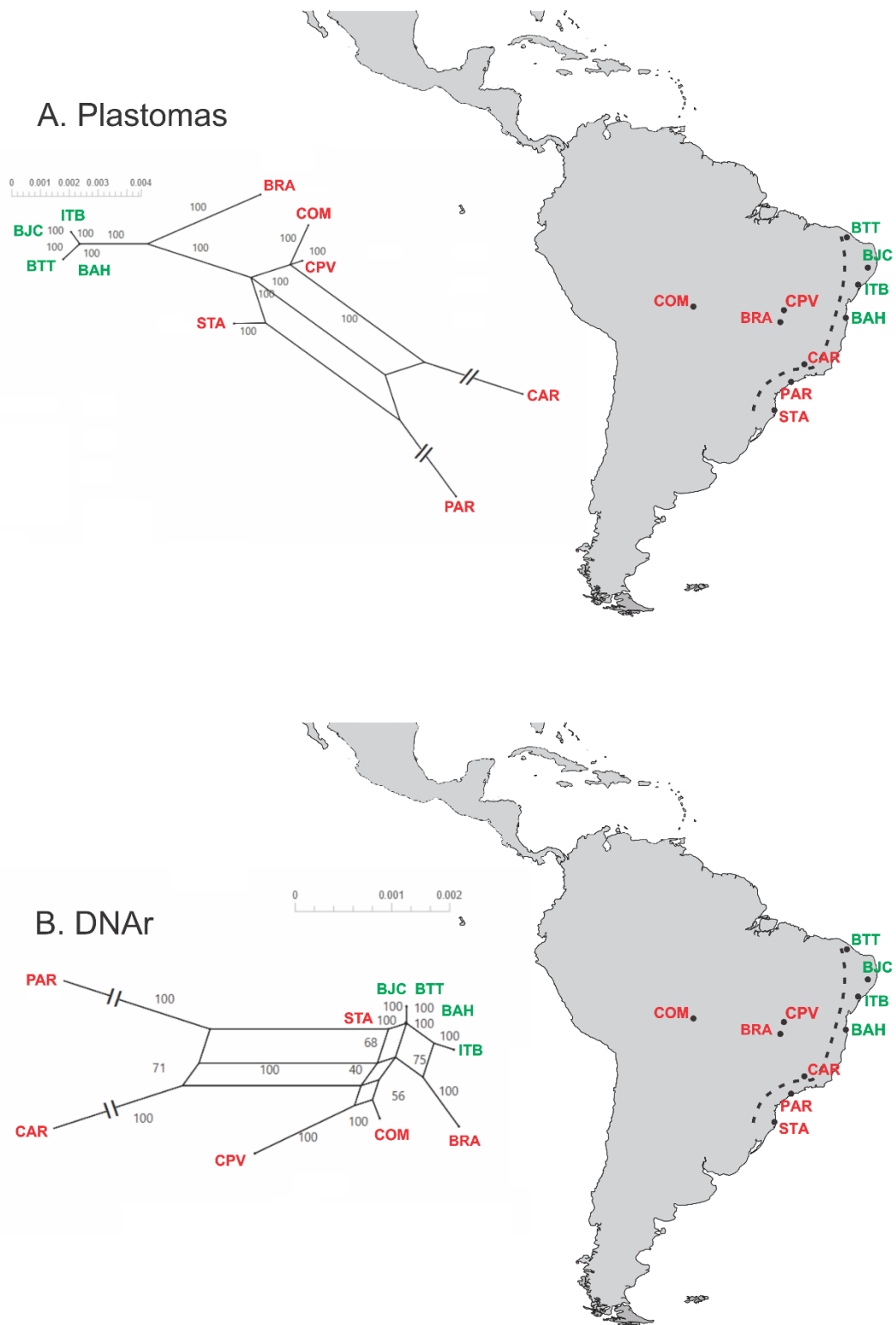


Figura 4. Redes filogenéticas inferidas para as 10 populações de *Podocarpus sellowii* (ITB, STA, BRA, BAH, BJC, BTT, CAR, COM, CPV, PAR), juntamente com os respectivos mapas de localização geográfica das populações amostradas. (A): a rede filogenética baseada em plastomas revela uma clara distinção entre as populações do Nordeste (BTT, BJC, ITB, BAH), além da população BRA, cujos ramos curtos indicam maior similaridade e sugerem um evento de divergência mais recente. (B): rede filogenética baseada em DNAr mostra populações do Centro-Oeste e Sudeste (PAR, CAR, CPV) com ramos mais longos e maior reticulação, sugerindo uma complexidade filogenética maior e um tempo de divergência mais longo em relação às populações do Nordeste. Os ramos com dois traços diagonais indicam ramos muito longos que foram reduzidos para melhorar a visualização. A escala na parte superior esquerda (Figura A) e superior direita (Figura B) representam a distância evolutiva. Fonte: O autor (2024).

Variabilidade ecológica das populações de *Podocarpus sellowii*

A distribuição das populações naturais de *Podocarpus sellowii* abrange um extenso gradiente latitudinal, de $-4,26$ (Baturité-CE) a $-27,72$ (Santo Amaro-SC) (Figura 3). Os mapas gerados no QGis (Figuras 2, 3 e 4) evidenciam a ampla dispersão geográfica da espécie, destacando sua ocorrência em diferentes zonas climáticas. A extração de 19 variáveis bioclimáticas permitiu identificar fatores ambientais associados à distribuição das populações (Tabela 2). Após testes de autocorrelação, as variáveis Variação da Temperatura Média Diurna (BIO2) e Precipitação do Mês Mais Seco (BIO14) surgiram como os preditores climáticos mais relevantes, otimizando a modelagem ambiental e proporcionando uma compreensão mais precisa dos fatores ecológicos que influenciam a variação genética e a adaptação das populações ao longo do gradiente latitudinal.

A variabilidade climática observada entre as populações de *P. sellowii* é significativa, com a Precipitação do Mês Mais Seco (BIO14) variando de 6 mm em Paranapiacaba (SP) a 110 mm em Santo Amaro (SC), e a Variação da Temperatura Média Diurna (BIO2) oscilando entre $6,94$ °C em Una (BA) e $13,06$ °C em Comodoro (MT). Essa variabilidade pode estar associada a adaptações ecológicas locais, uma hipótese corroborada por estudos com outras espécies que habitam gradientes amplos. De Sousa *et al.* (2020) investigaram populações de *Araucaria angustifolia* (Bert.) O. Ktze, uma conífera brasileira que, assim como *Podocarpus*, ocorre em uma faixa altitudinal e latitudinal diversa, mais precisamente ocorrendo nas regiões Sudeste, nos estados de Minas Gerais, Rio de Janeiro e São Paulo, e na região Sul, no Paraná, Santa Catarina e Rio Grande do Sul (Iganci & Dornelas, 2020). Os autores descobriram que fatores como a variação na temperatura e umidade foram determinantes para padrões de diferenciação genética local, corroborando os resultados apresentados aqui, nos quais a variação da temperatura média diurna foi identificada como um dos principais preditores (De Sousa *et al.*, 2020). Nesse sentido, a seleção de variáveis como Precipitação do Mês Mais Seco (BIO14) e a Variação da Temperatura Média Diurna (BIO2) otimiza a modelagem ambiental e revela padrões importantes de adaptação local. Iler *et al.* (2021) mostraram que a variação nas condições de precipitação, particularmente a precipitação do mês mais seco, está fortemente associada a mudanças na fenologia e na plasticidade fisiológica de plantas adaptadas a ambientes sazonais. O fato de a Precipitação do Mês Mais Seco variar de 6 mm em Paranapiacaba (SP) a 110 mm em Santo Amaro (SC) pode indicar adaptações diferenciadas em resposta às condições locais de umidade, tal como observado por Iler *et al.* (2021) em espécies mediterrâneas que enfrentam flutuações sazonais extremas.

Correlação genoma-ambiente

Foram utilizadas diferentes estratégias para estimar a relação genoma-ambiente em *Podocarpus sellowii*. As análises realizadas revelaram correlações significativas entre a abundância de repeats e as variáveis ambientais, indicando uma forte influência do ambiente sobre o genoma de *P. sellowii*. O correlograma (Figura 5) mostrou que diferentes repeats respondem de maneiras específicas às condições climáticas. Os elementos Ale e Reina apresentaram uma correlação negativa com a variável BIO5 (temperatura máxima do mês mais quente) e com a latitude, sugerindo que sua abundância diminui em regiões de maior temperatura e em localidades mais ao norte. Este padrão pode refletir uma influência das condições climáticas na regulação do elemento Ale, uma vez que temperaturas mais baixas e menores amplitudes térmicas podem favorecer sua estabilidade e mobilização no genoma, possivelmente devido a uma menor repressão epigenética e maior eficiência enzimática da transcriptase reversa (Ito *et al.*, 2011). Já o elemento Athila foi correlacionado negativamente com a BIO2 (amplitude térmica anual), o que indica uma menor abundância deste elemento em áreas com maior variação sazonal de temperatura. Da mesma forma, CACTA mostrou correlação negativa com BIO2, mas uma correlação positiva com BIO6 (temperatura mínima do mês mais frio). Isso sugere que esse elemento pode ser favorecido por climas com menores variações térmicas, mas com temperaturas mínimas relativamente mais altas. Outro resultado interessante foi observado no elemento Gymco-L, que apresentou uma correlação negativa com BIO2, indicando que ele pode ser sensível a variações térmicas mais acentuadas. hAT, por sua vez, apresentou forte correlação negativa com BIO2 e positiva com BIO14 (precipitação durante a estação seca), sugerindo uma abundância maior em ambientes com menor variação térmica e maior disponibilidade hídrica nas estações mais secas. O elemento Ivana também mostrou um padrão de correlação, apresentando uma correlação negativa com BIO13 (precipitação durante a estação mais úmida). Esses achados reforçam que a precipitação, tanto na estação seca quanto na úmida, exerce uma influência importante na distribuição destes elementos. Nessa perspectiva, os resultados sugerem uma correlação entre as variáveis climáticas e a abundância dos *repeats* em *P. sellowii*, indicando que fatores climáticos podem desempenhar um papel relevante nessa dinâmica. As correlações negativas observadas com variáveis associadas a temperaturas extremas e amplitudes térmicas, juntamente com as correlações positivas com a precipitação, sugerem que o ambiente climático pode influenciar a abundância de *repeats*. Assim, os resultados indicam uma possível associação entre fatores climáticos e a dinâmica dos *repeats*, que pode refletir

respostas adaptativas, mas essa hipótese requer investigações adicionais para confirmação. Comparativamente, estudos anteriores forneceram informações relevantes sobre essa relação. Um estudo com palmeiras (Arecaceae Schultz Sch.) observou que a aridez pode limitar a amplificação de *repeats*, com elementos como Ty1-copia sendo mais abundantes em condições áridas, enquanto Ty3-gypsy e TIR predominam em ambientes úmidos, sugerindo que condições abióticas do ambiente influenciam diretamente a composição do repeatoma e restringe a expansão repetitiva em condições extremas (Schley *et al.*, 2022). Outro trabalho explorou a dinâmica dos *repeats* em *Erythrostemon* Klotzsch (Fabaceae), mostrando que esses elementos não respondem apenas a pressões ambientais atuais, mas também são moldados por eventos históricos, como divergências geográficas. No caso de *P. sellowii*, as correlações observadas com variáveis climáticas contemporâneas indicam uma resposta mais imediata às condições ambientais. Em contraste, em espécies como *Erythrostemon*, fatores históricos parecem desempenhar um papel mais significativo na evolução do repeatoma (Castro *et al.*, 2024). Contudo, é importante ressaltar que a literatura sobre a família Podocarpaceae ainda é incipiente no que diz respeito à relação entre fatores ambientais e a evolução de *repeats*. Estudos mais robustos são necessários para confirmar esses padrões, já que atualmente as informações disponíveis para o gênero *Podocarpus* são limitadas. Essas variações refletem tanto a diversidade nas linhagens de *repeats* quanto as diferentes pressões ambientais enfrentadas por cada espécie, sugerindo que a influência do ambiente na dinâmica dos *repeats* pode variar substancialmente de acordo com o contexto ecológico e evolutivo.

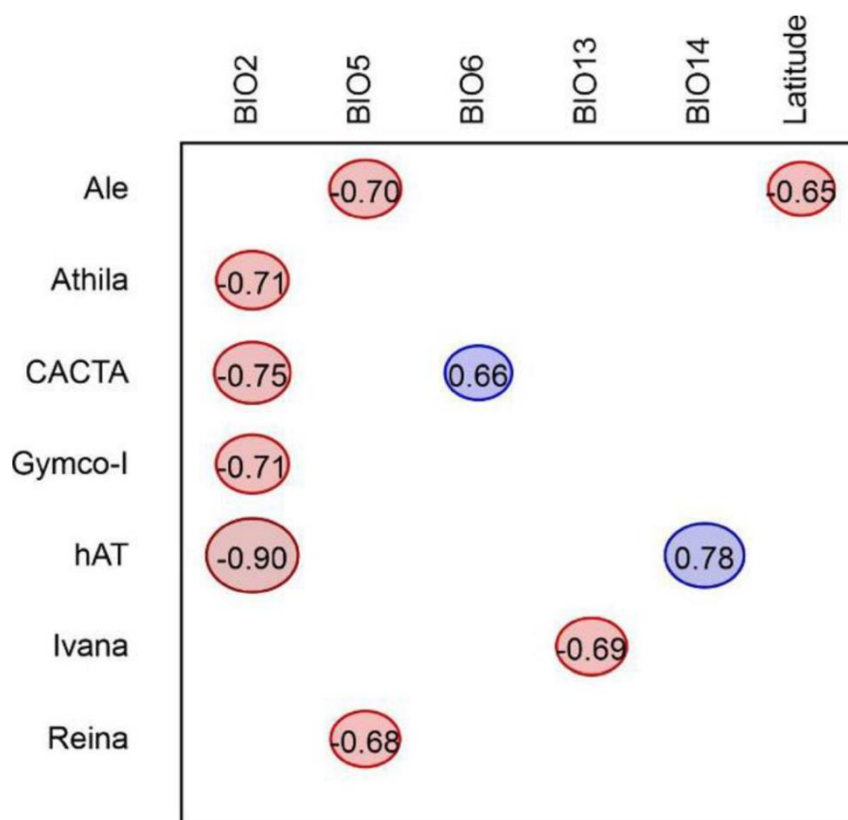


Figura 5. Correlograma gerado no PAST 4.16 mostrando as correlações entre os principais *repeats* e as variáveis climáticas em *Podocarpus sellowii*. *Repeats* com correlação negativa estão destacados em vermelho, como Ale com BIO5 (temperatura máxima do mês mais quente) e latitude, Athila com BIO2 (amplitude térmica anual), CACTA com BIO2, Gymco-L com BIO2, hAT com BIO2, Ivana com BIO13 (precipitação da estação mais úmida), e Reina com BIO5. As correlações positivas estão em azul, como CACTA com BIO6 (temperatura mínima do mês mais frio) e hAT com BIO14 (precipitação da estação seca). Fonte: O autor (2024).

Análise de Correlação (PGLS - *Phylogenetic Generalized Least Squares*)

A análise PGLS revelou associações significativas entre as variáveis geoclimáticas e a abundância de diferentes classes de *repeats* no genoma plastidial de *Podocarpus sellowii* (Figura 6). Esses resultados sugerem que o ambiente exerce uma pressão seletiva relevante sobre o genoma, influenciando diretamente a distribuição de *repeats* em diferentes populações. Entre os *repeats*, Reina mostrou uma correlação negativa com a variável BIO5 (temperatura máxima do mês mais quente), indicando que sua abundância tende a diminuir em regiões onde as temperaturas extremas são mais elevadas. Isso pode refletir uma adaptação desse elemento a climas mais moderados, onde as flutuações térmicas são menos drásticas. Similarmente, Ale também apresentou correlação negativa com a BIO5, reforçando a ideia de que variações extremas de temperatura podem exercer

pressão negativa sobre esse *repeat*. A variável BIO2 (variação da temperatura média diurna) também desempenhou um papel crucial nas correlações observadas. O elemento hAT apresentou uma forte correlação negativa com BIO2, sugerindo que maiores variações diurnas de temperatura afetam negativamente a abundância desse elemento, enquanto Cacta e Gymco também mostraram correlações negativas com a mesma variável. Esses resultados indicam que esses *repeats* são mais abundantes em locais com variações térmicas mais estáveis, sugerindo uma possível sensibilidade desses elementos a mudanças bruscas de temperatura ao longo do dia. Por outro lado, o elemento hAT demonstrou uma correlação positiva significativa com a variável BIO14 (precipitação do mês mais seco), o que sugere que esse elemento pode ser favorecido em ambientes onde a disponibilidade hídrica durante a estação seca é maior. Isso reflete uma adaptação às condições de maior umidade relativa nas regiões de coleta, onde a presença de água ao longo de períodos críticos pode influenciar positivamente a persistência de hAT. Já o elemento Ivana apresentou correlação negativa com BIO13 (precipitação do mês mais chuvoso), o que pode indicar que esse elemento é menos prevalente em áreas com maiores índices de precipitação durante a estação úmida. Esse padrão pode sugerir que o excesso de umidade em determinadas épocas do ano afeta negativamente a abundância desse Ivana. Esses resultados corroboram a crescente evidência de que os *repeats* desempenham um papel fundamental na resposta adaptativa de organismos às pressões ambientais (Canapa *et al.*, 2020). A seleção natural, induzida por condições climáticas extremas, como variações térmicas significativas e escassez hídrica, parece direcionar a composição dos genomas, favorecendo a proliferação de linhagens repetitivas que conferem vantagens adaptativas (Rogers *et al.*, 2017). Nessa perspectiva, em *Podocarpus sellowii*, assim como em outras plantas, como *Arabidopsis thaliana* e *Oryza sativa* (Zou *et al.*, 2021), a associação entre a abundância de determinados *repeats* e variáveis climáticas específicas sugere que esses elementos podem estar envolvidos em mecanismos moleculares que aumentam a resiliência ao estresse ambiental, especialmente em ecossistemas com amplitudes térmicas amplas e disponibilidade hídrica sazonal (Zou *et al.*, 2021).

PGLS

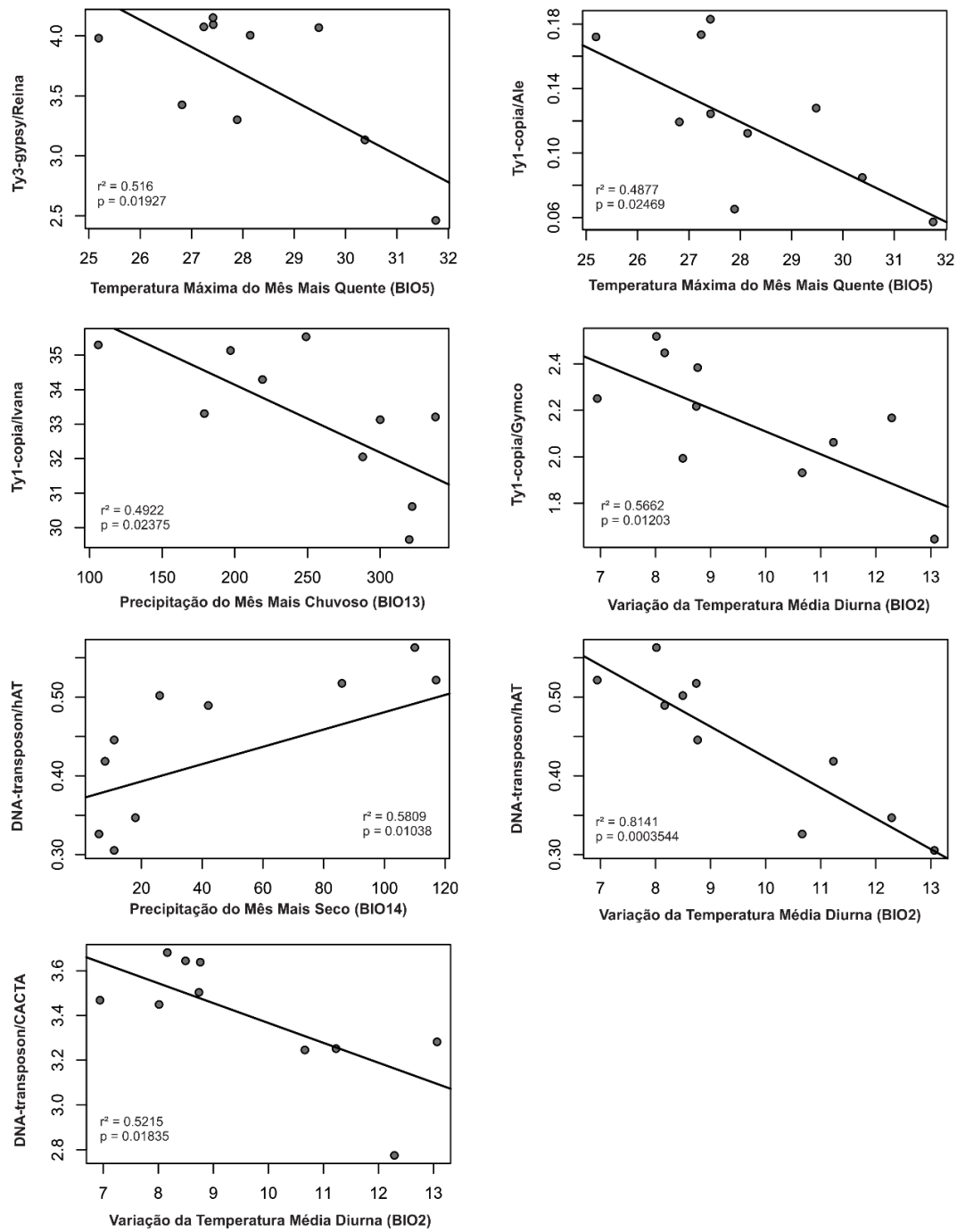


Figura 6. Análise de PGLS entre as variáveis ambientais: BIO2 (variação da temperatura média diurna), BIO5 (temperatura máxima do mês mais quente), BIO13 (precipitação do mês mais chuvoso) e BIO14 (precipitação do mês mais seco) e *repeats*: Ty3-gypsy/Reina, Ty1-copia/Ale, Ty1-copia/Ivana, Ty1-copia/Gymco, DNATransposon/hAT e DNATransposon/CACTA. Fonte: O autor (2024).

Tabela 1. Composição genômica de sequências repetitivas em diferentes populações de *Podocarpus sellowii*, expressa em proporção do genoma.. As sequências repetitivas estão classificadas em ordem, superfamília e linhagem/classe. Os valores representam a proporção de cada sequência repetitiva em relação ao genoma para dez populações de *P. sellowii* (ITB, STA, BRA, BAH, BJC, BTT, CAR, COM, CPV e PAR). Os retrotransposons são divididos em LTR e não-LTR. Os LTR incluem as superfamílias Ty3/Gypsy e Ty1/Copia, enquanto os não-LTR consistem na classe Non-chromovirus/OTA LINE. Os elementos de DNA transposon são PARA-RT, hAT, CACTA e Sola2.

Sequência repetitiva	Ordem	Super família	Linhagem/classe	ITB	STA	BRA	BAH	BJC	BTT	CAR	COM	CPV	PAR
Retro transposons	LTR	Ty3/gypsy	Tekay	0.959405	0.935996	0.9137	0.971897	0.94834	0.88486	1.040476	0.825255	0.83272	0.969789
			CRM	5.691163	5.710097	5.553663	5.565387	5.954329	6.020028	3.98849	6.005759	5.472875	5.505898
			Ogre	10.71009	11.2947	11.09201	11.03358	10.95514	10.87681	16.82102	10.2656	10.38391	11.05872
			Athila	0.332855	0.355574	0.321582	0.337606	0.399028	0.367863	0.173413	0.295756	0.270885	0.406686
			Retand	6.709307	5.877426	6.355064	6.296867	6.33781	6.566852	6.936505	5.399936	6.300578	5.78745
			OTA TatII	3.909411	3.953144	4.308172	4.47584	4.384127	4.329847	6.156148	3.529993	3.907763	4.019931
			Reina	3.13275	4.151847	4.004946	4.068509	4.093925	3.425103	4.075196	2.461455	3.300781	3.980827
		Ty1/Copia	SIRE	8.706435	8.695872	10.46416	9.110252	8.845987	8.948019	8.063687	8.906078	8.603098	9.267741
			Ale	0.084845	0.183016	0.112298	0.127881	0.124372	0.119307	0.173413	0.057243	0.065213	0.172059
			Tork	2.610625	2.787071	2.802354	2.424626	2.751739	2.88325	3.208133	2.537779	2.779077	2.424472
			Ivana	35.13249	34.29196	35.53221	33.30536	35.29585	33.12755	29.65356	30.61077	32.04968	33.20745
			Gymco-I	2.447461	2.217107	2.062206	2.250708	2.383804	1.99342	2.167658	1.64574	1.931308	2.518323
			OTA LINE	0.678763	0.894164	0.949431	0.8389	0.860242	0.919657	0.867063	1.020836	0.862818	0.954147
	Não LTR												
DNA transposon			PARA-RT	0.130531	0.115126	0.112702	0.11271	0.113414	0.110284	1.935167	0.105958	0.108746	0.166491

	hAT	0.489492	0.517674	0.418566	0.521755	0.445668	0.502083	0.346825	0.305297	0.326065	0.563103
	CACTA	3.680982	3.503448	3.251547	3.468137	3.637893	3.643832	2.774602	3.28194	3.2456	3.449007
	Sola2	1.44237	1.699434	1.715102	1.677801	1.818946	1.739893	1.213888	1.46447	1.514948	1.861369
DNAsat total		13.15102	12.81635	10.03028	13.41217	10.64939	13.54134	10.40476	21.28014	18.04393	13.68654
Não Classificado		18.45678	16.23145	15.68923	16.98754	17.54321	18.23467	18.87654	16.34567	18.12345	16.78923
Total		135.4567	130.1234	128.8765	132.5432	127.9876	138.2345	136.7892	125.6543	130.4567	134.1234

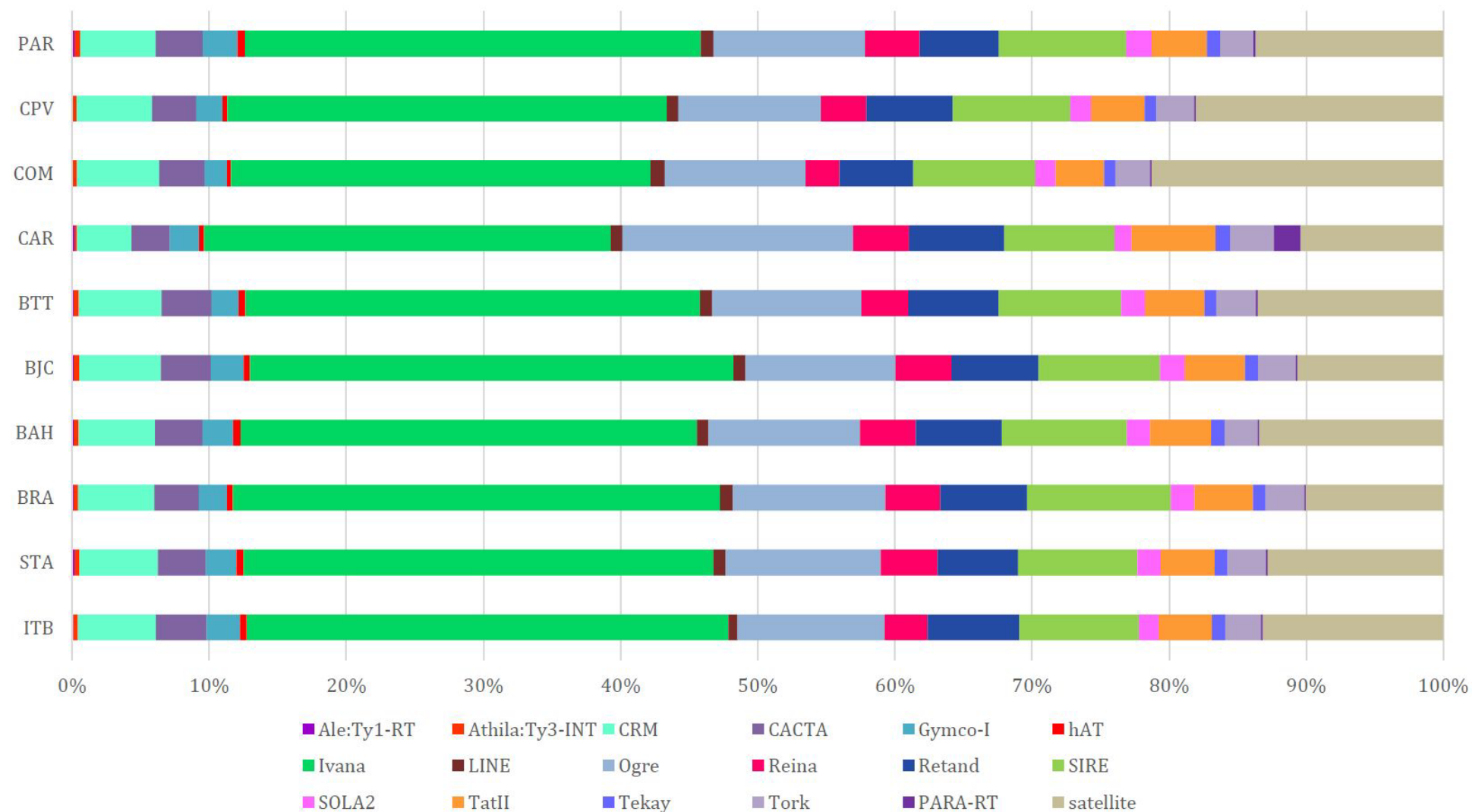
Fonte: O autor (2024).

Tabela 2. Variáveis geoclimáticas do período 1970 a 2000 (<https://worldclim.org/>) extraídas para as áreas de ocorrência das populações naturais de *Podocarpus sellowii*. *P. sellowii*. As variáveis incluem Longitude, Latitude, Variação da Temperatura Média Diurna (BIO2), Temperatura Máxima do Mês Mais Quente (BIO5), Temperatura Mínima do Mês Mais Frio (BIO6), Precipitação do Mês Mais Chuvoso (BIO13), Precipitação do Mês Mais Seco (BIO14) e Altitude para as seguintes localidades: Caruaru (PE), Itabaiana (SE), Baturité (CE), Una (BA), Brasília (DF), Comodoro (MT), Alto Paraíso de Goiás (GO), Carrancas (MG), Paranapiacaba (SP) e Santo Amaro (SC). As unidades de medida estão em graus para latitude, em milímetros para precipitação, e em graus Celsius para temperatura. A altitude está em metros.

Localização	Baturité (CE)	Caruaru (PE)	Itabaiana (SE)	Comodoro (MT)	Alto Paraíso de Goiás (GO)	Una (BA)	Brasília (DF)	Carrancas (MG)	Paranapiacaba (SP)	Santo Amaro - SC
	BTT	BJC	ITB	COM	CPV	BAH	BRA	CAR	PAR	STA
Latitude	-4.2628	-8.3692	-10.681	-13.66	-14.144	-15.181	-15.763	-21.496	-23.777	-27.718
Longitude	-38.9328	-36.0833	-37.4208	-59.7717	-47.4857	-39.0798	-47.9758	-44.706	-46.3079	-48.7793
BIO2	8,495833	8,764	8,1655	13,06583	10,664	6,94125	11,23108	12,2905	8,015583	8,74075
BIO5	30,377	27,418	28,143	29,478	27,423	26,815	27,239	31,755	27,889	25,192
BIO6	18,152	9,996	11,584	18,458	14,742	15,145	8,48	12,917	12,531	10,172
BIO13	197	219	249	179	106	300	320	322	288	338
BIO14	42	86	8	117	11	26	18	11	6	110
Altitude	178	209	1112	39	727	789	1075	612	1248	689

Fonte: O autor (2024).

Figura 1. Composição genômica de sequências repetitivas em populações de *Podocarpus sellowii* por meio de análise de clusterização.



Fonte: O autor (2024).

5. CONCLUSÕES

A análise do repeatoma de *Podocarpus sellowii* revelou um alto poder de informação desse tipo de dado, que apresentou um alto grau de polimorfismo mesmo analisado aqui no nível intraespecífico. Isso permitiu a caracterização detalhada da composição genômica repetitiva, destacando a predominância de elementos da classe Copia, em especial a linhagem Ivana, que apresenta maior abundância em todas as populações analisadas. Nossos dados sugerem que os *repeats* de *P. sellowii* não foram bons marcadores no estabelecimento de relações filogeográficas/estruturação populacional, fato esse evidenciado pela incongruência entre a topologia gerada pela abundância dos *repeats* e as filogenias baseadas em plastomas e DNAr.

A discrepância entre a árvore dos *repeats* e as redes filogenéticas sugere que a fração repetitiva do genoma pode estar mais fortemente influenciada por fatores ecológicos e pressões ambientais do que por relações de parentesco. Esse cenário é corroborado pelos resultados de correlação genoma-ambiente, que indicam que determinadas classes de *repeats*, como Ale, Athila e Ivana, estão significativamente associadas a variáveis climáticas, sugerindo uma pressão seletiva relevante exercida pelo ambiente sobre a abundância desses *repeats*.

Conclui-se que a evolução da fração repetitiva do genoma de *P. sellowii* é influenciada tanto por fatores ecológicos quanto por pressões seletivas locais, destacando a importância de abordagens integradas que considerem a interação entre variáveis ambientais e genômicas na diversificação e adaptação de espécies vegetais. Estudos mais aprofundados são necessários para esclarecer os mecanismos moleculares e ambientais que geraram essa plasticidade genômica e resiliência a nichos ecológicos específicos em plantas com disjunção relictual e divergências profundas.

REFERÊNCIAS

- AKAIKE, H. Maximum likelihood identification of Gaussian autoregressive moving average models. **Biometrika**, v. 60, n. 2, p. 255-265, 1973.
- ÁLVAREZ-YÉPIZ, J. C.; DOVČIAK, M.; BÚRQUEZ, A. Persistence of a rare ancient cycad: Effects of environment and demography. *Biological Conservation*, v. 144, n. 1, p. 122–130, 1 jan. 2011.
- ARAGÓN, E. *et al.*, Palaeogeography and palaeoenvironments of northern Patagonia from the Late Cretaceous to the Miocene: the Palaeogene Andean gap and the rise of the North Patagonian High Plateau. *Biological Journal of the Linnean Society*, v. 103, n. 2, p. 305–315, 2011.
- AVISE, J. C. *Phylogeography: retrospect and prospect*. 2009
- AVOLIO, Meghan L. *et al.* Measuring genetic diversity in ecological studies. **Plant Ecology**, v. 213, p. 1105-1115, 2012.
- BARTON, Kamil. MuMIn: multi-model inference. R package version 1.43. 17. 2020.
- BAUCOM, Regina S. *et al.* Natural selection on gene function drives the evolution of LTR retrotransposon families in the rice genome. **Genome Research**, v. 19, n. 2, p. 243-254, 2009.
- BENNETZEN, J. L. Patterns in grass genome evolution. *Current Opinion in Plant Biology*, v. 10, n. 2, p. 176–181, abr. 2007.
- BENNETZEN, Jeffrey L.; WANG, Hao. 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, 2014.
- BERENDSE, F.; SCHEFFER, M. The angiosperm radiation revisited, an ecological explanation for Darwin's 'abominable mystery'. *Ecology Letters*, v. 12, n. 9, p. 865–872, set. 2009.
- BIFFIN, E. *et al.*, Leaf evolution in Southern Hemisphere conifers tracks the angiosperm ecological radiation. *Proceedings of the Royal Society B: Biological Sciences*, v. 279, n. 1727, p. 341–348, 22 jan. 2012.
- BISCOTTI, Maria Assunta; OLMO, Ettore; HESLOP-HARRISON, J. S. Repetitive DNA in eukaryotic genomes. **Chromosome Research**, v. 23, p. 415-420, 2015.
- BISWAS, C.; JOHRI, B. M. *The Gymnosperms*. Berlin Heidelberg: Springer-Verlag, 1997.
- BOGAN, Samuel N.; YI, Soojin V. Potential role of DNA methylation as a driver of plastic responses to the environment across cells, organisms, and populations. **Genome biology and evolution**, v. 16, n. 2, p. evae022, 2024.

BOND, W. J. The tortoise and the hare: ecology of angiosperm dominance and gymnosperm persistence. *Biological Journal of the Linnean Society*, v. 36, n. 3, p. 227–249, 1989.

BORGES, Rui *et al.* Measuring phylogenetic signal between categorical traits and phylogenies. **Bioinformatics**, v. 35, n. 11, p. 1862-1869, 2019.

BOURGEOIS, Yann *et al.* Disentangling the determinants of transposable elements dynamics in vertebrate genomes using empirical evidences and simulations. **PLoS genetics**, v. 16, n. 10, p. e1009082, 2020.

BOURGEOIS, Yann XC; WARREN, Ben H. An overview of current population genomics methods for the analysis of whole-genome resequencing data in eukaryotes. **Molecular Ecology**, v. 30, n. 23, p. 6036-6071, 2021.

BOURGEOIS, Yann; BOISSINOT, Stéphane. On the population dynamics of junk: a review on the population genomics of transposable elements. **Genes**, v. 10, n. 6, p. 419, 2019.

BOURQUE, Guillaume *et al.* Ten things you should know about transposable elements. **Genome biology**, v. 19, p. 1-12, 2018.

BRENNER, E. D. *et al.*, EST analysis in Ginkgo biloba: an assessment of conserved developmental regulators and gymnosperm specific genes. *BMC genomics*, v. 6, p. 143, 15 out. 2005.

BRODRIBB, T. J.; FEILD, T. S. Leaf hydraulic evolution led a surge in leaf photosynthetic capacity during early angiosperm diversification. *Ecology Letters*, v. 13, n. 2, p. 175–183, fev. 2010.

BRODRIBB, T. J.; HILL, R. S. Southern Conifers in Time and Space. *Australian Journal of Botany*, v. 47, n. 5, p. 639, 1999.

BRODRIBB, Tim J.; FEILD, Taylor S.; SACK, Lawren. Viewing leaf structure and evolution from a hydraulic perspective. **Functional Plant Biology**, v. 37, n. 6, p. 488-498, 2010.

BRODRIBB, Timothy; HILL, R. S. The importance of xylem constraints in the distribution of conifer species. **The New Phytologist**, v. 143, n. 2, p. 365-372, 1999.

BRUMMITT, N. A. *et al.*, Green Plants in the Red: A Baseline Global Assessment for the IUCN Sampled Red List Index for Plants. *PLOS ONE*, v. 10, n. 8, p. e0135152, 7 ago. 2015.

BUERMANS, H. P. J.; DEN DUNNEN, J. T. Next generation sequencing technology: advances and applications. **Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease**, v. 1842, n. 10, p. 1932-1941, 2014.

BUREŠ, Petr *et al.* The global distribution of angiosperm genome size is shaped by climate. **New Phytologist**, v. 242, n. 2, p. 744-759, 2024.

CAIN, Amy K. *et al.* A decade of advances in transposon-insertion sequencing. **Nature Reviews Genetics**, v. 21, n. 9, p. 526-540, 2020.

CALCAGNO, Vincent; DE MAZANCOURT, Claire. glmulti: an R package for easy automated model selection with (generalized) linear models. **Journal of statistical software**, v. 34, p. 1-29, 2010.

CANAPA, A. *et al.* Shedding light upon the complex net of genome size, genome composition and environment in chordates. **The European Zoological Journal**, v. 87, n. 1, p. 192-202, 2020.

CARDOSO, Thiago Kossmann *et al.* Diversity and conservation of macrofungi from the Aparados da Serra with an emphasis on forest ecosystems. 2022.

CASTRO, Natália *et al.* Repeatoma evolution across space and time: Unravelling repeats dynamics in the plant genus *Erythrostemon* Klotzsch (Leguminosae Juss). **Molecular Ecology**, p. e17510, 2024.

CAVERS, S.; DICK, C. W. Phylogeography of Neotropical trees. *Journal of Biogeography*, v. 40, n. 4, p. 615–617, abr. 2013.

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, 2009.

CHAW, S. M. *et al.*, Molecular phylogeny of extant gymnosperms and seed plant evolution: analysis of nuclear 18S rRNA sequences. *Molecular Biology and Evolution*, v. 14, n. 1, p. 56–68, jan. 1997.

CHÉNAIS, Benoît *et al.* The impact of transposable elements on eukaryotic genomes: from genome size increase to genetic adaptation to stressful environments. **Gene**, v. 509, n. 1, p. 7-15, 2012.

CHÉNAIS, Benoît *et al.* The impact of transposable elements on eukaryotic genomes: from genome size increase to genetic adaptation to stressful environments. **Gene**, v. 509, n. 1, p. 7-15, 2012.

CHÉNAIS, Benoît. Transposable elements in cancer and other human diseases. **Current cancer drug targets**, v. 15, n. 3, p. 227-242, 2015.

CHRISTENHUSZ, M. J. M.; BYNG, J. W. The number of known plants species in the world and its annual increase. *Phytotaxa*, v. 261, n. 3, p. 201–217, 20 maio 2016a.

CHUMOVÁ, Zuzana *et al.* The relationship between transposable elements and ecological niches in the Greater Cape Floristic Region: A study on the genus *Pteronia* (Asteraceae). **Frontiers in Plant Science**, v. 13, p. 982852, 2022.

CLAPPERTON, C. M. Quaternary Geology and Geomorphology of South America. 1st edition ed. Amsterdam ; New York: Elsevier Science, 1993.

CNCFLORA. *Podocarpus sellowii*. Rio de Janeiro: Centro Nacional de Conservação da Flora, 2024. Disponível em: <https://cncflora.jbrj.gov.br/>. Acesso em: 4 set. 2024.

COLINVAUX, P. A. *et al.*, A Long Pollen Record from Lowland Amazonia: Forest and Cooling in Glacial Times. *Science*, v. 274, n. 5284, p. 85–88, 4 out. 1996.

COUPER R. A.; PANTIN CARL FREDERICK ABEL. Southern Hemisphere Mesozoic and Tertiary Podocarpaceae and Fagaceae and their palaeogeographic significance. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, v. 152, n. 949, p. 491–500, 12 jul. 1960.

CRANE, P. R. Vegetational consequences of the angiosperm diversification. In: Friis EM, Chaloner WG, Crane PR, eds. *The origins of the angiosperms and their biological consequences*. Cambridge, UK: Cambridge University Press, 107–144. 1987

DANTAS, Liliane G. *et al.*, Low genetic diversity and high differentiation among relict populations of the neotropical gymnosperm *Podocarpus sellowii* (Klotz.) in the Atlantic Forest. **Genetica**, v. 143, p. 21-30, 2015.

DE SOUSA, Valderês Aparecida *et al.* Genetic diversity and biogeographic determinants of population structure in *Araucaria angustifolia* (Bert.) O. Ktze. **Conservation Genetics**, v. 21, n. 2, p. 217-229, 2020.

DELSUC, Frédéric; BRINKMANN, Henner; PHILIPPE, Hervé. Phylogenomics and the reconstruction of the tree of life. **Nature Reviews Genetics**, v. 6, n. 5, p. 361-375, 2005.

DODSWORTH, S. *et al.*, Genome-wide *repeat* dynamics reflect phylogenetic distance in closely related allotetraploid *Nicotiana* (Solanaceae). *Plant Systematics and Evolution*, v. 303, n. 8, p. 1013–1020, out. 2017.

DODSWORTH, S. *et al.*, Potential of Herbariomics for Studying Repetitive DNA in Angiosperms. *Frontiers in Ecology and Evolution*, v. 6, 29 out. 2018.

DODSWORTH, S. *et al.*, Using genomic *repeats* for phylogenomics: a case study in wild tomatoes (*Solanum* section *Lycopersicon*: Solanaceae). *Biological Journal of the Linnean Society*, v. 117, n. 1, p. 96–105, 1 jan. 2016.

DODSWORTH, Steven *et al.* Genomic *repeat* abundances contain phylogenetic signal. **Systematic biology**, v. 64, n. 1, p. 112-126, 2015.

DODSWORTH, Steven *et al.* Hyb-Seq for flowering plant systematics. **Trends in Plant Science**, v. 24, n. 10, p. 887-891, 2019.

DODSWORTH, Steven. Genome skimming for next-generation biodiversity analysis. **Trends in Plant Science**, v. 20, n. 9, p. 525-527, 2015.

DRUMMOND, Alexei J.; RAMBAUT, Andrew. BEAST: Bayesian evolutionary analysis by sampling trees. **BMC evolutionary biology**, v. 7, n. 1, p. 1-8, 2007.

DU, Fang K.; PETIT, Remy J.; LIU, Jian Quan. More introgression with less gene flow: chloroplast vs. mitochondrial DNA in the *Picea asperata* complex in China, and comparison with other conifers. **Molecular Ecology**, v. 18, n. 7, p. 1396-1407, 2009.

DU, Xin-Yu *et al.* Structural variation of plastomes provides key insight into the deep phylogeny of ferns. **Frontiers in Plant Science**, v. 13, p. 862772, 2022.

EISEN, J. A. Phylogenomics: improving functional predictions for uncharacterized genes by evolutionary analysis. *Genome Research*, v. 8, n. 3, p. 163–167, mar. 1998.

FARJON, Aljos. The Kew review: conifers of the world. **Kew Bulletin**, v. 73, p. 1-16, 2018.

FICK, Stephen E.; HIJMANS, Robert J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. **International journal of climatology**, v. 37, n. 12, p. 4302-4315, 2017.

FINET, C. *et al.*, Multigene Phylogeny of the Green Lineage Reveals the Origin and Diversification of Land Plants. *Current Biology*, v. 22, n. 15, p. 1456–1457, 7 ago. 2012.

FOREST, F. *et al.*, Gymnosperms on the EDGE. *Scientific Reports*, v. 8, n. 1, dez. 2018.

FU, Ning *et al.* Testing plastomes and nuclear ribosomal DNA sequences as the next-generation DNA barcodes for species identification and phylogenetic analysis in *Acer*. **BMC Plant Biology**, v. 24, n. 1, p. 445, 2024.

GARAMSZEGI, László Zsolt (Ed.). **Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice**. Springer, 2014.

GARCIA, Lucinda Carneiro. Aspectos morfo-anatômicos e tolerância à dessecação de sementes de *Podocarpus lambertii* Klotz. e *Podocarpus sellowii* Klotz.(Podocarpaceae). 2003.

GARCIA, Lucinda Carneiro; NOGUEIRA, Antonio Carlos. Answer from the *Podocarpus lambertii* and *Podocarpus sellowii*–(Podocarpaceae) seeds to desiccation. **Ciência Florestal**, v. 18, n. 3, p. 347-352, 2008.

GAZDA, Małgorzata A. *et al.* A genetic mechanism for sexual dichromatism in birds. **Science**, v. 368, n. 6496, p. 1270-1274, 2020

GERNANDT, D. S.; WILLYARD, A.; SYRING, J. V.; LISTON, A. The conifers (Pinophyta). In C. Plomion, J. Bousquet, & C. Koe (Eds.), *Genetics, genomics and breeding of conifers* (pp. 1–39). New Hampshire: Science Publishers. 2011

GONÇALVES, F. R.; VIEIRA, F. A.; CARVALHO, D. Naturally fragmented but not genetically isolated populations of *Podocarpus sellowii* Klotzsch (Podocarpaceae) in southeast Brazil. **Genetics and Molecular Research**, v. 15, n. 4, p. 1-17, 2016.

GONZÁLEZ, Graciela Esther; POGGIO, Lidia. Intragenomic conflict between knob heterochromatin and B chromosomes is the key to understand genome size variation along altitudinal clines in maize. **Plants**, v. 10, n. 9, p. 1859, 2021.

HABERER, Georg *et al.* European maize genomes highlight intraspecies variation in repeat and gene content. **Nature Genetics**, v. 52, n. 9, p. 950-957, 2020.

HAMMER, Øyvind; HARPER, David AT. Past: paleontological statistics software package for education and data analysis. **Palaeontologia electronica**, v. 4, n. 1, p. 1, 2001.

HARTLEY, Gabrielle; O'NEILL, Rachel J. Centromere *repeats*: hidden gems of the genome. **Genes**, v. 10, n. 3, p. 223, 2019.

HARTLEY, Gabrielle; O'NEILL, Rachel J. Centromere repeats: hidden gems of the genome. **Genes**, v. 10, n. 3, p. 223, 2019.

HÉNAULT, Mathieu *et al.* The genomic landscape of transposable elements in yeast hybrids is shaped by structural variation and genotype-specific modulation of transposition rate. **Elife**, v. 12, p. RP89277, 2024.

HÉNAULT, Mathieu *et al.* The genomic landscape of transposable elements in yeast hybrids is shaped by structural variation and genotype-specific modulation of transposition rate. **Elife**, v. 12, p. RP89277, 2024.

HICKERSON, M. J. *et al.*, Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution*, v. 54, n. 1, p. 291–301, jan. 2010.

HIJMANS, Robert J. *et al.* Very high resolution interpolated climate surfaces for global land areas. **International Journal of Climatology: A Journal of the Royal Meteorological Society**, v. 25, n. 15, p. 1965-1978, 2005.

HILL, Robert S.; BRODRIBB, Tim J. Southern conifers in time and space. **Australian Journal of Botany**, v. 47, n. 5, p. 639-696, 1999.

HOLLAND, Barbara; MOULTON, Vincent. Consensus networks: a method for visualising incompatibilities in collections of trees. In: **International workshop on algorithms in bioinformatics**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003. p. 165-176.

HOLLINGSWORTH, Peter M. Refining the DNA barcode for land plants. **Proceedings of the National Academy of Sciences**, v. 108, n. 49, p. 19451-19452, 2011.

HOORN, C. *et al.*, Amazonia Through Time: Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity. *Science*, v. 330, n. 6006, p. 927–931, 12 nov. 2010.

HORVÁTH, Vivien; MERENCIANO, Miriam; GONZÁLEZ, Josefa. Revisiting the relationship between transposable elements and the eukaryotic stress response. **Trends in Genetics**, v. 33, n. 11, p. 832-841, 2017.

HOUSWORTH, Elizabeth A.; MARTINS, Emilia P.; LYNCH, Michael. The phylogenetic mixed model. **The American Naturalist**, v. 163, n. 1, p. 84-96, 2004.

HURVICH, Clifford M.; TSAI, Chih-Ling. Regression and time series model selection in small samples. **Biometrika**, v. 76, n. 2, p. 297-307, 1989.

HUSON, Daniel H.; BRYANT, David. Application of phylogenetic networks in evolutionary studies. **Molecular biology and evolution**, v. 23, n. 2, p. 254-267, 2006.

IGANCI, J.R.V.; DORNELES, M.P. *Araucariaceae in Flora e Funga do Brasil*. Jardim Botânico do Rio de Janeiro. Disponível em: <<https://floradobrasil.jbrj.gov.br/FB33971>>. Acesso em: 07 ago. 2024

ILER, Amy M. *et al.* Demographic consequences of phenological shifts in response to climate change. **Annual Review of Ecology, Evolution, and Systematics**, v. 52, n. 1, p. 221-245, 2021.

ITO, Hidetaka *et al.* An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. **Nature**, v. 472, n. 7341, p. 115-119, 2011.

JOHNSON, Norman A. Speciation: genomic sequence data and the biogeography of speciation. *National Science Review*, v. 9, n. 12, p. nwac294, 2022.

JOST, Matthias *et al.* Structural plastome evolution in holoparasitic Hydnoraceae with special focus on inverted and direct *repeats*. **Genome biology and evolution**, v. 14, n. 6, p. evac077, 2022.

KAMILAR, Jason M.; COOPER, Natalie. Phylogenetic signal in primate behaviour, ecology and life history. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 368, n. 1618, p. 20120341, 2013.

KARGER, Dirk N. *et al.* CHELSA-W5E5 v1. 0: W5E5 v1. 0 downscaled with CHELSA v2. 0. 2021.

KAZAZIAN JR, Haig H. Mobile elements: drivers of genome evolution. **science**, v. 303, n. 5664, p. 1626-1632, 2004.

KEPPEL, G.; HODGSKISS, P. D.; PLUNKETT, G. M. Cycads in the Insular South-West Pacific: Dispersal or Vicariance? *Journal of Biogeography*, v. 35, n. 6, p. 1004–1015, 2008.

KHAN, Raees *et al.* Diversity, distribution, systematics and conservation status of Podocarpaceae. **Plants**, v. 12, n. 5, p. 1171, 2023.

KNIGHT, Charles A.; ACKERLY, David D. Variation in nuclear DNA content across environmental gradients: a quantile regression analysis. **Ecology Letters**, v. 5, n. 1, p. 66-76, 2002.

KNOPF, Patrick *et al.* Relationships within Podocarpaceae based on DNA sequence, anatomical, morphological, and biogeographical data. **Cladistics**, v. 28, n. 3, p. 271-299, 2012.

LAUTERJUNG, Miguel Busarello *et al.* Phylogeography of Brazilian pine (*Araucaria angustifolia*): integrative evidence for pre-Columbian anthropogenic dispersal. **Tree Genetics & Genomes**, v. 14, p. 1-12, 2018.

LEDRU, M.-P. *et al.*, Regional assessment of the impact of climatic change on the distribution of a tropical conifer in the lowlands of South America: Tropical conifer and climatic change. *Diversity and Distributions*, v. 13, n. 6, p. 761–771, 26 jun. 2007.

LI, Hong-Tao *et al.* Plastid phylogenomic insights into relationships of all flowering plant families. **BMC biology**, v. 19, p. 1-13, 2021.

LIMA, T. V.; BARBOSA, D. C. A. Levantamento da distribuição espacial de plantas jovens não estabelecidas e estabelecidas de *Podocarpus sellowii* Klotzsch, em Serra dos Cavalos, Caruaru—PE. **Trabalho técnico, Imprensa Universitária, UFRPE, Recife**, v. 32, 1998.

LIN, C.-P. *et al.*, Comparative chloroplast genomics reveals the evolution of Pinaceae genera and subfamilies. *Genome Biology and Evolution*, v. 2, p. 504–517, 2010.

LOCKTON, Steven; GAUT, Brandon S. The evolution of transposable elements in natural populations of self-fertilizing *Arabidopsis thaliana* and its outcrossing relative *Arabidopsis lyrata*. *BMC Evolutionary Biology*, v. 10, p. 1-11, 2010.

LOSOS, Jonathan B. **Lizards in an evolutionary tree: ecology and adaptive radiation of anoles**. Univ of California Press, 2011.

LUPIA, R.; LIDGARD, S.; CRANE, P. R. Comparing palynological abundance and diversity: implications for biotic replacement during the Cretaceous angiosperm radiation. *Paleobiology*, v. 25, n. 3, p. 305–340, ed 1999.

LYNCH, Michael; WALSH, Bruce. **The origins of genome architecture**. Sunderland, MA: Sinauer associates, 2007.

MACAS, Jiří *et al.* In depth characterization of repetitive DNA in 23 plant genomes reveals sources of genome size variation in the legume tribe Fabaeae. **PloS one**, v. 10, n. 11, p. e0143424, 2015.

MAGALLÓN, S. Using fossils to break long branches in molecular dating: a comparison of relaxed clocks applied to the origin of angiosperms. *Systematic Biology*, v. 59, n. 4, p. 384–399, jul. 2010.

MATHEWS, S. Phylogenetic relationships among seed plants: Persistent questions and the limits of molecular data. *American Journal of Botany*, v. 96, n. 1, p. 228–236, jan. 2009.

MELO, Tiago Esposito Couceiro. Origem das disjunções sul-americanas de *Podocarpus sellowii* Klotzsch ex Endl (Podocarpaceae): filogeografia, modelagem de nicho e filogenômica. 2019. Tese (Doutorado em Biologia Vegetal) – Universidade Federal de Pernambuco, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; Orientador: Andrea Pedrosa Harand.

MICROSOFT CORPORATION. Microsoft Excel (Version 16.0). Redmond: Microsoft, 2021.

MILL, Robert R. Towards a biogeography of the Podocarpaceae. In: **IV International Conifer Conference 615**. 1999. p. 137-147.

MIOUSSE, Isabelle R. *et al.* Response of transposable elements to environmental stressors. **Mutation Research/Reviews in Mutation Research**, v. 765, p. 19-39, 2015.

MITCHELL, Sandra D. Perspectives, representation, and integration. **Understanding perspectivism: Scientific challenges and methodological prospects**, p. 178-193, 2020.

MITTELBACH, Gary G. *et al.* Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. **Ecology letters**, v. 10, n. 4, p. 315-331, 2007.

MITTEROECKER, Philipp; COLLYER, Michael L.; ADAMS, Dean C. Exploring Phylogenetic Signal in Multivariate Phenotypes by Maximizing Blomberg's K. *Systematic Biology*, p. syae035, 2024.

MORALES-CASTILLA, Ignacio *et al.* Inferring biotic interactions from proxies. **Trends in ecology & evolution**, v. 30, n. 6, p. 347-356, 2015.

MORGANTE, M.; DE PAOLI, E.; RADOVIC, S. Transposable elements and the plant pan-genomes. *Current Opinion in Plant Biology*, v. 10, n. 2, p. 149–155, abr. 2007.

MORLEY, R. J. Dispersal and Paleoecology of Tropical Podocarps. *Smithsonian Contributions to Botany*, n. 95, p. 21–41, 2011.

MORRONE, J. J. Biogeographic Areas and Transition Zones of Latin America and The Caribbean Islands Based on Panbiogeographic and Cladistic Analyses of the Entomofauna. *Annual Review of Entomology*, v. 51, n. 1, p. 467–494, jan. 2006.

MÜNKEMÜLLER, Tamara *et al.* How to measure and test phylogenetic signal. **Methods in Ecology and Evolution**, v. 3, n. 4, p. 743-756, 2012.

MUÑOZ-AMATRIAÍN, María *et al.* Distribution, functional impact, and origin mechanisms of copy number variation in the barley genome. **Genome biology**, v. 14, p. 1-17, 2013.

MURRAY, Brian G. Karyotype variation and evolution in gymnosperms. In: **Plant Genome Diversity Volume 2: Physical Structure, Behaviour and Evolution of Plant Genomes**. Vienna: Springer Vienna, 2012. p. 231-243.

NEALE, David B. *et al.* Noncoding and repetitive DNA. *The Conifers: Genomes, Variation And Evolution*, p. 61-74, 2019.

NEGI, Pooja; RAI, Archana N.; SUPRASANNA, Penna. Moving through the stressed genome: emerging regulatory roles for transposons in plant stress response. **Frontiers in plant science**, v. 7, p. 1448, 2016.

NIETO-BLÁZQUEZ, María Esther. **Historical biogeography of endemic plants in the Caribbean and *Podocarpus* as a case study**. 2019. Tese de Doutorado. Memorial University of Newfoundland.

NIETO-BLÁZQUEZ, María Esther; PEÑA-CASTILLO, Lourdes; RONCAL, Julissa. Historical biogeography of Caribbean *Podocarpus* does not support the progression rule. **Journal of Biogeography**, v. 48, n. 3, p. 690-702, 2021.

NOVÁK, P.; NEUMANN, P.; MACAS, J. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. *BMC bioinformatics*, v. 11, p. 378, 15 jul. 2010.

NOVÁK, Petr *et al.* RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. **Bioinformatics**, v. 29, n. 6, p. 792-793, 2013.

NOVÁK, Petr *et al.* TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. **Nucleic acids research**, v. 45, n. 12, p. e1111-e1111, 2017.

NOVÁK, Petr; NEUMANN, Pavel; MACAS, Jiří. Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. **Nature Protocols**, v. 15, n. 11, p. 3745-3776, 2020.

O'BRIEN, Stephen J.; STANYON, Roscoe. Ancestral primate viewed. **Nature**, v. 402, n. 6760, p. 365-366, 1999.

OHTA, Tomoko. The nearly neutral theory of molecular evolution. **Annual review of ecology and systematics**, p. 263-286, 1992.

OLIVEIRA, Maria Alice Silva *et al.*, High-throughput genomic data reveal complex phylogenetic relationships in *Stylosanthes* Sw (Leguminosae). **Frontiers in Genetics**, v. 12, p. 727314, 2021.

ORD, Terry J.; SUMMERS, Thomas C. Repeated evolution and the impact of evolutionary history on adaptation. **BMC Evolutionary Biology**, v. 15, p. 1-12, 2015.

ORME, David *et al.* The caper package: comparative analysis of phylogenetics and evolution in R. **R package version**, v. 5, n. 2, p. 1-36, 2013.

ORNELAS, Juan Francisco; RUIZ-SÁNCHEZ, Eduardo; SOSA, Victoria. Phylogeography of *Podocarpus matudae* (Podocarpaceae): pre-Quaternary relicts in northern Mesoamerican cloud forests. **Journal of Biogeography**, v. 37, n. 12, p. 2384-2396, 2010.

PAÇO, Ana; FREITAS, Renata; VIEIRA-DA-SILVA, Ana. Conversion of DNA sequences: from a transposable element to a tandem *repeat* or to a gene. **Genes**, v. 10, n. 12, p. 1014, 2019.

PAGE, Roderic DM. New Zealand and the new biogeography. **New Zealand Journal of Zoology**, v. 16, n. 4, p. 471-483, 1989.

PARADIS, Emmanuel. **Analysis of Phylogenetics and Evolution with R**. New York: Springer, 2012.

PASALA, Ratnakumar; PANDEY, Brij Bihari. Plant phenomics: High-throughput technology for accelerating genomics. **Journal of biosciences**, v. 45, n. 1, p. 111, 2020.

PAVOINE, Sandrine; RICOTTA, Carlo. Functional and phylogenetic similarity among communities. **Methods in Ecology and Evolution**, v. 5, n. 7, p. 666-675, 2014.

PENNINGTON, R. Toby *et al.* Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. **Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences**, v. 359, n. 1443, p. 515-538, 2004.

PIMPINELLI, Sergio; PIACENTINI, Lucia. Environmental change and the evolution of genomes: Transposable elements as translators of phenotypic plasticity into genotypic variability. **Functional Ecology**, v. 34, n. 2, p. 428-441, 2020.

PRICE, Morgan N.; DEHAL, Paramvir S.; ARKIN, Adam P. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. **Molecular biology and evolution**, v. 26, n. 7, p. 1641-1650, 2009.

QGIS DEVELOPMENT TEAM. QGIS Geographic Information System (Version 3.28). Open Source Geospatial Foundation Project, 2023. Disponível em: <https://qgis.org/>.

QUIROGA, M. Paula; PREMOLI, Andrea C. Genetic patterns in *Podocarpus parlatorei* reveal the long-term persistence of cold-tolerant elements in the southern Yungas. **Journal of Biogeography**, v. 34, n. 3, p. 447-455, 2007.

QUIROGA, María Paula *et al.* Molecular and fossil evidence disentangle the biogeographical history of *Podocarpus*, a key genus in plant geography. **Journal of Biogeography**, v. 43, n. 2, p. 372-383, 2016.

QUIROGA, María Paula *et al.*, Shrinking forests under warming: evidence of *Podocarpus parlatorei* (pino del cerro) from the subtropical Andes. **Journal of Heredity**, v. 103, n. 5, p. 682-691, 2012.

RAI, H. S. *et al.*, Inference of higher-order conifer relationships from a multi-locus plastid data set This paper is one of a selection of papers published in the Special Issue on Systematics Research. *Botany*, v. 86, n. 7, p. 658–669, 12 jun. 2008.

REINAR, William B. *et al.* Teleost genomic *repeat* landscapes in light of diversification rates and ecology. **Mobile DNA**, v. 14, n. 1, p. 14, 2023.

REVELL, Liam J. phytools: an R package for phylogenetic comparative biology (and other things). **Methods in ecology and evolution**, n. 2, p. 217-223, 2012.

ROGERS, Rebekah L.; SLATKIN, Montgomery. Excess of genomic defects in a woolly mammoth on Wrangel island. **PLoS genetics**, v. 13, n. 3, p. e1006601, 2017.

ROYAL BOTANIC GARDENS, K. State of the World's plants 2016.

ŠATOVIĆ-VUKŠIĆ, Eva; PLOHL, Miroslav. Satellite DNAs—From localized to highly dispersed genome components. **Genes**, v. 14, n. 3, p. 742, 2023.

SCHLEY, Rowan J. *et al.* The ecology of palm genomes: *repeat*-associated genome size expansion is constrained by aridity. **New Phytologist**, v. 236, n. 2, p. 433-446, 2022.

SCHRADER, Lukas; SCHMITZ, Jürgen. The impact of transposable elements in adaptive evolution. **Molecular Ecology**, v. 28, n. 6, p. 1537-1549, 2019.

SHAPIRO, J. A.; VON STERNBERG, R. Why repetitive DNA is essential to genome function. *Biological Reviews of the Cambridge Philosophical Society*, v. 80, n. 2, p. 227–250, maio 2005.

SIEPEL, A. Phylogenomics of primates and their ancestral populations. *Genome Research*, v. 19, n. 11, p. 1929–1941, nov. 2009.

SMITH, S. A.; BEAULIEU, J. M.; DONOGHUE, M. J. An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. *Proceedings of the National Academy of Sciences*, v. 107, n. 13, p. 5897–5902, 30 mar. 2010.

SMITH, Stephen A.; O'MEARA, Brian C. treePL: divergence time estimation using penalized likelihood for large phylogenies. **Bioinformatics**, v. 28, n. 20, p. 2689-2690, 2012.

SOLTIS, E. D.; SOLTIS, P. S. Contributions of plant molecular systematics to studies of molecular evolution. *Plant Molecular Biology*, v. 42, n. 1, p. 45–75, jan. 2000.

SYMONDS, Matthew RE; BLOMBERG, Simon P. A primer on phylogenetic generalised least squares. **Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice**, p. 105-130, 2014.

SZITENBERG, Amir *et al.* Genetic drift, not life history or RNAi, determine long-term evolution of transposable elements. **Genome biology and evolution**, v. 8, n. 9, p. 2964-2978, 2016.

TENAILLON, Maud I. *et al.* Genome size and transposable element content as determined by high-throughput sequencing in maize and *Zea luxurians*. **Genome biology and evolution**, v. 3, p. 219-229, 2011.

TONTI-FILIPPINI, J. *et al.*, What can we do with 1000 plastid genomes? *The Plant Journal*, v. 90, n. 4, p. 808–818, 1 maio 2017.

TURCHETTO-ZOLET, A. C. *et al.*, Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology*, v. 22, n. 5, p. 1193–1213, 2013.

TURNER, Benjamin L.; CERNUSAK, Lucas A. Ecology of the Podocarpaceae in tropical forests. **Smithsonian Contributions to Botany**, 2011.

UYEDA, Josef C.; ZENIL-FERGUSON, Rosana; PENNELL, Matthew W. Rethinking phylogenetic comparative methods. **Systematic Biology**, v. 67, n. 6, p. 1091-1109, 2018.

VAN-LUME, Brena 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, 2017.

VASCONCELLOS, Mariana M. *et al.* Evaluating the impact of historical climate and early human groups in the Araucaria Forest of eastern South America. **Ecography**, p. e06756, 2024.

VENNER, Samuel; FESCHOTTE, Cédric; BIEMONT, Christian. Dynamics of transposable elements: towards a community ecology of the genome. **Trends in Genetics**, v. 25, n. 7, p. 317-323, 2009.

VIEIRA, L. DO N. *et al.*, The Complete Chloroplast Genome Sequence of *Podocarpus lambertii*: Genome Structure, Evolutionary Aspects, Gene Content and SSR Detection. PLOS ONE, v. 9, n. 3, p. e90618, 4 mar. 2014.

WANG, Jie *et al.* The pan-plastome of *Prunus mume*: insights into *Prunus* diversity, phylogeny, and domestication history. **Frontiers in Plant Science**, v. 15, p. 1404071, 2024.

WEISS-SCHNEEWEISS, Hanna *et al.* Employing next generation sequencing to explore the *repeat* landscape of the plant genome. **Next generation sequencing in plant systematics. Regnum Vegetabile**, v. 157, p. 155-179, 2015.

WELLS, Jonathan N.; FESCHOTTE, Cédric. A field guide to eukaryotic transposable elements. **Annual review of genetics**, v. 54, n. 1, p. 539-561, 2020.

WICKER, Thomas *et al.* Impact of transposable elements on genome structure and evolution in bread wheat. **Genome biology**, v. 19, p. 1-18, 2018.

WILF, P.; ESCAPA, I. H. Green Web or megabiased clock? Plant fossils from Gondwanan Patagonia speak on evolutionary radiations. *New Phytologist*, v. 207, n. 2, p. 283–290, 2015.

WU, Jing-Yu *et al.* Fossil leaves of *Podocarpus* subgenus *Foliolatus* (Podocarpaceae) from the Pliocene of southwestern China and biogeographic history of *Podocarpus*. **Review of Palaeobotany and Palynology**, v. 287, p. 104380, 2021.

XIAO, L.-Q.; MÖLLER, M.; ZHU, H. High nDNAr ITS polymorphism in the ancient extant seed plant *Cycas*: incomplete concerted evolution and the origin of pseudogenes. *Molecular Phylogenetics and Evolution*, v. 55, n. 1, p. 168–177, abr. 2010.

ZGURSKI, J. M. *et al.*, How well do we understand the overall backbone of cycad phylogeny? New insights from a large, multigene plastid data set. *Molecular Phylogenetics and Evolution*, v. 47, n. 3, p. 1232–1237, jun. 2008.

ZHANG, Lei *et al.* Plastid genome data provide new insights into the dynamic evolution of the tribe Ampelopsideae (Vitaceae). *BMC genomics*, v. 25, n. 1, p. 247, 2024.