Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Chromosome-level genomes of a Neotropical mountain bromeliad reveal structural and gene family variation between contrasting elevations

Paulo Aecyo

pauloaecyo_1997@hotmail.com

Universidade Estadual de Campinas (UNICAMP)

Marilia Tavares

Universidade Estadual de Campinas (UNICAMP)

Jacqueline Mattos

Universidade Estadual de Campinas (UNICAMP)

Lucas Costa

Universidade Estadual de Campinas (UNICAMP)

Fernando Della-Rosa

Universidade de São Paulo

Alessandro Varani

Universidade Estadual Paulista (Unesp)

Drielli Canal

Universidade Estadual de Campinas (UNICAMP)

Fabio Pinheiro

Universidade Estadual de Campinas (UNICAMP)

Cleber Chaves

Universidade Estadual de Campinas (UNICAMP)

Aureliano Bombarely

Consejo Superior de Investigaciones Científicas

Diego Riaño-Pachón

Universidade de São Paulo

Clarisse Palma-Silva

Universidade Estadual de Campinas (UNICAMP)

Research Article

Keywords: gene ontology enrichment, local adaptation, neotropical biodiversity, structural variations

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11529366/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Chromosome-level genomes of a Neotropical mountain bromeliad reveal structural and gene family variation between contrasting elevations

Paulo Aecyo[1,2]*, Marilia M. Tavares[1,2], Jacqueline Salvi de Mattos[1,2,3], Lucas Costa[1,2] Fernando Della-Rosa[4], Alessandro M. Varani[2,5], Drielli Canal[1,2], Fabio Pinheiro[1,2], Cleber Juliano Neves Chaves[1,2], Aureliano Bombarely[6]*, Diego M. Riaño-Pachón[4,7], Clarisse Palma-Silva[1,2]

[1]Laboratory of Evolutionary Ecology and Genomics of Plants, Departamento de Botânica, Programa de Pós-Graduação em Biologia Vegetal, Instituto de Biologia, Universidade Estadual de Campinas, Campinas, Brazil

[2]Center for Research on Biodiversity Dynamics and Climate Change (CBioClima), Institute of Biosciences, São Paulo State University—UNESP, Av. 24A, 1515, Rio Claro 13506-900, São Paulo, Brazil

[3]Instituto Tecnológico Vale Desenvolvimento Sustentável, Belém - PA, Brazil 66055-090

[4]Laboratory of Computational, Evolutionary and Systems Biology, Center for Nuclear Energy in Agriculture, University of São Paulo, Piracicaba, Brazil

[5]Laboratory of Genomics and Bioinformatics, School of Agricultural and Veterinarian Sciences, Department of Agricultural and Environmental Biotechnology, São Paulo State University, Jaboticabal, São Paulo, Brazil

[6]Laboratory of Bioinformatics and Evolutionary Genomics, Instituto de Biologia Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas-Universidad Politécnica de Valencia (CSIC-UPV), Valencia, Spain

[7]National Institute for Science and Technology of Bioethanol (INCT - Bioetanol)

*Correspondence authours:
Paulo Aecyo - pauloaecyo_1997@hotmail.com
Aureliano Bombarely - abombarely@ibmcp.upv.es

Contact Information:
Paulo Aecyo - https://orcid.org/0000-0001-9254-5603 - pauloaecyo_1997@hotmail.com
Marilia M. Tavares - https://orcid.org/0000-0003-3199-5576 - m.manuppella@gmail.com
Jacqueline Salvi de Mattos - https://orcid.org/0000-0003-3104-0884 - jsalvidemattos@gmail.com
Lucas Costa - https://orcid.org/0000-0002-4798-097X - lucas.costa.18@hotmail.com
Alessandro M. Varani - https://orcid.org/0000-0002-8876-3269 - alessandro.varani@unesp.br
Drielli Canal - https://orcid.org/0000-0002-5771-3912 - drielliufes@gmail.com
Fabio Pinheiro - https://orcid.org/0000-0003-3243-2652 - biopin@unicamp.br
Cleber Juliano Neves Chaves - https://orcid.org/0000-0002-5960-7304 - cleberchaves@gmail.com
Aureliano Bombarely - https://orcid.org/0000-0001-6257-8914 - abombarely@ibmcp.upv.es
Diego M. Riaño-Pachón - https://orcid.org/0000-0001-9803-3465 - diego.riano@cena.usp.br
Clarisse Palma-Silva - https://orcid.org/0000-0003-0192-5489 - cpalma@unicamp.br

Abstract

Background: Mountainous regions harbor a high level of biodiversity world wide with locally adapted and highly isolated populations that provide natural models for studying adaptive evolution, yet the genomic architecture of local adaptation remains largely unexplored in many non-model plant lineages. Here, we sequenced, assembled, and annotated high-quality chromosome-level genomes for two individuals of *Pitcairnia flammea* (Bromeliaceae) sampled from contrasting elevations in the Brazilian Atlantic Forest. By comparing these genomes, we investigated structural genomic and gene family evolution within a Neotropical mountain species, providing a framework for future studies of the genomic basis of environmental differentiation.

Results: Both genomes showed similar genomic composition and PSMC-inferred demographic history, differing mainly in the relative abundance of some transposable element lineages. We identified six large chromosomal inversions, supported by independent long-read mapping, encompassing genes annotated with putative functions related to heat acclimation, transpiration, and stomatal morphogenesis, although no significant functional enrichment was detected. Gene family analyses further revealed expansions in the high-elevation genome significantly enriched for functions related to metabolic regulation, phosphorylation-mediated signaling, and lipid metabolism, including sphingolipid metabolism. Together with previously documented ecophysiological differences between high- and low-elevation populations, these genomic features provide a basis for future population and functional studies of adaptation to elevation.

Conclusion: The two genomes reveal substantial intraspecific variation despite broad syntenic conservation. These genomic resources provide a foundation for future

comparative, population genomic, and functional studies testing the potential role of genomic variation in elevational adaptation, while also contributing to our understanding of bromeliad genome evolution in one of the world's most biodiverse biomes.

Key-words: gene ontology enrichment, local adaptation, neotropical biodiversity, structural variations

Background

Environmental gradients in mountain systems impose strong ecological challenges on plant populations and provide natural settings for investigating biological differentiation across elevation **[1, 2]**. Plants occurring at high elevations often exhibit physiological and morphological differences, including the activation of multiple metabolic pathways, reductions in height and biomass, and changes in lipid composition **[3, 4]**. Lipid metabolism, in particular, plays a crucial role in plant responses to cold stress, contributing to membrane stabilization and fluidity, as well as to stress signaling pathways in both plastidial and extraplastidial membranes **[4]**.

Neotropical mountain systems harbor exceptional biodiversity shaped by the interplay of environmental heterogeneity, geological history, and historical connectivity **[2, 5, 6]**. Among these systems, the ancient mountain ranges of the Brazilian Atlantic Forest (BAF) occur within a highly heterogeneous landscape extending from sea level to nearly 2,900 m and encompassing diverse vegetation types and environmental conditions **[7–9]**. Such environmental heterogeneity can promote phenotypic and genomic differentiation among populations and create conditions for local adaptation and ecological divergence. At the genomic level, differentiation among populations

may involve chromosomal rearrangements, transposable element dynamics, and the expansion or contraction of gene families **[10–14]**. Such genomic variation can arise through neutral and selective processes and may contribute to evolutionary responses to environmental heterogeneity. Despite the ecological and evolutionary importance of these mountain systems, intraspecific genomic variation across elevational gradients in the BAF remains poorly characterized. High-quality chromosome-level genome assemblies therefore provide an important framework for characterizing such variation and investigating its potential evolutionary significance **[15, 16]**.

Among the species distributed across the BAF *Pitcairnia flammea* Lindl. (Bromeliaceae) is distributed across a broad elevational gradient in the BAF, encompassing substantial environmental heterogeneity **[17]**. During the sporophyte phase individuals are diploid (2*n* = 50; **[18]**) with a variation in the haploid genome size from ~528 Mbp (unpublished) to ~826 Mbp **[19]**. Populations of this species exhibit marked differences in ecophysiological and morphological traits associated with ecological strategies adopted by populations across the elevational gradient **[20, 21]**. In this study, using PacBio HiFi, Omni-C, and RNA-seq data, we assembled and annotated chromosome-level genomes from single individuals of *P. flammea* sampled from low- and high-elevation populations. We then compared their genomic architecture, focusing on structural variation, transposable element composition, gene family evolution, and demographic history. The two genomes were highly syntenic but differed in the relative abundance of some transposable element lineages and contained six large chromosomal inversions encompassing genes annotated with putative functions related to heat acclimation, transpiration, and stomatal morphogenesis. Gene family analyses further revealed expansions in the high-elevation genome enriched for functions related to metabolic regulation and lipid

metabolism. Together, these genomic resources provide an initial characterization of intraspecific genomic variation between contrasting elevations and establish a foundation for future population-level and functional studies testing the potential contribution of genomic variation to elevational adaptation.

Methods

Genomic DNA isolation and sequencing

Two individuals of *Pitcairnia flammea*, one from a population in Morro do Rangel in Rio de Janeiro city (23°02'00.6" S 43°29'17.1" W; UEC214365) at 25 m above sea level (a.s.l) (hereafter ‘low-elevation’ individual; sample name RAN170), and other from a population at Pico dos Marins in Piquete city (22°29'56.4" S 45°07'57.4" W) at 2,000 m a.s.l (hereafter ‘high-elevation’ individual; sample name MAR120) were collected and cultivated for two years in common garden conditions (average temperature of ~25°C) at a greenhouse in the Universidade Estadual de Campinas (UNICAMP). Plant material was formally identified by Dr. Cleber Juliano Neves Chaves and Prof. Dr. Clarisse Palma-Silva. A voucher specimen of the low-elevation individual (RAN170) was deposited in the herbarium of the Universidade Estadual de Campinas (UEC), Campinas, Brazil, under the deposition number UEC214365. No voucher specimen was deposited for the high-elevation individual (MAR120). Both individuals were collected and are maintained alive in the experimental garden of UNICAMP.

Two grams of newly developed fresh leaves were collected, frozen in liquid nitrogen and kept at -80ºC, for each individual. High molecular weight DNA isolation, PacBio SMRTbell WGS library preparation and sequencing were performed by Novogene Inc. The libraries were sequenced on a PacBio Sequel II platform. To aid

scaffolding, we prepared a high-resolution chromatin conformation capture library using the Omni-C kit (Dovetail Genomics Inc.), sequenced on a NovaSeq Illumina, for the low-elevation individual.

Quality assessment and Genome assembly

Genome size and statistics were verified with Jellyfish v. 2.3.0 **[22]** and GenomeScope2 **[23]** using raw PacBio HiFi reads and k = 21. Genome assembly was performed with Hifiasm v. 0.19.8-r603 **[24]** using PacBio HiFi and Omni-C reads for the low-elevation individual, and only PacBio HiFi reads for the high-elevation individual. Redundant contigs arising from the separate assembly of alternative alleles in heterozygous regions were removed with Purge_dups v. 1.2.5 **[25]**. Genome contiguity metrics were estimated using QUAST v. 5.2.0 **[26]**. Completeness was assessed with BUSCO v. 5.0 **[27]** using the Embryophyta and Liliopsida datasets from OrthoDB v. 10 **[28]**. Assembly accuracy and completeness were further evaluated with Merqury v. 1.3 **[29]**, and the Long Terminal Repeat Assembly Index (LAI) was estimated using LTR_Retriever v. 2.9.7 **[30, 31]**. Finally, potential contamination was identified and removed with Blobtools v. 1.1.1 **[32, 33]**, using similarity searches performed with Diamond v. 2.1.8 **[34]** against the UniProt TrEMBL database **[35]**.

To achieve chromosome-level assemblies, Omni-C reads from the low-elevation individual were used to scaffold the genomes of both individuals with YAHS v. 1.2 **[36]**. Although both individuals belong to the same species, we acknowledge that using Hi-C data from one individual to guide the scaffolding of another could, in principle, impose the chromatin interaction landscape of the reference individual onto the second genome, potentially masking true structural variation (e.g., inversions or translocations) or, conversely, generating artifactual rearrangements if polymorphic

structural variation exists between the two individuals. To mitigate these risks and validate the inferred chromosome structures, we performed a series of independent checks. First, telomeric sequences were identified with TIDK v. 0.2.41 **[37]** and used to guide manual curation of chromosome ends, ensuring that terminal contigs are correctly oriented and that no mis-joins are present at chromosome termini. Manual curation was performed using the Pretext suite (PretextMap, PretextGraph, and PretextView v. 0.1.9) to resolve any inconsistencies in the contact maps. Latter, breakpoints among the detected chromosomal rearrangements (see Synteny analysis section) were confirmed by mapping the long-read (PacBio) data from each individual against the genome assembly of the other using Minimap2 **[38]**, allowing us to identify regions where the long-read alignments support or contradict the scaffolding based on Omni-C data. All candidate breakpoints were then manually inspected in the Integrative Genomics Viewer (IGV v. 2.17.01) **[39, 40]** to verify that they are supported by multiple high-quality, non-chimeric alignments and to exclude regions showing anomalous mapping patterns indicative of scaffolding errors. The quality of the scaffolded assemblies was reassessed using QUAST, BUSCO, Merqury, and LTR Retriever to confirm high completeness and low error rates. By combining long-read validation with manual curation, we minimized the risk of imposing an incorrect structure onto the high-elevation individual while retaining the ability to detect genuine structural variants where long-read data provide independent support.

RNA isolation, sequencing and assembly

Total RNA was isolated from three individuals from the low-elevation population (including the sequenced individual). To have a wide view of different transcripts, we isolated the RNA by immediately freezing fresh tissues from three different sources

including newly developed fresh leaves, a mix of flowers from three different stages (3 cm floral bud, 6 cm floral bud and fully open flower) and unripe fruit.

The total RNA for each organ were isolated and sequenced separately. The RNeasy Plant Mini Kit (Qiagen) was used for total RNA isolation. RNA integrity was measured on a 1% agarose gel and its concentration was estimated using the Qubit4 fluorometer Invitrogen (Thermo Fisher Scientific, USA). RNA samples were sent to Macrogen Inc. (South Korea) to perform the RNA library construction using the TruSeq stranded mRNA library kit. A total of six libraries were constructed and 2x150 bp were sequenced on an Illumina NovaSeq 6000 platform.

Illumina raw reads were processed with the fastq-mcf tool v. 1.04.676 **[41, 42]** to remove adapter sequences, discard reads shorter than 50 bases and filter low-quality regions (QC < 30). For genome-guided assembly, STAR v. 2.7.11a **[43]** was used to map the sequences against each genome. StringTie v. 2.2.1 **[44]**, was used for transcriptome assembly, and TransDecoder (RRID:SCR_017647) v5.7.0 was used to identify putative protein-coding transcripts **[45]**.

Structural and functional annotation

To identify and classify the Transposable Elements (TE) sequences in the genome, we used the Plant_Annotation_TEs pipeline **[46]** that includes a modified version of EDTA **[30, 31]**, AnnoSINE (commit: 26301e9) **[47]** and MGEScan-non-LTR v3.0.0 **[48]**. This pipeline is tailored for non-model plant genomes and to remove potential false positive predictions to the identification and characterization of TE lineages. TEsorter v. 1.4.1 **[49]** was used for TE classification (following the nomenclature proposed by Orozco-Arias et al. **[50]**) and validation . This pipeline was also used to detect the Repeat Landscape, to estimate the time of insertion of LTR

elements and to investigate the evolutionary history of those elements. Further, the pipeline outputs a soft-masked genome file that was used for gene prediction.

Protein-coding genes were predicted following the protocol of Alves et al. **[46]**, which uses the soft-masked genome sequence to annotate the genes through the integration of gene predictors, combination tools and functional annotation software, including GeMoMa v. 1.9 **[51]**, BRAKER (RRID:SCR_018964) v. 3.0.4 **[52]**, EVidence Modeler v. 2.1.0 **[53]**, PASA **[53]**, and BLAST2GO Basic v6.0 **[54]**. The publicly available data from *Ananas comosus* (L.) Merr (GCA_001540865.1), *Tillandsia fasciculata* Sw. (GCA_029168755.2) and *T. leiboldiana* Schltdl. (GCA_029204045.2) were used as input for protein homology analysis. The output of each program was evaluated by BUSCO v. 5.0 **[27]** using the Embryophyta and Liliopsida lineages from OrthoDB v. 10 **[28]**.

The functional annotation was performed following the protocol by Alves et al. **[46]**. A BLAST search was performed against public plant databases including UniProt, NCBI RefSeq, PlantTFDB v5.0 **[55]**, NLR Atlas **[56]** and PRGdb v4.0 **[57]**. Gene Ontology, Enzyme Commission number (ECs), Carbohydrate-Active enZYme (CAZyme), PFAM, and InterPro IDs for each gene and detected isoform using EggNOG-mapper v2.1.4-2-4-gb493df4 **[58]** and InterProScan v5.61-93.0 **[59]**. PlantiSMASH v. 1.0 **[60]** was used to detect metabolic gene clusters. Blast2GO Basic v. 6.0 **[54]** was used to filter GOs irrelevant for Viridiplantae. The AGAT (Another Gtf/Gff Analysis Toolkit) package v1.2.0 (https://github.com/NBISweden/AGAT) was used to generate the final GFF3.

Gene family evolution

Orthogroup inference was performed with OrthoFinder v2.5.5 **[61]** using the longest protein isoform per gene from the deduced proteomes of both *Pitcairnia flammea* individuals and five chromosome-level Bromeliaceae genomes with publicly available structural annotations: *Ananas comosus* (L.) Merr (GCA_001540865.1, **[62]**), *Tillandsia fasciculata* Sw. (GCA_029168755.2, **[63]**), *T. leiboldiana* Schltdl. (GCA_029204045.2, **[63]**), *T. duratii* Vis. (GCA_042479765.1, **[64]**) and *Vriesea erythrodactylon* (É.Morren) É.Morren ex Mez (GCA_042480485.1, **[64]**). OrthoFinder was run with an e-value cut-off of $1e^{-5}$ and the default MCL inflation value of 1.5. A species tree was inferred from the concatenated alignment of all single-copy orthologs generated by the OrthoFinder pipeline, which uses MAFFT for alignment and IQ-TREE v2.3.6 **[65]** for tree inference. To obtain an ultrametric tree suitable for CAFE5, branch lengths were converted to absolute time using the '--date' option in IQ-TREE, with calibration points derived from the TimeTree 5 database **[66]**. For gene family evolution analysis, CAFE5 v5 **[67]** was employed. Orthogroups with more than 100 gene copies in any species were excluded globally to reduce computational burden, as recommended by the CAFE5 developers. To account for potential errors in gene family size estimates caused by assembly or annotation issues, CAFE5 was run with the error model ('-e'). We first estimated a global λ (the rate of gene family evolution), then used the error model to estimate assembly error rates, and finally re-estimated λ separately for each of the three Bromeliaceae subfamilies (Bromelioideae, Tillandsioideae, and Pitcairnioideae) while incorporating the estimated error rates. Families with significant expansions or contractions ($p < 0.05$ after Benjamini–Hochberg correction) were identified separately for each *P. flammea* individual. Gene Ontology enrichment analysis of these families was performed with goatools **[68]**, using the complete set of

annotated genes in the *P. flammea* genome as background; GO term assignments were obtained from the GFF3 annotation files.

Metabolic pathways analysis

Among the expanded genes we identified the orthologous in *Oryza sativa* subsp. *japonica* IRGSP 1.0 (GCA_001433935.1, **[69]**) using Diamond v. 2.1.8 **[34]**. Later we used the ShinyGO v. 0.82 **[70]** to identify the metabolic pathways influenced by the expansion of gene families in the genome of *Pitcairnia flammea* individuals. We considered only the pathways with an FDR cutoff of 0.05.

Synteny analysis

In order to investigate gene collinearity and infer macrosyntenic relationships among the *P. flammea* genomes, we employed the GeneSpace v. 1.2.3 pipeline **[71]**. Protein fasta files and corresponding gene annotations (GFF3 format) were used as input. An all-versus-all similarity search was performed using Diamond v. 2.1.12 **[34]** with an e-value cutoff of 1e-5 to identify homologous gene pairs across genomes. Syntenic blocks were then detected with MCScanX v. 1.0.0 software **[72]**, which identifies collinear gene pairs based on both sequence similarity and conserved gene order. Since we are working with genomes from individuals of the same species, which are expected to present variation only at finer scales, we lowered the default settings for minimum number of genes to define a syntenic block (from five to two). Syntenic orthogroups were inferred by integrating pairwise syntenic gene anchors across all genomes into a pangenomic synteny graph, highlighting inversion events.

PSMC inference of historical effective population size

For the inference of historical variation in effective population sizes (Ne) we used the Pairwise Sequential Markovian Chain (PSMC) model with the assembly of both individuals. We used Minimap2 v. 2.26 (r1175) **[38]** to align the PacBio HiFi sequences to the genome of each species, and then we used PSMC **[73]** with default parameters. To plot the graph with the estimation of time and Ne size, we used the *Arabidopsis thaliana* (L.) Heynh. substitution rate (μ) of 6 x 10^{-9} and a generation time of 5 years (personal observation).

Results

Evaluation of both Pitcairnia flammea *assembled genomes*

A total of 32 Gbp (48× coverage) and 27 Gbp (39× coverage) of PacBio HiFi sequences were generated for low-elevation and high-elevation individuals, respectively. Additionally, Omni-C data was also generated for the low-elevation individual, resulting in 37 Gbp (52× coverage) of sequences (Bioproject: PRJNA1345395). GenomeScope estimated a genome size of 476 Mbp for the low-elevation individual and 482 Mbp for the high-elevation individual (Fig S1A, B). Also, the high-elevation individual showed a higher heterozygosity (1.06) than the low-elevation individual (0.764).

We assembled a monoploid representation of the genome, with 1,039 contigs for the low-elevation individual and 491 for the high-elevation individual. The assemblies showed high completeness, with contig N50 values of 4.6 and 9.1 Mbp, and contained 98.70% and 99.10% of the Embryophyta single-copy BUSCO orthologs, respectively. Furthermore, both genomes contained 93.6% of the Liliopsida single-copy BUSCO orthologs (Table 1). After scaffolding and manual curation, we achieved chromosome-level genome assemblies, with telomere sequences at both

ends of all 25 pseudomolecules, except at one end of two chromosomes in the low-elevation individual (chromosome 13 and 25) and at one end of four chromosomes in the high-elevation individual (chromosomes 17, 20, 22 and 25) (Fig 1). The low-elevation individual presented 146 scaffolds, of which 23 comprised 90% of the total genome size (L90) of 498 Mbp. The N50 was 20 Mbp, with a Merqury completeness of 90.2% (Fig S1C) and 98.8% and 95.5% of Embryophyta and Liliopsida single-copy BUSCO ortholog genes, respectively. For the high-elevation individual, we recovered 146 scaffolds, with a L90 of 24 pseudomolecules and a total assembly size of 509 Mbp. The N50 was 19.06 Mbp, with a Merqury completeness of 87.63% (Fig S1D) and 99.1% and 95.60% of Embryophyta and Lilliopsida genes, respectively, detected among the assembled genomes (Table 1). These 25 pseudomolecules account for 97.9% (489,580,956 bp) and 94.1% (479,938,626 bp) of the total estimated genome size for the low-elevation and high-elevation individuals, respectively, with a LTR assembly index of 30.

Table 1. Statistics of the contigs and final genome assembly for both individuals of *Pitcairnia flammea* from populations distributed at low-elevation and at 2,000 m a.s.l.

	Low-elevation		**High-elevation**	
	Contigs	Scaffolds	Contig	Scaffolds
Total genome size (Mbp)	531	498	535	509
Total number of contigs/scaffolds	1,039	146	491	146
Longest contig/scaffold (Mbp)	15.99	25.15	21.99	26.95
N50 (Mb)	4.60	20.00	9.10	19.06
N90 (Mbp)	11.79	15.18	3.07	15.18
L50	35	12	17	12
L90	109	23	54	24
Ns/100 Kpb	0	11.11	0	4.04

Complete BUSCO - Embryophyta (%)	98.70	98.80	99.00	99.10
Complete BUSCO - Liliopsida (%)	93.60	95.50	93.60	95.60
Completeness - Merqury (%)	90.32	90.23	87.86	87.63
LAI	14.95	29.54	13.85	30.31

341

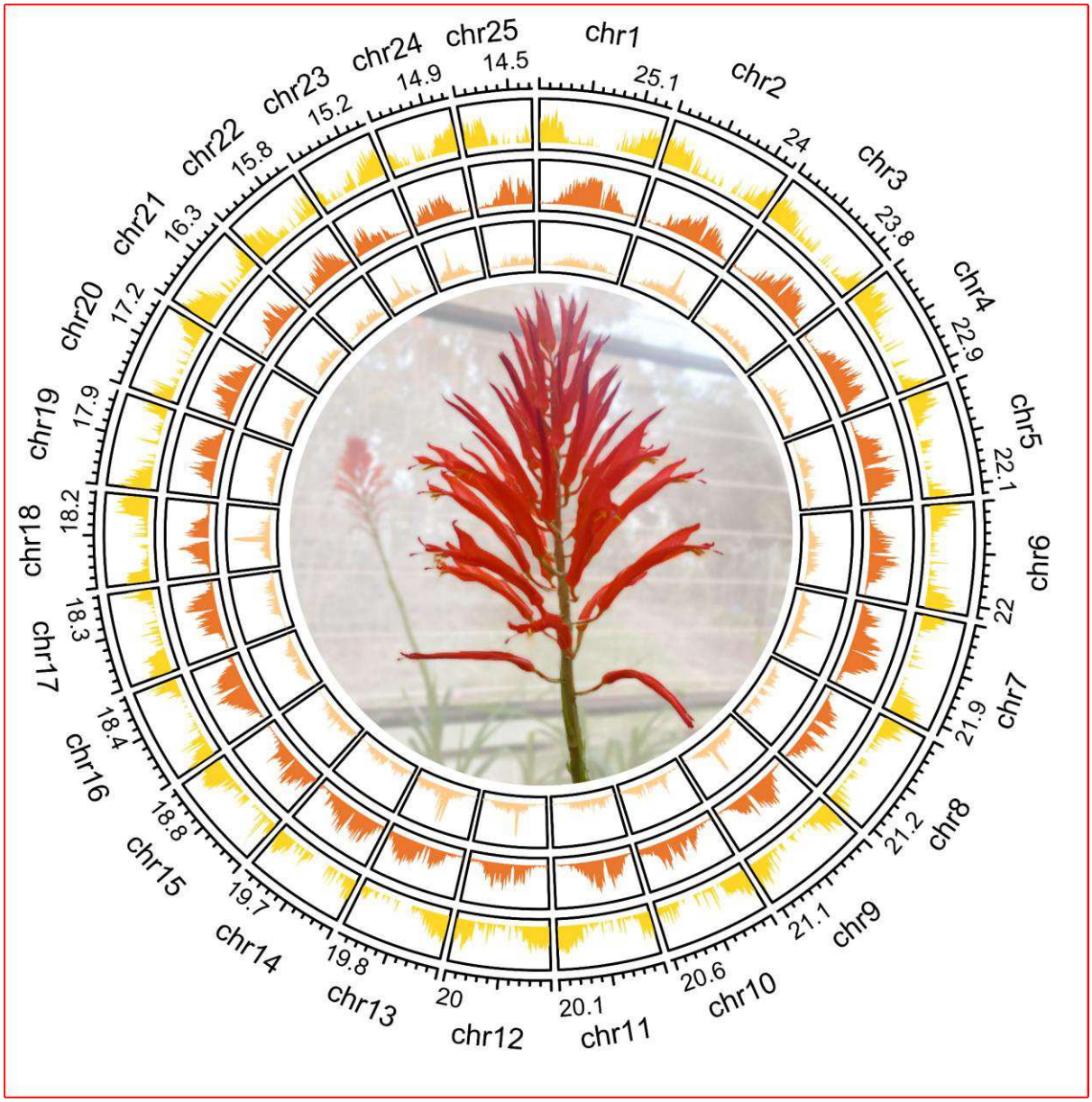

Fig 1. Circos plot showing the genomic features from the low-elevation individual of *Pitcairnia flammea*. The outer ruler indicates the size of the chromosome in megabase pairs (Mbp). Insider circles represent the estimated distribution and density: in yellow, Gene density; in red, LTR/Ty3 elements density; in orange, LTR/Copia elements density. The picture inside is the Inflorescence of a *P. flammea* individual cultivated in the green house at Universidade Estudadual de Campinas (UNICAMP).

Structural annotation and Spatial arrangement

We detected a total of 77.58% and 72.65% of TEs in the genome of the low-elevation and high-elevation individuals, respectively (Table 2). Both individuals exhibited a similar TE composition, with the LTR/Ty3 elements being the predominant group of repetitive elements present in the genome (25.12% in low-elevation and

20.56% in the high-elevation) (Fig 2). Six TEs lineages showed higher composition and similar expansion in both genomes. The LTR/Ty3/Chromovirus/Tekay lineage was the most prevalent, comprising 19.31% of the genome of the low-elevation individual and 15.97% of the genome of the high-elevation individual, followed by TR_GAG, Helitron, the non-autonomous LARDs element, LTR/Copia/SIRE and MITEs transposons (Fig. 2). Although the overall composition is similar, there are variations in the proportion of each lineage, particularly among the five main lineages that make up the majority of TE composition in both individuals.

Table 2. Statistics of the genome annotation for both individuals of *Pitcairnia flammea* from populations distributed at low-elevation (0,00 m a.s.l) and at high-elevation (2,000 m a.s.l).

Features		**Low-elevation**	**High-elevation**
Coding sequences			
	number of High Quality genes	23,181	23,737
	number of CDSs	57,780	55,158
	number of mRNAs with utr in both sides	40,194	33,901
	Number of mRNAs with at least one utr	43,309	37,414
	number of single exon genes	13,959	15,505
	number of overlapping genes	4,264	3,126
	mean gene length (bp)	3,839	3,573
	mean CDS length (bp)	1,161	1,088
	mean exons per mRNA	5.6	5.1
	mean introns in CDS per mRNA	4.2	3.8
	% of genome covered by protein coding genes	27.59	26.85
	% of genome covered by CDS	13.47	11.79
	BUSCO completeness - Embryophyta	98.8	99.1
	BUSCO completeness - Liliopsida	95.5	95.6
Transposable elements			
	% of genome covered by TEs	77.59	72.64
	% Total Class I Elements	61.27	61.09
	% Class I - LTR Ty3	25.12	20.56

% Class I - LTR Copia	17.66	14.04
% Class I - LTR nonautonomous	17.65	24.93
% Class I - non-LTR	0.54	0.50
% Total Class II Elements	16.47	12.08
% Class II - TIRs	6.34	2.51
% Class II - Helitron	10.13	9.57

Furthermore, when analyzing the timing of insertion of LTR/Copia and LTR/Ty3 elements in the genome of *Pitcairnia flammea*, a similar pattern is observed in both genomes. The increase in insertions of both LTR families began around 1.0 Mya. While the Ty3 elements are more frequent in the genome of *P. flammea*, they present a recent burst of insertions with a mean around 310 Kya and a peak at 48 kya. Copia elements, on the other hand, are older in both genomes, with a median around 665 Kya and a peak at 239 kya (Fig. 2).

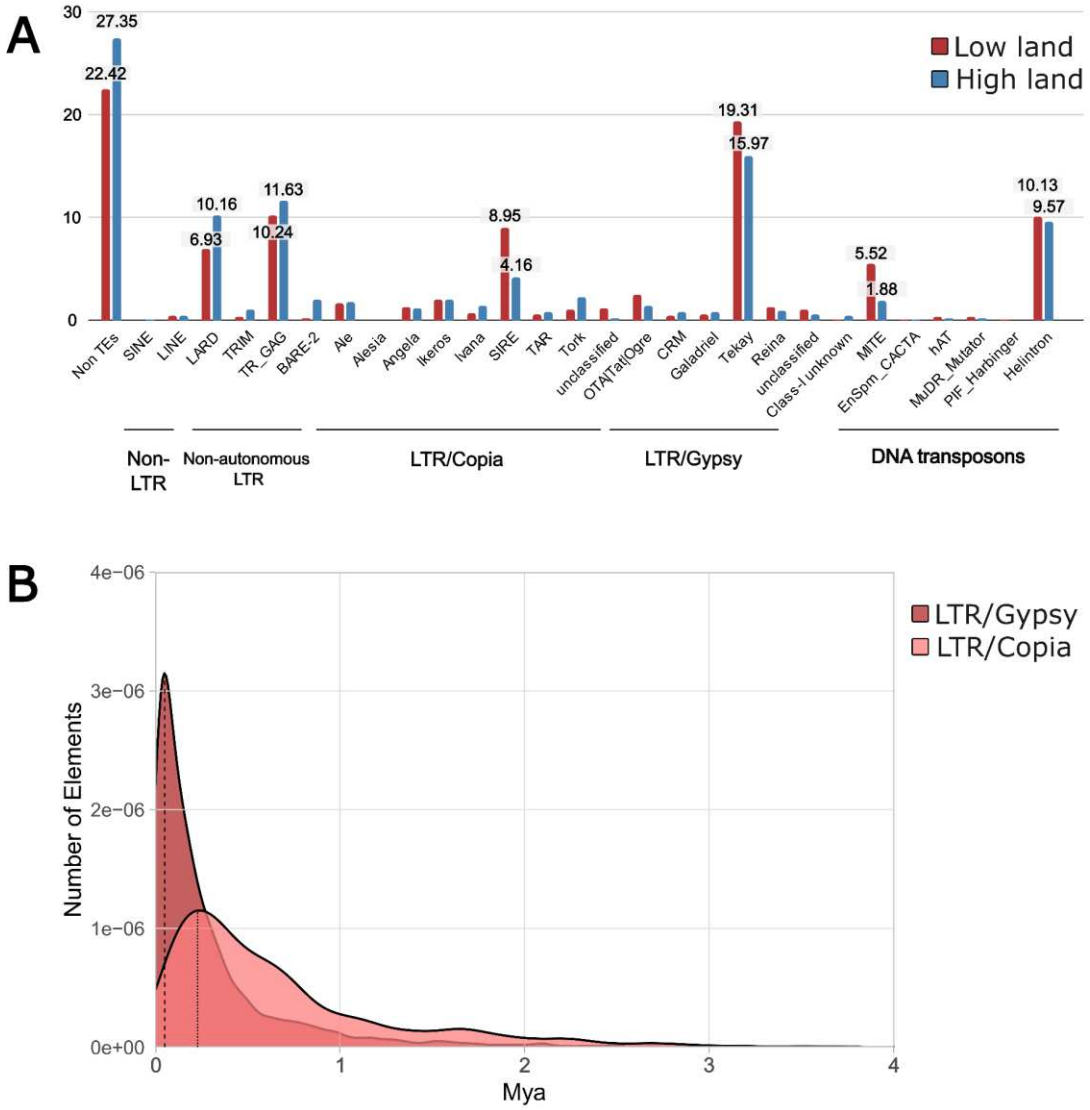

Fig. 2. **A** Proportion of unique sequence and Transposable Elements (TEs) lineages in the genome of the low-elevation (red) and high-elevation (blue) individuals. The classification of each lineage is based on Orozco-Arias et al. **[50]**. **B** Dating of LTR elements burst in the genome of the low-elevation individual of *Pitcairnia flammea*. Both individuals showed a similar age and pattern of burst, thus only the low-elevation individual is represented.

Genome annotation revealed a total of 23,181 and 23,737 genes in the genome of low-elevation and high-elevation individuals of *Pitcairnia flammea*, respectively, comprising circa 27% of the *P. flammea* genome. Structural gene annotation achieved a BUSCO completeness of 98.8% and 95.5% in the low-elevation individual, while the high-elevation individual achieved 99.1% and 95.6% for Embryophyta and Liliopsida

387 database, respectively. The average gene length was 3,8 Kbp, for the low-elevation
388 individual, and 3,6 Kbp for the high-elevation, average CDS length was 1,1 Kbp with
389 an average of approx. 5 exons per mRNA for both individuals (Table 2). The
390 distribution of gene sequences was similar to other species, primarily occupying distal
391 positions in the genome, while transposable elements were concentrated in proximal
392 regions (Fig. 1).
393
394 *Gene family evolution in* Pitcairnia flammea *and metabolic pathways affected*
395 A total of 220,719 protein sequences were deduced from the genome
396 annotation of both individuals of *Pitcairnia flammea*, along with the genomes of
397 *Ananas comosus, Tillandsia leiboldiana, T. fasciculata, T. duratii*, and *Vriesea*
398 *erythrodactylon*. These protein sequences formed 27,623 orthogroups, with an
399 average of 7.1 protein sequences per orthogroup (Fig. 3). The high-elevation *P.*
400 *flammea* genome contributed 38,246 proteins (30,869 assigned to orthogroups)
401 across 22,187 orthogroups, while the low-elevation genome contributed 35,788
402 proteins (31,147 assigned to orthogroups) across 22,012 orthogroups. *Pitcairnia*
403 *flammea* presented 1,517 unique clusters, with 497 being expanded exclusively in the
404 low-elevation individual and 637 in the high-elevation individual (Fig. 3)
405 In the low-elevation individual, 298 families were significantly expanded and
406 435 contracted, while the high-elevation individual showed 416 expanded and 335
407 contracted families (Fig. 3). Because the functional categories enriched among
408 contracted families in the low-elevation individual largely overlapped with those
409 enriched among expanded families in the high-elevation individual, we focus the
410 subsequent discussion on expanded gene families, acknowledging that contraction

events may also contribute to functional divergence in ways not captured here (Tables S1–S4).

Among the significantly enriched GO terms, the high-elevation individual showed categories primarily related to the regulation of metabolic processes, whereas the low-elevation individual was enriched for terms associated with reactive oxygen metabolism and stress response pathways (Tables S1 and S3). Metabolic pathway enrichment also differed between individuals (Tables S5–S6): the high-elevation individual displayed more enriched routes, including sphingolipid metabolism and the biosynthesis of secondary metabolites, while the low-elevation individual showed three enriched pathways, among them the MAPK signaling pathway, which was also enriched in the high-elevation individual, although with marginal statistical significance (FDR = 0.059).

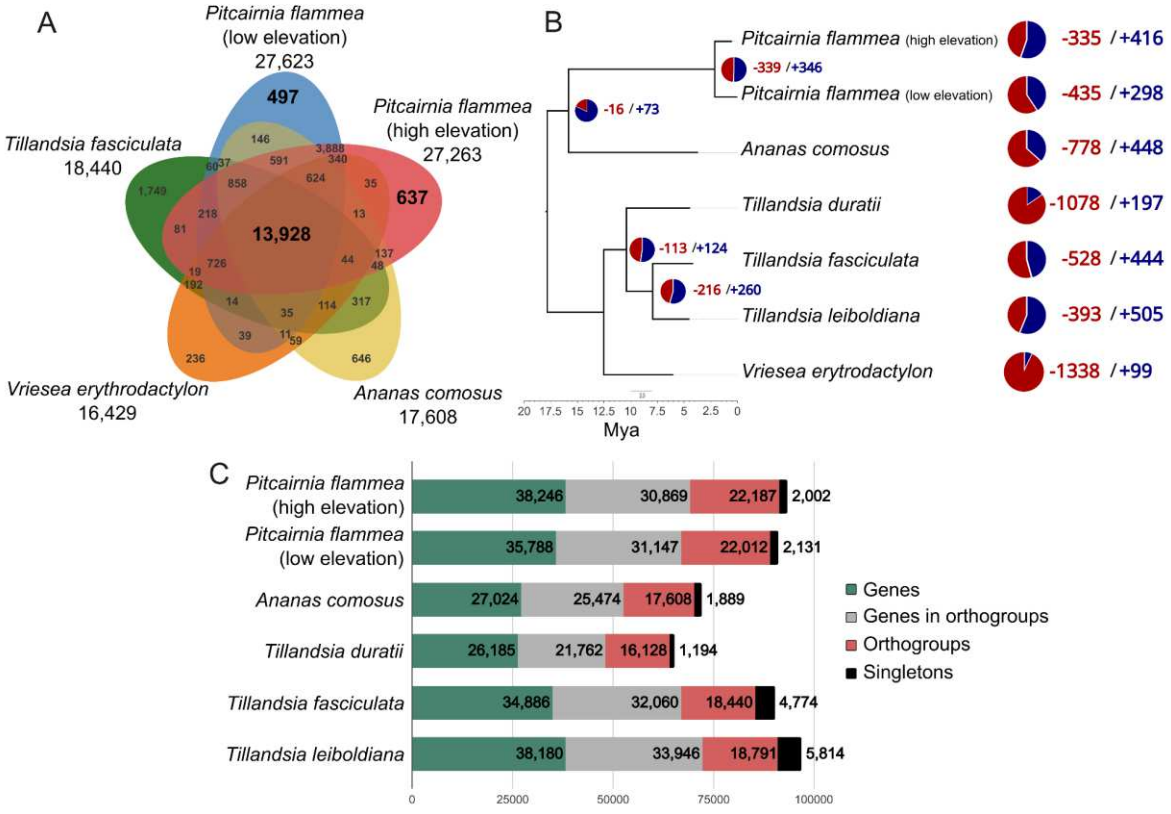

Fig 3. Orthogroups analysis among the genomes of both individuals of *Pitcairnia flammea* assembled in this study and other angiosperms. **A.** Orthougroups identified among *P. flammea* and three others Bromeliaceae species, indicating the number of shared orthogroups and singletons for each lineage. **B.** Analysis of gene families expansion and contracted among Bromeliaceae lineages. **C.** Total number of clusters and proteins identified within each lineage in the study.

Conservation of synteny among individuals of Pitcairnia flammea

Both individuals show a high degree of synteny conservation, however, structural inversions ranging from ~217 kbp to ~10 Mbp were observed in chromosomes 1, 7, 8, and 10 (Fig. 4, Fig. S2, Table S7). Besides, we also found a 13,170 bp translocation in chromosome 5. These structural variations encompassed approximately 110 genes in each genome (Tables S8–S9), including genes associated with heat acclimation, transpiration, and stomatal morphogenesis. However, no significantly enriched GO terms were identified among them.

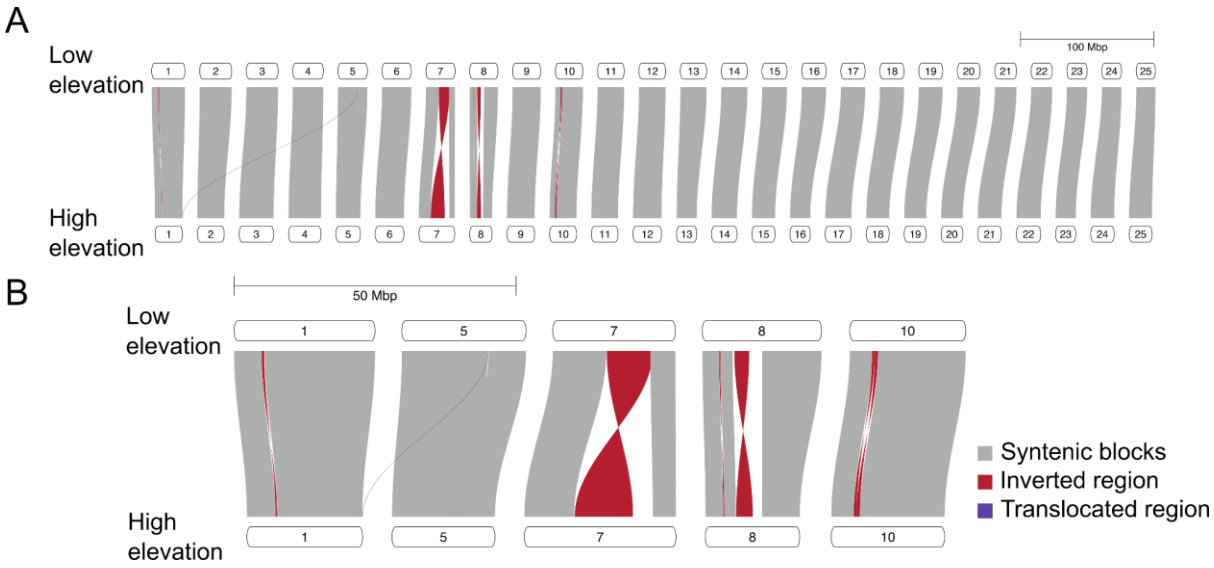

Fig 4. Synteny among the genomes of both *Pitcairnia flammea* individuals.. **A.** Overview of all chromosomes. **B.** Detailed view of rearrangement events, emphasizing inversions (red) and translocations (purple).

Historical effective population size inferred from individual genomes

The genomes of each individual indicated that both populations of *Pitcairnia flammea* experienced similar demographic histories (Fig. 5). Both populations exhibited a high effective population size around 10 Mya, followed by a substantial decline until approximately 3 Mya. After this period of sharp decline, the populations maintained a stable effective size for approximately 2 million years before beginning to decline again around 1 Mya, which has continued to the present day.

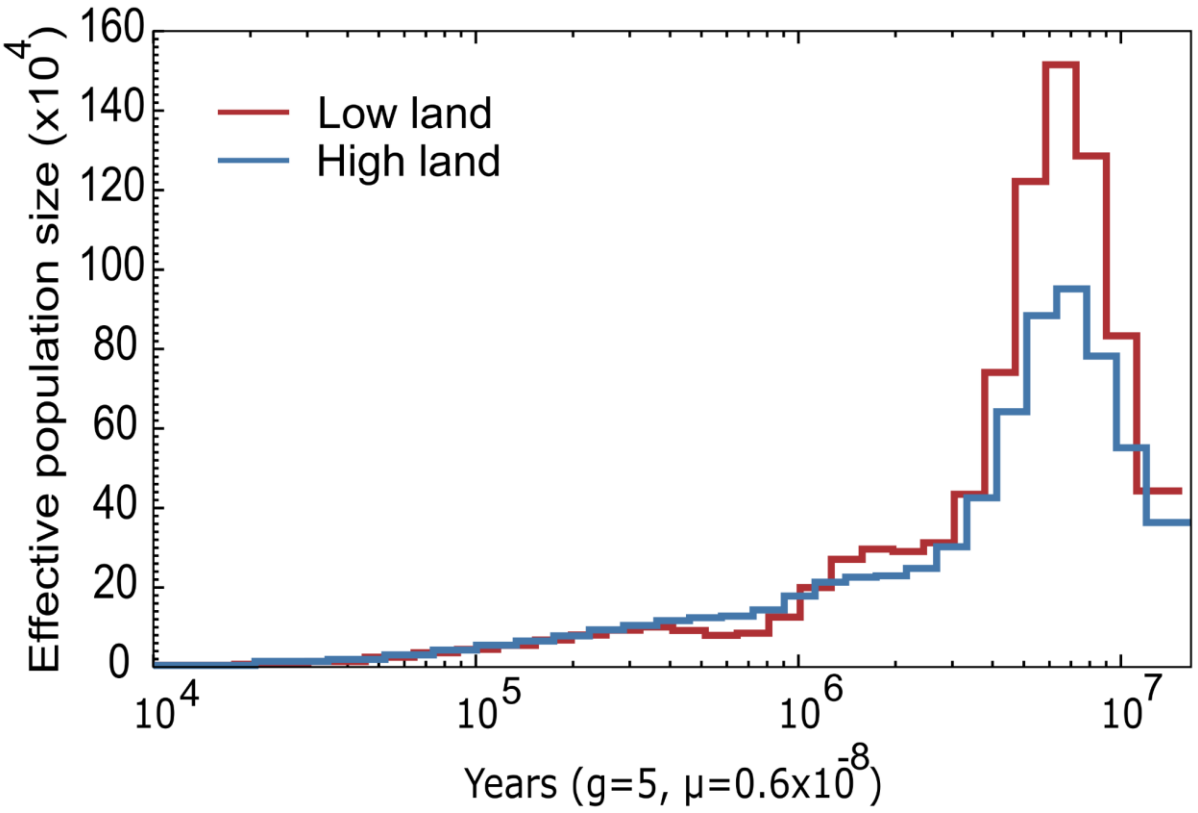

Fig 5. PSMC-inferred historical effective population size (Ne) trajectories for the low-elevation (red) and high-elevation (blue) individuals.

Discussion

We assembled and annotated high-quality chromosome-level genomes for two individuals of *Pitcairnia flammea* sampled from contrasting elevations, supported by consistently high contiguity and completeness metrics across multiple quality assessments. The two genomes showed a highly conserved macrosynteny and

genomic composition, while differing in six large chromosomal inversions and a translocation. Overall TE composition and patterns of LTR retrotransposon activity were broadly similar between genomes, although the relative abundance of some TE lineages differed. PSMC analyses revealed similar historical effective population size trajectories inferred from the two individual genomes. Together, these results reveal multiple dimensions of intraspecific genomic variation between individuals from contrasting elevations, providing a basis for investigating how individual genomic variation relates to population differentiation and elevational adaptation. These genomic resources provide a foundation for future population-level and functional studies testing the evolutionary significance of this variation, including its potential contribution to elevational adaptation, while also enabling broader comparative investigations of genome evolution and diversification in Bromeliaceae and Neotropical mountain plants.

Synteny conservation and structural variation between two Pitcairnia flammea *individuals*

We assembled 25 chromosome-level scaffolds for each of the two *Pitcairnia flammea* individuals (2*n* = 50), revealing broadly similar gene composition and genome architecture, including comparable overall distributions of transposable elements and telomeric regions, consistent with expectations for a monocentric species **[74]**. To date, 13 genomes representing six species have been assembled at the chromosome level for this family **[62–64, 75–77]**, highlighting a broad pattern of significant taxonomic sampling gaps **[78]**, given the existence of over 3,700 species within the family **[79]**. The genomes assembled here expand the genomic resources available for Bromeliaceae, contributing to broader efforts to reduce the

underrepresentation of tropical non-model plants in reference genome databases **[78, 80, 81]**. Within individual-heterozygosity also differed between the two individuals, with higher levels observed in the individual sampled at high elevation. This observation is consistent with previous population-level studies revealing substantial variation in genetic diversity, strong population structure, and restricted gene flow among *P. flammea* populations **[17, 82]**. Although the broader significance of this difference remains to be assessed with population-level sampling, the comparison of these chromosome-level genomes provides an initial perspective on intraspecific genomic variation and a valuable framework for investigating how such variation is structured across the species' broad elevational range.

Both genomes showed a high abundance of LTR/Ty3/Chromovirus/Tekay elements and broadly similar TE composition, although differences were observed in the relative abundance of some TE lineages. The inferred temporal patterns of LTR retrotransposon activity were also broadly similar between genomes, suggesting that much of their TE landscape may reflect a shared evolutionary history. Previous studies have shown that TE dynamics can be influenced by demographic and environmental processes across different evolutionary timescales **[10, 83–85]**. However, distinguishing the relative contributions of demographic history, environmental variation, and other evolutionary processes to TE dynamics in *P. flammea* will require broader population-level sampling and explicit demographic model testing to evaluate alternative histories of population size change and connectivity.

Despite the overall synteny between the two genomes, we identified six large inversions across four chromosomes and one translocation. Large chromosomal rearrangements, particularly inversions, can influence evolutionary divergence by reducing recombination and maintaining associations among alleles within rearranged

513 regions, and have been implicated in adaptation and speciation in several systems
514 **[12]**. Their frequency and persistence within populations may be influenced by
515 demographic processes, genetic drift, and natural selection **[11, 86]**. In *Pitcairnia*
516 *flammea*, the rearrangements identified between the two genomes reveal an additional
517 component of intraspecific genomic variation. Determining whether these variants are
518 individual, segregate within populations, or show consistent differentiation among
519 populations will require population-level analyses. The inversions encompass
520 genomic regions harboring genes annotated with putative functions related to heat
521 acclimation, stomatal morphogenesis, and water balance, traits previously associated
522 with contrasting thermal strategies among *P. flammea* populations, with lowland plants
523 exhibiting heat avoidance and water conservation and highland plants showing greater
524 thermal tolerance associated with membrane lipid adjustments **[20, 21]**. The co-
525 occurrence of these genes within rearranged regions raises the hypothesis that
526 structural variation may contribute to genomic differentiation associated with
527 contrasting ecophysiological strategies across the elevational gradient, which could
528 be tested by determining the frequency and distribution of these inversions across
529 populations and assessing their association with environmental and phenotypic
530 variation.
531
532 *Gene family expansion in two* Pitcairnia flammea *genomes*
533 Gene family expansion and contraction patterns differed between the two
534 *Pitcairnia flammea* individuals, suggesting metabolic pathways and candidate genes
535 that may reflect differences in the selective regimes experienced by their respective
536 populations. In the high-elevation individual, expanded gene families showed
537 enrichment of GO terms related to regulatory processes, while both individuals

exhibited enrichment of the MAPK signaling pathway, a pathway implicated in diverse stress responses, including cold tolerance **[87]**. The functional enrichment associated with gene family turnover therefore highlights pathways related to regulatory processes and stress responses as potentially relevant to the selective regimes experienced by these individuals. Although stochastic processes, including genetic drift and lineage-specific mutations, cannot be ruled out as alternative explanations, gene family turnover has been shown to be shaped by environmental pressures in other plant species **[14, 88]**. In this context, the enrichment of pathways related to regulatory processes and stress responses in the highland individual could be consistent with elevational adaptation.

Additionally, the high-elevation individual displayed enrichment of the ceramide metabolic pathway and lipid metabolism-related GO terms, including expanded gene families containing sphinganine C4-monooxygenase 2 (SUR2) and CTP synthase. These genes are involved in the metabolism of sphingolipids **[89]** and in the biosynthesis of phosphatidylcholine, the major phospholipid of extraplastidial membranes and a precursor of phosphatidylglycerol **[90, 91]**, and both have been associated with responses to cold stress **[89, 91]**. Notably, previous ecophysiological work on the same populations reported that the abundance of hexosylceramides and phosphatidylglycerols increases with elevation, suggesting a possible association between lipid composition and elevational gradients **[20]**. The correspondence among ecophysiological traits with the genomic expansions observed here raises the hypothesis that gene family variation may contribute to membrane-related processes among populations. The expansion of these lipid metabolism gene families may represent one possible genomic basis for the observed variation in lipid composition, potentially influencing membrane fluidity and cell signaling. Future studies integrating

gene-family copy-number variation, gene expression, and lipidomic profiles across populations could directly test this association.

Conclusions

Altogether, the high-quality chromosome-level genomes assembled in this study for both individuals of *Pitcairnia flammea* represent a valuable contribution to the genomic resources available for Bromeliaceae, one of the best-known examples of adaptive radiation in plants. These assemblies help to fill a significant taxonomic gap in a family with over 3,700 described species. Beyond their utility as genomic resources, the two genomes reveal substantial intraspecific variation despite broad syntenic conservation. Notably, large structural rearrangements were identified in regions harboring genes related to water-saving traits and thermal responses, raising the possibility that architectural variation may accompany ecological differentiation across elevations.

In parallel, we observed expansions of gene families associated with stress signaling and lipid metabolic pathways in the high-elevation individual, particularly those involved in ceramide, very-long-chain fatty acid, and phosphatidylglycerol metabolism. These observations align with previously documented differences in membrane lipid composition and thermal tolerance strategies among populations of *P. flammea* along the altitudinal gradient **[20, 21]**. Although neutral processes, including drift and lineage-specific stochastic events, cannot be excluded, the correspondence between genomic variation, metabolic pathways, and ecophysiological traits previously associated with elevational gradients suggests that structural variants and gene family expansions may contribute to ecological differentiation in this system. These findings provide a foundation for future research

in population genomic and functional studies investigating how genomic variation interacts with environmental heterogeneity in the Brazilian Atlantic Forest mountains and whether such variation contributes to population responses and resilience to environmental change.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The genome assemblies, the raw DNA, RNA and Omni-C sequences generated for this study were deposited in the National Center of Biotechnology Information (NCBI) under the Bioproject PRJNA1345395.

The raw data is available to the reviewers at https://dataview.ncbi.nlm.nih.gov/object/PRJNA1345395?reviewer=90d8501lv4ahs825b04574bas2

Final annotation files can be found on the Github page: https://github.com/pauloaecyo/Genome-Assembly-Pitcairnia-flammea/tree/main

Competing interests

The authors declare no conflict of interest.

Funding

This work and authors has been supported by the following Brazilian research agencies: "This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, São Paulo Research Foundation (FAPESP Grants: 2020/14805-4, 2021/10639-5, 2022/07480-7, 2022/04700-6, 2022/09509-2, 2023/05776-9, 2023/18306-0, 2024/13924-0, 2024/01417-7, 2024/08569-7, 2024/18110-1), and the National Council for Scientific and Technological Development (CNPq, Brazil; process number: 141427/2021-3).

C.P-S. is a fellow of CNPq (302962/2022-0). D.M.R.P. is a fellow of CNPq (311558/2021-6). F.P. is a fellow of CNPq (302849/2021-1). A.M.V. is a fellow of CNPq (304367/2022-2).

Authors' contributions

PA designed the project, collected the data, analyzed the data, performed research, and wrote the paper. **MMT**, **JSM, LC, FDR, DC and AV** analyzed data and reviewed and edited versions of the manuscript. **FP** reviewed and edited versions of the manuscript. **CJNC** performed field work, conceptualization, data curation, and reviewed and edited versions of the manuscript. **AB** and **DMRP** contributed with resources, formal analysis, conceptualization, supervision, and reviewed and edited versions of the manuscript. **CPS** designed the research, acquired the funding, collected the data, supervised the study, and reviewed and edited versions of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We would like to thank the Universidade Estadual de Campinas (UNICAMP), the Center for Nuclear Energy in Agriculture (CENA/USP), the São Paulo State University (UNESP) and the Instituto de Biología Molecular y Celular de Plantas for the infrastructure used. This research was performed with the support of the Núcleo Avançado de Computação de Alto Desempenho (NACAD) from COPPE, Universidade Federal do Rio de Janeiro (UFRJ). We also would like to thank the Serviço de Processamento Científico em Cluster de Alto Desempenho – Instituto de Biologia/UNICAMP.

References

1. Janzen DH. Why mountain passes are higher in the tropics. Am Nat. 1967;101:233–49. https://doi.org/10.1086/282487
2. Rahbek C, Borregaard MK, Colwell RK, Dalsgaard B, Holt BG, Morueta-Holme N, Nogues-Bravo D, Whittaker RJ, Fjeldså J. Humboldt's enigma: what causes global patterns of mountain biodiversity? Science. 2019;365(6458):1108–13. https://doi.org/10.1126/science.aax0149
3. Halbritter AH, Fior S, Keller I, Billeter R, Edwards PJ, Holderegger R, Karrenberg S, Pluess AR, Widmer A, Alexander JM. Trait differentiation and

655 adaptation of plants along elevation gradients. J Evol Biol. 2018;31(6):784–800.
656 https://doi.org/10.1111/jeb.13262
657 4. Bhattacharya A. Physiological processes in plants under low temperature
658 stress. Singapore: Springer; 2022. p. 1–106. https://doi.org/10.1007/978-981-
659 16-9037-2
660 5. Antonelli A, Ariza M, Albert J, Andermann T, Azevedo J, Bacon C, Faurby S,
661 Guedes T, Hoorn C, Lohmann LG, Matos-Maraví P, Ritter CD, Sanmartín I,
662 Silvestro D, Tejedor M, ter Steege H, Tuomisto H, Werneck FP, Zizka A,
663 Edwards SV. Conceptual and empirical advances in Neotropical biodiversity
664 research. PeerJ. 2018;6:e5644. https://doi.org/10.7717/peerj.5644
665 6. Rahbek C, Borregaard MK, Antonelli A, Colwell RK, Holt BG, Nogues-Bravo D,
666 Rasmussen CMØ, Richardson K, Rosing MT, Whittaker RJ, Fjeldså J. Building
667 mountain biodiversity: geological and evolutionary processes. Science.
668 2019;365(6458):1114–9. https://doi.org/10.1126/science.aax0151
669 7. Tabarelli M, Pinto LP, Silva JMC, Hirota MM, Bedê LC. Desafios e
670 oportunidades para a conservação da biodiversidade na Mata Atlântica
671 brasileira. Megadiversidade. 2005;1(1):132–8.
672 8. Martinelli G. Mountain biodiversity in Brazil. Braz J Bot. 2007;30(4):587–97.
673 https://doi.org/10.1590/S0100-84042007000400005
674 9. Dantas-Queiroz MV, Hurbath F, de Russo Godoy FM, Lanna FM, Versieux LM,
675 Palma-Silva C. Comparative phylogeography reveals the demographic patterns
676 of neotropical ancient mountain species. Mol Ecol. 2023;32(12):3165–81.
677 https://doi.org/10.1111/mec.16929
678 10. Bilinski P, Albert PS, Berg JJ, Birchler JA, Grote MN, Lorant A, Quezada J,
679 Swarts K, Yang J, Ross-Ibarra J. Parallel altitudinal clines reveal trends in
680 adaptive evolution of genome size in Zea mays. PLoS Genet.
681 2018;14(5):e1007162. https://doi.org/10.1371/journal.pgen.1007162
682 11. Hager ER, Harringmeyer OS, Wooldridge TB, Theingi S, Gable JT, McFadden
683 S, Neugeboren B, Turner KM, Jensen JD, Hoekstra HE. A chromosomal
684 inversion contributes to divergence in multiple traits between deer mouse
685 ecotypes. Science. 2022;377(6604):399–405.
686 https://doi.org/10.1126/science.abg0718

12. Huang K, Rieseberg LH. Frequency, origins, and evolutionary role of chromosomal inversions in plants. Front Plant Sci. 2020;11:296. https://doi.org/10.3389/fpls.2020.00296
13. Liu L, Tumi L, Suni ML, Arakaki M, Wang ZF, Ge XJ. Draft genome of Puya raimondii (Bromeliaceae), the Queen of the Andes. Genomics. 2021;113(4):2537–46. https://doi.org/10.1016/j.ygeno.2021.05.042
14. Xu S, Shao S, Feng X, Li S, Zhang L, Wu W, Liu M, Tracy ME, Zhong C, Guo Z, Wu CI, Shi S, He Z. Adaptation in unstable environments and global gene losses: small but stable gene networks by the May–Wigner theory. Mol Biol Evol. 2024;41(4):msae059. https://doi.org/10.1093/molbev/msae059
15. Formenti G, Theissinger K, Fernandes C, Bista I, Bombarely A, Bleidorn C, Ciofi C, Crottini A, Godoy JA, Höglund J, Malukiewicz J, Mouton A, Oomen RA, Paez S, Palsbøll PJ, Pampoulie C, Ruiz-López MJ, Svardal H, Theofanopoulou C, de Vries J, Waldvogel AM, Zhang G, Mazzoni CJ, Jarvis ED, Bálint M. The era of reference genomes in conservation genomics. Trends Ecol Evol. 2022;37(3):197–202. https://doi.org/10.1016/j.tree.2021.11.008
16. Theissinger K, Fernandes C, Formenti G, Bista I, Berg PR, Bleidorn C, Bombarely A, Crottini A, Gallo GR, Godoy JA, Jentoft S, Malukiewicz J, Mouton A, Oomen RA, Paez S, Palsbøll PJ, Pampoulie C, Ruiz-López MJ, Secomandi S, Svardal H, Theofanopoulou C, de Vries J, Waldvogel AM, Zhang G, Jarvis ED, Bálint M, Ciofi C, Waterhouse RM, Mazzoni CJ, Höglund J. How genomics can help biodiversity conservation. Trends Genet. 2023;39(7):545–59. https://doi.org/10.1016/j.tig.2023.01.005
17. Mota MR, Pinheiro F, Leal BSS, Sardelli CH, Wendt T, Palma-Silva C. From micro- to macroevolution: insights from a Neotropical bromeliad with high population genetic structure adapted to rock outcrops. Heredity. 2020;125(5):353–70. https://doi.org/10.1038/s41437-020-0342-8
18. Gitaí J, Paule J, Zizka G, Schulte K, Benko-Iseppon AM. Chromosome numbers and DNA content in Bromeliaceae: additional data and critical review. Bot J Linn Soc. 2014;176(3):349–68. https://doi.org/10.1111/boj.12211
19. Moura M, Forzza R, Cristiano M. Reconstruction of ancestral genome size in Pitcairnioideae (Bromeliaceae): what can genome size tell us about the evolutionary history of its five genera? Bot J Linn Soc. 2018;186(3):321–33. https://doi.org/10.1093/botlinnean/box101

20. Chaves CJN, Cacossi TC, Matos TS, Bento JP, Gonçalves LN, Silva SF, Silva-Ferreira MV, Barbin DF, Mayer JLS, Sussulini A, Ribeiro R, Palma-Silva C. Bromeliad populations perform distinct ecological strategies across a tropical elevation gradient. Funct Ecol. 2024;38:910–25. https://doi.org/10.1111/1365-2435.14528
21. Chaves CJN, Tavares DU, Lemos-Silva IV, Bento JPSP, Vilela-Bianchini HM, Aecyo P, Cacossi TC, Tavares MM, Sabino GP, Kamimura VA, Dos Santos WL, Gonçalves LN, Silva KT, Mayer JLS, Ribeiro RV, Escobar-Escobar D, Feeley KJ, Palma-Silva C. Canalized to heat, plastic to cold: adaptive coordination of leaf and seed strategies in populations spanning an elevational gradient. New Phytol. 2026;250(2):873–84. https://doi.org/10.1111/nph.70912
22. Marçais G, Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics. 2011;27(6):764–70. https://doi.org/10.1093/bioinformatics/btr011
23. Ranallo-Benavidez TR, Jaron KS, Schatz MC. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat Commun. 2020;11(1):1432. https://doi.org/10.1038/s41467-020-14998-3
24. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat Methods. 2021;18(2):170–5. https://doi.org/10.1038/s41592-020-01056-5
25. Guan D, McCarthy SA, Wood J, Howe K, Wang Y, Durbin R. Identifying and removing haplotypic duplication in primary genome assemblies. Bioinformatics. 2020;36(9):2896–8. https://doi.org/10.1093/bioinformatics/btaa025
26. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. Versatile genome assembly evaluation with QUAST-LG. Bioinformatics. 2018;34(13):i142–50. https://doi.org/10.1093/bioinformatics/bty266
27. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. arXiv:2106.11799 [q-bio]. 2021. http://arxiv.org/abs/2106.11799
28. Kriventseva EV, Kuznetsov D, Tegenfeldt F, Manni M, Dias R, Simão FA, Zdobnov EM. OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations

of orthologs. Nucleic Acids Res. 2019;47(D1):D807–11. https://doi.org/10.1093/nar/gky1053

29. Rhie A, Walenz BP, Koren S, Phillippy AM. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. Genome Biol. 2020;21:1–27. https://doi.org/10.1186/s13059-020-02134-9
30. Ou S, Jiang N. LTR_retriever: a highly accurate and sensitive program for identification of long terminal repeat retrotransposons. Plant Physiol. 2018;176(2):1410–22. https://doi.org/10.1104/pp.17.01310
31. Ou S, Chen J, Jiang N. Assessing genome assembly quality using the LTR Assembly Index (LAI). Nucleic Acids Res. 2018;gky730. https://doi.org/10.1093/nar/gky730
32. Laetsch DR, Blaxter ML. BlobTools: interrogation of genome assemblies. F1000Res. 2017;6:1287. https://doi.org/10.12688/f1000research.12232.1
33. Laetsch DR, Koutsovoulos G, Booth T, Stajich J, Kumar S. DRL/blobtools: BlobTools v1.0.1. Zenodo; 2017. https://doi.org/10.5281/zenodo.845347
34. Buchfink B, Reuter K, Drost HG. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods. 2021;18:366–8. https://doi.org/10.1038/s41592-021-01101-x
35. The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. 2023;51(D1):D523–31. https://doi.org/10.1093/nar/gkac1052
36. Zhou C, McCarthy SA, Durbin R. YaHS: yet another Hi-C scaffolding tool. Bioinformatics. 2023;39(1):btac808. https://doi.org/10.1093/bioinformatics/btac808
37. Brown MR, González de la Rosa PM, Blaxter M. tidk: a toolkit to rapidly identify telomeric repeats from genomic datasets. Bioinformatics. 2025;btaf049. https://doi.org/10.1093/bioinformatics/btaf049
38. Li H. New strategies to improve minimap2 alignment accuracy. Bioinformatics. 2021;37(23):4572–4. https://doi.org/10.1093/bioinformatics/btab705
39. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative Genomics Viewer. Nat Biotechnol. 2011;29(1):24–6. https://doi.org/10.1038/nbt.1754

40. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013;14(2):178–92. https://doi.org/10.1093/bib/bbs017
41. Aronesty E. ea-utils: command-line tools for processing biological sequencing data. 2011. https://github.com/ExpressionAnalysis/ea-utils
42. Aronesty E. Comparison of sequencing utility programs. Open Bioinform J. 2013;7:1–8. https://doi.org/10.2174/1875036201307010001
43. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21. https://doi.org/10.1093/bioinformatics/bts635
44. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33(3):290–5. https://doi.org/10.1038/nbt.3122
45. Haas BJ. TransDecoder [source code]. 2023. https://github.com/TransDecoder/TransDecoder
46. Alves RM, de Abreu VA, Oliveira RP, Almeida JVDA, de Oliveira MDM, Silva SR, Paschoal AR, Almeida SS, Souza PAF, Ferro JA, Miranda VFO, Figueira A, Domingues DS, Varani AM. Genomic decoding of Theobroma grandiflorum (cupuassu) at chromosomal scale: evolutionary insights for horticultural innovation. GigaScience. 2024;13. https://doi.org/10.1093/gigascience/giae027
47. Li Y, Jiang N, Sun Y. AnnoSINE: a short interspersed nuclear elements annotation tool for plant genomes. Plant Physiol. 2022. https://doi.org/10.1093/plphys/kiab524
48. Rho M, Tang H. MGEScan-non-LTR: computational identification and classification of autonomous non-LTR retrotransposons in eukaryotic genomes. Nucleic Acids Res. 2009. https://doi.org/10.1093/nar/gkp752
49. Zhang RG, Li GY, Wang XL, Dainat J, Wang ZX, Ou S, Ma Y. TEsorter: an accurate and fast method to classify LTR-retrotransposons in plant genomes. Hortic Res. 2022;9:uhac017. https://doi.org/10.1093/hr/uhac017
50. Orozco-Arias S, Isaza G, Guyot R. Retrotransposons in plant genomes: structure, identification, and classification through bioinformatics and machine learning. Int J Mol Sci. 2019. https://doi.org/10.3390/ijms20153837

51. Keilwagen J, Hartung F, Grau J. GeMoMa: homology-based gene prediction utilizing intron position conservation and RNA-seq data. Methods Mol Biol. 2019;1962:161–77. https://doi.org/10.1007/978-1-4939-9173-0_9
52. Gabriel L, Brůna T, Hoff KJ, Ebel M, Lomsadze A, Borodovsky M, Stanke M. BRAKER3: fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS, and TSEBRA. Genome Res. 2024;34(5):769–77. https://doi.org/10.1101/gr.278090.123
53. Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol. 2008;9:R7. https://doi.org/10.1186/gb-2008-9-1-r7
54. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 2005. https://doi.org/10.1093/bioinformatics/bti610
55. Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, Gao G. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Res. 2017. https://doi.org/10.1093/nar/gkw982
56. Martin EC, Ion CF, Ifrimescu F, Spiridon L, Bakker J, Goverse A, Petrescu AJ. NLRscape: an atlas of plant NLR proteins. Nucleic Acids Res. 2023. https://doi.org/10.1093/nar/gkac1014
57. Calle García J, Guadagno A, Paytuvi-Gallart A, Saera-Vila A, Amoroso CG, D'Esposito D, Andolfo G, Aiese Cigliano R, Sanseverino W, Ercolano MR. PRGdb 4.0: an updated database dedicated to genes involved in plant disease resistance process. Nucleic Acids Res. 2022;50(D1):D1483–90. https://doi.org/10.1093/nar/gkab1087
58. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, von Mering C, Bork P. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 2019. https://doi.org/10.1093/nar/gky1085
59. Mulder N, Apweiler R. InterPro and InterProScan: tools for protein sequence classification and comparison. In: Comparative genomics. 2007. p. 59–70. https://doi.org/10.1007/978-1-59745-515-2_5

60. Kautsar SA, Suarez Duran HG, Medema MH. Genomic identification and analysis of specialized metabolite biosynthetic gene clusters in plants using plantiSMASH. Methods Mol Biol. 2018. https://doi.org/10.1007/978-1-4939-7874-8_15
61. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019;20:1–14. https://doi.org/10.1186/s13059-019-1832-y
62. Ming R, VanBuren R, Wai CM, Tang H, Schatz MC, Bowers JE, Lyons E, Wang ML, Chen J, Biggers E, Zhang J, Huang L, Zhang L, Miao W, Zhang J, Ye Z, Miao C, Lin Z, Wang H, Zhou H, Yim WC, Priest HD, Zheng C, Woodhouse M, Edger PP, Guyot R, Guo HB, Guo H, Zheng G, Singh R, Sharma A, Min X, Zheng Y, Lee H, Gurtowski J, Sedlazeck FJ, Harkess A, McKain MR, Liao Z, Fang J, Liu J, Zhang X, Zhang Q, Hu W, Qin Y, Wang K, Chen LY, Shirley N, Lin YR, Liu LY, Hernandez AG, Wright CL, Bulone V, Tuskan GA, Heath K, Zee F, Moore PH, Sunkar R, Leebens-Mack JH, Mockler T, Bennetzen JL, Freeling M, Sankoff D, Paterson AH, Zhu X, Yang X, Smith JAC, Cushman JC, Paull RE, Yu Q. The pineapple genome and the evolution of CAM photosynthesis. Nat Genet. 2015;47(12):1435–42. https://doi.org/10.1038/ng.3435
63. Groot Crego C, Hess J, Yardeni G, de La Harpe M, Priemer C, Beclin F, Saadain S, Cauz-Santos LA, Temsch EM, Weiss-Schneeweiss H, Barfuss MHJ, Till W, Weckwerth W, Heyduk K, Lexer C, Paun O, Leroy T. CAM evolution is associated with gene family expansion in an explosive bromeliad radiation. Plant Cell. 2024;koae130. https://doi.org/10.1093/plcell/koae130
64. Lyu X, Li P, Jin L, Yang F, Pucker B, Wang C, Liu L, Zhao M, Shi L, Zhang Y, Yang Q, Xu K, Li X, Hu Z, Yang J, Yu J, Zhang M. Tracing the evolutionary and genetic footprints of atmospheric tillandsioids transition from land to air. Nat Commun. 2024;15(1):9599. https://doi.org/10.1038/s41467-024-53756-7
65. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020;37(5):1530–4. https://doi.org/10.1093/molbev/msaa015
66. Kumar S, Suleski M, Craig JM, Kasprowicz AE, Sanderford M, Li M, Stecher G, Hedges SB. TimeTree 5: an expanded resource for species divergence times. Mol Biol Evol. 2022;39(8):msac174. https://doi.org/10.1093/molbev/msac174

67. Mendes FK, Vanderpool D, Fulton B, Hahn MW. CAFE 5 models variation in evolutionary rates among gene families. Bioinformatics. 2020;36(22–23):5516–8. https://doi.org/10.1093/bioinformatics/btaa1022
68. Klopfenstein DV, Zhang L, Pedersen BS, Ramírez F, Warwick Vesztrocy A, Naldi A, Mungall CJ, Yunes JM, Botvinnik O, Weigel M, Dampier W, Dessimoz C, Flick P, Tang H. GOATOOLS: a Python library for Gene Ontology analyses. Sci Rep. 2018;8(1):10872. https://doi.org/10.1038/s41598-018-28948-z
69. Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S. A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science. 2002;296(5565):92–100. https://doi.org/10.1126/science.1068275
70. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics. 2020;36(8):2628–9. https://doi.org/10.1093/bioinformatics/btz931
71. Lovell JT, Sreedasyam A, Schranz ME, Wilson M, Carlson JW, Harkess A, Emms D, Goodstein DM, Schmutz J. GENESPACE tracks regions of interest and gene copy number variation across multiple genomes. eLife. 2022;11:e78526. https://doi.org/10.7554/eLife.78526
72. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, Kissinger JC, Paterson AH. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49. https://doi.org/10.1093/nar/gkr1293
73. Li H, Durbin R. Inference of human population history from individual whole-genome sequences. Nature. 2011;475(7357):493–6. https://doi.org/10.1038/nature10231
74. Hofstatter PG, Thangavel G, Lux T, Neumann P, Vondrak T, Novak P, Zhang M, Costa L, Castellani M, Scott A, Toegelová H, Fuchs J, Mata-Sucre Y, Dias Y, Vanzela ALL, Huettel B, Almeida CCS, Šimková H, Souza G, Pedrosa-

921 Harand A, Macas J, Mayer KFX, Houben A, Marques A. Repeat-based
922 holocentromeres influence genome architecture and karyotype evolution. Cell.
923 2022;185(17):3153–68. https://doi.org/10.1016/j.cell.2022.06.045
924 75. Chen LY, VanBuren R, Paris M, Zhou H, Zhang X, Wai CM, Yan H, Chen S,
925 Alonge M, Ramakrishnan S, Liao Z, Liu J, Lin J, Yue J, Fatima M, Lin Z, Zhang
926 J, Huang L, Wang H, Hwa TY, Kao SM, Choi JY, Sharma A, Song J, Wang L,
927 Yim WC, Cushman JC, Paull RE, Matsumoto T, Qin Y, Wu Q, Wang J, Yu Q,
928 Wu J, Zhang S, Boches P, Tung CW, Wang ML, Coppens d'Eeckenbrugge G,
929 Sanewski GM, Purugganan MD, Schatz MC, Bennetzen JL, Lexer C, Ming R.
930 The bracteatus pineapple genome and domestication of clonally propagated
931 crops. Nat Genet. 2019;51(10):1549–58. https://doi.org/10.1038/s41588-019-
932 0506-8
933 76. Yow AG, Bostan H, Castanera R, Ruggieri V, Mengist MF, Curaba J, Young R,
934 Gillitt N, Iorizzo M. Improved high-quality genome assembly and annotation of
935 pineapple (Ananas comosus) cultivar MD2 revealed extensive haplotype
936 diversity and diversified FRS/FRF gene family. Genes (Basel). 2021;13(1):52.
937 https://doi.org/10.3390/genes13010052
938 77. Li Z, Wang J, Zhang X, Zhu G, Fu Y, Jing Y, Huang B, Wang X, Meng C, Yang
939 Q, Xu L. The genome of Aechmea fasciata provides insights into the evolution
940 of tank epiphytic habits and ethylene-induced flowering. Commun Biol.
941 2022;5(1):920. https://doi.org/10.1038/s42003-022-03918-4
942 78. Marks RA, Hotaling S, Frandsen PB, VanBuren R. Representation and
943 participation across 20 years of plant genome sequencing. Nat Plants.
944 2021;7(12):1571–8. https://doi.org/10.1038/s41477-021-01031-8
945 79. Gouda EJ, Butcher D, Gouda CS. Encyclopaedia of Bromeliads, version 4.
946 Utrecht University Botanic Gardens. http://bromeliad.nl/encyclopedia/.
947 Accessed 31 May 2024.
948 80. Song B, Ning W, Wei D, Jiang M, Zhu K, Wang X, Edwards D, Odeny DA,
949 Cheng S. Plant genome resequencing and population genomics: current status
950 and future prospects. Mol Plant. 2023;16(8):1252–68.
951 https://doi.org/10.1016/j.molp.2023.07.009
952 81. Varani AM, Domingues DS, Palma-Silva C, Van Sluys MA. Operationalizing
953 tropical plant genomics in Brazil. Trends Plant Sci. 2026.
954 https://doi.org/10.1016/j.tplants.2026.01.006

82. Cacossi TdC, Chaves CJN, Tavares MM, Aecyo P, Leal BSS, da Silva R, Pinheiro F, Sanín MJ, Palma-Silva C. Population genomic differentiation and demographic history across an elevational gradient in an endemic bromeliad of the Atlantic Forest mountains. Heredity. 2026. https://doi.org/10.1038/s41437-026-00903-z
83. Lyu H, He Z, Wu CI, Shi S. Convergent adaptive evolution in marginal environments: unloading transposable elements as a common strategy among mangrove genomes. New Phytol. 2018;217(1):428–38. https://doi.org/10.1111/nph.14784
84. Van-Lume B, Mata-Sucre Y, Báez M, Ribeiro T, Huettel B, Gagnon E, Leitch IJ, Pedrosa-Harand A, Lewis GP, Souza G. Evolutionary convergence or homology? Comparative cytogenomics of Caesalpinia group species (Leguminosae) reveals diversification in the pericentromeric heterochromatic composition. Planta. 2019;250:2173–86. https://doi.org/10.1007/s00425-019-03287-z
85. Mata-Sucre Y, Sader M, Van-Lume B, Gagnon E, Pedrosa-Harand A, Leitch IJ, Lewis GP, Souza G. How diverse is heterochromatin in the Caesalpinia group? Cytogenomic characterization of *Erythrostemon hughesii* Gagnon & GP Lewis (Leguminosae: Caesalpinioideae). Planta. 2020;252:1–14. https://doi.org/10.1007/s00425-020-03453-8
86. Faria R, Johannesson K, Butlin RK, Westram AM. Evolving inversions. Trends Ecol Evol. 2019;34(3):239–48. https://doi.org/10.1016/j.tree.2018.12.005
87. Manna M, Rengasamy B, Sinha AK. Revisiting the role of MAPK signalling pathway in plants and its manipulation for crop improvement. Plant Cell Environ. 2023;46(8):2277–95. https://doi.org/10.1111/pce.14606
88. Ma D, Ding Q, Guo Z, Xu C, Liang P, Zhao Z, Song S, Zheng HL. The genome of a mangrove plant, Avicennia marina, provides insights into adaptation to coastal intertidal habitats. Planta. 2022;256(1):6. https://doi.org/10.1007/s00425-022-03916-0
89. Chen LH, Dong J, Fan BL, Chen B, Huang Y, Yang L, Guo H, Cai W, Chen LL. Integrated transcriptomic and metabolomic analysis reveals key pathways and metabolites underlying cold tolerance in passion fruit. Hortic Res. 2026;13(9):uhag164. https://doi.org/10.1093/hr/uhag164

90. Nakamura Y. Plant phospholipid diversity: emerging functions in metabolism and protein–lipid interactions. Trends Plant Sci. 2017;22(12):1027–40. https://doi.org/10.1016/j.tplants.2017.09.002
91. Liu X, Giarola V, Quan W, Song X, Bartels D. Identification and characterization of CTP: phosphocholine cytidylyltransferase CpCCT1 in the resurrection plant Craterostigma plantagineum. Plant Sci. 2021;302:110698. https://doi.org/10.1016/j.plantsci.2020.110698

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- submitedSupplMaterialBMCplantBioGenomePflammea.docx