


Article

Transposable Element Landscape in the Monotypic Species *Barthea barthei* (Hance) Krass (Melastomataceae) and Its Role in Ecological Adaptation

Wei Wu ^{1,*},, Yuan Zeng ^{1,†}, Zecheng Huang ¹, Huiting Peng ¹, Zhanghai Sun ¹ and Bin Xu ^{2,*}

¹ College of Horticulture and Landscape Architecture, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China; zengyuan025@gmail.com (Y.Z.); luckzechengh@gmail.com (Z.H.); penghuiting.001@gmail.com (H.P.); sunzhanghai.17@gmail.com (Z.S.)

² Guangdong Provincial Key Laboratory of Silviculture, Protection and Utilization, Guangdong Academy of Forestry, Guangzhou 510520, China

* Correspondence: wuei06@zhku.edu.cn (W.W.); xubin@sinogaf.cn (B.X.)

† These authors contributed equally to this work.

Abstract: Transposable elements (TEs) are crucial for genome evolution and ecological adaptation, but their dynamics in non-model plants are poorly understood. Using genomic, transcriptomic, and population genomic approaches, we analyzed the TE landscape of *Barthea barthei* (Melastomataceae), a species distributed across tropical and subtropical southern China. We identified 64,866 TE copies (16.76% of a 235 Mb genome), dominated by Ty3/Gypsy retrotransposons (8.82%) and DNA/Mutator elements (2.7%). A genome-wide analysis revealed 13 TE islands enriched in genes related to photosynthesis, tryptophan metabolism, and stress response. We identified 3859 high-confidence TE insertion polymorphisms (TIPs), including 29 fixed insertions between red and white flower ecotypes, affecting genes involved in cell wall modification, stress response, and secondary metabolism. A transcriptome analysis of the flower buds identified 343 differentially expressed TEs between the ecotypes, 30 of which were near or within differentially expressed genes. The non-random distribution (primarily within 5 kb of genes) and association with adaptive traits suggest a significant role in *B. barthei*'s successful colonization of diverse habitats. Our findings provide insights into how TEs contribute to plant genome evolution and ecological adaptation in tropical forests, particularly through their influence on regulatory networks governing stress response and development.

Keywords: transposable elements; ecological adaptation; genome evolution; *Barthea barthei*; TE insertion polymorphisms



Academic Editor: Jürg Bähler

Received: 21 January 2025

Revised: 21 February 2025

Accepted: 25 February 2025

Published: 27 February 2025

Citation: Wu, W.; Zeng, Y.; Huang, Z.; Peng, H.; Sun, Z.; Xu, B. Transposable Element Landscape in the Monotypic Species *Barthea barthei* (Hance) Krass (Melastomataceae) and Its Role in Ecological Adaptation. *Biomolecules* **2025**, *15*, 346. <https://doi.org/10.3390/biom15030346>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Transposable elements (TEs) constitute a substantial proportion of most plant genomes. Active TEs mobilize within the host genome through either ‘copy and paste’ (Retrotransposons, Class I) or ‘cut and paste’ (DNA transposons, Class II) mechanisms [1]. Retrotransposons utilize an RNA intermediate, while DNA transposons employ a DNA intermediate. Both classes are further categorized as autonomous (self-mobilizing) or non-autonomous (dependent on other TEs). A hierarchical classification system organizes TEs into orders, superfamilies, families, and subfamilies based on their sequence structure and phylogenetic relationships [2]. In plant genomes, the long terminal repeat (LTR) retrotransposon order, comprising the Copia (RLC) and Gypsy (RLG) superfamilies, are typically the most abundant and range in size from a few hundred base pairs to 25 kilobases (kb) [3]. Other

retrotransposon orders include non-LTR elements such as long interspersed sequences (LINEs), short interspersed sequences (SINEs), dictyostelium intermediate repeat sequences (DIRS), and Penelope-like elements (PLEs) [3]. Class II elements are classified into the terminal inverted repeat (TIR) and Helitron orders. TIR transposons are further divided into superfamilies based on target site duplications (TSDs) and terminal inverted sequences, including Tc1–Mariner (DTT), hAT (DTA), Mutator (DTM), P (DTP), PIF–Harbinger (DTH), and CACTA (DTC) [3]. Miniature inverted-repeat transposable elements (MITEs) are non-autonomous DNA transposons characterized by a short length (<500 bp) and a high copy number. While common in plant genomes (e.g., *Stowaway* and *Tourist* families in rice [4]), MITEs share structural features with diverse TEs; thus, their designation lacks taxonomic significance [3]. Helitrons, another Class II order, utilize a unique rolling-circle mechanism for transposition. They are defined by a 5′-TC and 3′-CTRR motif, a short hairpin structure 15–20 bp upstream of the 3′ end, and the frequent duplication of passenger genes [3].

Once dismissed as ‘junk DNA’ or ‘parasites’, TEs are now recognized as key drivers of genome architecture and evolution [5,6]. Genome size variation across plant species is largely attributable to TE content, ranging from <3% in the small genome of *Utricularia gibba* (77 Mb) [7] to >70% in the massive *Picea abies* genome (19.6 Gb) [8]. While a positive correlation between genome size and TE diversity was initially expected, the relationship only holds true for genomes smaller than 500 Mb [9]. For example, despite a low TE family diversity, the *P. abies* genome is dominated by a few LTR families with high copy numbers that are estimated to be 5–60 million years old (MYA). Conversely, medium-sized genomes with high TE diversity often exhibit rapid turnover, with most TEs being <5 MYA [8,10]. The factors driving this variability in TE diversity and the success of specific TE superfamilies/families in certain lineages remain poorly understood [11]. However, the growing availability of plant genome sequences offers the potential to characterize TE landscapes across broader phylogenetic scales and elucidate the environmental and genetic mechanisms underlying the observed variations.

Transposable elements are a significant source of genetic variation, contributing to phenotypic innovation and adaptation through mechanisms such as domestication, exaptation, host gene regulation, retrogene formation, and enhanced genomic plasticity [12–14]. While TEs are typically silenced via DNA methylation, histone modification, small RNA-based silencing, and chromatin modifications [15], they can be reactivated by stress or developmental cues [16,17]. Diverse TEs have been shown to regulate various agronomically and ecologically relevant traits, including flower color in morning glories [18], fruit coloration in apples [19], sex determination in melons [20], drought tolerance in maize seedlings [21], and photoperiod sensitivity in maize [22]. Although these TE-mediated trait variations are primarily documented in crops and model species, their impact in natural populations, where diverse abiotic and biotic factors influence TE expression, may be underestimated. The growing availability of non-model plant genomes promises to reveal further TE diversity.

Barthea barthei, an evergreen shrub and the sole member of the monotypic genus *Barthea* (Melastomataceae), is endemic to southern mainland China as well as Taiwan Province [23]. While two varieties were previously distinguished based on capsule wing width, a recent population genetic analysis does not support this classification [24]. Unlike the uniformly red flowers of *Melastoma* species, *B. barthei* exhibits flower color polymorphisms, with red and white flowers observed in natural populations spanning altitudes of 400–2500 m (Figure 1A). Red-flowered populations typically inhabit open forest areas or mountaintops, while white-flowered populations are usually found in the shaded understory. These two groups are designated the ‘red ecotype’ and ‘white ecotype’, reflecting their adaptation to differing light intensities. While the role of TEs in genome evolution is well-established,

their contribution to adaptive traits, particularly in non-model plant species, remains poorly understood. In particular, the link between TE variation and ecologically relevant traits, such as flower color polymorphisms, requires further investigation. *B. barthei*, with its distinct red and white flower ecotypes adapted to different light environments, provides an excellent system to explore this connection. Therefore, this study aims to address the following key questions: (1) What is the TE landscape of *B. barthei*, including TE abundance, types, and genomic distribution? (2) Are there differences in TE content or expression between the red and white ecotypes? (3) Do specific TE insertions correlate with differential gene expression and potentially contribute to the observed flower color polymorphism? To answer these questions, we characterize the *B. barthei* TE landscape using a published genome assembly. Integrating RNA-seq and population genome resequencing data from the two ecotypes, we investigate the potential role of TEs in adaptation to contrasting light environments (Figure S1).

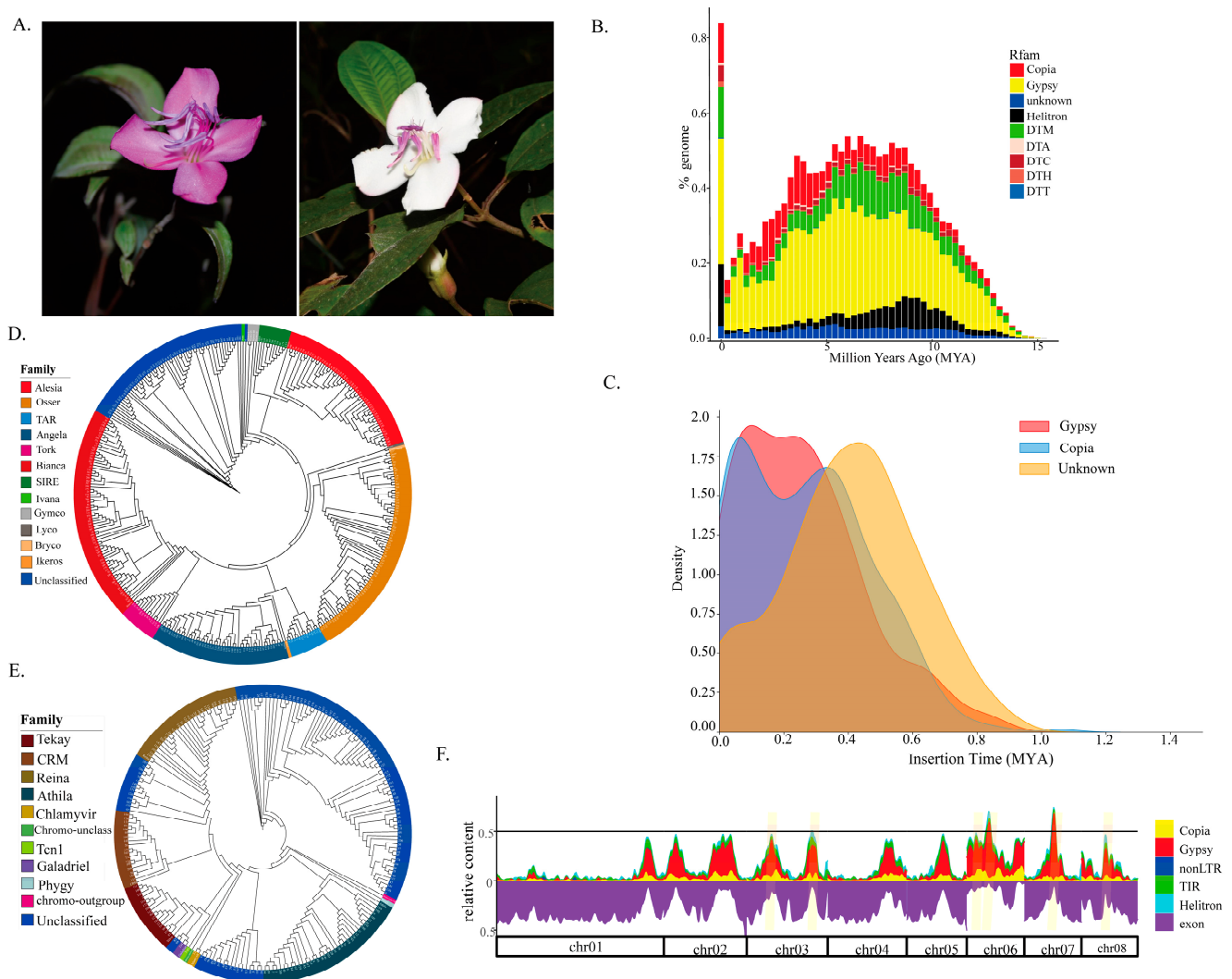


Figure 1. Transposable element landscape of *Barthea barthei*. (A) Photographs of the two ecotypes of *B. barthei*: white ecotype (right) and red ecotype (left); (B) the age distributions of different superfamilies measured by divergence with corresponding consensus sequences, including superfamily Copia (red), Gypsy (yellow), DTM (green), Helitron (black), DTA (light pink), DTC (dark red), DTH (brown), DTT (light blue), and Unknown (blue); (C) insertion time distribution for intact LTR retrotransposons, including superfamily Copia (blue), Gypsy (red), and Unknown (unclassified LTR, yellow); (D) classifications for LTR retrotransposons based on phylogenies of reverse transcriptase

(RT) domains for superfamily Ty1/Copia, including family *Alesia* (red), *Osser* (orange), *TAR* (light blue), *Angela* (dark blue), *Tork* (magenta), *Bianca* (bright red), *SIRE* (green), *Ivana* (light green), *Gymco* (gray), *Lyco* (dark gray), *Bryco* (light orange), *Ikeros* (dark orange), and Unclassified (blue); (E) classification for superfamily Ty3/Gypsy, including family *Tekay* (dark red), *CRM* (brown), *Reina* (dark brown), *Athila* (teal), *Chlamyvir* (olive green), *Chromo-unclass* (light green), *Tcn1* (bright green), *Galadriel* (purple), *Phyggy* (light blue), *chromo-outgroup* (pink), and Unclassified (blue); (F) relative content of exonic and TE-derived sequences along the eight largest scaffolds of the *B. barthei* genome. Shown are DNA transposons (DNA), long interspersed nuclear element (LINE) and LTR retrotransposons, as well as other TEs (other). The genome is well structured into TE-poor regions ('low-density regions', LDRs), TE-rich regions ('TE islands', orange highlights), and genome features such as Copia (yellow), Gypsy (red), non-LTR (blue), TIR (green), Helitron (cyan), and exon (purple).

2. Materials and Methods

2.1. Plant Material and Sequencing

Flower buds from two *B. barthei* ecotypes were collected from natural populations. One ecotype with white petals was sampled from Erhuangzhang Nature Reserve, Yangcun County, Guangdong, China (YC population, 111°25'22.59" E, 21°53'14.37" N, alt. 611 m) and the other red ecotype from Lianhua Mt. Huidong county, Guangdong, China (HD population, 115°13'56.10" E, 23°3'42.43" N, alt. 1280 m). Flower buds from each ecotype for three individual plants were sampled, immediately frozen in liquid nitrogen in the field, and taken back for total RNA isolation using the RNeasy Pure Plant Kit (TIANGEN Biotech Co., Ltd., Beijing, China). The quality of total RNA was determined using Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). The qualified RNAs were subject to paired-end library construction using Illumina TruSeq RNA Sample Preparation Kit (Illumina, San Diego, CA, USA), and sequenced on the Illumina HiSeq2500 platform using 150 bp × 2 paired-end reads. Additionally, we utilized existing genome resequencing data from nine HD and eleven YC individuals generated in a previous study [25].

2.2. Transposable Element Annotation of the Genome *B. barthei* and Phylogenetic Analysis

The chromosome-level *B. barthei* genome assembly, generated using PacBio Sequel II platform and high-throughput chromatin conformation capture mapping [25], was used for TE annotation. A non-redundant TE library was curated and the genome annotated using the Extensive de novo TE Annotator (EDTA) package [26]. Unknown superfamily LTR and non-LTR retrotransposons were classified using DeepTE with a default probability threshold of 0.6 [27]. Retroelements were searched against NCBI's conserved domain database (CDD, <http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml> (accessed on 22 January 2024)) [28] using RPSBLASTN v.2.11.0+. Elements matching Pfam00078 (RVT_1, Ty3/Gypsy) or Pfam07727 (RVT_2, Ty1/Copia) with a minimum of E-value of 0.001 were retained. For family-level classification of these filtered retroelements, one RT domain protein sequence from each family in the REXdb (Viridiplante v3.0) [29] was randomly selected. Open reading frames (ORFs) within the corresponding nucleotide sequences were identified using Open Reading Frame Finder program (RRID:SCR_016643, <https://www.ncbi.nlm.nih.gov/orffinder> (accessed on 23 January 2024)), and amino acid sequences > 50 residues were aligned using MAFFT v7.407 with default settings [30]. ModelTest-NG v0.1.7 [31] was used to determine the appropriate amino acid substitution model before maximum likelihood phylogeny inference using RAxML v8.2.12 [32] with 100 nonparametric bootstrap replicates. Custom scripts are available on GitHub (https://github.com/altingia/Barthea_TE_manuscript accessed on 23 January 2023).

2.3. Insertion Time Calculation of Intact LTRs and History of TE Proliferation Inference

The insertion time of intact LTR retrotransposons was estimated by calculating the divergence between their initially identical long terminal repeats. LTRpred v1.1.3 [33], optimized for intact LTR retrotransposon detection, was used to calculate insertion ages based on a mutation rate of 1.3×10^{-8} substitutions per site per year [34]. Kimura genetic distances between individual TE family copies and their consensus sequences were retrieved from RepeatMasker-v4.1.6 output using the ‘parseRM.pl’ script. These distances were used to summarize and visualize the accumulated coverage of TE classes and subclasses across different age ranges, providing insights into TE proliferation history. Custom R scripts used for this analysis are available on GitHub (https://github.com/altingia/Barthea_TE_manuscript/blob/main/03.TE_distributions/3.2.TE_islands/500k_sliding/Copia.windows.coverage.bed (accessed on 23 January 2024)).

2.4. Relationship Between Transposable Elements and Adjacent Genes and Transposable Element Islands in the Host Genome

To identify TEs located near gene promoters, a *Barthea*-specific transcription start site (TSS) model was trained using TSSFinder v1.0.0 [35] based on the Arabidopsis model. TSSs with a TATA-box motif in the core promoter region were predicted, and TEs located ≤ 1000 bp upstream of a TSS were considered promoter insertions. A custom pipeline was developed to determine the relationship between each TE and its nearest gene, classifying them as ‘upstream (promoter)’, ‘5-overlap’, ‘5-UTR’, ‘CDS’, ‘intron’, ‘3-UTR’, ‘3-overlap’, or ‘downstream’ (details seen in Figure S2). For nested TEs, spanned regions were determined by integrating TE coordinates into the gene feature coordinates array. This allowed precise mapping of TE and gene feature relationships (https://github.com/altingia/Barthea_TE_manuscript/tree/main/04.TE_gene_relationships accessed on 22 January 2024).

The frequency of each category and their associated distances/overlap lengths were summarized for different TE classes. TE abundance was analyzed using 500 kb sliding windows (100 kb step size) along each linkage group. Windows with $>50\%$ TE content were defined as TE islands, and consecutive islands were merged. Gene ontology (GO) terms and KEGG pathways were assigned to genes within TE islands (foreground) and the entire genome (background) using KOBAS v3.03 [36] with an E-value threshold of 1×10^{-3} . Enrichment analysis was performed, and significance was assessed using the Benjamini–Hochberg (BH) test and False Discovery Rate (FDR) correction [37].

2.5. Transcriptome Profiles for Transposable Elements and Genes Between White and Red Ecotypes

RNA-seq reads from flower buds (three biological replicates per ecotype) were trimmed and filtered using Trimmomatic v0.39 [38] with default parameters (ILLUMINACLIP: TruSeq3-PE.fa:2:30:10, LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, MINLEN:36, TOPHRED33). Trimmed reads were mapped to the reference genome using STAR 2.7.9 [39] with parameters settings ‘—outFilterMultimapNmax 100—winAnchorMultimapNmax 200—outSAMtype BAM SortedByCoordinate’, allowing for multiple mapped reads, as is recommended for TE analyses. Both TE and gene expression levels were normalized to counts per million (CPM). Differential expression analysis between ecotypes was performed using TEdrscripts v2.0.3 [40] with the following parameters: ‘-stranded no-mode multi -p 0.05 -i 10’. Only genes or TEs with reads spanning their entire length were considered transcribed.

2.6. Transposable Element Insertion Polymorphisms Among *B. barthei* Populations

TE insertion polymorphisms (TIPs) were identified using SPLITREADER beta-1.2 [41] and TEPID v0.10 [42], following established recommendations for specificity and sensitivity [43]. We used the programs to identify non-reference insertion sites and reference absence variants. SPLITREADER identifies potential non-reference insertions, which are

then merged and filtered by TE family, retaining sites with ≥ 3 supporting reads in at least one individual. Both pipelines examine negative coverage patterns in flanking regions (100 bp upstream and downstream) to distinguish presence and absence variants. True non-reference insertions exhibit reduced coverage, while true absence variants show coverage drops at their edges. Low-coverage regions lacking sufficient evidence for either presence (positive coverage) or absence (negative coverage) are marked as NA. Single nucleotide polymorphisms (SNPs) were called using the GATK pipeline (The Broad Institute, Cambridge, Massachusetts, USA) following best practices [44,45]. Reads were aligned with BWA-MEM, duplicates were marked with Picard, and base quality scores were recalibrated with GATK's BaseRecalibrator. GATK HaplotypeCaller (The Broad Institute, Cambridge, MA, USA) was used for variant calling, and SNPs were filtered ($QD < 2.0$, $FS > 60.0$, $MQRankSum < -12.5$, $ReadPosRankSum < -8.0$, and $QUAL < 30.0$). Principal component analysis (PCA) of TIPs was performed using the 'prcomp' function from the stats package v3.2.3 in R v3.4.4 [46], retaining the first two eigenvectors. Linkage disequilibrium (LD) for each TIP was analyzed using PLINK v1.9 [47]. The LD coefficient (r^2) was calculated between each TIP and 300 flanking SNPs (upstream and downstream), as well as pairwise r^2 values among all 600 SNPs. The proportion of TIP-SNP pairs with high LD ($r^2 > 0.4$) was compared to the proportion of high LD SNP-SNP pairs. Similar patterns were observed using a less stringent threshold of $r^2 > 0.2$. To account for regional LD variation, TE variants were categorized as 'low', 'medium', or 'high' LD based on their ranked of TIP-SNP r^2 values relative to the median ranked r^2 of SNP-SNP pairs in the same region.

3. Results

3.1. Composition and Proliferation History of Transposable Elements in the *B. barthei* Genome

Using EDTA v2.0.1, we curated a *B. barthei*-specific, non-redundant library of 1081 consensus TE sequences. The RepeatMasker analysis revealed a total of 64,866 TE copies, comprising 16.76% of the genome (40% divergence threshold, and 19,316 copies of other repeat elements (low-complexity regions, microsatellites), comprising 2.1%) (Table 1). DeepTE reclassified 5665 of the 5923 initially unclassified LTRs, assigning 3768 to the Ty3/Gypsy and 1897 to the Ty1/Copia superfamilies (Table 1). The unclassified non-LTRs were categorized as 224 PLEs, 95 LINEs, 38 SINE/tRNAs, and 1 DIR. In total, we identified 24,296 Ty3/Gypsy (8.82%) and 9602 Ty1/Copia (2.47%) copies (11.29% total), non-LTRs including 588 LINEs/L1, 8 LINEs/I, and 278 Penelope copies, and 220 tRNA copies (Table 1).

The Class II DNA transposons constituted 5.28% of the genome (3.44% TIRs, 1.63% Helitrons, and 0.21% MITEs) (Table 1). Among the TIRs, Mutator was most abundant (13,868 copies, 2.7%), followed by CACTA (2820 copies, 0.49%), PIF-Harbinger (747 copies, 0.12%), hAT (516 copies, 0.09%), and Tc1-Mariner (127 copies, 0.04%). Of the 64,864 TEs, 1108 (2.1%) were intact, including 120 LTR/Copia, 220 LTR/Gypsy, 1 nonLTR/PLE, 426 TIR/DTM, 118 TIR/DTC, 81 TIR/DTA, 38 TIR/DTH, 15 TIR/DTT, and 89 Helitron. The remaining 14.66% consisted of truncated or fragmented TEs (Table 1).

The TE age distribution peaked sharply at zero, indicating recent proliferation, followed by a steep decline and a broader distribution from 3 to 10 million years ago (MYA) (Figure 1B). The superfamily proliferation, in decreasing order, was as follows: LTR/Gypsy, DNA/Helitron, DNA/DTM, LTR/Copia, DNA/DTC, and DNA/DTH, with most elements < 15 MYA. Of the 782 LTRs with intact flanking long terminal repeats (188 Ty1/Copia, 381 Ty3/Gypsy, and 213 Unclassified), the insertions occurred between 0.001 and 1.064 MYA and were concentrated within the last 1.00 MYA (Figure 1C). The Copia and Gypsy insertion times largely overlapped, peaking < 0.5 MYA, while Unclassified LTRs had an older peak (Figure 1C). These recent, intact LTRs may still be mobile and functional.

Table 1. Summaries of repeat contents in the genome *Barthea barthei*. Intact transposable elements are shown in the parentheses.

Class	Order	Superfamily	Number	Total Length (bp)	Percentage of the Genome (%)
Retrotransposon	LTR	Copia	9602 (120)	5,809,421 (665,156)	2.47 (0.28)
		Gypsy	24,293 (220)	20,729,316 (2,036,809)	8.82 (0.87)
	DIRs	DIRs	1	254	0
	LINE	L1	588	277,520	0.12
		I	8	13,274	0.01
	PLE	Penelope	278 (1)	176,212 (4083)	0.07 (0.00)
	SINE	tRNA	220	21,226	0.01
DNA Transposon	TIRs	hAT	516 (56)	222,301 (98,002)	0.09 (0.04)
		CACTA	2820 (112)	1,144,474 (310,909)	0.49 (0.13)
		PIF_Harbinger	747 (30)	282,584 (92,465)	0.12 (0.04)
		Mutator	13,868 (296)	6,337,070 (729,432)	2.7 (0.31)
		Tc1-Mariner	127 (14)	97,949 (41,156)	0.04 (0.02)
	MITEs	DTA	455 (25)	107,235 (8582)	0.05 (0.00)
		DTC	7 (6)	2317 (2196)	0 (0.00)
		DTH	70 (8)	11,773 (2513)	0.01 (0.00)
		DTM	1769 (130)	360,581	0.15
		DTT	1 (1)	221 (221)	0 (0.00)
	Helitron	Helitron	9494 (89)	3,834,769 (984,799)	1.63 (0.42)
Other_repeats					
	Other_repeats	Other_repeats	19,316	4,942,020	2.1
Total					18.88

3.2. Retrotransposon Classification Based on RT Domains

The phylogenetic analysis of the RT domains identified 342 Ty1/Copia (RVT_2) and 214 Ty3/Gypsy (RVT_1) domains. Seventeen Ty1/Copia and twelve Ty3/Gypsy family representatives were used to construct a maximum likelihood phylogenetic tree (PROTGAM-MAJTT model, rapid bootstrap). Nine distinct Ty1/Copia lineages were identified, including the abundant *Osser* (74 copies), *Bianca* (73 copies), and *Angela* (47 copies), and less prevalent families such as *Ale* (20 copies), *Sire* (10 copies), *Tork* (12 copies), and *Tar* (12 copies) (Figure 1D). Among the Ty3/Gypsy elements, 108 remained unclassified, while the rest were assigned to *Athila* (36 copies), *Renia* (28 copies), *CRM* (18 copies), *Tekay* (16 copies), *Galadriel* (3 copies), and *Chlamyvir* (1 copy) (Figure 1E).

3.3. Identification and Characterization of Transposable Element Islands in Gene Spaces

The TEs were unevenly distributed across the 20 linkage groups. Thirteen TE islands, ranging from 20.9 kb to 1.4 Mb and encompassing 11.7% of the total TE content, were identified (Figure 1E, Table S1, Figure S2). The largest TE island, located on chromosome 6, spanned 1.4 Mb. These islands contained 1037 protein-encoding genes, exhibiting a significantly lower gene density than the TE-poor regions (Fisher's exact test, $p < 2.2 \times 10^{-16}$). However, these regions were enriched in genes involved in photosynthesis (ath00195,

FDR = 2.2×10^{-9} , BH test) and tryptophan metabolism (ath00380, FDR = 0.04, BH test). Twenty-nine GO terms, particularly those related to abiotic stress (e.g., ‘response to hydroxyurea’, ‘response to aluminum ion’, and ‘response to UV-B’), were also significantly enriched in these TE islands (Table S2). This suggests that *B. barthei* genetic evolution has been shaped by adaptation to the acidic, aluminum-rich soils and variable light conditions of South China.

3.4. Genome-Wide Survey of Transposable Element Insertion Preferences in the *B. barthei* Genome

A genome-wide survey mapped TE insertion sites in *B. barthei* using a custom pipeline (https://github.com/altinia/Barthea_TE_manuscript/tree/main/04.TE_gene_relationships (accessed on 23 January 2024)) (Figure S3, Table S3). Of the 64,864 identified TE copies, 63,665 were mapped relative to neighboring genes, and 1199 were found on unannotated scaffolds (Table 2). The TEs preferentially inserted upstream (45.8%; 29,144 copies) compared to downstream (41.7%; 26,562 copies). The LTR retrotransposons were frequently located >5 kb from neighboring genes (Figure S4A,B), while the DNA transposons (TIRs and Helitrons) were predominantly within 5 kb of genes (Figure S4C,D). A total of 17,991 upstream and 14,845 downstream TEs were within 5 kb of neighboring genes. Additionally, 577 TE copies were identified within gene promoters, affecting 431 genes (Table 2). Approximately 26.1% of *B. barthei* genes had at least one TE copy within 1 kb, lower than *Arabidopsis* (36%) or maize (86%) [48,49]. The most common TE types near genes were LTR/Gypsy, DNA/TIRs, LTR/Copia, DNA/Helitron, and nLTR (Table 2). Only a small proportion of the genes with adjacent TEs showed a significant GO term enrichment, including ‘ADP binding (GO0043531)’ and ‘response to ethylene (GO0009723)’ (Tables S4 and S5). Additionally, 760 TE copies overlapped with gene 5’ ends and 630 with 3’ ends.

Table 2. The summaries of coordinate relationships between different orders or superfamilies of transposon elements and the nearest host genes in the genome *Barthea barthei*.

		Copia	Gypsy	TIR	Helitron	nLTR	Total	Proportion (%)
Upstream	Promoter	64	95	254	161	3	577	0.9
	other	4480	11,422	8599	3970	96	28,567	44.9
Downstream		4185	10,611	7999	3635	132	26,562	41.7
5'-Overlap		71	246	281	159	3	760	1.2
3'-Overlap		61	249	209	111	0	630	1.0
Nested							6569	10.3
	Single intron	543	1333	2347	1213	18	5454	8.6
	Single 5' UTR	43	47	59	41	0	190	0.3
	Single 3' UTR	61	112	119	78	0	370	0.6
	Single CDS	3	23	37	12	0	75	0.1
	Exon/intron	91	155	123	108	3	480	0.7
Total		9602	24,293	20,027	9488	255	63,665	

A substantial number of TEs (6569 copies, 10.3%) were nested within host genes (intronic, coding sequence (CDS), or untranslated region (UTR)) (Table 2). Single-intron nested TEs were the most prevalent (5454 copies in 2560 genes). *Barthea36295* (homologous to *Vitis vinifera* LUTEIN DEFICIENT 5, a cytochrome P450 member [50]) contained the most nested TEs (37 copies, 23 in its third intron) (Figure S5A). The genes with single-intron nested TEs were significantly enriched for ‘cytosol’ (GO:0005829, FDR = 3.3×10^{-6}) and ‘cytoplasm’ (GO:0005737, FDR = 3.2×10^{-4}) (Table S6).

Exon-nested TEs (635 copies in 227 genes) were predominantly in UTRs (190 in 5' UTR, 370 in 3' UTR) rather than CDS regions (75 copies in 43 genes). *Barthea44534* (ho-

mologous to beta-glucosidase 12-like [51]) contained the most UTR-nested TEs (22 copies) (Figure S5B). *Barthea28554* (homologous to the RNA polymerase beta subunit) contained the most DNA/DTM copies (Figure S5C). While eight genes lacked significant hits, 35 others resembled conserved proteins (ribonuclease, reverse transcriptase, plant disease resistance polypeptide, heat shock protein, etc.) (Table S7). The genes with exon-nested TEs were associated with cellular and stress responses, including protein folding chaperone activity and response to temperature stimulus (Table S8). Additionally, 480 TE copies spanning multiple introns/exons were found in 618 genes, and these were linked to developmental processes and epigenetic modification (Table S9).

3.5. Expression Profiles of Genes and Transposable Elements During Flower Bud Development Between Contrasting Ecotypes

The RNA-sequencing sample quality was validated by hierarchical clustering, demonstrating a high reproducibility within the ecotypes (Figure S6). The transcriptome analysis identified 40,538 expressed genes and 4495 transcribed TE copies (Figure 2, Table S10). The differentiation expression analysis (baseMean > 10, $|\log_2\text{FoldChange}| \geq 2$, and $\text{Padj} < 0.05$) identified 2544 differentially expressed genes (DEGs), comprising 1357 up-regulated and 1187 down-regulated genes between the two ecotypes (Table S11).

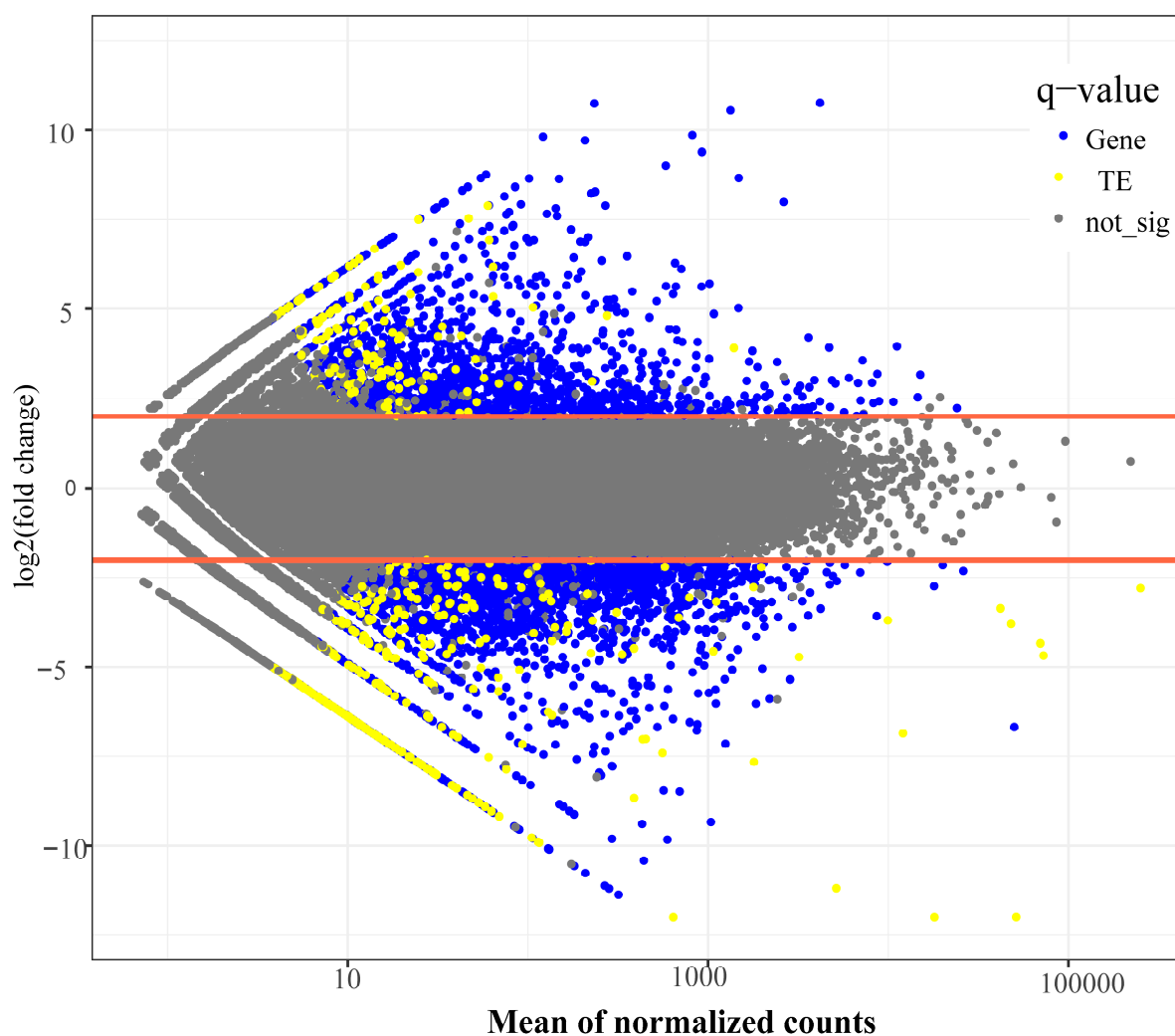


Figure 2. The distribution of differentially expressed genes and transposable elements between the white ecotype and red ecotype during flower bud development of *Barthea barthei*. The significant levels were determined using a $|\log_2\text{FoldChange}| \geq 2$ (as delimited by the red lines) and an adjusted p -value of 0.05.

The KEGG pathway enrichment analysis of the DEGs revealed three significant pathways (Table S12). ‘Photosynthesis-antenna proteins’ (KO00196, FDR = 3.23×10^{-8} , BH test) included 20 DEGs (12 Arabidopsis homologs, including LHCA1~LHCA6 and LHCB1~LHCB6) (Table S13), which are crucial for energy equilibrium under variable light [52]. ‘Phenylpropanoid biosynthesis’ (KO00940, FDR = 1.50×10^{-7} , BH test) comprised 51 DEGs (30 Arabidopsis homologs) (Tables S14 and S15), which are involved in lignin and flavonoid synthesis and the light-induced stress response [53]. ‘Cutin, suberine, and wax biosynthesis’ (KO00073, FDR = 2.09×10^{-4} , BH test) contained 15 DEGs (*Arabidopsis* homologs, including HXXXD-type acyl-transferase family protein, FAR2, CYP86A1, CER1, CYP704B1, and CYP86B1) (Table S16), which are involved in protective cuticular wax production against environmental stressors [54,55].

Of the transcribed TEs, retrotransposons were abundant (1457 transcripts): 1059 LTR/Gypsy, 206 LTR/Copia, 148 LTR/unknown, and 44 LINEs/unknown. These were frequently near genes or within single introns, with LTR retroelements averaging 5738.5 bp from adjacent genes (Tables 3 and S17). DNA transposon transcripts (1938, 43.1% of expressed TEs) were also detected, including 403 DNA/DTC, 626 DNA/DTM, 48 DNA/DTH, 45 DNA/DTA, 27 DNA/DTT, 140 MITEs, and 649 Helitrons. Of these, 855 (19.0%) were within intron/UTR/CDS regions, and the rest averaged 1881 bp from genes (Tables 3 and S17). The consistent detection of DNA transposon transcripts across the samples suggests they are not mere DNA contaminants, as previously suggested in other RNA-seq studies [56,57]. Their distribution (42.8% nested, 57.2% within 2000 bp of host genes) indicates probable passive co-transcription. Furthermore, 1083 unclassified TE transcripts were identified, with 391 nested and 605 averaging 7995.8 bp from genes (Table 3).

Table 3. Coordinated expression of transposable element transcripts and nearby/nested genes during flower bud development in *Barthea barthei*.

Class	Order	Superfamily	Copy Number	Average BaseMean	^a Instances of Coordination
Class I (Retransposons)	LTR		1457	9.85	
		Copia	206	5.62	69:7:2:54:2:7:1:64
		Gypsy	1059	10.55	392:18:4:207:13:0:4:24
		Unknown	148	9.37	33:4:0:55:2:0:0:4
	LINE	Unknown	44	14.53	4:1:0:30:3:0:0:6
Class II (DNA transposons)	TIR		1149	16.67	
		DTA	45	13.1	11:5:0:19:0:0:1:9
		DTC	403	18.43	57:30:0:235:7:2:16:56
		DTH	48	13.14	8:10:0:17:0:0:2:11
		DTM	626	8.43	174:25:4:216:3:0:38:165
		DTT	27	17.67	9:2:0:2:0:0:3:11
	MITE		140	6.17	
		DTA	68	8.15	5:0:0:58:1:0:0:4
		DTC	3	2.1	0:0:0:1:0:0:0:2
		DTH	13	4.51	2:1:0:8:0:0:0:2
		DTM	56	4.37	22:2:0:18:0:0:2:12
		DTT	0	0	
	Helitron		649	24	136:47:2:264:8:3:39:150
			1083	573	289:70:7:360:31:6:34:316
Unknown pararetrovirus			17	5.24	4:0:0:5:1:0:0:7

^a Upstream:5'-overlap:5'UTR: intron:intron-CDS (CDS):3'UTR:3'-overlap:downstream.

The differential expression analysis of the TEs identified 343 significantly differentially expressed elements between the ecotypes (90 upregulated and 253 downregulated) (Figure 2, Table S18). These included 76 LTR/Gypsy, 16 LTR/Copia, 9 LTR/unknown, 125 DNA transposons, and 117 unknown elements. Thirty of these differentially expressed TEs were located near or within genes (Table S19). Overall, the TE expression levels were significantly lower than the gene expression levels (t -test, $p = 1.91 \times 10^{-7}$).

3.6. Transposable Element Insertion Polymorphisms Among the Two Ecotypes

Stringent filtering identified 3859 high-confidence TIPs, including 3615 non-reference TE insertions and 244 reference absences (Table S20). Most of the TIPs were attributed to Gypsy (2764) and Copia (231) retrotransposons and the DNA transposon superfamilies DNA/DTM (488) and DNA/DTA (128) (Figure 3A). The site frequency spectrum (SFS) analysis showed 1.62% of the TIPs had a minor allele frequency < 0.05 , though this may be underestimated due to the sample size (Figure 3B).

The TIP distribution analysis revealed significantly higher variant frequencies in intergenic regions (1699), compared to introns (415), UTRs (264), and CDSs (95). Across each category, the variant frequency exhibited a consistent decrease as the MAF bins increased from 0–0.1 to 0.4–0.5 (Figure 3C). Notably, specific superfamily groups exhibited similar patterns: 2764 TIPs in LTR/Gypsy, 488 in DNA/DTM, 231 in LTR/Copia, 156 in Helitron, 128 in DNA/DTA, 56 in DNA/DTC, and 28 in LTR/Unknown (TIP counts > 10 for each superfamily) (Figure 3D). Across all the superfamilies, a consistent negative correlation between the variant frequency and the MAF bins was observed, indicating a purifying selection on these TIPs (Figure 3D). The PCA of the TIPs demonstrated a significant genetic differentiation between the HD and YC populations, highlighting the role of the TIPs in distinguishing genetic ancestry (Figure 3E). The LD analysis with the nearby SNPs revealed that 51.4%, 39.0%, and 9.6% of the TE variants exhibited low, intermediate, and high LD, respectively (Figure 3F). A positive correlation was observed between the MAF and the LD, with the higher MAF variants more frequently exhibiting a high LD (Figure 3F). This pattern was consistent for both TE insertions and deletions (Figure 3G). Notably, the LD category assignments were largely consistent between insertions and deletions (Figure 3G), supporting the observation that common alleles tend to be in high-LD states.

Twenty-nine fixed TIPs associated with 28 genes were identified between the red ecotype and white ecotype populations (Table S21). These TIPs were distributed as follows: 22 intergenic, 3 intronic, and 2 each in CDS and UTR regions. The affected genes included homologs of transcription factors (AP2/ERF, bHLH, MYB-like, and C2H2) and enzymes (Homocysteine S-methyltransferase (HMT), alpha/beta hydrolase fold, and 2OG-Fe (II) oxygenase superfamily) (Table S22). Three genes with fixed TIPs were differentially expressed between the white and red color ecotypes. *Barthea13541* (\log_2 fold change = -2.70 , p -adj = 0.0003), containing a non-reference DNA/DTM insertion in the white ecotype population, is homologous to the Pmr5/Cas1p GDSL/SGNH-like acyl-esterase family protein, influencing cell wall modification, plant–pathogen interactions, stress responses, and development [58]. *Barthea35147* (\log_2 Fold Change = 3.90, p -adj = 0.002), with a fixed DNA/DTM present in HD but absent in YC, is homologous to the AP2/ERF transcription factor, regulating plant morphogenesis, stress responses, hormone signaling, and metabolism [59,60]. *Barthea36291* (\log_2 Fold Change = 4.05, p -adj = 0.003), with a fixed DNA/DTA present in the white ecotype but absent in the red ecotype, is homologous to Cytochrome P450 (CYP), synthesizing compounds crucial for membrane structure, hormones, UV protection, pigments, signaling, and volatile compounds mediating biotic and abiotic interactions [61].

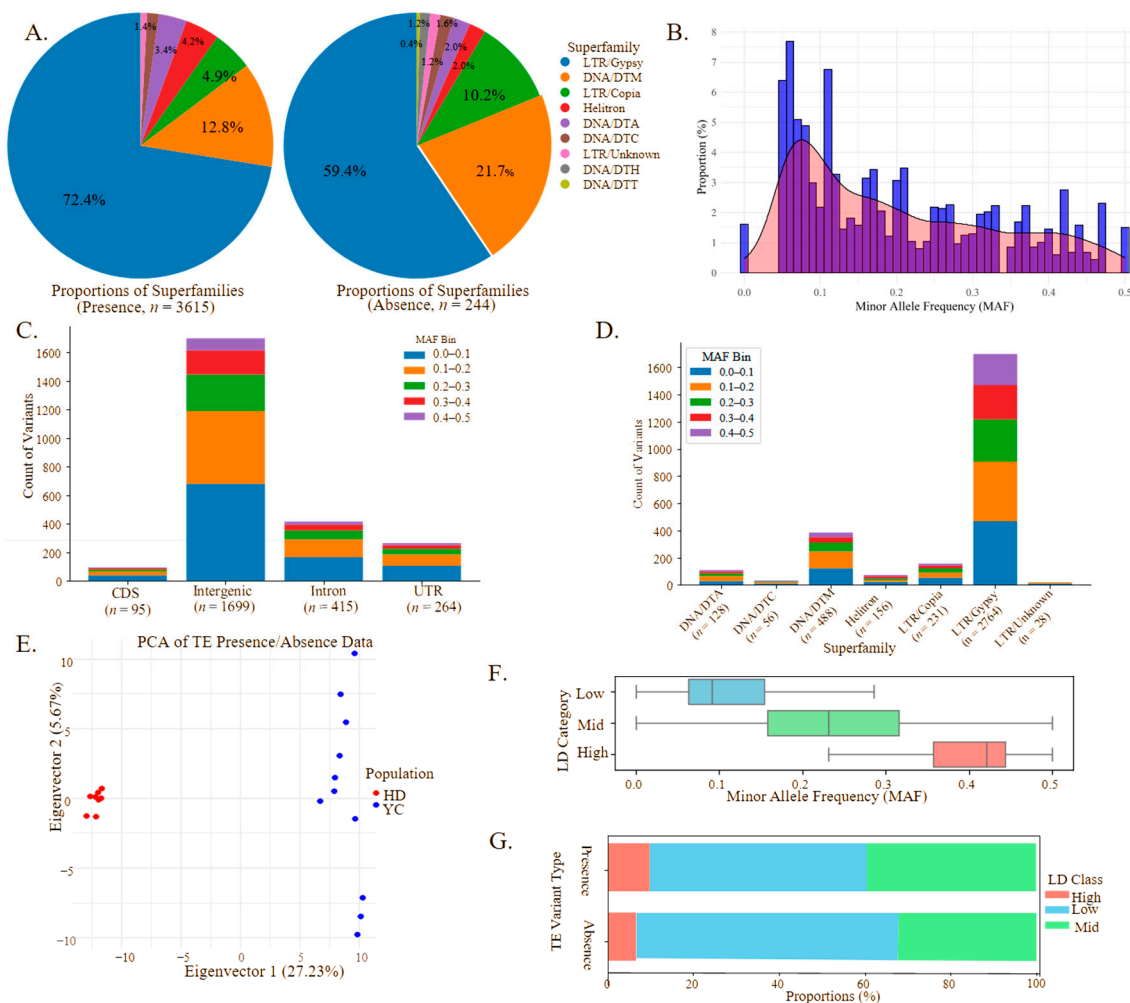


Figure 3. Transposable element insertion polymorphism patterns between the white ecotype from the YC population and the red ecotype from the HD population of *Barthea barthei*. **(A)** Superfamily components and proportions of non-reference transposon element insertion variants (**Left**) and reference absence variants (**Right**). The different superfamily color codes are as follows: LTR/Gypsy (dark blue), LTR/Copia (green), LTR/Unknown (pink), DNA/DTM (orange), DNA/DTA (purple), DNA/DTC (brown), DNA/DTH (gray), DNA/DTT (olive green), and Helitron (red). **(B)** The minor allele frequency (MAF) distribution of transposable element insertion polymorphisms for *B. barthei*. **(C)** Counts of TE variants with different minor allele frequencies within each genomic feature classified as coding regions (CDS), intergenic regions, introns, and untranslated regions (UTRs). Color coding for MAF bins are as follows: 0.0–0.1 (blue), 0.1–0.2 (orange), 0.2–0.3 (green), 0.3–0.4 (red), and 0.4–0.5 (purple). **(D)** Counts of TE variants with different minor allele frequencies within each TE superfamily including LTR/Gypsy, LTR/Copia, LTR/Unknown, DNA/DTM, DNA/DTA, DNA/DTC, and Helitron. Color coding for MAF bins is as in Panel C. **(E)** Principal component analysis for the samples of red ecotype (HD population, red dots) and white ecotype (YC population, blue dots) based on non-reference TE insertion variants and reference TE absence variants. **(F)** Minor allele frequency distribution by relative TE-SNP linkage disequilibrium. Boxplots of minor allele frequency (MAF) for genetic variants are grouped by the relative linkage disequilibrium (LD) of nearby transposable elements (TEs) with SNPs. LD categories (Low, Mid, High) are based on the ranking of TIP-SNP r^2 values relative to the median ranked SNP-SNP r^2 within the same region. Boxes show interquartile range (IQR), with median indicated; whiskers extend to $1.5 \times$ IQR. **(G)** TE variant LD class distribution by presence/absence. Stacked bar chart showing the proportion of transposable element (TE) variants (TIPs) in each relative linkage disequilibrium (LD) class (High, Low, Mid) for two TIP states: Presence and Absence. LD classes are defined relative to regional SNP-SNP LD. Bars represent 100% for each TIP state.

4. Discussion

4.1. Impact and Distribution of Transposable Elements in the Compact Genome of *B. barthei*

The *B. barthei* genome (235.03 Mb) exhibits a compact architecture with a low transposable element content, consistent with other streamlined plant genomes. Its TE proportion aligns with the broad spectrum observed across plant genomes, ranging from 2.5% in *U. gibba* [7] to 85% in maize [62]. Within this compact genome, the Ty1/Copia and Ty3/Gypsy superfamilies dominate, distinguished by divergent reverse transcriptase (RT) and integrase (INT) domain arrangements in their POL polypeptides. The predominance of Ty3/Gypsy elements across Viridiplantae [63] may reflect their distinct genomic dynamics: older Ty3/Gypsy elements typically localize heterochromatin, whereas younger Ty1/Copia elements preferentially occupy euchromatic regions with a reduced recombination suppression [64,65].

The Class II DNA transposons in the *B. barthei* genome encompass five major superfamilies but lack the P superfamily. Their abundance varies markedly among plant species [66], with TIR and Helitron transposons showing greater interspecific variability than LTR retrotransposons. For instance, *Arabidopsis* harbors more Helitron elements than TIR elements, and the DNA transposon content can diverge significantly even between closely related taxa [67]. Although DNA transposons typically constitute a small fraction of plant genomes, their frequent gene-proximal localization and mutagenic potential profoundly influence genome architecture [66]. Notably, Mutator-like elements (MULEs), which elevate mutation rates up to 50-fold in plants [68], dominate *B. barthei*'s DNA transposon repertoire. Pack-MULEs further drive biased gene modifications through targeted insertion and sequence capture [4], positioning these elements as key contributors to *B. barthei*'s genomic plasticity.

Most of the TEs in *B. barthei* reside in intergenic regions (Table 2), with merely 577 copies (0.9%) localized to promoters including putative cis-regulatory elements. While the majority appear linearly distal from genes, chromatin spatial organization (e.g., 3D folding) may mediate their physical proximity to transcriptional units [69]. Intragenic TEs (6569 copies, 10.3%) likely reshape regulatory networks by introducing cryptic splicing sites, facilitating exonization, or embedding novel regulatory motifs [70,71]. Collectively, TEs orchestrate gene expression through both sequence-encoded features and chromatin-mediated spatial interactions, forming a dynamic regulatory framework that amplifies transcriptional diversity.

Our genome-wide TE characterization reveals their dual evolutionary role in *B. barthei* as generators of adaptive variation and potential genomic destabilizers. The compact genome architecture may reflect an evolutionary equilibrium between TE-driven adaptability and stability maintenance. While this study delineates the TE landscape, the ecological relevance of these elements—particularly under environmental stressors or niche-specific pressures—remains an open question, warranting functional studies to dissect their contribution to adaptive evolution.

4.2. TE Islands Facilitate *B. barthei* Adaptation to Tropical Forests

Despite their sparse gene distribution, TE islands in *B. barthei* exhibit non-random gene enrichment, notably in photosynthesis, tryptophan metabolism, and abiotic stress tolerance pathways. Photosynthesis, which is essential for carbon fixation, is highly sensitive to abiotic stressors such as ultraviolet (UV) radiation, fluctuating light intensity, and oxygen deprivation [72]. Tryptophan metabolism plays a key defensive role against pathogens, as shown in rice and *Arabidopsis* [73,74]. Its enrichment in *B. barthei*'s TE islands may provide an evolutionary advantage against pathogens and herbivores, consistent with the Janzen–Connell hypothesis regarding tropical forest diversity [75].

B. barthei's distribution, extending into the subtropical regions of southern China, aligns with MacArthur's (1972) ecological paradigm, which posits that species' ranges are limited by abiotic factors at one extreme and biotic pressures at the other [76]. *B. barthei* has evolved adaptations to overcome these constraints, including enhanced photosynthetic capacity, polymorphic flower colors, and diverse growth forms (personal observations). High-altitude populations, including the red-flowered ecotype from Huidong, Guangdong, exemplify these adaptations. Adapted to mountain mists, variable rainfall, strong winds, and intense sunlight, this ecotype displays deep red flowers, thick coriaceous leaves, and a compact growth structure.

While the adaptive significance of TE islands is documented in species such as the invasive ant *C. obscurior* [13], where they facilitate a relaxed selection and genetic modifications for enhanced chemical perception, learning, and insecticide resistance, their role in plant adaptation is less clear. In *B. barthei*, the observed flower color and leaf texture diversity suggests that the TE islands may be crucial for niche expansion. Further research is needed to understand how the TEs adjacent to candidate genes regulate these adaptive traits, revealing how TE islands contribute to *B. barthei*'s successful adaptation to tropical forest environments.

4.3. Transposable Element Insertion Preferences and Polymorphisms Associated with Ecological Divergence Between Ecotypes

The distribution of TEs in *B. barthei* reveals distinct insertion patterns and their potential role in ecological adaptation. While only 26.1% of *B. barthei* genes are associated with TEs, compared to 78% in maize [77], both species share similar TE distribution patterns, with most copies within 5 kb of genes. Our survey identified 3859 high-confidence TIPs, primarily Gypsy (2764) and Copia (231) retrotransposons and DNA transposon superfamilies DNA/DTM (488) and DNA/DTA (128). The TIP distribution mirrors TE insertion preferences, with most in intergenic regions (1699) followed by introns (415), UTRs (264), and CDS regions (95).

The non-random TE distribution reflects the interplay of natural selection and genetic drift [78]. The decreasing TIP variant frequency with increasing MAF suggests an ongoing selection, especially against deleterious gene body insertions. This is supported by the LD patterns: 51.4% of the TE variants showed a low LD with nearby SNPs, and only 9.6% exhibited a high LD. The positive MAF–LD correlation suggests that beneficial or neutral TE insertions are more likely to persist and become common variants.

The TE insertion impact varies by genomic location [79]. Regulatory region (promoter, enhancer) insertions can result in gene expression nullification, enhancement (via new cis-regulatory sites), or silencing (via repressive chromatin marks) [80]. Examples include a Copia-like element insertion upstream of the *ruby* gene, a key MYB transcriptional activator that affects anthocyanin production in blood oranges [81] and two independent insertions by a MITE and an LTR element at the *teosinte branched1 (tb1)* locus altering maize ear morphology during domestication [82].

The significant genetic differentiation between the red ecotype and the white ecotype populations (TIP-based PCA) highlights the role of TE polymorphisms in ecological divergence. The 29 fixed TIPs associated with 28 genes, including three differentially expressed genes, may contribute to ecotype-specific adaptations. The fixed DNA/DTM insertion in *Barthea13541* (YC) potentially affects cell wall modification and stress responses. Similarly, the ecotype-specific insertions near *Barthea35147* (AP2/ERF homolog) and *Barthea36291* (Cytochrome P450 homolog) likely contribute to divergent morphogenesis, stress responses, and secondary metabolism between the white and red ecotypes.

While intron-nested insertions (8.6% of total) are linked to phenotypic variation (e.g., double flower in morning glories [83] and yellowhorns [84]), CDS insertions are rare due to

a strong negative selection. The few surviving CDS-nested TEs, such as the Copia insertion affecting soybean salt sensitivity [85], are typically ancient and conserved. The insertion distribution, combined with fixed inter-ecotype polymorphisms, suggests that TEs have significantly contributed to *B. barthei* adaptive divergence across ecological niches.

5. Conclusions

Our comprehensive analysis of the *B. barthei* TE landscape reveals key insights into genome organization and ecological adaptation. The relatively compact *B. barthei* genome (235 Mb) has a low TE proportion with distinct distribution patterns dominated by Ty3/Gypsy retrotransposons and Mutator-like DNA transposons. Despite their limited abundance, these TEs exhibit a non-random distribution, with most within 5 kb of genes, suggesting their potential regulatory roles in genome evolution and adaptation.

The identification of TE islands enriched in photosynthesis, tryptophan metabolism, and stress response genes provides compelling evidence for their role in ecological adaptation. These genomic features likely facilitated *B. barthei*'s successful colonization of diverse tropical and subtropical habitats in southern China, as evidenced by distinct ecotypes including the red-flowered variant in Huidong.

An analysis of 3859 high-confidence TIPs revealed a significant genetic differentiation between the HD and YC ecotypes. The presence of twenty-nine fixed TIPs in 28 genes (involved in cell wall modification, stress responses, and secondary metabolism) suggest that TE-mediated genetic variation has substantially contributed to the ecological divergence of *B. barthei*. The observed TE distribution and polymorphism patterns (preferential intergenic insertion, strong selection against CDS insertions) reflect the balance between adaptive variation and genome stability.

These findings advance our understanding of the contribution of TEs to plant genome evolution and adaptation, particularly in tropical forests. Future research should investigate the specific regulatory mechanisms by which TE insertions influence adaptive traits, potentially revealing the role of mobile genetic elements in plant speciation and ecological divergence.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biom15030346/s1>, Figure S1: Analytical pipeline for transposable element dynamics and ecological adaptation in *Barthea barthei*; Figure S2: The distribution of transposable element landscapes shown for the genome assembly of *Bathea barthei* across the chromosomes between chr09 and chr20 (continued from Figure 1F); Figure S3: Schematic representation of the relationships between transposable elements and the nearest flanking host genes in the genome assembly of *Barthea barthei*; Figure S4: The distance distributions between the transposable elements and their nearest flanking host genes in the genome assembly of *Barthea barthei* for four categories of transposable elements; Figure S5: The Illustrations of different scenarios for host genes nested with distinct transposable elements in the genome assembly of *Barthea barthei*; Figure S6: Heatmap of samples similarities based on Euclidean distances calculated using the normalized counts of gene/transposable element for the leaf transcriptomes of *Barthea barthei*; Supplementary Tables S1–S22: separate Excel sheets.

Author Contributions: Conceptualization, W.W. and B.X.; methodology, W.W., Y.Z., B.X., Z.H., H.P., and Z.S.; software, W.W., Y.Z., B.X., Z.H., H.P., and Z.S.; validation, W.W., Y.Z., and Z.H.; formal analysis, W.W., Y.Z., and B.X.; investigation, W.W. and Y.Z.; resources, W.W. and B.X.; data curation, W.W. and Y.Z.; writing—original draft preparation, W.W. and Y.Z.; writing—review and editing, W.W., Y.Z., B.X., Z.H., H.P., and Z.S.; visualization, W.W. and Y.Z.; supervision, W.W. and B.X.; project administration, W.W. and B.X.; funding acquisition, W.W. and B.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was financially supported by the Science and Technology Program from the Forestry Administration of Guangdong Province, grant number 2020KJCX002.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data of whole genome sequencing and RNA sequencing are deposited in the Genome Sequence Archive (GSA accession: CRA012896 under the project PR-JCA020264) in the China National Genomics Data Center (NGDC) database. The genome assemblies and annotations are available at [https://github.com/alingia/Barthea_TE_manuscript accessed on 23 January 2023] under accession number GWHUDUN00000000.

Acknowledgments: We are grateful to Yubing Zhou for technical assistance.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DIRS	DIRS-like elements
DTA	hAT
DTC	CACTA
DTH	PIF-Harbinger
DTM	Mutator
DTT	Tc1-Mariner
GO	Gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
LD	Linkage disequilibrium
LINEs	Long interspersed sequences
LTRs	Long terminal repeats
MITEs	Miniature inverted repeats
MYA	Million years ago
P	DTP
PCA	Principal component analysis
PLEs	Penelope-like elements
RLC	LTR/Copia
RLG	LTR/Gypsy
SINEs	Short interspersed sequences
SNPs	Single nucleotide polymorphisms
TE	Transposable element
TIRs	Terminal inverted repeats
TSDs	Target site duplications
TSS	Transcription start site
UTR	Untranslated region

References

1. Finnegan, D.J. Eukaryotic transposable elements and genome evolution. *Trends Genet.* **1989**, *5*, 103–107. [[CrossRef](#)] [[PubMed](#)]
2. Arkhipova, I.R. Using bioinformatic and phylogenetic approaches to classify transposable elements and understand their complex evolutionary histories. *Mob. DNA* **2017**, *8*, 19. [[CrossRef](#)]
3. Wicker, T.; Sabot, F.; Hua-Van, A.; Bennetzen, J.L.; Capy, P.; Chalhoub, B.; Flavell, A.; Leroy, P.; Morgante, M.; Panaud, O.; et al. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **2007**, *8*, 973–982. [[CrossRef](#)] [[PubMed](#)]
4. Jiang, N.; Bao, Z.; Zhang, X.; Eddy, S.R.; Wessler, S.R. Pack-MULE transposable elements mediate gene evolution in plants. *Nature* **2004**, *431*, 569–573. [[CrossRef](#)] [[PubMed](#)]
5. Lisch, D. How important are transposons for plant evolution? *Nat. Rev. Genet.* **2013**, *14*, 49–61. [[CrossRef](#)] [[PubMed](#)]
6. Ariel, F.D.; Manavella, P.A. When junk DNA turns functional: Transposon-derived non-coding RNAs in plants. *J. Exp. Bot.* **2021**, *72*, 4132–4143. [[CrossRef](#)]

7. Ibarra-Laclette, E.; Lyons, E.; Hernández-Guzmán, G.; Pérez-Torres, C.A.; Carretero-Paulet, L.; Chang, T.-H.; Lan, T.; Welch, A.J.; Juárez, M.J.A.; Simpson, J.; et al. Architecture and evolution of a minute plant genome. *Nature* **2013**, *498*, 94–98. [[CrossRef](#)]
8. Nystedt, B.; Street, N.R.; Wetterbom, A.; Zuccolo, A.; Lin, Y.-C.; Scofield, D.G.; Vezzi, F.; Delhomme, N.; Giacomello, S.; Alexeyenko, A.; et al. The Norway spruce genome sequence and conifer genome evolution. *Nature* **2013**, *497*, 579–584. [[CrossRef](#)] [[PubMed](#)]
9. Elliott, T.A.; Gregory, T.R. Do larger genomes contain more diverse transposable elements? *BMC Evol. Biol.* **2015**, *15*, 69. [[CrossRef](#)]
10. Stitzer, M.C.; Anderson, S.N.; Springer, N.M.; Ross-Ibarra, J. The genomic ecosystem of transposable elements in maize. *PLOS Genet.* **2021**, *17*, e1009768. [[CrossRef](#)]
11. Wells, J.N.; Feschotte, C. A field guide to eukaryotic transposable elements. *Annu. Rev. Genet.* **2020**, *54*, 539–561. [[CrossRef](#)] [[PubMed](#)]
12. Catlin, N.S.; Josephs, E.B. The important contribution of transposable elements to phenotypic variation and evolution. *Curr. Opin. Plant Biol.* **2022**, *65*, 102140. [[CrossRef](#)]
13. Schrader, L.; Kim, J.W.; Ence, D.; Zimin, A.; Klein, A.; Wyschetzki, K.; Weichselgartner, T.; Kemena, C.; Stökl, J.; Schultner, E.; et al. Transposable element islands facilitate adaptation to novel environments in an invasive species. *Nat. Commun.* **2014**, *5*, 5495. [[CrossRef](#)] [[PubMed](#)]
14. Wei, L.; Cao, X. The effect of transposable elements on phenotypic variation: Insights from plants to humans. *Sci. China Life Sci.* **2016**, *59*, 24–37. [[CrossRef](#)] [[PubMed](#)]
15. Lisch, D. Epigenetic Regulation of Transposable Elements in Plants. *Annu. Rev. Plant Biol.* **2009**, *60*, 43–66. [[CrossRef](#)] [[PubMed](#)]
16. Casacuberta, E.; González, J. The impact of transposable elements in environmental adaptation. *Mol. Ecol.* **2013**, *22*, 1503–1517. [[CrossRef](#)] [[PubMed](#)]
17. Rey, O.; Danchin, E.; Mirouze, M.; Loot, C.; Blanchet, S. Adaptation to global change: A transposable element–epigenetics perspective. *Trends Ecol. Evol.* **2016**, *31*, 514–526. [[CrossRef](#)]
18. Iida, S.; Morita, Y.; Choi, J.-D.; Park, K.-I.; Hoshino, A. Genetics and epigenetics in flower pigmentation associated with transposable elements in morning glories. *Adv. Biophys.* **2004**, *38*, 141–159. [[CrossRef](#)] [[PubMed](#)]
19. Zhang, L.; Hu, J.; Han, X.; Li, J.; Gao, Y.; Richards, C.M.; Zhang, C.; Tian, Y.; Liu, G.; Gul, H.; et al. A high-quality apple genome assembly reveals the association of a retrotransposon and red fruit colour. *Nat. Commun.* **2019**, *10*, 1494. [[CrossRef](#)]
20. Martin, A.; Troadec, C.; Boualem, A.; Rajab, M.; Fernandez, R.; Morin, H.; Pitrat, M.; Dogimont, C.; Bendahmane, A. A transposon-induced epigenetic change leads to sex determination in melon. *Nature* **2009**, *461*, 1135–1138. [[CrossRef](#)] [[PubMed](#)]
21. Mao, H.; Wang, H.; Liu, S.; Li, Z.; Yang, X.; Yan, J.; Li, J.; Tran, L.-S.P.; Qin, F. A transposable element in a NAC gene is associated with drought tolerance in maize seedlings. *Nat. Commun.* **2015**, *6*, 8326. [[CrossRef](#)] [[PubMed](#)]
22. Yang, Q.; Li, Z.; Li, W.; Ku, L.; Wang, C.; Ye, J.; Li, K.; Yang, N.; Li, Y.; Zhong, T.; et al. CACTA-like transposable element in ZmCCT attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16969–16974. [[CrossRef](#)] [[PubMed](#)]
23. Chen, C.; Renner, S. Melastomataceae. In *Flora of China*; Wu, Z., Raven, P., Eds.; Science Press: Beijing, China; Missouri Botanical Garden Press: St. Louis, MO, USA, 2007.
24. Huang, G.; Liu, H.; Sun, H.; Liu, Y.; Zhou, R.; Liao, W.; Fan, Q. Development and characterization of 18 polymorphic SSR markers for *Barthea barthei* (Melastomataceae). *Appl. Plant Sci.* **2017**, *5*, 1600149. [[CrossRef](#)]
25. Huang, W.; Xu, B.; Guo, W.; Huang, Z.; Li, Y.; Wu, W. De novo genome assembly and population genomics of a shrub tree *Barthea barthei* (Hance) krass provide insights into the adaptive color variations. *Front. Plant Sci.* **2024**, *15*, 1365686. [[CrossRef](#)] [[PubMed](#)]
26. Ou, S.; Su, W.; Liao, Y.; Chougule, K.; Agda, J.R.A.; Hellinga, A.J.; Lugo, C.S.B.; Elliott, T.A.; Ware, D.; Peterson, T.; et al. Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.* **2019**, *20*, 275. [[CrossRef](#)] [[PubMed](#)]
27. Yan, H.; Bombarely, A.; Li, S. DeepTE: A computational method for de novo classification of transposons with convolutional neural network. *Bioinformatics* **2020**, *36*, 4269–4275. [[CrossRef](#)] [[PubMed](#)]
28. Marchler-Bauer, A.; Derbyshire, M.K.; Gonzales, N.R.; Lu, S.; Chitsaz, F.; Geer, L.Y.; Geer, R.C.; He, J.; Gwadz, M.; Hurwitz, D.I. CDD: NCBI’s conserved domain database. *Nucleic Acids Res.* **2015**, *43*, D222–D226. [[CrossRef](#)]
29. Neumann, P.; Novák, P.; Hošťáková, N.; Macas, J. Systematic survey of plant LTR-retrotransposons elucidates phylogenetic relationships of their polyprotein domains and provides a reference for element classification. *Mob. DNA* **2019**, *10*, 1. [[CrossRef](#)] [[PubMed](#)]
30. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
31. Darriba, D.; Posada, D.; Kozlov, A.M.; Stamatakis, A.; Morel, B.; Flouri, T. ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. *Mol. Biol. Evol.* **2020**, *37*, 291–294. [[CrossRef](#)]
32. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)] [[PubMed](#)]

33. Drost, H.-G. LTRpred: De novo annotation of intact retrotransposons. *J. Open Source Softw.* **2020**, *5*, 2170. [\[CrossRef\]](#)
34. Ma, J.; Bennetzen, J.L. Rapid recent growth and divergence of rice nuclear genomes. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12404–12410. [\[CrossRef\]](#) [\[PubMed\]](#)
35. de Medeiros Oliveira, M.; Bonadio, I.; Lie de Melo, A.; Mendes Souza, G.; Durham, A.M. TSSFinder—Fast and accurate ab initio prediction of the core promoter in eukaryotic genomes. *Brief. Bioinform.* **2021**, *22*, bbab198. [\[CrossRef\]](#)
36. Bu, D.; Luo, H.; Huo, P.; Wang, Z.; Zhang, S.; He, Z.; Wu, Y.; Zhao, L.; Liu, J.; Guo, J. KOBAS-i: Intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Res.* **2021**, *49*, W317–W325. [\[CrossRef\]](#)
37. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Methodol.)* **1995**, *57*, 289–300. [\[CrossRef\]](#)
38. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, *29*, 15–21. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Jin, Y.; Tam, O.H.; Paniagua, E.; Hammell, M. TETranscripts: A package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* **2015**, *31*, 3593–3599. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Quadrana, L.; Bortolini Silveira, A.; Mayhew, G.F.; LeBlanc, C.; Martienssen, R.A.; Jeddelloh, J.A.; Colot, V. The Arabidopsis thaliana mobilome and its impact at the species level. *eLife* **2016**, *5*, e15716. [\[CrossRef\]](#)
42. Stuart, T.; Eichten, S.R.; Cahn, J.; Karpievitch, Y.V.; Borevitz, J.O.; Lister, R. Population scale mapping of transposable element diversity reveals links to gene regulation and epigenomic variation. *eLife* **2016**, *5*, e20777. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Baduel, P.; Quadrana, L.; Colot, V. Efficient detection of transposable element insertion polymorphisms between genomes using short-read sequencing data. In *Plant Transposable Elements: Methods and Protocols*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 157–169.
44. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernysky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Van der Auwera, G.A.; Carneiro, M.O.; Hartl, C.; Poplin, R.; Del Angel, G.; Levy-Moonshine, A.; Jordan, T.; Shakir, K.; Roazen, D.; Thibault, J. From FastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinform.* **2013**, *43*, 11.10.1–11.10.33. [\[CrossRef\]](#) [\[PubMed\]](#)
46. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria, 2021; Available online: <https://www.R-project.org/>.
47. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.; Bender, D.; Maller, J.; Sklar, P.; De Bakker, P.I.; Daly, M.J. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [\[CrossRef\]](#)
48. Li, Q.; Gent, J.I.; Zynda, G.; Song, J.; Makarevitch, I.; Hirsch, C.D.; Hirsch, C.N.; Dawe, R.K.; Madzima, T.F.; McGinnis, K.M.; et al. RNA-directed DNA methylation enforces boundaries between heterochromatin and euchromatin in the maize genome. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 14728–14733. [\[CrossRef\]](#)
49. Saze, H.; Kitayama, J.; Takashima, K.; Miura, S.; Harukawa, Y.; Ito, T.; Kakutani, T. Mechanism for full-length RNA processing of Arabidopsis genes containing intragenic heterochromatin. *Nat. Commun.* **2013**, *4*, 2301. [\[CrossRef\]](#)
50. Ma, G.; Zhang, L.; Yungyuen, W.; Tsukamoto, I.; Iijima, N.; Oikawa, M.; Yamawaki, K.; Yahata, M.; Kato, M. Expression and functional analysis of citrus carotene hydroxylases: Unravelling the xanthophyll biosynthesis in citrus fruits. *BMC Plant Biol.* **2016**, *16*, 148. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Kumar, M.; Yusuf, M.A.; Yadav, P.; Narayan, S.; Kumar, M. Overexpression of Chickpea Defensin Gene Confers Tolerance to Water-Deficit Stress in Arabidopsis thaliana. *Front. Plant Sci.* **2019**, *10*, 290. [\[CrossRef\]](#)
52. Blankenship, R.E. Structural and functional dynamics of photosynthetic antenna complexes. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 13751–13752. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Vogt, T. Phenylpropanoid Biosynthesis. *Mol. Plant* **2010**, *3*, 2–20. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Lewandowska, M.; Keyl, A.; Feussner, I. Wax biosynthesis in response to danger: Its regulation upon abiotic and biotic stress. *New Phytol.* **2020**, *227*, 698–713. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Shepherd, T.; Wynne Griffiths, D. The effects of stress on plant cuticular waxes. *New Phytol.* **2006**, *171*, 469–499. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Helmprobst, F.; Kneitz, S.; Klotz, B.; Naville, M.; Dechaud, C.; Volff, J.-N.; Scharl, M. Differential expression of transposable elements in the medaka melanoma model. *PLoS ONE* **2021**, *16*, e0251713. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Lanciano, S.; Cristofari, G. Measuring and interpreting transposable element expression. *Nat. Rev. Genet.* **2020**, *21*, 721–736. [\[CrossRef\]](#)
58. Vogel, J.P.; Raab, T.K.; Somerville, C.R.; Somerville, S.C. Mutations in PMR₅ result in powdery mildew resistance and altered cell wall composition. *Plant J.* **2004**, *40*, 968–978. [\[CrossRef\]](#) [\[PubMed\]](#)

59. Gutterson, N.; Reuber, T.L. Regulation of disease resistance pathways by AP₂/ERF transcription factors. *Curr. Opin. Plant Biol.* **2004**, *7*, 465–471. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Mizoi, J.; Shinozaki, K.; Yamaguchi-Shinozaki, K. AP₂/ERF family transcription factors in plant abiotic stress responses. *Biochim. Et Biophys. Acta* **2012**, *1819*, 86–96. [\[CrossRef\]](#)
61. Hansen, C.C.; Nelson, D.R.; Møller, B.L.; Werck-Reichhart, D. Plant cytochrome P₄₅₀ plasticity and evolution. *Mol. Plant* **2021**, *14*, 1244–1265. [\[CrossRef\]](#)
62. Schnable, P.S.; Ware, D.; Fulton, R.S.; Stein, J.C.; Wei, F.; Pasternak, S.; Liang, C.; Zhang, J.; Fulton, L.; Graves, T.A. The B₇₃ maize genome: Complexity, diversity, and dynamics. *Science* **2009**, *326*, 1112–1115. [\[CrossRef\]](#) [\[PubMed\]](#)
63. Sun, H.-Y.; Dai, H.-Y.; Zhao, G.-L.; Ma, Y.; Ou, C.-Q.; Li, H.; Li, L.-G.; Zhang, Z.-H. Genome-wide Characterization of Long Terminal Repeat-retrotransposons in Apple Reveals the Differences in Heterogeneity and Copy Number between Ty₁-copia and Ty₃-gypsy Retrotransposons. *J. Integr. Plant Biol.* **2008**, *50*, 1130–1139. [\[CrossRef\]](#)
64. Xu, Y.; Du, J. Young but not relatively old retrotransposons are preferentially located in gene-rich euchromatic regions in tomato (*Solanum lycopersicum*) plants. *Plant J.* **2014**, *80*, 582–591. [\[CrossRef\]](#) [\[PubMed\]](#)
65. Zavallo, D.; Crescente, J.M.; Gantuz, M.; Leone, M.; Vanzetti, L.S.; Masuelli, R.W.; Asurmendi, S. Genomic re-assessment of the transposable element landscape of the potato genome. *Plant Cell Rep.* **2020**, *39*, 1161–1174. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Feschotte, C.; Pritham, E.J. DNA Transposons and the Evolution of Eukaryotic Genomes. *Annu. Rev. Genet.* **2007**, *41*, 331–368. [\[CrossRef\]](#) [\[PubMed\]](#)
67. Quesneville, H. Twenty years of transposable element analysis in the Arabidopsis thaliana genome. *Mob. DNA* **2020**, *11*, 28. [\[CrossRef\]](#)
68. Robertson, D.S. Characterization of a mutator system in maize. *Mutat. Res. Fundam. Mol. Mech. Mutagen.* **1978**, *51*, 21–28. [\[CrossRef\]](#)
69. Raviram, R.; Rocha, P.P.; Luo, V.M.; Swanzey, E.; Miraldi, E.R.; Chuong, E.B.; Feschotte, C.; Bonneau, R.; Skok, J.A. Analysis of 3D genomic interactions identifies candidate host genes that transposable elements potentially regulate. *Genome Biol.* **2018**, *19*, 216. [\[CrossRef\]](#) [\[PubMed\]](#)
70. Arribas, Y.A.; Baudon, B.; Rotival, M.; Suárez, G.; Bonté, P.-E.; Casas, V.; Roubert, A.; Klein, P.; Bonnin, E.; Mchich, B. Transposable element exonization generates a reservoir of evolving and functional protein isoforms. *Cell* **2024**, *187*, 7603–7620.e22. [\[CrossRef\]](#) [\[PubMed\]](#)
71. Rebollo, R.; Romanish, M.T.; Mager, D.L. Transposable elements: An abundant and natural source of regulatory sequences for host genes. *Annu. Rev. Genet.* **2012**, *46*, 21–42. [\[CrossRef\]](#)
72. Muhammad, I.; Shalmani, A.; Ali, M.; Yang, Q.-H.; Ahmad, H.; Li, F.B. Mechanisms Regulating the Dynamics of Photosynthesis Under Abiotic Stresses. *Front. Plant Sci.* **2021**, *11*, 615942. [\[CrossRef\]](#) [\[PubMed\]](#)
73. Hiruma, K.; Fukunaga, S.; Bednarek, P.; Piślewska-Bednarek, M.; Watanabe, S.; Narusaka, Y.; Shirasu, K.; Takano, Y. Glutathione and tryptophan metabolism are required for Arabidopsis immunity during the hypersensitive response to hemibiotrophs. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 9589–9594. [\[CrossRef\]](#) [\[PubMed\]](#)
74. Ishihara, A.; Hashimoto, Y.; Tanaka, C.; Dubouzet, J.G.; Nakao, T.; Matsuda, F.; Nishioka, T.; Miyagawa, H.; Wakasa, K. The tryptophan pathway is involved in the defense responses of rice against pathogenic infection via serotonin production. *Plant J.* **2008**, *54*, 481–495. [\[CrossRef\]](#)
75. Bagchi, R.; Gallery, R.E.; Gripenberg, S.; Gurr, S.J.; Narayan, L.; Addis, C.E.; Freckleton, R.P.; Lewis, O.T. Pathogens and insect herbivores drive rainforest plant diversity and composition. *Nature* **2014**, *506*, 85–88. [\[CrossRef\]](#)
76. MacArthur, R.H. *Geographical Ecology: Patterns in the Distribution of Species*; Princeton University Press: Princeton, NJ, USA, 1972.
77. Anderson, S.N.; Stitzer, M.C.; Brohammer, A.B.; Zhou, P.; Noshay, J.M.; O'Connor, C.H.; Hirsch, C.D.; Ross-Ibarra, J.; Hirsch, C.N.; Springer, N.M. Transposable elements contribute to dynamic genome content in maize. *Plant J.* **2019**, *100*, 1052–1065. [\[CrossRef\]](#) [\[PubMed\]](#)
78. Lynch, M.; Walsh, B. *The Origins of Genome Architecture*; Sinauer Associates: Sunderland, MA, USA, 2007; Volume 98.
79. Sigman, M.J.; Slotkin, R.K. The first rule of plant transposable element silencing: Location, location, location. *Plant Cell* **2016**, *28*, 304–313. [\[CrossRef\]](#) [\[PubMed\]](#)
80. Dubin, M.J.; Mittelsten Scheid, O.; Becker, C. Transposons: A blessing curse. *Curr. Opin. Plant Biol.* **2018**, *42*, 23–29. [\[CrossRef\]](#) [\[PubMed\]](#)
81. Butelli, E.; Licciardello, C.; Zhang, Y.; Liu, J.; Mackay, S.; Bailey, P.; Reforgiato-Recupero, G.; Martin, C. Retrotransposons Control Fruit-Specific, Cold-Dependent Accumulation of Anthocyanins in Blood Oranges. *Plant Cell* **2012**, *24*, 1242–1255. [\[CrossRef\]](#) [\[PubMed\]](#)
82. Clark, R.M.; Wagler, T.N.; Quijada, P.; Doebley, J. A distant upstream enhancer at the maize domestication gene tb₁ has pleiotropic effects on plant and inflorescent architecture. *Nat. Genet.* **2006**, *38*, 594–597. [\[CrossRef\]](#)
83. Nitasaka, E. Insertion of an En/Spm-related transposable element into a floral homeotic gene DUPLICATED causes a double flower phenotype in the Japanese morning glory. *Plant J.* **2003**, *36*, 522–531. [\[CrossRef\]](#) [\[PubMed\]](#)

84. Wang, H.; Lu, Y.; Zhang, T.; Liu, Z.; Cao, L.; Chang, Q.; Liu, Y.; Lu, X.; Yu, S.; Li, H.; et al. The double flower variant of yellowhorn is due to a LINE₁ transposon-mediated insertion. *Plant Physiol.* **2023**, *191*, 1122–1137. [[CrossRef](#)] [[PubMed](#)]
85. Guan, R.; Qu, Y.; Guo, Y.; Yu, L.; Liu, Y.; Jiang, J.; Chen, J.; Ren, Y.; Liu, G.; Tian, L. Salinity tolerance in soybean is modulated by natural variation in G m SALT₃. *Plant J.* **2014**, *80*, 937–950. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.