Check for updates

Low-coverage whole genome sequencing of diverse *Dioscorea bulbifera* accessions for plastome resource development, polymorphic nuclear SSR identification, and phylogenetic analyses

Ruisen Lu[1,2], Ke Hu[1,2], Xiaoqin Sun[1,2,3] and Min Chen[1,2]*

[1]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China, [2]Jiangsu Key Laboratory for the Research and Utilization of Plant Resources, Nanjing, China, [3]Jiangsu Provincial Science and Technology Resources Coordination Platform (Agricultural Germplasm Resources) Germplasm Resources Nursery of Medicinal Plants, Nanjing, China

Dioscorea bulbifera (Dioscoreaceae), a versatile herbaceous climber native to Africa and Asia, holds significant nutritional and medicinal value. Despite extensive characterization and genetic variability analyses of African accessions, studies on the genetic variation of this species in China are limited. To address this gap, we conducted low-coverage whole genome sequencing on *D. bulbifera* accessions from diverse regions across mainland China and Taiwan island. Our initial investigation encompassed comprehensive comparative plastome analyses of these *D. bulbifera* accessions, and developing plastome resources (including plastome-derived repetitive sequences, SSRs, and divergent hotspots). We also explored polymorphic nuclear SSRs and elucidated the intraspecific phylogeny of these accessions. Comparative plastome analyses revealed that *D. bulbifera* plastomes exhibited a conserved quadripartite structure with minimal size variation mainly attributed to intergenic spacer regions, reinforcing prior observations of a high degree of conservation within a species. We identified 46 to 52 dispersed repeats and 151 to 163 plastome-derived SSRs, as well as highlighted eight key divergent hotspots in these *D. bulbifera* accessions. Furthermore, we developed 2731 high-quality candidate polymorphic nuclear SSRs for *D. bulbifera*. Intraspecific phylogenetic analysis revealed three distinct clades, where accessions from Southeast China formed a sister group to those from South China and Taiwan island, and collectively, these two clades formed a sister group to the remaining accessions, indicating potential regional genetic divergence. These findings not only contributed to the understanding of the genetic variation of *D. bulbifera*, but also offered valuable resources for future research, breeding efforts, and utilization of this economically important plant species.

1 Introduction

Dioscorea bulbifera L., commonly referred to as the air potato, air yam, bitter yam, cheeky yam, potato yam, is a dioecious herbaceous climber belonging to the yam family, Dioscoreaceae (Coursey, 1967; Kundu et al., 2021). This species is native to Africa and Asia, but has widely naturalized and is cultivated across various regions, including Central and South America, Nepal, China, the Americas, the West Indies, Pacific Islands, Southeast Asia, and even parts of Australia (Coursey, 1967; Guan et al., 2017; Kundu et al., 2021; Kuncari, 2022). This species is characterized by a twining stem with a sleek surface and alternately arranged vibrant green leaves in a broadly cordate shape (Burkill, 1960; Ding and Gilbert, 2000). The emergence of purplish-brown bulbils (aerial tubers) from leaf axils is particularly noteworthy, as these serve as the primary organ for asexual propagation of the species (Terauchi et al., 1991; Kundu et al., 2021). Additionally, the plant generates underground tubers that bear a resemblance to petite potatoes (Terauchi et al., 1991; Kundu et al., 2021).

Dioscorea bulbifera is a rich source of primary metabolites, encompassing carbohydrates, starch, sugars, proteins, lipids, vitamins, minerals, and fibers (Abara et al., 2011). Its tubers offer versatile culinary possibilities, being adaptable to roasting and cooking as a vegetable, providing sustenance for tribal communities during food crises (Dutta, 2015; Ojinnaka et al., 2017). Significantly, the presence of essential amino acids, such as threonine and phenylalanine, coupled with significant mineral content, notably iron, enhances its nutritional importance (Ezeocha et al., 2014; Otegbayo et al., 2018). Beyond its nutritional richness, *D. bulbifera* holds a profound place in traditional medicine serving as a purgative, anthelmintic, diuretic, rejuvenating tonic, and exhibiting aphrodisiac qualities (Kumar et al., 2017). In traditional Chinese medicine, *D. bulbifera* is employed to address conditions such as cough, pharyngitis, skin infections, piles, hemoptysis, and goiter (Guan et al., 2017). Recent studies have highlighted the potency of *D. bulbifera* against cancer, demonstrating its efficacy in inhibiting tumor growth in various cells, including colon and liver cancer (Guan et al., 2017).

The vast nutritional and medicinal benefits of *Dioscorea bulbifera* have triggered substantial characterization and genetic variability analyses in diverse regions like Brazil (Silva et al., 2016), Ethiopia (Beyene, 2013; Mulualem and Weldemichel, 2013), Nigeria (Jayeola and Oyebola, 2013), Uganda (Ikiriza et al., 2023), and West Africa (Osuagwu and Edem, 2020). These investigations have significantly propelled the advancement of breeding techniques aimed at enhancing its desirable traits for both food and medicinal purposes in these areas (Osuagwu and Edem, 2020). Nevertheless, despite the abundance of resources in China (Guan et al., 2017), the characterization and genetic variation analyses of this species lag far behind. It remains underutilized, marginalized, and less cultivated in China. Consequently, there exists an urgent necessity for comprehensive characterization, particularly at the molecular level, and genetic variation analysis of this plant in China.

Nowadays, the application of low-coverage whole genome sequencing has emerged as a cost-effective and efficient strategy for selectively capturing high-copy elements such as the plastome, ribosomal DNA, and SSRs across diverse plant species (Straub et al., 2012;

Dodsworth, 2015; Lu et al., 2022). Among these molecular markers, whole plastome sequences, have demonstrated immense value in plant phylogenetic studies owing to their distinctive traits, including the absence of recombination, small effective population sizes, low nucleotide substitution rates, and typically uniparental inheritance (Birky et al, 1983). Additionally, conducting comparative plastomes analyses can facilitate the identification of regions with sequence variation, thereby aiding in accurate and rapid species discrimination—a critical element for the optimal utilization and conservation of plant species (Lu et al., 2021). More significantly, utilizing assembled nuclear sequences derived from low-coverage whole genome sequencing data has successfully enabled the extensive exploration of polymorphic nuclear SSRs (nSSRs) in non-model plant species (Liu et al., 2018, Liu et al., 2021), which play pivotal roles in population genetic analyses and marker-assisted selection (Kumar et al., 2015).

In this study, we conducted low-coverage whole genome sequencing of *Dioscoea bulbifera* accessions spanning diverse regions across mainland China and Taiwan island. Our objectives were to: (1) explore and compare *D. bulbifera* plastomes to unravel their evolutionary patterns; (2) pinpoint plastome-derived markers including repetitive sequences, plastomic SSRs, and divergent hotspots; (3) develop polymorphic nuclear SSRs using multiple assembled nuclear sequences; and (4) reconstruct the phylogenetic relationships among *D. bulbifera* accessions based on plastome data. These discoveries will not only broaden our understanding of the genetic variations within *D. bulbifera* accessions but also furnish essential genetic resources pivotal for advancing molecular characterization and commercial breeding schemes of this species.

2 Materials and methods

2.1 Plant materials, DNA extraction and genome sequencing

We gathered 10 *Dioscoea bulbifera* accessions from various regions across mainland China, spanning Anhui (AHLA), Fujian (FJZZ), Guangdong (GDSG), Guangxi (GXQZ), Henan (HeNXY), Hunan (HuNXX), Jiangsu (JSYX), Sichuan (SCEM), Zhejiang (ZJLS) provinces, along with Taiwan island (TWXB) (Table 1), for comprehensive analysis. For each accession, the pristine, fresh green leaves were harvested from a wild mature individual, and then preserved by desiccation with silica gel. Voucher specimens were deposited at the Herbarium of Institute of Botany, Jiangsu Province and the Chinese Academy of Sciences (NAS). Genomic DNA was then extracted from ~50 mg of silica-dried leaves using the DNAsecure Plant Kit (Tiangen Biotech, Beijing, China), with the purity and integrity of the isolated genomic DNA evaluated through agarose gel electrophoresis and spectrophotometric analysis.

Approximately 1 μg genomic DNA was broken into small fragments using the Covaris E210 sonicator (Covaris Inc., MA, USA). Fragments were then size-selected selected by Agencourt AMPure XP-Medium kit (Thermo Fisher Scientific, USA) to attain sizes ranging from 200 to 400 bp. Following end repair, 3'adenylation, adaptor ligation, PCR amplification, and purification, the resulting double-stranded PCR products were

TABLE 1 Summary of plastome characteristics of 10 *Dioscorea bulbifera* accessions.

Characteristics	AHLA	FJZZ	GDSG	GXQZ	HeNXY	HuNXX	JSYX	SCEM	TWXB	ZJLS
Locality	Lu'an, Anhui	Zhangzhou, Fujian	Shaoguan, Guangdong	Qinzhou, Guangxi	Xinyang, Henan	Xiangxi, Hunan	Yixing, Jiangsu	Emeishan, Sichuan	Xinbei, Taiwan	Lishui, Zhejiang
Clean Reads	45,979,402	51,380,484	37,387,226	40,006,520	43,568,546	39,480,040	49,054,050	45,186,582	53,141,458	41,480,186
Latitude (N)/ Longitude (E)	31°25′57″/ 116°8′31″	24°33′06″/ 117°20′07″	24°55′33″/ 113°01′22″	21°59′06″/ 108° 42′16″	30°59′45″/ 116° 04′49″	29°07′40″/ 110° 27′48″	31°10′05″/ 119° 40′55″	29°35′48″/ 103°22′13″	24°52′03″/ 121° 24′51″	27°54′50″/ 118° 55′23″
Total plastome length (bp)	153,074	153,002	153,099	153,002	152,970	153,093	153,074	153,093	153,002	153,074
LSC	83,225	83,152	83,249	83,152	83,120	83,240	83,225	83,240	83,152	83,225
SSC	18,851	18,852	18,852	18,852	18,852	18,855	18,851	18,855	18,852	18,851
IR	25,499	25,499	25,499	25,499	25,499	25,499	25,499	25,499	25,499	25,499
Total GC content (%)	37.0	37.0	37.0	37.0	37.0	37.0	37.0	37.0	37.0	37.0
LSC	34.8	34.8	34.8	34.8	34.8	34.8	34.8	34.8	34.8	34.8
SSC	30.8	30.8	30.8	30.8	30.8	30.8	30.8	30.8	30.8	30.8
IR	43.0	43.0	43.0	43.0	43.0	43.0	43.0	43.0	43.0	43.0
Total number of genes	113	113	113	113	113	113	113	113	113	113
PCGs	79	79	79	79	79	79	79	79	79	79
tRNA genes	30	30	30	30	30	30	30	30	30	30
rRNA genes	4	4	4	4	4	4	4	4	4	4
Duplicated genes	19	19	19	19	19	19	19	19	19	19

transformed into single-stranded circular DNA (ssCir DNA), through heat denaturation and circularization using a splint oligo sequence. The ssCir DNA was formatted as the final library, and sequenced on DNBSEQ-T7 sequencing platform to generate 150-bp paired-end reads.

2.2 Plastome assembly and annotation

After preprocessing raw data with Trimmomatic v.0.36 (Bolger et al., 2014) to remove adaptor sequences, contamination, and low-quality reads, each accession yielded a total of 39,480,040–53,141,458 clean reads utilized for subsequent plastome assembly. *De novo* assembly of the plastome was executed using GetOrganelle v.1.7.6 (Jin et al., 2020), with recommended parameters: -R 15 -k 21,45,65,85,105 -F embplant_pt. Plastome sequences of all 10 accessions were assembled in complete circular sequences. Initial annotation was conducted with the MAFFT v.7 plugin (Katoh and Standley, 2013) integrated within Geneious Prime® 2022.0.1 (https://www.geneious.com), by aligning them with the previously published ones of *Dioscoea bulbifera* (MG805604) and *D. nipponica* (OQ525997) as references. Reference annotations were then transferred to these newly assembled plastomes, followed by meticulous manual verification to ensure precision of exon/intron boundaries and accurate start/stop codon placement. To present comprehensive insights of *D. bulbifera* plastomes, high-resolution

circular plastome maps were generated using the web-based tool OrganellarGenomeDRAW (OGDRAW) v.1.3.1 (Greiner et al., 2019).

2.3 Whole plastome sequence comparison

The mVISTA program (http://genome.lbl.gov/vista/mvista/submit.shtml) was employed to assess the structural resemblance of complete plastome sequences among *Dioscoea bulbifera* accessions, with the annotation from the AHLA plastome sequence serving as the reference. Plastome sequences were aligned using the Shuffle-LAGAN mode with default parameters (Brudno et al., 2003), and the resulting alignments were displayed using the VISTA program (Frazer et al., 2004). Additionally, to detect potential expansions or contractions in the inverted repeat (IR) regions within *D. bulbifera* plastomes, a comparison and visualization of the four junctions between the inverted repeat (IR) and the large single copy (LSC)/small single copy (SSC) regions were conducted using IRscope (https://irscope.shinyapps.io/irapp/, Amiryousefi et al, 2008).

2.4 Plastome-derived markers development

Repetitive sequences, comprising forward (direct), reverse, complement, and palindromic repeats within *Dioscoea bulbifera*

plastomes, were identified through the online tool REPuter (Kurtz et al., 2001). The parameters for repetitive sequences identification involved a minimum repeat size of 30 bp, a sequence identity of at least 90%, and a hamming distance of 3. Simple sequence repeats (SSRs) within the 10 *D. bulbifera* plastome sequences were identified using the MISA-web application (Beier et al., 2017). The SSR search criteria specified a minimum of 10 repeat units for mononucleotide SSRs, 5 for dinucleotide SSRs, 4 for trinucleotide SSRs, and 3 for tetra-, penta-, and hexanucleotide SSRs, respectively.

For an in-depth exploration of divergent hotspots within *Dioscoea bulbifera* plastomes, the 10 newly assembled plastome sequences were aligned using MAFFT v.7 (Katoh and Standley, 2013) in Geneious Prime® 2022.0.1. Regions within this alignment, including protein coding areas, intergenic spacers, and introns, displaying a total mutation count above 0 and an aligned length exceeding 200 bp, were systematically extracted from the alignment matrix. Subsequently, the nucleotide diversity (π) of these extracted regions was computed using DnaSP v.6.12.03 (Rozas et al., 2017).

2.5 Polymorphic nuclear SSRs identification

Low-coverage whole genome sequencing data from three geographically distinct *Dioscoea bulbifera* accessions (AHLA, SCEM, and TWXB) were employed to develop polymorphic nuclear SSR markers. Clean reads of these accessions were aligned to the genome sequences of *D. alata* (Bredeson et al., 2022) and *D. zingiberensis* (Li et al., 2022) to exclude mitochondria and chloroplast reads using BWA-MEM v.0.7.17 (Li, 2013). The resulting alignment files were sorted and converted into Binary Alignment/Map (BAM) format using SAMtools v.1.9 (Li et al., 2009). Subsequently, BAM files were transformed into FastQ files with SAMtools bam2fq. The obtained reads were then *de novo* assembled into contigs using CLC Genomics Workbench v.23.0.4 (CLC bio, Aarhus, Denmark) with default settings. Utilizing these assembled nuclear sequences, CandiSSR (Xia et al., 2016) was employed to identify polymorphic nuclear SSRs (nSSRs) within *D. bulbifera*, using the default parameters except for specifying a flanking sequence length of 200 bp.

2.6 Phylogenetic analyses within *Dioscoea bulbifera*

Phylogenetic analyses were performed using two datasets: complete plastome sequences and 79 protein coding genes shared across all 10 *Dioscoea bulbifera* accessions (Table 1), employing *D. nipponica* (OQ525997), *D. elephantipes* (EF380353) and *D. alata* (OP787123) as outgroups, based on Noda et al. (2020). For the complete plastome dataset, both maximum likelihood (ML) and Bayesian inference (BI) analyses were conducted employing two partitioning scenarios: (1) unpartitioned, and (2) partitioned by each gene and intergenic region (265 partitions). In contrast, the protein coding gene dataset underwent four partitioning scenarios: (1) unpartitioned, (2) partitioned by codon position (three partitions), (3) partitioned by each gene (79 partitions), and (4) partitioned by PartitionFinder v.2.1.1 (Lanfear et al., 2017) (19

partitions). Alignments of both complete plastome sequences and protein coding sequences were executed using the MAFFT v.7 plugin (Katoh and Standley, 2013) within Geneious Prime® 2022.0.1. Except for the partition scenario determined through PartitionFinder, the optimal substitution model was obtained using PartitionFinder, while the remaining partition schemes selected GTR + I + G as the optimal substitution based on the Akaike Information Criterion (AIC) computed by jModelTest v.2.1.4 (Darriba et al., 2012). Maximum Likelihood (ML) analyses were carried out using RAxML v.8.2.12 (Stamatakis, 2014) via the CIPRES Science Gateway v.3.3 (http://www.phylo.org/portal2/), utilizing 1000 bootstrap replications. Bayesian Inference (BI) analyses were conducted using MrBayes v.3.2.7 (Ronquist et al., 2012), utilizing Markov Chain Monte Carlo (MCMC) runs for 1×10^6 generations with a sampling frequency of 1000 trees. The initial 1000 trees were discarded as 'burn-in', and the remaining trees were employed to generate a majority-rule consensus tree and estimate posterior probabilities (PPs).

3 Results

3.1 Plastome characteristics of *Dioscoea bulbifera*

The plastome sizes exhibited narrow variation among these 10 *Dioscoea bulbifera* accessions: GDSG displayed a length of 153,099 bp, HuNXX and SCEM had a size of 153,093 bp, AHLA, JSYX, and ZJLS shared a size of 153,074 bp, FJZZ, GXQZ, and TWXB showcased a size of 153,002 bp, while HeNXY exhibited a size of 152,970 bp (Figure 1; Table 1). Each of these *D. bulbifera* plastomes maintained the typical circular quadripartite structure, comprising a pair of inverted repeat (IR) regions (25,499 bp) separated by a large single copy (LSC) region ranging between 83,120–83,249 bp, and a small single copy (SSC) region varying from 18,851–18,855 bp. The GC content across the entire plastome sequence (37.0%), as well as in the LSC (34.8%), SSC (30.8%), and IR (43.0%) regions, remained consistent among all 10 *D. bulbifera* accessions.

All 10 *Dioscoea bulbifera* plastomes shared an exact set of 113 unique genes, encompassing 79 protein-coding genes (PCGs), 30 transfer RNA (tRNA) genes, and four ribosomal RNA (rRNA) genes (Figure 1; Table 1). Out of these, 19 genes (seven PCGs, eight tRNA genes, and all four rRNA) were duplicated within the inverted repeats (IRs), resulting in a cumulative count of 132 genes (Table 1; Supplementary Table S1). Among these unique genes, eight PCGs and six tRNAs contained a single intron, while three PCGs harbored two introns each (see details in Supplementary Table S1). An intact gene encoding initiation factor IF1 (*infA*) was identified, while the *rps16* gene was independently lost in *D. bulbifera* plastomes (Figure 1).

3.2 Plastome comparison within *Dioscoea bulbifera*

Using the accession AHLA as the reference, mVISTA results indicated high sequence similarity among these *Dioscoea bulbifera*
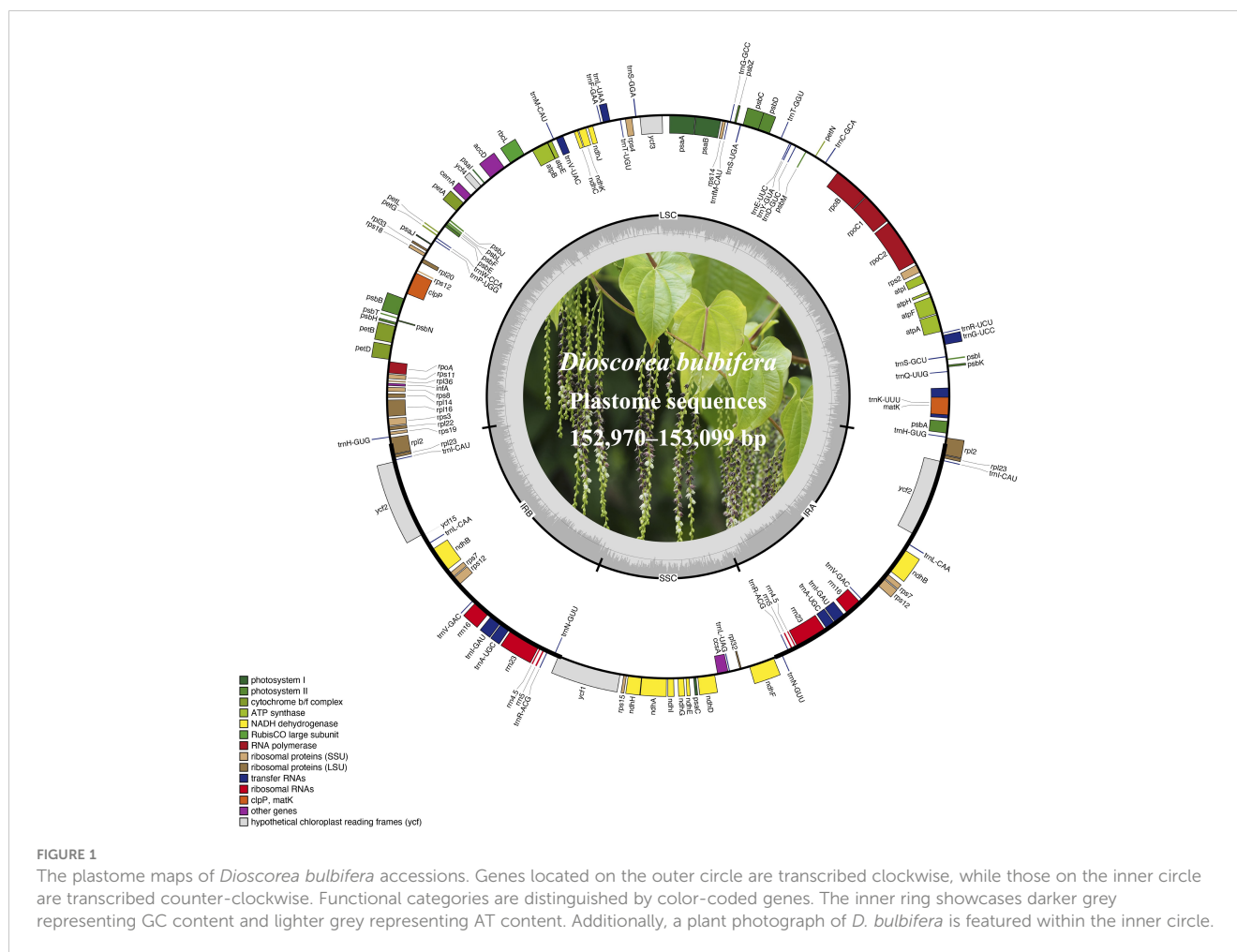
FIGURE 1
The plastome maps of *Dioscorea bulbifera* accessions. Genes located on the outer circle are transcribed clockwise, while those on the inner circle are transcribed counter-clockwise. Functional categories are distinguished by color-coded genes. The inner ring showcases darker grey representing GC content and lighter grey representing AT content. Additionally, a plant photograph of *D. bulbifera* is featured within the inner circle.

plastomes, especially in the IR regions (Figure 2). Moreover, the coding regions demonstrated notably higher similarity levels compared to non-coding regions, including introns and intergenic spacers. Notably, the intergenic spacers, specifically *trnK–trnQ* and *psbM–trnD*, displayed the lowest sequence similarity (Figure 2). Within the *trnK–trnQ* region, AHLA, JSYX, HeNXY, HuNXX, SCEM, and ZJLS showed a 24 bp gap compared to FJZZ, GDSC, GXQZ, and TWXB. Similarly, in the *psbM–trnD* region, FJZZ, GXQZ, and TWXB presented a more substantial 96 bp gap compared to the other accessions.

Comparative analysis of IR/SC junctions underscored the stability of *Dioscoea bulbifera* plastomes, revealing no expansion or contraction in the IR regions among these accessions (Figure 3). Across these 10 plastomes, the LSC/IRa junction (JLA) was situated in the *psbA–trnH* intergenic spacer region, 87 bp away from the adjacent gene *psbA*. Concurrently, the LSC/IRb junction (JLB) was positioned within the intergenic spacer of *rps19–trnH*, maintaining an 8 bp distance from the *rps19* gene (Figure 3). Notably, the *ycf1* gene traversed the LSC/IRb junction (JLB), maintaining a consistent length of 365 bp within the IRb region and extending to 5199 bp in the SSC region. Meanwhile, the *ndhF* gene was completely located in the *SSC* region, merely 4 bp away from the SSC/IRa junction (JSA) (Figure 3).

3.3 Plastome-derived markers of *Dioscoea bulbifera*

The types and lengths of dispersed repeats, including forward (direct), reverse, complement, and palindromic repeats, as well as simple sequence repeats (SSRs) were detected and analyzed within these 10 *Dioscoea bulbifera* plastomes. A total of 250 dispersed repeats were detected across all 10 *D. bulbifera* plastomes, comprising three repeat types: forward (120), reverse (10) and palindromic (120) repeats (Figure 4A). Among all 10 plastomes, FJZZ, GDSG, GXQZ and TWXB displayed the highest count of repeats (total: 26, forward: 13, reverse: 1, and palindromic: 12), followed by AHLA, JSYX and ZJLS (total: 25, forward: 12, reverse: 1, and palindromic: 12), and HuNXX and SCEM (total: 24, forward: 11, reverse: 1, and palindromic: 12), while HeNXY exhibited the fewest (total: 23, forward: 10, reverse: 1, and palindromic: 12) (Figure 4A). Across each *D. bulbifera* plastome, a substantial majority of repeats (61.5% in HeNXY to 76.0% in AHLA, JSYX, and ZJLS) ranged in size between 30 and 40 bp (Figure 4B).

The count of plastome-derived SSRs within the *Dioscoea bulbifera* plastomes varied from 151 (AHLA, JSYX, and ZJLS) to 163 (HuNXX and SCEM) (Figure 5; Supplementary Table S2).
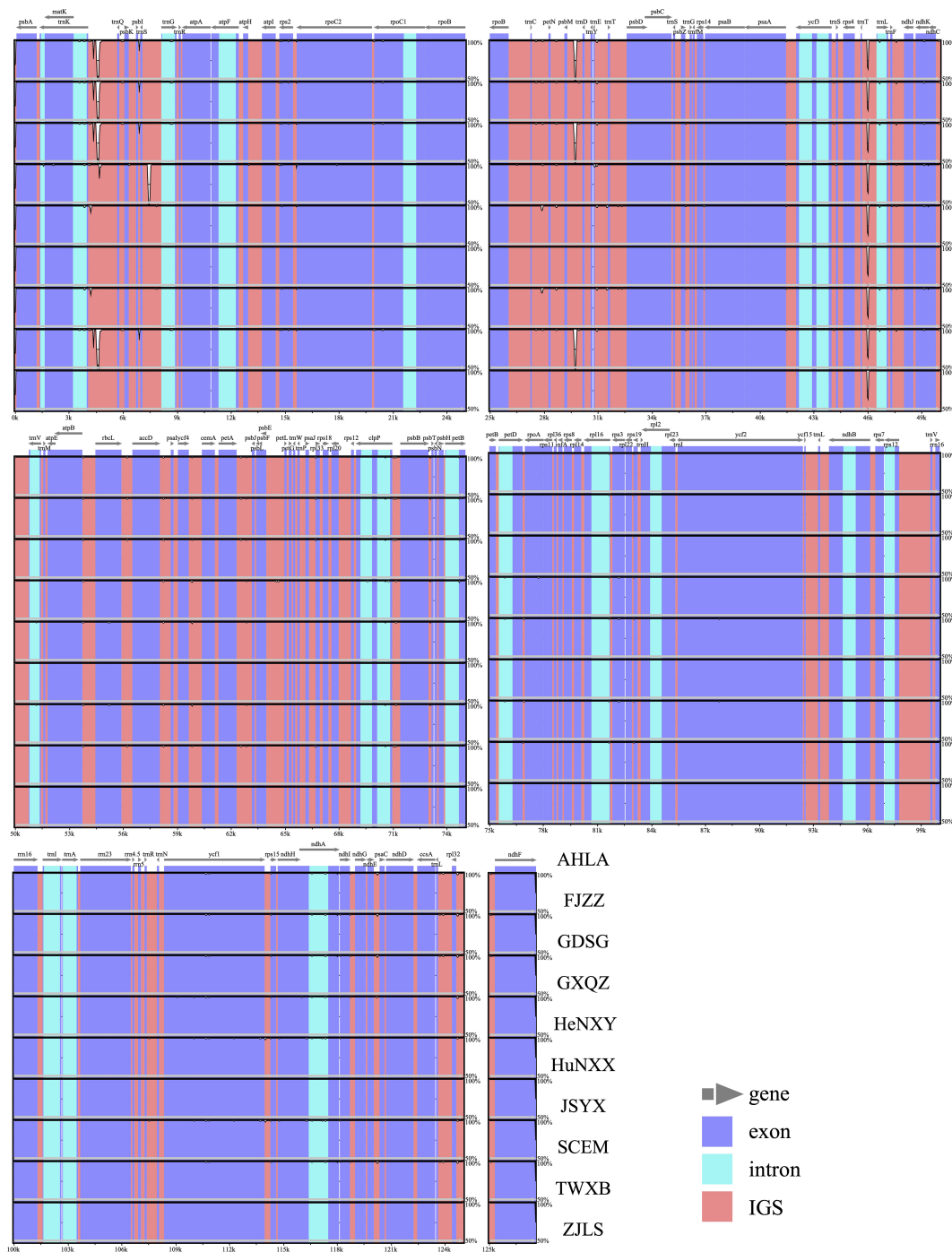
FIGURE 2
Sequence similarity plots among *Dioscorea bulbifera* plastomes, using the accession AHLA as a reference. Annotated genes are shown along the top. The vertical scale indicates percent identity, ranging from 50% to 100%. Genome regions are color-coded, distinguishing between exons, introns, and intergenic spacers (IGS).

Among these, dinucleotides emerged as the most prevalent SSR type, ranging from 62 in AHLA, JSYX, and ZJLS to 68 in HuNXX and SCEM. Following dinucleotides were mononucleotides, which varied from 46 in AHLA, JSYX, and ZJLS to 52 in HuNXX and SCEM. Trinucleotides and tetranucleotides exhibited a same pattern and were next in frequency, ranging from 17 in FJZZ, GDSG, GXQZ, HeNXY, and TWXB to 19 in HuNXX and SCEM.

Conversely, pentanucleotides were observed in only 4 instances (HeNXY, HuNXX, SCEM) or 5 in the remaining accessions, while hexanucleotides were least common, occurring once in FJZZ, GDSG, GXQZ, HeNXY, HuNXX, SCEM, and TWXB, or twice in the other accessions, (Figure 5; Supplementary Table S2). Among the motifs in the SSRs, A/T and AA/TT were the most frequently occurring motifs, followed by AT/AT, AAA/TTT and AAAA/
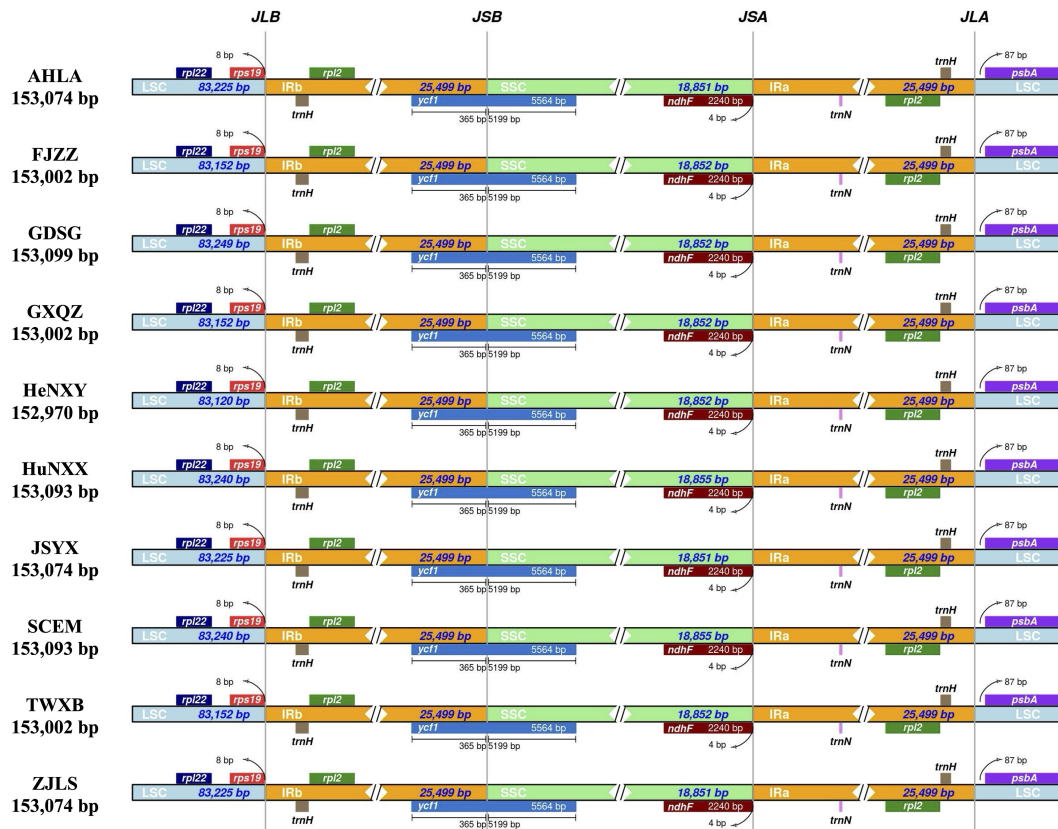
FIGURE 3
Comparison of inverted repeat/single copy (IR/SC) junction regions in *Dioscorea bulbifera* plastomes.

TTTT, while the remaining types appeared relatively infrequently (Figure 5; Supplementary Table S2). Moreover, a specific set of at least eight plastome-derived SSRs—$(A/T)_{16,17}$, $(C/G)_{10,11,12}$, $(AT/AT)_{6,7}$, $(AT/AT)_{8,9}$, $(CC/GG)_{5,6}$, $(CCC/GGG)_{4}$, $(ATAT/ATAT)_{3}$, $(CGCG/CGCG)_{3}$, $(ATATAT/ATATAT)_{3}$ could effectively distinguish these *D. bulbifera* accessions into two to three distinct groups (Supplementary Table S2).

Although a total of 133 regions (62 CDS, 52 IGS, 13 introns and 6 tRNAs) showed an aligned length exceeding 200 bp, only 38 regions (12 CDS, 20 IGS, 3 introns and 3 tRNAs) had a mutation count greater than zero. Consequently, these 38 regions were selected from the alignment of all 10 *Dioscoea bulbifera* plastomes to identify divergent hotspots (Figure 6). These 38 regions displayed π values ranging from 0.000053 (CDS *rpoC2*) to 0.0036 (IGS *ndhE–*
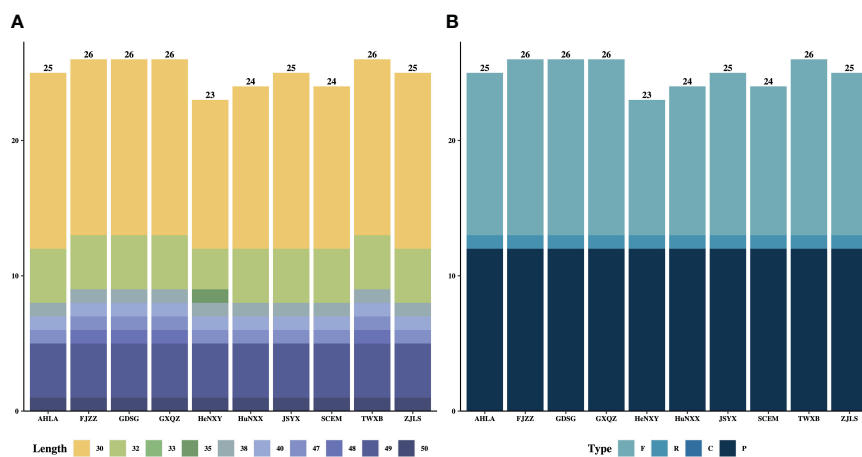

FIGURE 4
The **(A)** lengths and **(B)** types of dispersed repeats in the 10 *Dioscorea bulbifera* plastomes.

FIGURE 5
The lengths and motif types of plastome-derived SSRs in the 10 *Dioscorea bulbifera* accessions.

psaC) (Figure 6). Among protein-coding regions (CDS), π values for each region ranged from 0.000053 (*rpoC2*) to 0.0014 (*rps2*), with only two CDS regions (*rps2* and exon 1 of *ndhA*) exhibiting notably high values (π > 0.001). For the non-coding regions (IGS, introns and tRNAs), π values ranged from 0.00020 (IGS *psbE–petL*) to 0.0036 (IGS *ndhE–psaC*). The top six most variable non-coding regions (π > 0.001) were IGS *ndhE–psaC*, *trnL-UAA*, IGS *trnF–ndhJ*, *trnG-UCC*, IGS *psaJ–rpl33*, and IGS *trnC–petN* (Figure 6). These six non-protein-coding regions in conjunction with the two protein-coding regions showed significant potential as highly informative molecular markers for *D. bulbifera*.

3.4 Polymorphic nuclear SSRs for *Dioscoea bulbifera*

Utilizing low-coverage whole-genome sequencing data, a total of 159,233, 164,875, and 192,193 nuclear contigs were generated for AHLA, SCEM, and TWXB, respectively, each having N50 lengths of 1,629 bp, 1,595 bp, and 1,650 bp. Based on these nuclear contigs, a pool of 2731 candidate polymorphic nSSRs (PolynSSRs) was identified for *Dioscoea bulbifera* (Supplementary Table S3). The average similarity of the flanking sequences of these polymorphic nSSRs was 0.96, with 71.95% (1965/2731) exhibiting a similarity

FIGURE 6
Nucleotide variability (π) values of 38 regions (12 CDS, 20 IGS, three introns and three tRNAs) extracted from the alignment matrix of 10 *Dioscorea bulbifera* plastomes.

above 0.98, indicating their high transferability across *D. bulbifera* accessions. Subsequent filtration, eliminating low-quality PolynSSRs with transferability (similarity) < 95% and a missing rate (MR) ≥ 0.5, resulted in a collection of 2433 high-quality candidate PolynSSRs (Supplementary Table S3). Out of these, 2331 high-quality PolynSSRs could be designed for primers, encompassing 95.81% of the total (Supplementary Table S3). Within this set of high-quality candidate PolynSSRs, tetranucleotide repeats comprised the majority at 1041 (42.79%), followed by tri-, penta-, and hexanucleotide repeats, accounting for 32.71%, 12.33%, and 12.17% of the total, respectively (Supplementary Table S3).

3.5 Intraspecific phylogeny of *Dioscoea bulbifera*

Both the ML and BI analyses, employing complete plastome sequences and 79 shared protein coding genes under different partitioning strategies, yielded identical tree topologies. Consequently, only the phylogenetic trees based on complete plastome sequences are presented here (Figure 7). Phylogenetic analyses revealed three main distinct clades within these *Dioscoea bulbifera* accessions. Specifically, AHLA, ZJLS, and JSYX from Southeast China formed a distinctive monophyletic clade (Clade I), that is sister to the Clade II encompassing accessions from South China (FJZZ, GDSG, and GXQZ) and TWXB from Taiwan island. These two clades collectively form a sister group to the remaining accessions (HeNXY, HuNXX, and SCEM) (Clade III) (Figure 7).

4 Discussion

4.1 Plastome characteristics and evolution of *Dioscoea bulbifera*

The comprehensive exploration of 10 *Dioscoea bulbifera* plastomes from diverse geographic regions across mainland China and Taiwan island unveiled intriguing insights into the plastome structure, genetic composition, and variation of this species. Across all 10 accessions, the plastomes maintained a conserved quadripartite structure, with minimal size variation predominantly residing within the single copy regions (Figure 2; Table 1). The consistency in gene content (including unique and duplicated genes), gene order, and GC content of these *D. bulbifera* plastomes reinforced earlier findings that highlighted a high degree of conservation among plastomes within a species in terms of structure, gene composition, and gene order synteny (Muraguri et al., 2020; Lu et al., 2022). The absence of the *rps16* gene in *D. bulbifera* plastomes corroborated our prior observation that this gene may be absent across *Dioscorea* clades, excluding the *Stenophora* clade (= *D.* sect. *Stenophora*) (Jansen et al., 2007; Hu et al., 2023a). Conversely, the intact presence of the initiation factor IF1 gene (*infA*) in *D. bulbifera* plastomes aligned with its occurrence in other *Dioscorea* species (Lu et al., 2023; Hu et al., 2023a). Notably, among monocots, the depletion of *infA* genes is particularly concentrated (>70%) in Alismatales, Commelinales, Liliales, Pandanales, with minimal loss occurrences of

7.69% observed in Dioscoreales, including 12.50% within Dioscoreaceae (Lu et al., 2021).

Plastome size variations typically arise from two primary factors: i) the dynamic changes in the junctions between the inverted repeat (IR) and single copy (SC) regions (Kim and Lee, 2004; Lu et al., 2016), and ii) the variability of gene spacer regions, and the presence or absence of genes and introns (Jansen et al., 2007; Li et al., 2021). Our investigation of IR/SC junctions across *Dioscoea bulbifera* accessions unveiled a consistent structural configuration without observable expansions or contractions (Figure 3). Considering the conserved nature of gene content and structure, it appeared that differences in plastome sizes among these accessions mainly result from alterations in gene spacer regions. Significantly, the mVISTA analysis underscored that specific intergenic spacers, particularly *trnK–trnQ* and *psbM–trnD*, exhibited low sequence similarity, displaying gaps among the identified clades (Figure 2). This observation, coupled with our previous studies on plastomes of *D. alata* (Lu et al., 2023) and *D. nipponica* (Hu et al., 2023b), suggested that variations in intergenic spacers, particularly *trnK–trnQ*, could be contributing to the diversity in plastome sizes within individual *Dioscorea* species. However, further investigation is warranted to comprehensively understand these variations.

4.2 Molecular markers for *Dioscoea bulbifera*

The recognition of the importance of conserving medicinal plants, enhancing cultivars with desirable traits, and comprehending germplasm diversity has witnessed significant growth in recent years (Baruah et al., 2017; Marakli, 2018). This growing emphasis has led to the utilization of various molecular markers that offer elaborate genomic insights surpassing the capabilities of phenotypic methods (Marakli, 2018). Plastome-derived markers have emerged as valuable assets, enabling the identification of germplasm resources and contributing to their conservation and breeding efforts (Daniell et al., 2016). Despite the high conservation of plastome sequences within *Dioscoea bulbifera*, nucleotide substitutions, SSRs, and indels could serve as valuable markers to elucidate the genetic diversity and guide molecular breeding of this medically important plant (Li et al., 2020). In this study, we successfully identified a remarkable array of plastome-derived SSRs, ranging from 151 (AHLA, JSYX, and ZJLS) to 163 (HuNXX and SCEM) (Figure 5; Supplementary Table S2), and revealed at least eight potentially polymorphic SSRs, highlighting their utility in marker-assisted studies and population genetics.

Previous phylogenetic studies in *Disocorea* primarily relied on *matK*, *rbcL*, and *trnL–F* genes, which often lacked sufficient phylogenetic resolution within closely related species and within a single species (Hu et al., 2023a). Recent comparative plastome studies have emphasized the concentration of divergent hotspot regions in non-protein-coding areas across *Dioscorea* species. Notable examples included six IGS regions (i.e., *ndhD–ccsA*, *petA–psbJ*, *psbZ–trnG*, *rpl32–ndhF*, *trnD–trnY*, and *trnL–rpl32*) and the *rps16* intron sequence emerged as potential molecular markers for species within
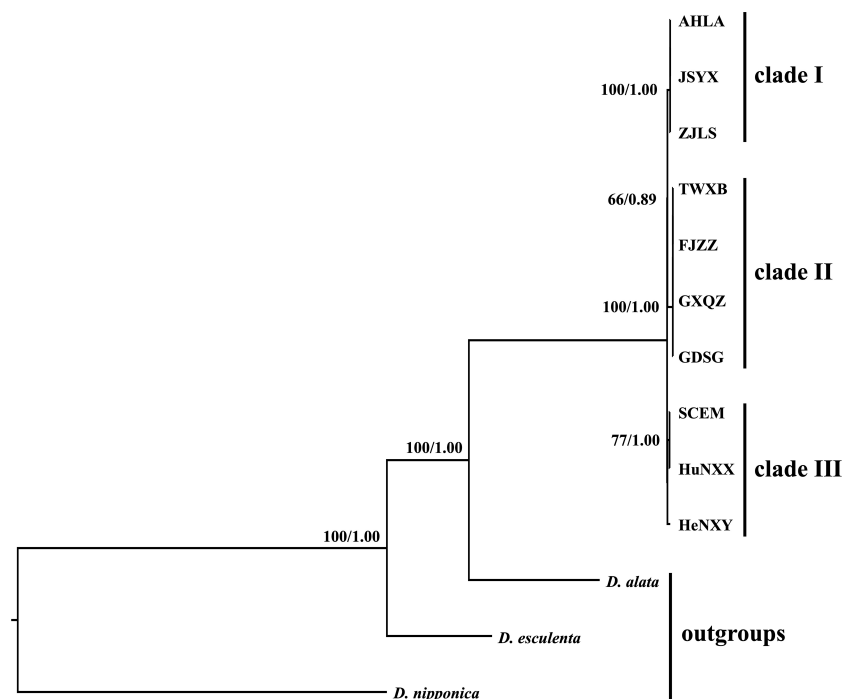
FIGURE 7
Intraspecific phylogenetic relationships among the 10 *Dioscorea bulbifera* accessions originating from different regions across mainland China and Taiwan island, inferred from the methods of maximum likelihood (ML) and Bayesian inference (BI). The ML bootstrap values/BI posterior probabilities are displayed above the lines.

D. sect. *Stenophora* (Hu et al., 2023a). Similarly, regions such as *ndhD–ccsA*, *trnC–petN*, and *trnL–rpl32* were identified as potential DNA barcodes for species within *D.* sect. *Enantiophyllum* (Lu et al., 2023). Notably, regions like three intergenic spacers (*rps16–trnQ*, *trnE–trnT*, and *trnL–rpl32*) and two intron regions (intron 1 of *clpP* and intron *trnG*) promised substantial insights into assessing intraspecific genetic variability of *D. nipponica* (Hu et al., 2023b). Given these findings, further exploration of molecular markers, particularly in non-coding regions, becomes imperative in *Dioscorea* species. The comparative analysis across 10 *D. bulbifera* plastomes unveiled four IGS regions (*ndhE–psaC*, *trnF–ndhJ*, *psaJ–rpl33*, and *trnC–petN*), two tRNA regions (*trnL-UAA*, *trnG-UCC*), and two CDS regions (*rps2* and exon 1 of *ndhA*) with notably high values of nucleotide diversity (Figure 6), holding substantial promise for population genetic and intraspecific phylogenetic studies of *D. bulbifera*.

Nuclear SSR markers (nSSRs) have demonstrated their value in diverse applications, including population genetic analyses, cultivar and germplasm identification, and marker-assisted selection, due to their high polymorphism and co-dominant inheritance in a Mendelian fashion (Kaldate et al., 2017). Recent advancements in sequencing technologies and bioinformatic analyses have opened an unprecedented window for identifying high-quality, polymorphic nuclear SSR markers in non-model organisms, offering effective results within optimized cost and time frames (Xia et al., 2016). In this study, a significant discovery of 2433 high-quality candidate PolynSSRs was made (Supplementary Table S3), providing potent tools for conducting population genetic studies of *Dioscorea bulbifera*. In summary, the identified intraspecific plastome-

derived and nuclear markers could offer complementary insights into the genetic structure, differentiation, and gene flow among *D. bulbifera* populations, being crucial for their conservation and efficient management. Furthermore, these markers can be used to develop genetic maps and conduct marker assisted breeding.

4.3 Phylogenetic relationships of *Dioscorea bulbifera*

Nowadays, the utilization of whole plastome sequences has become widespread in elucidating the phylogenetic relationships among plant species (Lu et al., 2016; Hu et al., 2023b). Within the *Dioscorea* genus, the application of whole plastome sequences has significantly clarified previously ambiguous phylogenetic aspects in certain taxa. For instance, Magwé-Tindo et al. (2018) utilized whole plastomes to construct a robust and well-supported phylogenetic tree of West African *Dioscorea* species, revealing six monophyletic groups within them. Additionally, Hu et al. (2023a) conducted phylogenetic analyses for *D.* sect. *Stenophora* using plastome sequences, suggesting that *D. biformifolia* and *D. banzhuana* represent successive sister species to the remaining *Stenophora* species. Despite these advancements, limited research has explored intraspecific variation and phylogeny of *Dioscorea* species, using whole plastome data. In this study, phylogenetic analyses based on plastome sequences delineated three distinct clades among *D. bulbifera* accessions originating from diverse regions across mainland China and Taiwan island (Figure 7). The identification of three distinct clades implied potential genetic

divergence among populations from different geographic regions. It is noteworthy that accession TWXB unexpectedly clustered with accessions from southern mainland China (FJZZ, GDSG, and GXQZ), displaying no mutations in their plastomes. This finding is surprising given that Taiwan Island became isolated around 10,000 years ago due to rising sea levels (Voris, 2000), leading to disrupted gene flow in many species through the formation of the Taiwan Strait (Lin et al., 2014). One plausible explanation for these results is that glaciations may have caused lowered sea levels, facilitating dispersal between Taiwan and mainland China and thereby obscuring the genetic endemism of Taiwanese accessions (Qu et al., 2015).

Overall, these findings underscored the significance of plastome phylogenomics in resolving intraspecific variation and phylogenetic relationships within *Dioscoea bulbifera*. Moving forward, it is imperative to acquire additional plastomes from *D. bulbifera* accessions in tropical Asia, Northern Australia, America, and sub-Saharan Africa (Kundu et al., 2021). This expansive dataset will provide a comprehensive perspective on the evolutionary relationships and processes of *D. bulbifera*, laying a robust foundation for further exploration of this economically significant species.

5 Conclusions

In conclusion, this study presented a comprehensive analysis of *Dioscorea bulbifera*, a versatile herbaceous climber with substantial nutritional and medicinal importance, through low-coverage whole genome sequencing. The investigation covered diverse accessions from mainland China and Taiwan, shedding light on the genetic variation within this species. Comparative plastome analysis revealed conserved structural features across accessions, with variations mainly attributed to intergenic spacer regions. The identification of plastome-derived markers, including dispersed repeats, SSRs, and divergent hotspots, along with high-quality polymorphic nuclear SSRs, provided valuable tools for population genetic studies and molecular breeding of *D. bulbifera*. The phylogenetic analysis revealed three distinct clades in these *D. bulbifera* accessions, indicating potential genetic divergence among populations from different geographic regions. Overall, this study not only addressed the existing gap in genetic variation studies of *D. bulbifera* in China but also laid the groundwork for further exploration and utilization of this valuable plant species.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: All newly generated plastome sequences were deposited in GenBank (accession numbers: PP130724-PP130733). The low-coverage whole genome

sequencing data of 10 Dioscorea bulbifera accessions generated in this study have been submitted to the NCBI SRA database (https://www.ncbi.nlm.nih.gov/sra/), under accession numbers: SRR27556260–SRR27556269.

Author contributions

RL: Software, Writing – original draft. KH: Software, Writing – original draft. XS: Resources, Writing – review & editing. MC: Conceptualization, Funding acquisition, Resources, Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1373297/full#supplementary-material

References

Abara, A., Tawo, E., Obi-Abang, M., and Obochi, G. (2011). Dietary fiber components of four common Nigerian *Dioscorea* species. *Pakistan J. Nutr.* 10, 383–387. doi: 10.3923/pjn.2011.383.387

Amiryousefi, A., Hyvönen, J., and Poczai, P. (2018). IRscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* 34, 3030–3031. doi: 10.1093/bioinformatics/bty220

Baruah, J., Gogoi, B., Das, K., Ahmed, N. M., Sarmah, D. K., Lal, M., et al. (2017). Genetic diversity study amongst *Cymbopogon* species from NE-India using RAPD and ISSR markers. *Ind. Crop Prod.* 95, 235–243. doi: 10.1016/j.indcrop.2016.10.022

Beier, S., Thiel, T., Münch, T., Scholz, U., and Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33, 2583–2585. doi: 10.1093/bioinformatics/btx198

Beyene, T. M. (2013). Genetic diversity of aerial yam (*Dioscorea bulbifera* L.) accessions in Ethiopia based on agronomic traits. *Agric. For. Fish* 2, 67–71. doi: 10.11648/j.aff.20130202.12

Birky, C. W., Maruyama, T., and Fuerst, P. (1983). An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. *Genetics* 103, 513–527. doi: 10.1093/genetics/103.3.513

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bredeson, J. V., Lyons, J. B., Oniyinde, I. O., Okereke, N. R., Kolade, O., Nnabue, I., et al. (2022). Chromosome evolution and the genetic basis of agronomically important traits in greater yam. *Nat. Commun.* 13, 2001. doi: 10.1038/s41467-022-29114-w

Brudno, M., Do, C. B., Cooper, G. M., Kim, M. F., Davydov, E., Green, E. D., et al. (2003). LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res.* 13, 721–731. doi: 10.1101/gr.926603

Burkill, I. H. (1960). The organography and evolution of Dioscoreaceae, the family of yams. *Bot. J. Linn. Soc* 56, 319–412. doi: 10.1111/j.1095-8339.1960.tb02508.x

Coursey, D. G. (1967). *Yams: an account of the nature, origins, cultivation and utilization of the useful members of the Dioscoreaceae* (London, UK: Longgmans, Green and Co), 230.

Daniell, H., Lin, C. S., Yu, M., and Chang, W. J. (2016). Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol.* 17, 1–29. doi: 10.1186/s13059-016-1004-2

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772–772. doi: 10.1038/nmeth.2109

Ding, Z. Z., and Gilbert, M. G. (2000). "Dioscoreaceae," in *Flora of China*, vol. 24. Eds. Z. Y. Wu and P. H. Raven (Science Press & St. Louis: Missouri Botanical Garden Press, Beijing), 287.

Dodsworth, S. (2015). Genome skimming for next-generation biodiversity analysis. *Trends Plant Sci.* 20, 525–527. doi: 10.1016/j.tplants.2015.06.012

Dutta, B. (2015). Food and medicinal values of certain species of *Dioscorea* with special reference to Assam. *J. Pharmacognosy Phytochem.* 3, 15–18.

Ezeocha, V., Nwogha, J., Ohuoba, A., and Chukwu, L. (2014). Evaluation of poultry manure application rates on the nutrient composition of *Dioscorea bulbifera* (Aerial yam). *Nigerian Food J.* 32, 92–96. doi: 10.1016/S0189-7241(15)30122-3

Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M., and Dubchak, I. (2004). VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* 32, 273–279. doi: 10.1093/nar/gkh458

Greiner, S., Lehwark, P., and Bock, R. (2019). OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* 47, 59–64. doi: 10.1093/nar/gkz238

Guan, X. R., Zhu, L., Xiao, Z. G., Zhang, Y. L., Chen, H. B., and Yi, T. (2017). Bioactivity, toxicity and detoxification assessment of *Dioscorea bulbifera* L.: a comprehensive review. *Phytochem. Rev.* 16, 573–601. doi: 10.1007/s11101-017-9505-5

Hu, K., Chen, M., Li, P., Sun, X. Q., and Lu, R. S. (2023b). Intraspecific phylogeny and genomic resources development for an important medical plant *Dioscorea nipponica*, based on low-coverage whole genome sequencing data. *Front. Plant Sci.* 14, 1320473. doi: 10.3389/fpls.2023.1320473

Hu, K., Sun, X. Q., Chen, M., and Lu, R. S. (2023a). Low-coverage whole genome sequencing of eleven species/subspecies in *Dioscorea* sect. *Stenophora* (Dioscoreaceae): comparative plastome analyses, molecular markers development and phylogenetic inference. *Front. Plant Sci.* 14, 1196176. doi: 10.3389/fpls.2023.1196176

Ikiriza, H., Okella, H., Tuyiringiye, N., Milton, A., Catherine, N., Wangalwa, R., et al. (2023). Diversity of *Dioscorea bulbifera* Linn in Uganda assessed by morphological markers and genotyping-by-sequencing technology (GBS). *J. Plant Breed. Crop Sci.* 15, 74–85. doi: 10.5897/JPBCS2023.1013

Jansen, R. K., Cai, Z., Raubeson, L. A., Daniell, H., Depamphilis, C. W., Leebens-Mack, J., et al. (2007). Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *P. Natl. Acad. Sci. U.S.A.* 104, 19369–19374. doi: 10.1073/pnas.0709121104

Jayeola, A., and Oyebola, T. (2013). Morpho-molecular studies in the natural populations of *Dioscorea bulbifera* Linn. in Nigeria. *J. Exp. Mol. Biol.* 14, 19.

Jin, J. J., Yu, W. B., Yang, J. B., Song, Y., DePamphilis, C. W., Yi, T. S., et al. (2020). GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. *Genome Biol.* 21, 1–31. doi: 10.1186/s13059-020-02154-5

Kaldate, R., Rana, M., Sharma, V., Hirakawa, H., Kumar, R., Singh, G., et al. (2017). Development of genome-wide SSR markers in horsegram and their use for genetic diversity and cross-transferability analysis. *Mol. Breed.* 37, 1–10. doi: 10.1007/s11032-017-0701-1

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kim, K. J., and Lee, H. L. (2004). Complete chloroplast genome sequences from Korean ginseng (*Panax schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants. *DNA Res.* 11, 247–261. doi: 10.1093/dnares/11.4.247

Kumar, M., Choi, J. Y., Kumari, N., Pareek, A., and Kim, S. R. (2015). Molecular breeding in *Brassica* for salt tolerance: importance of microsatellite (SSR) markers for molecular breeding in *Brassica*. *Front. Plant Sci.* 6, 688. doi: 10.3389/fpls.2015.00688

Kumar, S., Das, G., Shin, H. S., and Patra, J. K. (2017). *Dioscorea* spp. (a wild edible tuber): a study on its ethnopharmacological potential and traditional use by the local people of Similipal Biosphere Reserve, India. *Front. Pharmacol.* 8, 52. doi: 10.3389/fphar.2017.00052

Kuncari, E. S. (2022). Nutrition value and phytochemical screening of gembolo (*Dioscorea bulbifera* L.) bulbils and tubers from Bogor, West Java. *Jurnal Ilmu Pertanian Indonesia* 28, 18–25. doi: 10.18343/jipi.28.1.18

Kundu, B. B., Vanni, K., Farheen, A., Jha, P., Pandey, D. K., and Kumar, V. (2021). *Dioscorea bulbifera* L. (Dioscoreaceae): a review of its ethnobotany, pharmacology and conservation needs. *S. Afr. J. Bot.* 140, 365–374. doi: 10.1016/j.sajb.2020.07.028

Kurtz, S., Choudhuri, J. V., Ohlebusch, E., Schleiermacher, C., Stoye, J., and Giegerich, R. (2001). REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* 29, 4633–4642. doi: 10.1093/nar/29.22.4633

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., and Calcott, B. (2017). PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* 34, 772–773. doi: 10.1093/molbev/msw260

Li, B., Lin, F., Huang, P., Guo, W., and Zheng, Y. (2020). Development of nuclear SSR and chloroplast genome markers in diverse *Liriodendron chinense* germplasm based on low-coverage whole genome sequencing. *Biol. Res.* 53, 1–12. doi: 10.1186/s40659-020-00289-0

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv* 1303, 3997.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and samtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, X., Zhao, Y., Tu, X., Li, C., Zhu, Y., Zhong, H., et al. (2021). Comparative analysis of plastomes in Oxalidaceae: phylogenetic relationships and potential molecular markers. *Plant Diversity* 43, 281–291. doi: 10.1016/j.pld.2021.04.004

Li, Y., Tan, C., Li, Z., Guo, J., Li, S., Chen, X., et al. (2022). The genome of *Dioscorea zingiberensis* sheds light on the biosynthesis, origin and evolution of the medicinally important diosgenin saponins. *Hortic. Res.* 9, uhac165. doi: 10.1093/hr/uhac165

Lin, A. Q., Csorba, G., Li, L. F., Jiang, T. L., Lu, G. J., Thong, V., et al. (2014). Phylogeography of *Hipposideros armiger* (Chiroptera: Hipposideridae) in the Oriental Region: the contribution of multiple Pleistocene glacial refugia and intrinsic factors to contemporary population genetic structure. *J. Biogeogr.* 41, 317–327. doi: 10.1111/jbi.12163

Liu, L., Wang, Y., He, P., Li, P., Lee, J., Soltis, D. E., et al. (2018). Chloroplast genome analyses and genomic resource development for epilithic sister genera *Oresitrophe* and *Mukdenia* (Saxifragaceae), using genome skimming data. *BMC Genomics* 19, 235. doi: 10.1186/s12864-018-4633-x

Liu, L., Zhang, Y., and Li, P. (2021). Development of genomic resources for the genus *Celtis* (Cannabaceae) based on genome skimming data. *Plant Diversity* 43, 43–53. doi: 10.1016/j.pld.2020.09.005

Lu, R. S., Chen, M., Feng, Y., Yuan, N., Zhang, Y. M., Cao, M., et al. (2022). Comparative plastome analyses and genomic resource development in wild rice (*Zizania* spp., Poaceae) using genome skimming data. *Ind. Crop Prod.* 186, 115224. doi: 10.1016/j.indcrop.2022.115244

Lu, R. S., Hu, K., Zhang, F. J., Sun, X. Q., Chen, M., and Zhang, Y. M. (2023). Pan-plastome of greater yam (*Dioscorea alata*) in China: intraspecific genetic variation, comparative genomics, and phylogenetic analyses. *Int. J. Mol. Sci.* 24, 3341. doi: 10.3390/ijms24043341

Lu, R. S., Li, P., and Qiu, Y. X. (2016). The complete chloroplast genomes of three *Cardiocrinum* (Liliaceae) species: comparative genomic and phylogenetic analyses. *Front. Plant Sci.* 7, 2054. doi: 10.3389/fpls.2016.02054

Lu, R. S., Yang, T., Chen, Y., Wang, S. Y., Cai, M. Q., Cameron, K. M., et al. (2021). Comparative plastome genomics and phylogenetic analyses of Liliaceae. *Bot. J. Linn. Soc* 196, 279–293. doi: 10.1093/botlinnean/boaa109

Magwé-Tindo, J., Wieringa, J. J., Sonké, B., Zapfack, L., Vigouroux, Y., Couvreur, T. L., et al. (2018). Guinea yam (*Dioscorea* spp., Dioscoreaceae) wild relatives identified using whole plastome phylogenetic analyses. *Taxon* 67, 905–915. doi: 10.12705/675.4

Marakli, S. (2018). A brief review of molecular markers to analyze medicinally important plants. *Int. J. Life Sci. Biotechnol.* 1, 29–36. doi: 10.38001/ijlsb.438133

Mulualem, T., and Weldemichel, G. (2013). Agronomical evaluation of aerial yam (*Dioscorea bulbifera*) accessions collected from South and Southwest Ethiopia. *Greener J. Agric. Sci.* 3, 693–704. doi: 10.15580/GJAS.2013.3.073113767

Muraguri, S., Xu, W., Chapman, M., Muchugi, A., Oluwaniyi, A., Oyebanji, O., et al. (2020). Intraspecific variation within Castor bean (*Ricinus communis* L.) based on chloroplast genomes. *Ind. Crop Prod.* 155, 112779. doi: 10.1016/j.indcrop.2020.112779

Noda, H., Yamashita, J., Fuse, S., Pooma, R., Poopath, M., Tobe, H., et al. (2020). A large-scale phylogenetic analysis of *Dioscorea* (Dioscoreaceae), with reference to character evolution and subgeneric recognition. *Acta Phytotax. Geobot.* 71, 103–128. doi: 10.18942/apg.201923

Ojinnaka, M., Okudu, H., and Uzosike, F. (2017). Nutrient composition and functional properties of major cultivars of aerial yam (*Dioscorea bulbifera*) in Nigeria. *Food Sci. Qual. Manage.* 62, 1–2.

Osuagwu, A., and Edem, U. (2020). Evaluation of genetic diversity in Aerial Yam (*Dioscorea bulbifera* L.) using simple sequence repeats (SSR) markers. *Agrotechnology* 9, 202. doi: 10.35248/2168-9881.20.9.202

Otegbayo, B., Oguniyan, D., Olunlade, B., Oroniran, O., and Atobatele, O. (2018). Characterizing genotypic variation in biochemical composition, anti-nutritional and mineral bioavailability of some Nigerian yam (*Dioscorea* spp.) land races. *J. F. Sci. Tech.* 55, 205–216. doi: 10.1007/s13197-017-2913-0

Qu, Y., Song, G., Gao, B., Quan, Q., Ericson, P. G., and Lei, F. (2015). The influence of geological events on the endemism of East Asian birds studied through comparative phylogeography. *J. Biogeogr.* 42, 179–192. doi: 10.1111/jbi.12407

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/sys029

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299–3302. doi: 10.1093/molbev/msx248

Silva, D., Siqueira, M., Carrasco, N., Mantello, C., Nascimento, W., and Veasey, E. A. (2016). Genetic diversity among air yam (*Dioscorea bulbifera*) varieties based on single sequence repeat markers. *Genet. Mol. Res.* 15, 1–12. doi: 10.4238/gmr.15027929

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Straub, S. C., Parks, M., Weitemier, K., Fishbein, M., Cronn, R. C., and Liston, A. (2012). Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *Am. J. Bot.* 99, 349–364. doi: 10.3732/ajb.1100335

Terauchi, R., Terachi, T., and Tsunewaki, K. (1991). Intraspecific variation of chloroplast DNA in *Dioscorea bulbifera* L. *Theor. Appl. Genet.* 81, 461–470. doi: 10.1007/BF00219435

Voris, H. K. (2000). Maps of Pleistocene sea levels in South-east Asia: shorelines, river systems and time durations. *J. Biogeogr.* 27, 1153–1167. doi: 10.1046/j.1365-2699.2000.00489.x

Xia, E. H., Yao, Q. Y., Zhang, H. B., Jiang, J. J., Zhang, L. P., and Gao, L. Z. (2016). CandiSSR: an efficient pipeline used for identifying candidate polymorphic SSRs based on multiple assembled sequences. *Front. Plant Sci.* 6, 1171. doi: 10.3389/fpls.2015.01171