Journal of Tea Science Research, 2024, Vol.14, No.1, 64-78
http://hortherbpublisher.com/index.php/jtsr

Systematic Review Open Access

Genomic Insights into the Evolutionary History of the *Camellia* Genus: Comprehensive Analysis of Phylogenetic Relationships, Speciation, and Adaptive Evolution

Xinzhuan Yao [1], Hu Tang [1], Yujie Jiao [1], Yumei He [1] ✉, Litang Lu [2] ✉
1 College of Tea Sciences, Institute of Plant Health & Medicine, Guizhou University, Guiyang, 550025, Guizhou, China
2 College of Life Sciences/Key Laboratory of Plant Resource Conservation and Germplasm Innovation in the Mountainous Region (Ministry of Education), Guizhou University, Guiyang, 550025, Guizhou, China
✉ Corresponding authors: yumeih2010@163.com; ltlv@gzu.edu.cn

Journal of Tea Science Research, 2024, Vol.14, No.1 doi: 10.5376/jtsr.2024.14.0007
Received: 07 Jan., 2024
Accepted: 13 Feb., 2024
Published: 29 Feb., 2024

Preferred citation for this article:
Yao X.Z., Tang H., Jiao Y.J., He Y.M., and Lu L.T., 2024, Genomic insights into the evolutionary history of the *Camellia* genus: comprehensive analysis of phylogenetic relationships, speciation, and adaptive evolution, Journal of Tea Science Research, 14(1): 64-78 (doi: 10.5376/jtsr.2024.14.0007)

Abstract This study aims to synthesize current genomic research to elucidate the evolutionary history of the *Camellia* genus. By integrating various studies, it provides comprehensive insights into the genetic and evolutionary mechanisms that have shaped the diversity and adaptation of *Camellia* species. Genomic research has significantly advanced the understanding of the *Camellia* genus, revealing the genetic basis of adaptive traits and the mechanisms by which *Camellia* plants thrive in diverse ecological niches. Comparative chloroplast genomics has identified sequence polymorphisms and divergent hotspots that are valuable for phylogenetic analysis and species identification. The draft genome of tea (*Camellia sinensis* var. *sinensis*) highlighted two whole-genome duplications and the evolution of gene families critical for tea quality. Transcriptomic analysis of 116 *Camellia* plants provided evidence of a recent whole-genome duplication and identified gene families associated with stress resistance and secondary metabolism. The study found that hybridization events have significantly contributed to increased genetic diversity and adaptability. Additionally, the practical applications of genomic research in breeding programs have been demonstrated, leading to the development of new cultivars with improved traits. The integration of genomic, transcriptomic, and chloroplast data provides profound insights into the evolutionary history of the *Camellia* genus. These findings are crucial for developing effective conservation strategies and optimizing breeding programs to ensure the sustainability and economic viability of *Camellia* species, promoting the conservation and utilization of *Camellia* plants.

Keywords *Camellia* genus; Genomic insights; Evolutionary history; Phylogenetic relationships; Conservation genetics

Introduction

The genus *Camellia*, belonging to the family Theaceae, comprises over 200 species. These plants are primarily known for their commercial products, including tea leaves, ornamental flowers, and high-quality edible oils (Yang et al., 2013; Wu et al., 2022; Cheng et al., 2022). Particularly, the tea plant (*Camellia sinensis*) holds a crucial position due to its extensive use in tea production. Tea, a beverage widely consumed globally for its unique flavors and health benefits, makes significant contributions to the agricultural economy and cultural practices of many regions (Wei et al., 2018; Cheng et al., 2022). Additionally, ornamental species in the *Camellia* genus, such as *Camellia japonica* and *Camellia sasanqua*, are highly valued in horticulture for their beautiful flowers. The *Camellia* genus also makes important contributions to agriculture, with certain species used for oil production, such as *Camellia oleifera* Abel (Wu et al., 2022; Gao et al., 2023). The diverse applications of *Camellia* plants underscore their substantial impact on agriculture, the economy, and culture. Despite their significance, the phylogenetic relationships within the genus Camellia have been challenging to resolve due to frequent interspecific hybridization and polyploidization events (Yang et al., 2013; Huang et al., 2013; Li et al., 2019).

Genomics has emerged as a powerful tool in plant evolutionary studies, offering unprecedented insights into the genetic architecture and evolutionary dynamics of species. By analyzing the complete DNA sequences of organisms, genomic studies enable researchers to identify genetic variations, trace lineage relationships, and understand the mechanisms driving evolution (Xu et al., 2015). Genomic studies play a crucial role in understanding the evolutionary history of *Camellia* plants. By analyzing genomic data, researchers can uncover the genetic basis of important traits, such as stress resistance, secondary metabolite production, and adaptation to different environments (Zhang et al., 2021; Wu et al., 2022; Cheng et al., 2022). Whole-genome sequencing, transcriptomic analyses, and chloroplast genome studies have provided valuable insights into the phylogenetic relationships and evolutionary dynamics within the genus (Wei et al., 2018; Li et al., 2021; Cheng et al., 2022). These studies have revealed significant events such as whole-genome duplications and introgressions, which have contributed to the genetic diversity and adaptation of *Camellia* species (Wei et al., 2018; Li et al., 2021; Cheng et al., 2022).

The primary objective of this study is to synthesize current genomic research to elucidate the evolutionary history of the *Camellia* genus. By summarizing recent genomic research findings, we aim to provide a comprehensive understanding of the phylogenetic relationships, speciation events, and adaptive evolution within the *Camellia* genus. Identifying key evolutionary events, including highlighting significant genomic events such as whole-genome duplications, introgressions, and selection pressures that have shaped the evolution of *Camellia* species. Promoting future research and breeding programs, laying a foundation for future studies on the functional and morphological evolution of *Camellia*. By achieving these goals, we hope to gain a better understanding of the evolutionary history of *Camellia* plants and enhance their conservation and utilization in agriculture and horticulture.

1 Literature Review and Theoretical Framework

1.1 Classification and grading of the genus *Camellia*

The genus *Camellia*, belonging to the family Theaceae, comprises over 200 species, making it the largest and most significant genus within this family (Li et al., 2019). The classification of *Camellia* species has been challenging due to frequent interspecific hybridization and polyploidization, which complicate phylogenetic and taxonomic resolutions (Huang et al., 2013; Li et al., 2019). Recent advances in chloroplast genome sequencing have provided new insights into the phylogenetic relationships among *Camellia* species, aiding in more accurate classification (Huang et al., 2013; Yang et al., 2013; Li et al., 2019).

Among the major species within the genus, *Camellia sinensis*, *Camellia japonica*, and *Camellia sasanqua* are the most well-known. The distribution and diversity of these species are influenced by their evolutionary history and environmental adaptations (Huang et al., 2013; Wu et al., 2022). *Camellia sinensis* is widely cultivated in subtropical and tropical regions across Asia, particularly in China, India, and Sri Lanka, due to its significance in tea production. *Camellia japonica* and *Camellia sasanqua* are predominantly found in Japan and China, where they are highly valued as ornamental plants. Other notable species, such as *Camellia oleifera*, known for its oil-rich seeds, are distributed in various parts of East Asia (Yang et al., 2013; Li et al., 2019; Wu et al., 2022).

1.2 Morphological and ecological characteristics of *Camellia* genus

Camellia species exhibit a wide range of morphological traits, including variations in leaf shape, flower color, and seed structure. These traits not only aid in species identification but also reflect their adaptation to different ecological environments. Studies have found that *Camellia sinensis* typically has small, leathery leaves and white flowers, whereas *Camellia japonica* is known for its large, glossy leaves and a variety of flower colors, including red, pink, and white (Majumder et al., 2020). These morphological differences are often used to distinguish between species and are supported by genetic studies (Huang et al., 2013; Yang et al., 2013; Wu et al., 2022). For instance, chloroplast genome analysis has revealed specific genetic markers associated with morphological traits, aiding in species identification and classification (Li et al., 2019; Lin et al., 2022).

Camellia species occupy diverse ecological niches, ranging from subtropical forests to mountainous regions (Wu et al., 2022). *Camellia sinensis* thrives in humid, subtropical climates with well-drained soils, often found in regions with ample rainfall and moderate temperatures. In contrast, *Camellia japonica* prefers cooler, temperate climates and is commonly found in shaded woodland areas (Majumder et al., 2020). *Camellia sasanqua*, known for its ability to withstand drought and poorer soil conditions, often occupies more challenging habitats compared to its counterparts.

1.3 Economic and cultural importance

Camellia species hold significant cultural value in various regions, particularly in East Asia. In China and Japan, *Camellia japonica* have been celebrated in art, literature, and horticulture for centuries (Yang et al., 2013; Wu et al., 2022). Their flowers are often associated with beauty and resilience, symbolizing these virtues in cultural narratives and practices. Similarly, the *Camellia sinensis* holds deep cultural significance in countries such as China, India, and Japan, where tea ceremony and related traditions are deeply rooted in the social and cultural fabric (Wei et al., 2018; Zhang et al., 2021; Wu et al., 2022).

The economic value of *Camellia* species is substantial, with applications in the production of tea, ornamental plants, and edible oils. *Camellia sinensis* is perhaps the most economically valuable, as it is the primary source of tea, one of the most consumed beverages worldwide. The tea industry generates significant revenue and provides livelihoods for millions of people in tea-producing countries. Beyond its use in tea production, *Camellia sinensis* also has applications in traditional medicine and cosmetics, attributed to its antioxidant and anti-inflammatory properties (Wei et al., 2018; Zhang et al., 2021; Wu et al., 2022). *Camellia oleifera* is another economically valuable species. Its seeds can produce high-quality tea oil, which is rich in unsaturated fats and antioxidants. This oil is also used in cosmetics and skincare products (Wu et al., 2022). The genetic and genomic research on *Camellia* species has the potential to enhance their economic value by improving breeding programs and developing new varieties with desirable traits (Yan et al., 2018; Zhang et al., 2020; Zhang et al., 2021).

2 Genomic Approaches in *Camellia* Research

2.1 Genomic technologies

High-throughput sequencing (HTS) technologies, also known as next-generation sequencing (NGS), have revolutionized genomic research by enabling the rapid and cost-effective sequencing of entire genomes. In *Camellia* research, HTS technologies such as Illumina sequencing, Pacific Biosciences (PacBio) single-molecule real-time (SMRT) sequencing, and Oxford Nanopore sequencing are widely used. These technologies facilitate the assembly of high-quality reference genomes, identification of genetic variants, and comprehensive analysis of gene expression patterns (Zhang et al., 2020). Illumina sequencing technology has been extensively used to sequence complete chloroplast genomes of various *Camellia* species, providing valuable phylogenetic information and aiding in species identification (Yang et al., 2013; Huang et al., 2014).

Additionally, both Illumina and PacBio sequencing technologies have been employed to generate high-quality genome assemblies, such as the chromosome-level genome sequence of *Camellia chekiangoleosa* Hu (CCH). The study evaluated the genome size of CCH using two methods, generating 51.09 Gb of PacBio HiFi reads (approximately 19-fold genome coverage) and 283.40 Gb of Illumina Hi-C reads (approximately 102-fold genome coverage). Using the hifiasm software for assembly, a 2.73 Gb CCH genome was obtained, covering 97.40% of the scaffolds and consisting of 15 pseudochromosomes (scaffold N50 = 185.30 Mb) (Figure 1) (Shen et al., 2022). These studies have facilitated the study of tea oil biosynthesis and tea quality traits, respectively (Wei et al., 2018; Shen et al., 2022).

Bioinformatics tools and techniques are crucial for processing and analyzing the vast amounts of data generated by HTS. Several bioinformatics tools and techniques are employed in *Camellia* genomic research to ensure accurate data interpretation. For instance, PCR-based product sequencing and de novo assembly methods have been used to validate genome assemblies and identify genetic variations in *Camellia* species (Yang et al., 2013; Huang et al., 2014). Furthermore, bioinformatics pipelines have been developed to process raw HTS reads, generate consensus sequences, and recover alleles from heterozygous regions, which are essential for phylogenetic analyses and understanding evolutionary patterns (Uribe-Convers et al., 2016).

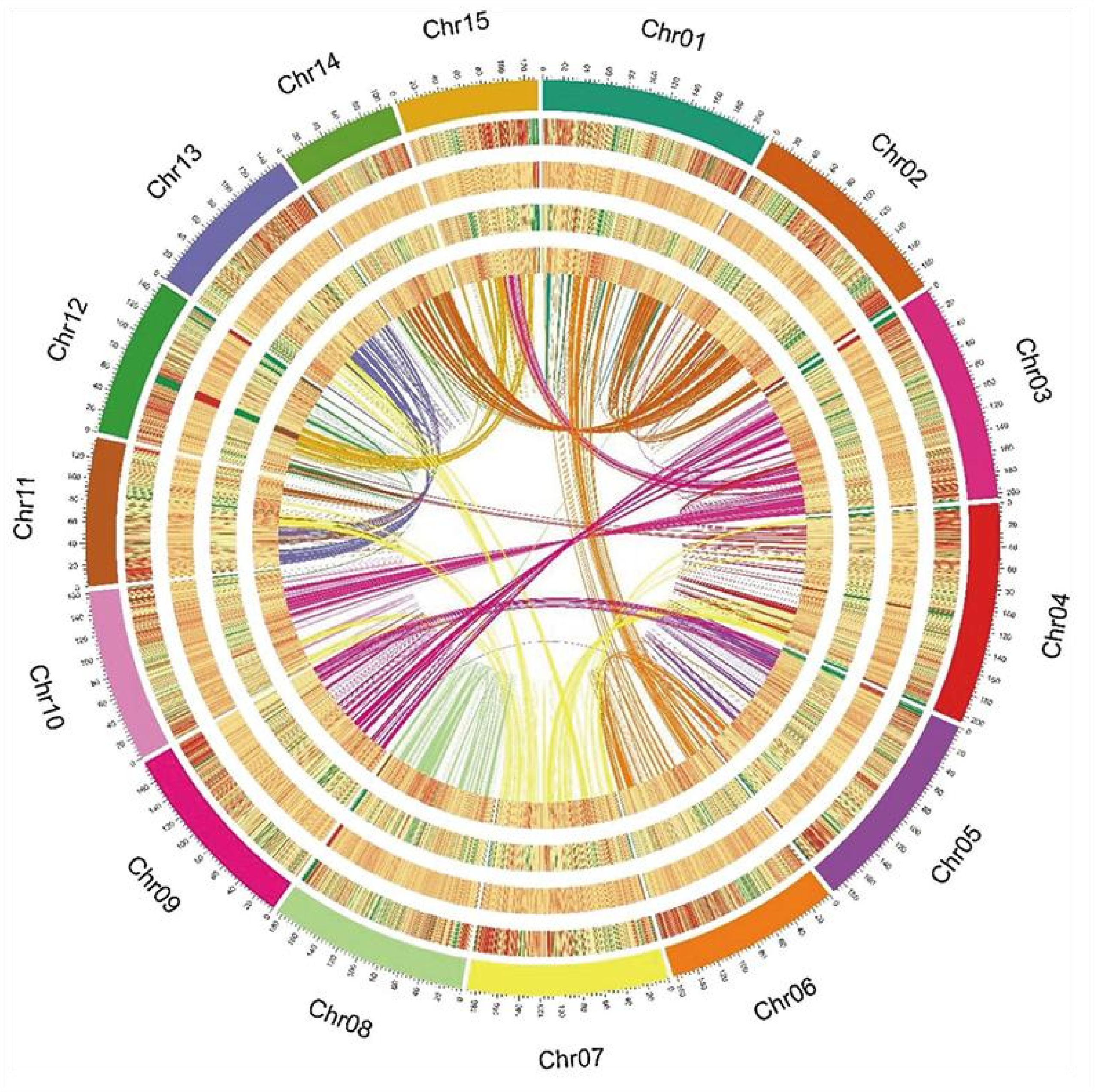

Figure 1 Features of the CCH (*Camellia chekiangoleosa* Hu) Genome (Adopted from Shen et al., 2022)

Image caption: The outermost ring of the figure shows the 15 pseudochromosomes of CCH. The second to fifth rings represent the genes, total TEs, Copia element distribution, and Gypsy element distribution, respectively. Green indicates lower density, while red indicates higher density. The innermost area represents the collinear regions within the CCH genome. The genome size of CCH was 2.73 Gb, covering 97.40% of the scaffolds (scaffold N50 = 185.30 Mb) (Adapted from Shen et al., 2022)

2.2 Genomic resources

Several reference genomes for *Camellia* species have been made available through recent genomic studies. The first chromosome-level genome sequence of *Camellia chekiangoleosa* provides insights into *Camellia* evolution and tea oil biosynthesis (Shen et al., 2022). Similarly, the draft genome sequence of *Camellia sinensis* var. *sinensis* offers valuable information on the evolution of the tea genome and the biosynthesis of key tea metabolites (Wei et al., 2018). These reference genomes serve as essential resources for further genomic and evolutionary studies in the *Camellia* genus.

Genomic data for *Camellia* species are stored in various databases and repositories, facilitating easy access and retrieval for researchers. For example, National Center for Biotechnology Information (NCBI) provides access to genomic sequences, gene annotations, and related metadata through platforms such as GenBank and the Sequence Read Archive (SRA) (Yi et al., 2024). The Tea Plant Information Archive (TPIA) is a specialized database that offers comprehensive genomic and transcriptomic data specific to *Camellia sinensis*. Additionally, Wu et al. (2022) have created a web-accessible database to store transcriptome sequences of 116 *Camellia* plants, enabling efficient conservation and utilization of *Camellia* germplasm for breeding programs. Chloroplast genome sequences of multiple *Camellia* species are also available, providing robust evidence for taxonomic studies and species identification (Huang et al., 2014).

2.3 Data analysis

Common data analysis methods in *Camellia* genomic research include sequence alignment, phylogenetic analysis, and comparative genomics. Sequence alignment and phylogenetic analysis have been used to resolve interspecies relationships and identify rapidly evolving regions in chloroplast genomes (Yang et al., 2013; Huang et al., 2014). Comparative genomics methods are used to identify orthologous genes and conserved regions across different *Camellia* species, providing insights into their evolutionary relationships. The results showed that the comparative genomic analysis has revealed whole-genome duplication events and gene family expansions, providing insights into the evolutionary history and trait adaptations in *Camellia* species (Wei et al., 2018; Shen et al., 2022).

Data processing and interpretation involve several steps, including quality control, sequence assembly, and functional annotation. Quality control measures, such as using propidium iodide (PI) flow cytometry analysis, ensure high-quality results with low coefficient of variation values (Huang et al., 2013). Sequence assembly techniques, such as de novo and reference-guided assembly, are employed to construct complete and draft genome sequences (Huang et al., 2014). Functional annotation helps identify genes related to key economic traits, such as flower and fruit development and stress tolerances, which are crucial for molecular breeding programs (Yan et al., 2018).

The integration of high-throughput sequencing technologies, bioinformatics tools, and comprehensive genomic resources has significantly advanced our understanding of the evolutionary history and genomic diversity of the *Camellia* genus. These approaches provide a solid foundation for future research and breeding efforts aimed at improving economically important traits in *Camellia* species.

3 Evolutionary History of the *Camellia* Genus

3.1 Phylogenetic studies

Phylogenetic studies aim to reconstruct the evolutionary relationships among species by analyzing genetic data. Phylogenetic studies of the *Camellia* genus have employed various molecular techniques to resolve the complex interspecies relationships. Chloroplast genome sequencing has been a prominent method, providing high-resolution data for phylogenetic analysis. For instance, complete chloroplast genomes of multiple *Camellia* species have been sequenced using Illumina technology, revealing conserved yet sufficiently variable regions for phylogenetic inference (Yang et al., 2013; Li et al., 2019; Lin et al., 2022). Additionally, nuclear DNA sequences such as ITS and *waxy* genes have been utilized, although with varying degrees of success due to issues like low divergence rates and amplification difficulties (Li et al., 2021).

Phylogenetic analyses have revealed significant insights into the evolutionary relationships within the *Camellia* genus. Studies have shown that the genus is monophyletic, but the interspecies relationships remain complex due to frequent hybridization and polyploidization events (Li et al., 2021; Li et al., 2019; Lin et al., 2022). Comparative genomic analyses have identified specific regions in the chloroplast genome that serve as effective barcode markers for species identification and phylogenetic resolution (Lin et al., 2022). Furthermore, transcriptomic data from 116 *Camellia* plants have provided robust phylogenetic trees supported by nuclear gene trees and morphological traits (Figure 2), highlighting a recent whole genome duplication event in the genus (Wu et al., 2022).

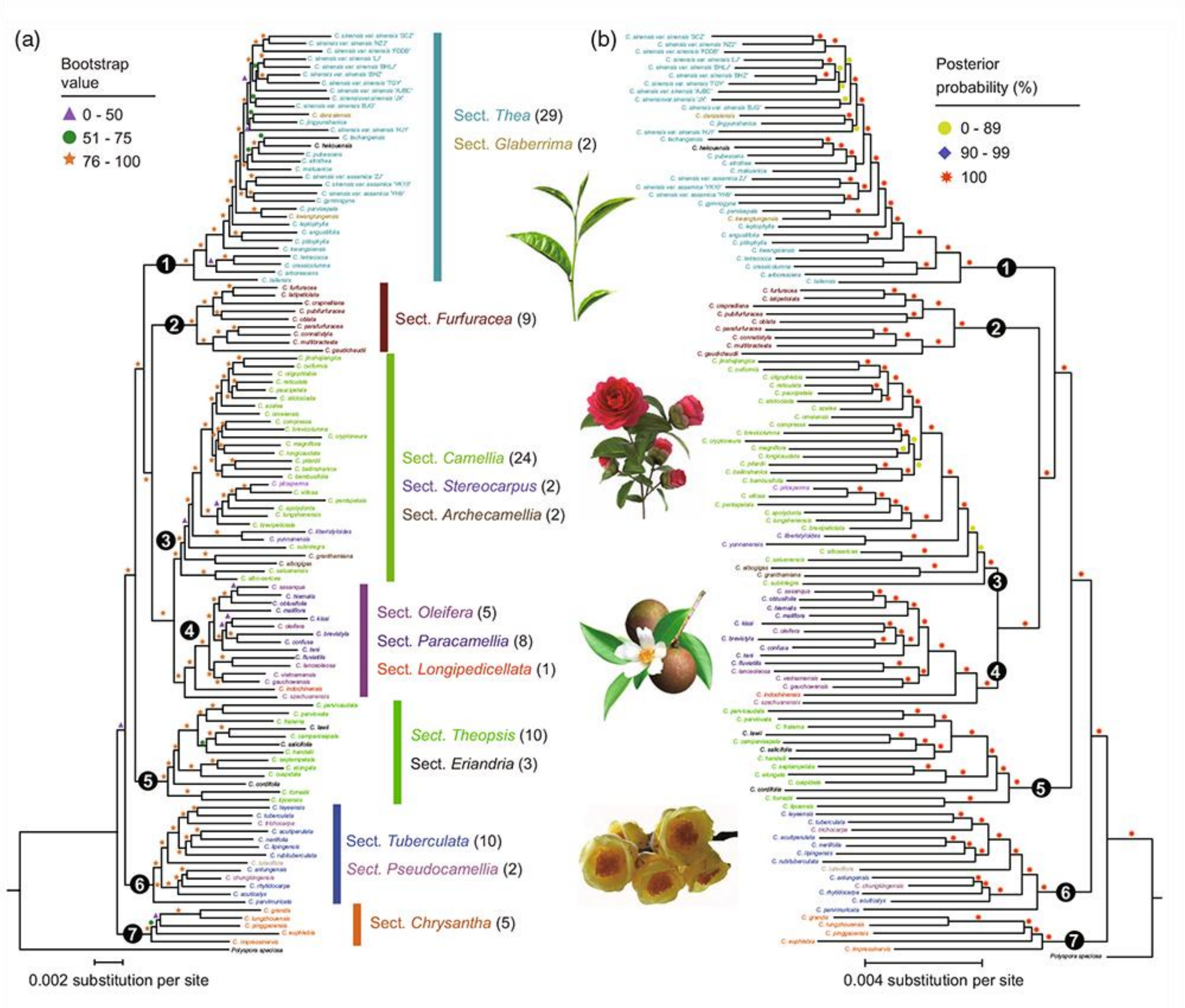

Figure 2 Phylogeny of 116 *Camellia* plants constructed using maximum likelihood and Bayesian methods (Adopted from Wu et al., 2022)

Image caption: (a) Maximum likelihood tree. (b) Bayesian tree. The bootstrap values and posterior probabilities are colored and displayed on the branches of the phylogeny. Major sections of Camellia are marked beside the tree. Numbers in brackets indicate the species numbers used in the present study. The figure shows that *Camellia* plants can be divided into seven major clusters. Most sections, such as Thea and Furfuracea, have 100% support rates. There are close relationships between sections Glaberrima and Thea, and between sections Paracamellia and Oleifera. Species from sections Pseudocamellia and Tuberculata cluster toether, while *Camellia lawii* from section Eriandria is most closely related to species from section Theopsis. Chrysantha is positioned at the base, indicating it is the oldest section. Unlike traditional classification systems, *C. luteoflora* from section Luteoflora is closer to species in Tuberculata. The phylogenetic tree supports revisions of many *Camellia* species in the Flora of China, indicating its accuracy and robustness (Adapted from Wu et al., 2022)

3.2 Speciation and divergence

Speciation in the *Camellia* genus is driven by a combination of hybridization, polyploidization, and environmental adaptation. Frequent interspecific hybridization has been documented, contributing to the genetic diversity and complexity of the genus (Li et al., 2019; Bajpai and Chaturvedi, 2023). Polyploidization events have also played a crucial role by creating reproductive barriers and promoting genetic divergence, further facilitating speciation, as evidenced by the significant variation in genome sizes among different *Camellia* species (Bajpai and Chaturvedi, 2023). Additionally, selection pressures from different environmental conditions have influenced the evolutionary trajectories of various species, leading to speciation (Li et al., 2019; Wu et al., 2022).

Estimating the divergence times of *Camellia* species provides insights into their evolutionary history. Molecular clock analyses, which use the rate of genetic mutations to estimate the timing of evolutionary events, have been employed to date the divergence of major *Camellia* clades. Results suggest that the genus *Camellia* began to diversify during the late Cretaceous to early Paleogene periods, approximately 60-70 million years ago (Cheng et al., 2022). Whole genome duplication events, critical to the evolution of key traits, have been dated to around 30 to 40 and 90 to 100 million years ago (Wei et al., 2018). The divergence between two major lineages, *Camellia sinensis* var. *sinensis* and *Camellia sinensis* var. *assamica*, is estimated to have occurred approximately 0.38 to 1.54 million years ago (Wei et al., 2018). These events have significantly impacted the genetic makeup and evolutionary history of the genus.

3.3 Hybridization events

Hybridization events are well-documented in the *Camellia* genus, contributing to its genetic diversity. Comparative genomic studies have identified numerous instances of interspecific hybridization, which complicates the phylogenetic resolution of the genus (Li et al., 2019; Lin et al., 2022). The presence of hybridization is further supported by the detection of sequence polymorphisms and structural variations in the chloroplast genomes of different species (Lin et al., 2022).

Hybridization has had a profound impact on the genetic diversity and adaptability of *Camellia* species. It has facilitated the exchange of genetic material between species, leading to the emergence of new traits and increased genetic variability (Li et al., 2019; Lin et al., 2022). This genetic diversity enhances the adaptability of *Camellia* species to various environmental conditions, contributing to their widespread distribution and ecological success (Li et al., 2019; Wu et al., 2022). Moreover, hybridization events have played a role in the evolution of economically important traits, such as stress resistance and secondary metabolite production, which are crucial for the cultivation and utilization of Camellia plants (Yan et al., 2018; Wu et al., 2022).

The evolutionary history of the *Camellia* genus is characterized by complex phylogenetic relationships, driven by mechanisms such as hybridization, polyploidization, and environmental adaptation. These processes have shaped the genetic diversity and adaptability of the genus, making it a valuable resource for both ecological and economic purposes.

4 Genomic Insights into Adaptation and Evolution

4.1 Adaptive traits

The adaptability of *Camellia* species to diverse environments is underpinned by specific genomic traits that confer resistance to pests, tolerance to environmental stresses, and other adaptive characteristics. High-throughput sequencing and genome-wide association studies (GWAS) have identified numerous genes associated with these adaptive traits. For instance, results showed that the cold tolerance in *Camellia japonica* var. *decumbens* has been linked to specific gene expression changes under cold stress. Transcriptome analysis identified differentially expressed genes (DEGs) involved in cold response, including transcription factors and genes related to signal transduction and plasma membrane stabilization (Wu et al., 2019).

Additionally, the tea plant *Camellia sinensis* has been found to possess lineage-specific genes (LSGs) that contribute to stress resistance and secondary metabolite production, such as catechins, which have antioxidant properties (Zhao and Ma, 2019). These findings highlight the genomic basis of key adaptive traits in *Camellia* species, providing valuable insights for breeding programs aimed at enhancing stress tolerance and other desirable traits.

4.2 Gene flow and genetic variation

Gene flow, the transfer of genetic material between populations, is a critical factor influencing genetic diversity and adaptation in *Camellia* species. Gene flow within the *Camellia* genus has been studied to understand the evolutionary dynamics and genetic diversity among populations. The chloroplast genome analysis of *Camellia japonica* revealed significant genetic variation and evolutionary routes influenced by environmental pressures (Li et al., 2019). This study demonstrated that different *Camellia* species exhibit distinct patterns of gene flow, which are crucial for maintaining genetic diversity and facilitating adaptation to varying environmental conditions.

Genetic variation plays a pivotal role in the adaptive evolution of *Camellia* species. Comparative genomics and transcriptomic analyses have shown that *Camellia* plants have undergone whole genome duplication events, leading to the expansion of transcription factor families associated with stress resistance and secondary metabolism (Wu et al., 2022). This genetic variation enables *Camellia* species to adapt to diverse environmental conditions and enhances their resilience to biotic and abiotic stresses. Furthermore, the presence of polymorphic ribosomal DNA (rDNA) in *Camellia* species suggests a mixture of concerted and birth-and-death evolution, contributing to the genetic diversity and potential adaptive functions of rRNA pseudogenes (Zhang et al., 2020).

4.3 Comparative genomics

Comparative genomics involves analyzing the genomes of different species to identify conserved and divergent genetic elements, providing insights into evolutionary relationships and functional adaptations. Comparative genomics studies have provided insights into the evolutionary history and adaptive mechanisms of the *Camellia* genus. For example, the comparative analysis of chloroplast genomes among different *Camellia* species has shed light on their phylogenetic relationships and evolutionary routes (Li et al., 2019). These studies have revealed that *Camellia* species exhibit unique evolutionary patterns compared to other related genera, highlighting the importance of chloroplast genome sequences in understanding the genetic basis of adaptation and evolution.

Comparative genomics has also uncovered significant findings related to the adaptive evolution of *Camellia* species. The analysis of transcriptomes from 116 *Camellia* plants revealed a recent whole genome duplication event and the expansion of transcription factor families associated with stress resistance and secondary metabolism (Wu et al., 2022). These genomic insights have provided a deeper understanding of the molecular mechanisms underlying the adaptation of *Camellia* species to their environments. Additionally, the identification of lineage-specific genes in *Camellia sinensis* has highlighted the role of gene duplication in generating new functions and adaptive traits (Figure 3) (Zhao and Ma, 2019). These findings underscore the importance of comparative genomics in elucidating the evolutionary history and adaptive strategies of the *Camellia* genus.

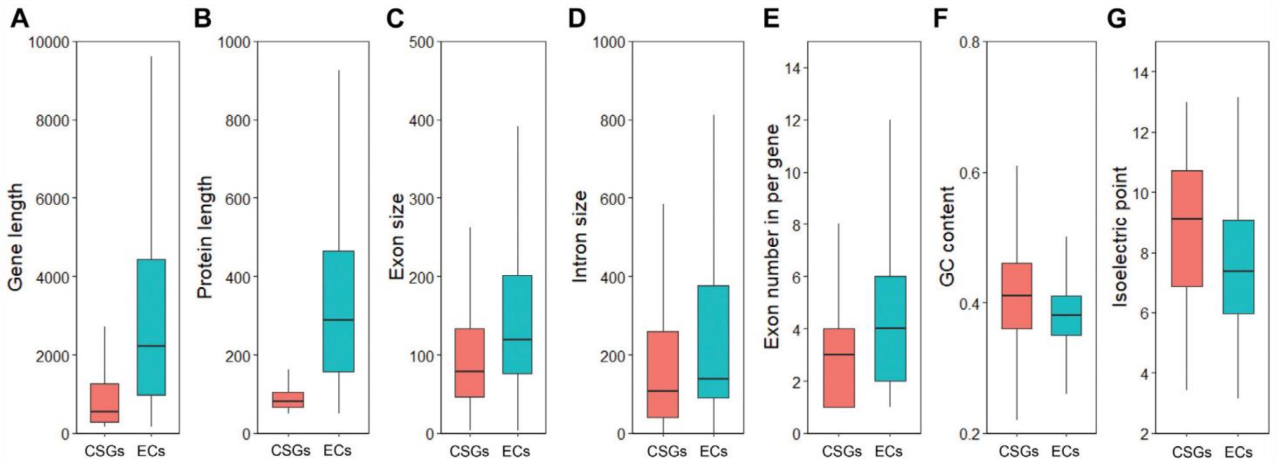

Figure 3 Analyze and compare the structural characteristics of *Camellia*-specific genes (CSGs) and evolutionarily conserved genes (ECs) (Adopted from Zhao and Ma, 2021)

Image caption: (A) Box-plot comparisons of gene length. (B) protein length. (C) exon size. (D) intron size. (E) exon number in per gene. (F) GCs content. (G) isoelectric point. The analysis of Figure 3 indicates that CSGs are generally shorter, with fewer and smaller exons and introns, but have higher GC content and isoelectric points compared to ECs. These structural differences suggest that CSGs are simpler and potentially more specialized, reflecting their specific evolutionary roles in *Camellia* species (Adapted from Zhao and Ma, 2021)

5 Case Studies

5.1 Case study: Tea plant (*Camellia sinensis*)

Camellia sinensis, commonly known as the tea plant, is one of the most economically significant species within the *Camellia* genus. It is the primary source of tea, a globally consumed beverage with a rich cultural heritage spanning thousands of years. The economic value of tea is immense, contributing substantially to the GDP of several tea-producing countries, including China, India, Sri Lanka, and Kenya (Wei et al., 2018; Xia et al., 2020; Wang et al., 2020). The tea industry provides employment to millions of people worldwide, from cultivation and harvesting to processing and marketing. In addition to its economic importance, tea is also valued for its health benefits, which include antioxidant properties and potential protective effects against various diseases.

Recent genomic research has provided extensive insights into the tea plant's genetic makeup. High-quality genome assemblies have been developed for various tea plant cultivars, such as *Camellia sinensis* var. *sinensis* and *Camellia sinensis* var. *assamica*, using advanced sequencing technologies like Illumina and PacBio (Wei et al., 2018; Xia et al., 2020; Zhang et al., 2021). These studies have identified key gene families involved in the biosynthesis of important tea metabolites, such as catechins, theanine, and caffeine, which are crucial for tea quality and health benefits (Wei et al., 2018; Yu et al., 2020; Zhang et al., 2021).

Additionally, population genomic analyses have revealed significant genetic diversity and evolutionary history within the tea plant species (Wang et al., 2020; Zhang et al., 2021). Zhang et al. (2021) revealed the evolutionary history of tea plant varieties by analyzing the haplotype-resolved genome assembly of Tieguanyin oolong tea. The analysis found that Tieguanyin tea has accumulated numerous mutations during long-term asexual reproduction, leading to potential gene expression differences and evolutionary selection (Figure 4) (Zhang et al., 2021). A population genomics analysis of 190 tea plant samples revealed the independent evolutionary histories and parallel domestication processes of the broad leaf variety (var. *assamica*) and the small-leaf variety (var. *sinensis*). The results indicate that extensive intra- and interspecific gene introgression has occurred throughout the evolution of tea plants. This genetic diversity plays a significant role in the flavor characteristics and resistance traits of modern tea plant varieties (Zhang et al., 2021).

5.2 Evolutionary insights from the tea plant

Phylogenetic analyses have placed *Camellia sinensis* within the broader *Camellia* genus, highlighting its evolutionary relationships with other *Camellia* species. Comparative studies of chloroplast and mitochondrial genomes have provided further insights into the evolutionary dynamics and phylogenetic clustering of different tea plant varieties (Li et al., 2021; Li et al., 2023). These analyses have also supported the hypothesis of multiple domestication events and the distinct evolutionary paths of *Camellia sinensis* var. *sinensis* and *Camellia sinensis* var. *assamica* (An et al., 2020; Li et al., 2021).

Genomic studies have identified several adaptive traits in the tea plant, including genes associated with stress resistance, flavor biosynthesis, and environmental adaptability. For instance, genes involved in terpene biosynthesis, which contribute to tea aroma, have been significantly amplified through recent tandem duplications (Xia et al., 2020; Wang et al., 2020). Additionally, selection for disease resistance and flavor has been stronger in *Camellia sinensis* var. *sinensis* populations compared to *Camellia sinensis* var. *assamica* populations, indicating adaptive evolution in response to domestication pressures (Wang et al., 2020).

5.3 Implications for breeding and conservation

The availability of high-quality genome sequences and the identification of simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs) have facilitated the development of molecular markers for tea plant breeding. These markers are valuable for genetic diversity assessment, QTL mapping, and marker-assisted selection, enabling the breeding of improved tea varieties with desirable traits such as enhanced flavor, disease resistance, and stress tolerance (Liu et al., 2018; An et al., 2020). Furthermore, haplotype-resolved genome assemblies provide a foundation for gene editing to enhance specific traits in tea crops (Zang et al., 2021).

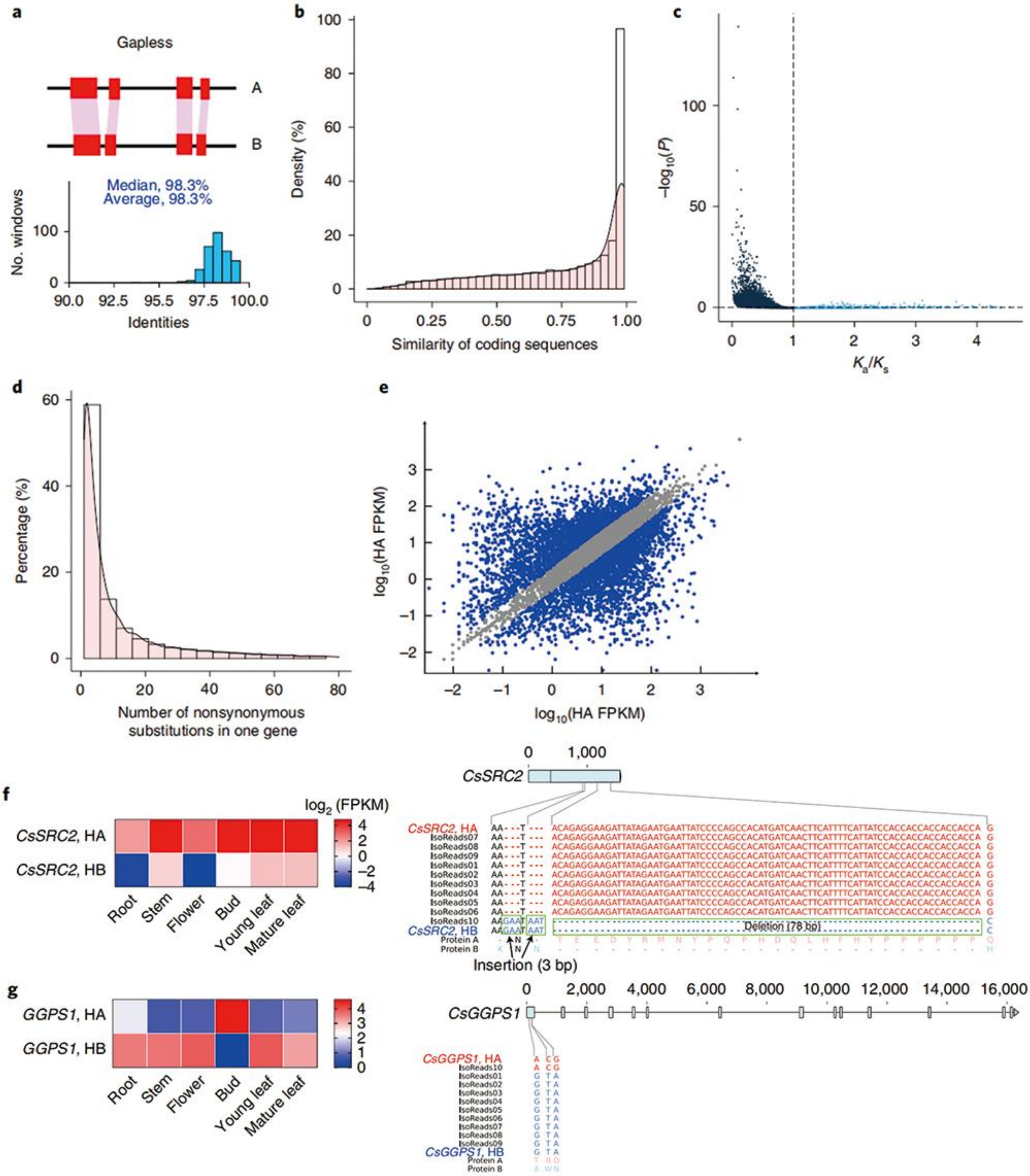

Figure 4 Genetic variations between haplotypes and allelic imbalance in *C. sinensis* (Adopted from Zhang et al., 2021)

Image caption: picture shows that the haplotype-resolved genome assembly of Tieguanyin tea reveals significant allelic expression imbalance and high genomic heterogeneity. Specifically, part a presents a whole-genome comparison of the two haplotypes, part b compares the coding sequences of alleles, part c shows the distribution of Ka/Ks ratios for allelic genes, part d displays the numerical distribution of nonsynonymous mutations, part e identifies genes with significant allele-specific expression (ASEGs) in tea leaves, and part f demonstrates the expression patterns of these genes in different tissues through an example. These data indicate that Tieguanyin tea has accumulated numerous mutations during long-term asexual reproduction, leading to gene expression differences and functional variations (Adapted from Zhang et al., 2021)

Genomic data have also provided critical insights for the conservation of tea plant genetic resources. Understanding the genetic diversity and evolutionary history of tea plant populations helps in the identification of unique and rare alleles that are important for maintaining genetic diversity. Conservation strategies can be developed to protect these genetic resources, ensuring the long-term sustainability of tea cultivation (Yu et al., 2020; Wang et al., 2020; Xia et al., 2020). Additionally, the identification of genes associated with environmental adaptability can inform conservation efforts in the face of climate change (An et al., 2020).

In summary, genomic research on *Camellia sinensis* has significantly advanced our understanding of its evolutionary history, adaptive traits, and genetic diversity. These insights are crucial for improving tea plant breeding and conservation strategies, ensuring the continued economic and ecological importance of this valuable crop.

6 The Significance of Conservation and Breeding

6.1 Conservation genetics

Genetic diversity is fundamental to the survival and adaptability of plant species, including those in the *Camellia* genus. High genetic diversity within populations allows for greater resilience against environmental changes, diseases, and pests. Conservation genetics focuses on understanding and preserving this diversity to ensure the long-term sustainability of species. The genus *Camellia*, known for its economic and ecological importance, exhibits significant genetic diversity across its species. This diversity is crucial for the resilience and adaptability of *Camellia* plants to various environmental stresses. Recent genomic studies have highlighted the extensive genetic polymorphism within the genus, particularly in the ribosomal DNA (rDNA) regions, which are subject to both concerted and birth-and-death evolution processes (Zhang et al., 2020).

To effectively conserve genetic resources of *Camellia* genus, it is imperative to develop comprehensive conservation strategies. These strategies should include the identification and preservation of genetically diverse populations, as well as the establishment of gene banks and in situ conservation areas. The use of genomic data can greatly enhance these efforts by providing detailed insights into the genetic structure and diversity of *Camellia* populations (Supple and Shapiro, 2018).

Preserving the genetic resources of *Camellia* is vital for several reasons. It ensures the availability of genetic material for future breeding programs aimed at improving desirable traits such as stress tolerance, disease resistance, and quality of tea and ornamental flowers (Yan et al., 2018; Wu et al., 2022). It helps maintain the ecological balance and biodiversity of the regions where *Camellia* species are native. And the conservation of genetic resources supports ongoing research efforts to understand the evolutionary history and adaptive mechanisms of *Camellia* plants (Zhang et al., 2021; Lin et al., 2022).

6.2 Breeding programs

Advancements in genomic technologies have revolutionized breeding programs for *Camellia*. Genomic-assisted breeding techniques, such as marker-assisted selection (MAS) and genomic selection (GS), enable the identification and utilization of specific genetic markers associated with desirable traits. For instance, the development of simple sequence repeat (SSR) markers from RNA-seq data has provided valuable tools for molecular breeding in *Camellia japonica* and other species within the genus (Li et al., 2021).

These techniques allow breeders to accelerate the selection process and improve the efficiency of breeding programs by targeting genes related to key economic traits such as flower and fruit development, stress tolerance, and secondary metabolite production (Yan et al., 2018; Wei et al., 2018). The integration of large-scale population and quantitative genetics data further enhances the precision and effectiveness of genomic-assisted breeding (Yan et al., 2018).

Genomic research have significantly contributed to the enhancement of desirable traits in *Camellia* breeding programs. For example, the identification of gene families involved in the biosynthesis of key tea metabolites, such as catechins and theanine, has provided valuable targets for improving tea quality (Wei et al., 2018).

Additionally, the discovery of genes associated with stress resistance and secondary metabolism has facilitated the development of *Camellia* varieties with enhanced tolerance to environmental stresses (Wu et al., 2020). The use of haplotype-resolved genome assemblies and population genomic analyses has also revealed the genetic basis of flavor characteristics and other important traits in tea plants, enabling breeders to select for these traits more effectively (Zhang et al., 2021). These genomic insights are instrumental in developing improved *Camellia* cultivars with superior agronomic and quality traits.

6.3 Practical applications

Several practical applications of genomic research in *Camellia* breeding have been documented. For instance, the construction of a pan-transcriptome for *Camellia* plants has facilitated the identification of gene expression patterns associated with tea quality and stress resistance, providing valuable information for breeding programs (Wu et al., 2020). Additionally, the development of a comprehensive tea plant information archive (TPIA2) has integrated large-scale genomic, transcriptomic, and metabolic data, offering a valuable resource for functional genomics and population genetic studies in *Camellia* (Gao et al., 2023).

To effectively conserve the genetic resources of *Camellia* species, specific measures should be followed. First, comprehensive genetic surveys are necessary to assess the genetic diversity of wild and cultivated *Camellia* populations and to identify key conservation targets. Second, habitat protection measures should be implemented, including the establishment of protected areas and land use management to reduce habitat destruction. *Ex-situ* conservation is equally important, involving the creation and maintenance of seed banks, living collections, and cryopreservation facilities to safeguard genetic material. Additionally, community involvement is crucial; local communities should be engaged in conservation efforts through education and the promotion of sustainable harvesting practices. Policy support is essential; policies that support conservation efforts should be developed and enforced, including land use regulations and incentives for preserving genetic diversity. By adhering to these guidelines, we can ensure the long-term preservation and sustainable use of *Camellia* genetic resources for future generations (Supple and Shapiro, 2018).

7 Challenges and Future Directions

7.1 Emerging technologies

The potential of new genomic technologies in further research on the genus *Camellia* is immense. High-throughput sequencing technologies, such as Illumina and PacBio, have already provided significant insights into the genomic structure and evolutionary history of various Camellia species (Wei et al., 2018; Li et al., 2019; Shen et al., 2022). The application of haplotype-resolved genome assembly has revealed allele-specific expressions and genetic diversity, which are crucial for understanding the evolutionary mechanisms and domestication processes in Camellia (Zhang et al., 2021). Additionally, the development of a web-accessible database for *Camellia* transcriptomes facilitates efficient data retrieval and utilization, promoting further research and breeding programs (Wu et al., 2022).

7.2 Research gaps and opportunities

Despite the advancements, several research gaps remain. One major gap is the limited understanding of the functional roles of rRNA pseudogenes and their evolutionary patterns within the *Camellia* genus (Zhang et al., 2020). Another gap is the need for more comprehensive studies on the chloroplast and mitochondrial genomes, which are essential for phylogenetic analyses and species identification (Li et al., 2019; Lin et al., 2022; Lu et al., 2022). Furthermore, the genetic basis of key economic traits, such as oil biosynthesis and stress tolerance, requires further investigation to enhance breeding programs (Yan et al., 2018; Shen et al., 2022).

Opportunities for future research include the integration of multi-omics approaches, such as genomics, transcriptomics, and metabolomics, to identify candidate genes associated with desirable traits (Yan et al., 2018; Wei et al., 2018). Additionally, exploring the genetic diversity and evolutionary history of less-studied *Camellia* species can provide valuable insights into the genus's adaptation mechanisms and potential for crop improvement (Zhang et al., 2019; Zhang et al., 2021).

7.3 Future directions

In the long term, genomic research on the genus *Camellia* is expected to revolutionize the understanding and utilization of this economically and ecologically valuable genus. The continuous development of high-throughput sequencing technologies and bioinformatics tools will enable more detailed and accurate genomic analyses, facilitating the identification of key genetic factors involved in trait development and adaptation (Wei et al., 2018; Li et al., 2021). Possible future developments include the creation of comprehensive genomic databases and resources that integrate data from various *Camellia* species, enhancing the ability to conduct comparative studies and identify evolutionary trends (Wu et al., 2022; Lin et al., 2022). Advancements in gene editing technologies, such as CRISPR/Cas9, also hold promise for targeted genetic modifications to improve desirable traits in *Camellia* crops, such as tea quality and oil content (Zhang et al., 2021; Shen et al., 2022). Overall, the future of genomic research in the *Camellia* genus looks promising, with the potential to significantly impact agriculture, horticulture, and conservation efforts.

8 Concluding Remarks

This study has synthesized the current genomic researches into the evolutionary history of the *Camellia* genus, highlighting several key findings and insights. The comparative analysis of chloroplast genomes in *Camellia sinensis* revealed that repeat-induced and indel-induced mutations are crucial for the diversification of the chloroplast genome. The study also highlighted the different selection pressures faced by Chinese and Indian tea varieties, suggesting distinct evolutionary paths. Multiple studies identified whole-genome duplication events in the *Camellia* genus, which have played a significant role in the expansion of gene families associated with secondary metabolism and stress resistance. These events have been crucial in shaping the genetic landscape of *Camellia* species. Phylogenetic analyses using nuclear and chloroplast genomes have provided a clearer picture of the relationships within the *Camellia* genus. These studies have identified several clades and suggested that regional constraints significantly influence the sequence evolution of chloroplast genomes. The availability of high-quality genome sequences and transcriptomes has facilitated the identification of key genes involved in tea quality, such as those responsible for catechin and caffeine biosynthesis. These resources are invaluable for breeding programs aimed at improving tea quality and stress resistance.

The importance of continuing genomic research on the *Camellia* genus cannot be overstated. Understanding the evolutionary history of *Camellia* species is not only crucial for taxonomy and phylogenetics but also for the conservation and utilization of these valuable plants. Genomic studies provide the tools needed to identify and preserve genetic diversity, which is essential for adapting to changing environmental conditions and improving crop resilience. Moreover, the insights gained from these studies can lead to the development of superior *Camellia* cultivars with enhanced qualities, such as improved tea flavor and increased resistance to pests and diseases.

There is a pressing need for further research and collaboration in the field of *Camellia* genomics. Researchers should focus on expanding the genomic data available for a broader range of *Camellia* species, particularly those that are underrepresented in current studies. Collaborative efforts between institutions, countries, and disciplines are crucial for sharing knowledge, resources, and technological advancements. Future research should also prioritize the functional characterization of genes identified in genomic studies, integrating multi-omics approaches to provide a comprehensive understanding of the molecular mechanisms underlying key traits. Additionally, exploring the genetic basis of hybridization and its evolutionary consequences will enhance our understanding of speciation and genetic diversity within the genus.

This study underscores the importance of continued genomic research and collaboration to advance our understanding of the *Camellia* genus. By building on the current genomic insights, researchers can develop innovative strategies for conservation, breeding, and sustainable utilization of *Camellia* species, ensuring their continued significance for generations to come.

Conflict of Interest Disclosure

The authors affirm that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

An Y., Mi X., Zhao S., Guo R., Xia X., Liu S., and Wei C., 2020, Revealing distinctions in genetic diversity and adaptive evolution between two varieties of Camellia sinensis by whole-genome resequencing, Frontiers in Plant Science, 11: 603819.
https://doi.org/10.3389/fpls.2020.603819

Bajpai R., and Chaturvedi R., 2023, In vitro production of doubled haploid plants in *Camellia* spp. and assessment of homozygosity using microsatellite markers, Journal of Biotechnology, 361: 89-98.
https://doi.org/10.1016/j.jbiotec.2022.11.019

Cheng L., Li M., Han Q., Qiao Z., Hao Y., Balbuena T.S., and Zhao Y., 2022, Phylogenomics resolves the phylogeny of Theaceae by using low-copy and multi-copy nuclear gene makers and uncovers a fast radiation event contributing to tea plants diversity, Biology, 11(7): 1007.
https://doi.org/10.3390/biology11071007

Gao Q., Tong W., Li F., Wang Y., Wu Q., Wan X., and Xia E., 2023, TPIA2: an updated tea plant information archive for Camellia genomics, Nucleic Acids Research, 52(D1): D1661-D1667.
https://doi.org/10.1093/nar/gkad701

Huang H., Shi C., Liu Y., Mao S., and Gao L., 2014, Thirteen Camelliachloroplast genome sequences determined by high-throughput sequencing: genome structure and phylogenetic relationships, BMC Evolutionary Biology, 14: 1-17.
https://doi.org/10.1186/1471-2148-14-151

Huang H., Tong Y., Zhang Q., and Gao L., 2013, Genome size variation among and within Camellia species by using flow cytometric analysis, PLoS One, 8(5): e64981.
https://doi.org/10.1371/journal.pone.0064981

Li J., Tang H., Luo H., Tang J., Zhong N., and Xiao L., 2023, Complete mitochondrial genome assembly and comparison of *Camellia sinensis* var. *Assamica* cv. Duntsa, Frontiers in Plant Science, 14: 1117002.
https://doi.org/10.3389/fpls.2023.1117002

Li L., Hu Y., He M., Zhang B., Wu W., Cai P., Huo D., and Hong Y., 2021a, Comparative chloroplast genomes: insights into the evolution of the chloroplast genome of Camellia sinensis and the phylogeny of Camellia. BMC genomics, 22: 1-22.
https://doi.org/10.1186/s12864-021-07427-2

Li Q., Su X., Ma H., Du K., Yang M., Chen B., Fu S., Fu T., Xiang C., Zhao Q., and Xu L., 2021b, Development of genic SSR marker resources from RNA-seq data in Camellia japonica and their application in the genus Camellia. Scientific Reports, 11(1): 9919.
https://doi.org/10.1038/s41598-021-89350-w

Li W., Zhang C., Guo X., Liu Q., and Wang K., 2019, Complete chloroplast genome of Camellia japonica genome structures, comparative and phylogenetic analysis, PLoS ONE, 14(5): e0216645.
https://doi.org/10.1371/journal.pone.0216645

Lin P., Yin H., Wang K., Gao H., Liu L., and Yao X., 2022, Comparative genomic analysis uncovers the chloroplast genome variation and phylogenetic relationships of *Camellia* species. Biomolecules, 12(10): 1474.
https://doi.org/10.3390/biom12101474

Liu S., An Y., Li F., Li S., Liu L., Zhou Q., Zhao S., and Wei C., 2018, Genome-wide identification of simple sequence repeats and development of polymorphic SSR markers for genetic studies in tea plant (*Camellia sinensis*), Molecular Breeding, 38: 1-13.
https://doi.org/10.1007/s11032-018-0824-z

Lu C., Gao L., and Zhang Q., 2022, A high-quality genome assembly of the mitochondrial genome of the oil-tea tree *Camellia gigantocarpa* (Theaceae). Diversity, 14(10): 850.
https://doi.org/10.3390/d14100850

Majumder S., Ghosh A., and Bhattacharya M., 2020, Natural anti-inflammatory terpenoids in *Camellia japonica* leaf and probable biosynthesis pathways of the metabolome. Bulletin of the National Research Centre, 44: 1-14.
https://doi.org/10.1186/s42269-020-00397-7

Shen T., Huang B., Xu M., Zhou P., Ni Z., Gong C., Wen Q., Cao F., and Xu L., 2022, The reference genome of *Camellia chekiangoleosa* provides insights into *Camellia* evolution and tea oil biosynthesis. Horticulture Research, 9: uhab083.
https://doi.org/10.1093/hr/uhab083

Supple M., and Shapiro B., 2018, Conservation of biodiversity in the genomics era. Genome Biology, 19: 1-12.
https://doi.org/10.1186/s13059-018-1520-3

Uribe-Convers S., Settles M., and Tank D., 2016, A phylogenomic approach based on PCR target enrichment and high throughput sequencing: Resolving the diversity within the South American species of *Bartsia* L.(Orobanchaceae), PLoS One, 11(2): e0148203.
https://doi.org/10.1371/journal.pone.0148203

Wang X., Feng H., Chang Y., Ma C., Wang L., Hao X., Li A., Cheng H., Wang L., Cui P., Jin J., Wang X., Wei K., Ai C., Zhao S., Wu Z., Li Y., Liu B., Wang G., Chen L., Ruan J., and Yang Y., 2020, Population sequencing enhances understanding of tea plant evolution, Nature Communications, 11(1): 4447.

Wei C., Yang H., Wang S., Zhao J., Liu C., Gao L., Xia E., Lu Y., Tai Y., She G., Sun J., Cao H., Tong W., Gao Q., Li Y., Deng W., Jiang X., Wang W., Chen Q., Zhang S., Li H., Wu J., Wang P., Li P., Shi C., Zheng F., Jian J., Huang B., Shan D., Shi M., Fang C., Yue Y., Li F., Li D., Wei S., Han B., Jiang C., Yin Y., Xia T., Zhang Z., Bennetzen J., Zhao S., and Wan X., 2018, Draft genome sequence of *Camellia sinensis* var. *sinensis* provides insights into the evolution of the tea genome and tea quality. Proceedings of the National Academy of Sciences of the United States of America, 115(18): E4151-E4158.
https://doi.org/10.1073/pnas.1719622115

Wu Q., Tong W., Zhao H., Ge R., Li R., Huang J., Li F., Wang Y., Ali I., Deng W., Wang W., Wan X., Zhang Z., and Xia E., 2022, Comparative transcriptomic analysis unveils the deep phylogeny and secondary metabolite evolution of 116 *Camellia* plants, The Plant Journal, 111(2): 406-421.
https://doi.org/10.1111/tpj.15799

Wu Y., Müller M., Bai T., Yao S., Gailing O., and Liu Z., 2019, Transcriptome profiling in *Camellia japonica* var. *decumbens* for the discovery of genes involved in chilling tolerance under cold stress, Annals of Forest Research.
https://doi.org/10.15287/AFR.2018.1311

Xia E., Tong W., Hou Y., An Y., Chen L., Wu Q., Liu Y., Yu J., Li F., Li R., Li P., Zhao H., Ge R., Huang J., Mallano A., Zhang Y., Liu S., Deng W., Song C., Zhang Z., Zhao J., Wei S., Zhang Z., Xia T., Wei C., and Wan X., 2020, The reference genome of tea plant and resequencing of 81 diverse accessions provide insights into genome evolution and adaptation of tea plants, Molecular Plant, 13(7): 1013-1026.
https://doi.org/10.1016/j.molp.2020.04.010

Xu J., Xu Y., Yonezawa T., Li L., Hasegawa M., Lu F., Chen J., and Zhang W., 2015, Polymorphism and evolution of ribosomal DNA in tea (*Camellia sinensis*, Theaceae), Molecular Phylogenetics and Evolution, 89: 63-72.
https://doi.org/10.1016/j.ympev.2015.03.020

Xu Y., Liu Y., Yu Z., and Jia X., 2023, Complete Chloroplast Genome Sequence of the Long Blooming Cultivar Camellia 'Xiari Qixin': Genome Features, Comparative and Phylogenetic Analysis. Genes, 14(2): 460.
https://doi.org/10.3390/genes14020460

Yan C., Lin P., Lyu T., Hu Z., Fan Z., Li X., Yao X., Li J., and Yin H., 2018, Unraveling the Roles of Regulatory Genes during Domestication of Cultivated Camellia: Evidence and Insights from Comparative and Evolutionary Genomics, Genes, Unraveling the roles of regulatory genes during domestication of cultivated Camellia: evidence and insights from comparative and evolutionary genomics, Genes, 9(10): 488.
https://doi.org/10.3390/genes9100488

Yang J., Yang S., Li H., Yang J., and Li D., 2013, Comparative Chloroplast Genomes of Camellia Species. PLoS ONE, 8(8): e73053.

Yi L., Wang Y., Li Y., Zhang D., and Tong W., 2024, Characterization of the complete chloroplast genome sequence of Camellia tetracocca (Theaceae). Mitochondrial DNA Part B, 9(4), 461-464.

Yu X., Xiao J., Chen S., Yu Y., Ma J., Lin Y., Li R., Lin J., Fu Z., Zhou Q., Chao Q., Chen L., Yang Z., and Liu R., 2020, Metabolite signatures of diverse *Camellia sinensis* tea populations, Nature Communications, 11(1): 5586.
https://doi.org/10.1038/s41467-020-19441-1

Zhang M., Tang Y., Xu Y., Yonezawa T., Shao Y., Wang Y., Song Z., Yang J., and Zhang W., 2020, Concerted and birth-and-death evolution of 26S ribosomal DNA in *Camellia* L., Annals of Botany, 127(1): 63-73.
https://doi.org/10.1093/aob/mcaa169

Zhang W., Zhao Y., Yang G., Peng J., Chen S., and Xu Z., 2019, Determination of the evolutionary pressure on *Camellia oleifera* on Hainan Island using the complete chloroplast genome sequence, PeerJ, 7: e7210.
https://doi.org/10.7717/peerj.7210

Zhang X., Chen S., Shi L., Gong D., Zhang S., Zhao Q., Zhan D., Vasseur L., Wang Y., Yu J., Liao Z., Xu X., Qi R., Wang W., Ma Y., Wang P., Ye N., Ma D., Shi Y., Wang H., Ma X., Kong X., Lin J., Wei L., Ma Y., Li R., Hu G., He H., Zhang L., Ming R., Wang G., Tang H., and You M., 2021, Haplotype-resolved genome assembly provides insights into evolutionary history of the tea plant *Camellia sinensis*. Nature Genetics, 53: 1250-1259.
https://doi.org/10.1038/s41588-021-00895-y

Zhao D., Hodkinson T., and Parnell J., 2023, Phylogenetics of global Camellia (Theaceae) based on three nuclear regions and its implications for systematics and evolutionary history, Journal of Systematics and Evolution, 61(2): 356-368.
https://doi.org/10.1111/jse.12837

Zhao Z., and Ma D., 2021, Genome-wide identification, characterization and function analysis of lineage-specific genes in the tea plant *Camellia sinensis*, Frontiers in Genetics, 12: 770570.
https://doi.org/10.3389/fgene.2021.770570