Advances in Agriculture and Biology (2025) 8(1): 1-13
DOI: https://doi.org/10.63072/aab.25001
https://aabinternational.com/
ISSN: 3079-174X

RESEARCH PAPER

# RNA-seq-based SNP discovery and functional characterization reveal genetic variation and evolutionary insights in Cinnamomum species

**Gülistan Genli[1] and Mehmet Alp Furan[1]***

[1]Department of Agricultural Biotechnology, Agriculture Faculty, Van Yuzuncu Yil University, 65090 Van, Turkey

*Corresponding author's email: alpfuran@gmail.com

**Key Message:** RNA-Seq-based SNP profiling of *Cinnamomum camphora* and *C. verum* uncovered extensive genomic variation, functional enrichment in major metabolic and regulatory pathways, and evidence of evolutionary selection, offering new insights into genetic diversity, adaptation, and molecular evolution within the Cinnamomum genus.

## Abstract

Genetic differences are essential for the adaptability, resilience, and evolutionary success of plant species. Cinnamomum, a genus of commercial and therapeutic significance, remains little studied at the genomic level. In this study, we conducted a comprehensive RNA-Seq-based single nucleotide polymorphism (SNP) analysis on six transcriptome samples from *C. camphora* and *C. verum*, aligned with the *C. kanehirae* reference genome. We identified a total of 184,532 high-confidence SNPs, comprising 410,247 missense SNPs and 397,053 synonymous SNPs, yielding a missense-to-synonymous ratio of 1.03. Chromosome-wide analysis revealed that chromosome QPKB01000001.1 contained the highest SNP count (12,458 SNPs), whereas QPKB01000005.1 exhibited the strongest population differentiation, with a mean FST value of 0.342. A strong positive correlation ($r = 0.79$, $p < 0.001$) was observed between SNP density and chromosomal length. Functional enrichment analysis demonstrated that SNP-associated genes were significantly enriched (FDR < 0.05) in metabolic processes, cellular architecture, and regulatory pathways. KEGG annotation linked SNPs to significant metabolic pathways, including purine metabolism, phenylpropanoid biosynthesis ($-\log10(p) > 5$), and lignin biosynthesis. Analysis of linkage disequilibrium (LD) decay showed that the mean $r^2$ value dropped below 0.2 within approximately 50 kilobases, indicating high recombination rates and low genome-wide LD. The allele frequency spectrum revealed a predominance of low-frequency variants (MAF < 0.1), constituting over 60% of all SNPs, which suggests the influence of genetic drift and recent population expansion. Notably, over 20 genes exhibited an accumulation of more than 50 missense SNPs, highlighting potential targets of evolutionary selection. These results clarify the molecular and evolutionary dynamics of Cinnamomum species and provide a robust basis for additional research into plant adaptation, genetic conservation, and trait-based selection in non-model plant genomes. 

**Keywords:** *Cinnamomum*, Functional annotation, Genetic variation, Population genetics, RNA-Seq, SNP analysis

**Abbreviations:** **AFS**: Allele Frequency Spectrum; **BAM**: Binary Alignment Map; **BCFtools**: Variant Calling and Manipulation Tools; **BWA-MEM**: Burrows-Wheeler Aligner - Maximal Exact Matches; **FDR-adjusted p-value:** False Discovery Rate-adjusted (p-values); **FastQ:** FASTA Qualty scores; **FastQC**: Quality Control Tool for High Throughput Sequence Data; **Fst**: Wright's Fixation Index; **GO**: Gene Ontology; **Kb:** Kilobase; **KEGG**: Kyoto Encyclopedia of Genes and Genomes; **LD**: Linkage Disequilibrium; **MAF**: Minor Allele Frequency; **MultiQC**: Aggregate Reports for Bioinformatics Analyses; **NCBI**: National Center for Biotechnology Information; **NumPy:** Numerical Python; **PASS:** Passed all filters; **PCA**: Principal Component Analysis; **PLINK**: Whole Genome Association Analysis Toolset; **RNA-Seq**: RNA Sequencing; **SAM**: Sequence Alignment/Map Format; **SciPy:** Scientific Python; **SNP**: Single Nucleotide Polymorphism; **SnpEff**: SNP Effect Predictor (SNP anotasyonu için kullanılan yazılım); **VCF**: Variant Call Format; **VCFtools**: Variant Call Format Data Analysis

**Citation:** Genli, G., & Furan, M. A. (2025). RNA-seq-based SNP discovery and functional characterization reveal genetic variation and evolutionary insights in Cinnamomum species. *Advances in Agriculture and Biology*, *8*(1), 1–13. https://doi.org/10.63072/aab.25001

## Introduction

Cinnamomum is a genus of evergreen aromatic trees and shrubs belonging to the Lauraceae family. Cinnamomum species are economically important; their leaf extracts are used to treat wounds, fever, intestinal worms, and headaches, and exhibit high antimicrobial activity (Bharathi et al., 2025). They also hold significant commercial and pharmacological value (Zhang et al., 2025). *Cinnamomum camphora* and *Cinnamomum verum* have drawn significant interest owing to their essential oil constituents and therapeutic attributes (Lopez-Maestre et al., 2016). Nonetheless, the genetic diversity and evolutionary adaptation mechanisms of these species remain inadequately comprehended. Recent research has emphasized the antioxidant, antibacterial, anti-inflammatory, and anticancer characteristics of *Cinnamomum kanehirae* leaf extracts, indicating prospective uses in pharmacology and phytotherapy (Ma et al., 2016). Comprehending the genetic distinctions across

*Cinnamomum* species can enhance biotechnological applications, agricultural breeding research, and species conservation through the analysis of both intra- and interspecies variants (Micheel et al., 2021; Manzoor et al., 2025).

Genetic variety is the main process by which organisms adapt to environmental conditions and biodiversity (Andrews, 2010; Rafique et al., 2025). Understanding the genetic variances can help to conserve and breed genetic resources for plant species that are vital for agriculture and pharmaceuticals (Booy et al., 2000; Ajayi et al., 2024). The United Nations Sustainable Development Goals (SDGs), Aichi Biodiversity Targets, and IPBES Global Assessment Reports all stress the value of genetic variety for ecosystem health and human well-being (Danecek et al., 2011; Cingolani et al., 2012; Ewels et al., 2016; Hoban et al., 2022; Huong et al., 2023). A plant's adaptive response to a changing environment depends on its level of genetic diversity. Assessing genetic diversity and understanding population indices are therefore pivotal for deciphering plant adaptation to dynamic environmental stressors (Abdullah et al., 2025). Various molecular markers have been employed to evaluate genetic diversity, examine population structure, and delineate clustering patterns within and among populations (Bidyananda et al., 2024).

Molecular markers are widely used in plant genetic diversity and population genetics studies, playing an essential role in breeding and improvement, conservation, preservation, promotion, and reintroduction of endangered and valuable plant species (Tikendra et al., 2019; Hussain & Nisar, 2020; Tuvesson et al., 2021). By enabling the identification of new plant varieties and the detection of genetic changes in existing ones, they provide valuable insights into genetic variation within and between plant populations (Amom et al., 2023). Genetic diversity is vital for plants to adapt to environmental changes (Pereira-Dias et al., 2019). Plant responses to climate change and their adaptive capacity depend largely on their genetic diversity (Apana et al., 2021). Molecular markers also play a key role in optimizing plant use and ensuring effective management by providing essential information on genetic composition and population structure (Gyani et al., 2020). Single nucleotide polymorphisms (SNPs) are one of the most common genomic variations that provide critical information about the evolutionary processes, environmental adaptations, and functional genetic structures of plant species (Jain et al., 2022).

In recent years, RNA-Seq-based SNP analysis has emerged as a powerful method for identifying genome-wide genetic variations and correlating them with gene expression (Jehl et al., 2021). Recent studies have demonstrated the application of RNA-Seq in identifying SNPs linked to functional traits in various plant species, such as *Camellia sinensis* (Leinonen et al., 2010; Kardos et al., 2021). RNA-Seq plays a critical role in understanding the genetic structures of plants, determining their phylogenetic relationships, and population genetics studies by detecting SNPs at the transcriptome level (Li, 2013). This study employed RNA-Seq-based SNP analysis to investigate genetic variations in *Cinnamomum camphora* and *Cinnamomum verum*, aiming to elucidate the biological significance of these polymorphisms (Nevo, 2001; Nonić & Šijačić-Nikolić, 2020). Genome-wide SNP distribution was assessed, followed by functional annotation and population genetic analysis (Pathak & Sharma, 2021; Li et al., 2023). Furthermore, the study explores the genetic architecture of *Cinnamomum* species, emphasizing the computational and statistical approaches utilized in SNP analysis (Purcell et al., 2007; Schmidt et al., 2022). Advanced bioinformatics and computational biology techniques were applied for SNP identification, functional characterization, and large-scale RNA-Seq data processing.

Despite their economic and ecological importance, Cinnamomum species are among the many non-model plants that suffer from a significant genomic data deficiency (Zhang et al., 2021). A comprehensive catalog of Single Nucleotide Polymorphisms (SNPs) for *C. camphora* and *C. verum* is not available, hindering population genetic and association studies (Gong et al., 2021; Thanthirige et al., 2025). The specific genes, metabolic pathways, and genomic regions underpinning their adaptation to diverse environments and contributing to their valuable secondary metabolite profiles (e.g., phenylpropanoids) are largely unknown (Ninkuu et al., 2025). In addition to genomic approaches, studies on plant bioactive compounds highlight the wider significance of understanding species-specific chemical diversity, with comparative *in vitro* analyses demonstrating how leaf extracts from different Artemisia species exhibit broad antimicrobial potential against infection-causing pathogens (Fazal et al., 2025). Moreover, research on medicinal plants such as *Panax ginseng* highlights the complex interplay between genetic background and functional traits by detailing the chemical characteristics, functional properties, and health benefits of ginseng bioactives, which may parallel the metabolic variation observed in other plant taxa (Bo et al., 2025). Integrating insights from phytochemical and genomic variation reinforces the importance of uncovering both genetic polymorphisms and their potential phenotypic and functional correlates in non-model plant species (Khan et al., 2020).

Fundamental population genetic parameters such as linkage disequilibrium decay rates, demographic history (e.g., population expansion), and the signatures of natural selection are yet to be characterized for these species (Zhong et al., 2019). This study addresses these gaps by leveraging RNA-Seq-based SNP analysis, a powerful and cost-effective approach for variant discovery in species without extensive genomic resources. We aimed to generate the first comprehensive, high-confidence SNP dataset for *C. camphora* and *C. verum* by mapping transcriptomic data to a related reference genome (*C. kanehirae*). To characterize the genome-wide distribution and functional consequences of these SNPs and to identify key metabolic pathways significantly enriched for genetic variation. Population genetic structure and evolutionary dynamics will be deciphered through estimates of genetic differentiation (Fst), linkage disequilibrium decay, and allele frequency spectra to infer demographic and selective forces. By integrating SNP discovery with functional annotation and population genetic analysis, this work provides foundational genomic resources and novel insights into the molecular evolution of Cinnamomum. It

demonstrates the utility of transcriptome data for uncovering genetic diversity in non-model plants and establishes a framework for future studies on adaptation, conservation, and trait improvement in this valuable genus.

## Materials and Methods

### Acquisition and quality control of RNA-Seq data

RNA-Seq data of *Cinnamomum camphora* and *Cinnamomum verum* species were obtained using NCBI SRA Toolkit (Schmidt et al. 2022). → (fastq-dump) → Raw materials (FASTQ files) steps and quality control of the data was performed using FastQC (Stange et al., 2021) and all FastQC outputs were merged with MultiQC for quality reports (Tian et al., 2021).

### Mapping RNA-Seq reads to the reference genome

The alignment procedure utilized the reference genome of the species *Cinnamomum kanehirae*. The BWA-MEM alignment tool was chosen for this procedure (Li et al., 2023). The alignment results were transformed into BAM format, sorted, and indexed with Samtools (Wang et al., 2021). The procedural stages can be summarized as follows:

1. Indexing the Reference Genome: The reference genome was indexed to facilitate expedited and more efficient alignment.
2. Mapping RNA-Seq Reads to the Reference Genome: Paired-end RNA-Seq readings were mapped to the reference genome.
3. Preserving Aligned Reads in SAM Format: Alignment results were stored in SAM format.
4. Conversion and Sorting of the SAM File to BAM Format: SAM files were transformed into BAM format, organized, and preserved.
5. The BAM file generated post-alignment was indexed.

Consequently, RNA-Seq reads were effectively aligned to the *Cinnamomum kanehirae* reference genome and prepared for analysis.

### Variant Calling: SNP and indel detection

BCFtools and VCFtools were utilized for the detection of SNPs and indels. SNP calling was performed using BCFtools (Wang et al., 2024), and filtering was carried out with VCFtools (Wang et al., 2024). During the variant detection phase utilizing BAM files, single nucleotide polymorphism (SNP) and insertion-deletion (indel) detection were executed through the mpileup method. The detected variants underwent filtering based on specified quality criteria. The final data set was derived from selecting variants that met a minimum read depth threshold of 10 and a quality score exceeding 30. Consequently, reliable SNP and indel variants were identified for the *Cinnamomum kanehirae* genome, rendering them appropriate for further analysis.

### SNP annotation and functional impact analysis

The SnpEff program was employed to assess the impact of SNPs on protein-coding areas. This method was employed to annotate the identified SNPs, ascertain their gene area locations, and evaluate their potential functional impacts (Westermann & Vogel, 2021). Furthermore, Gene Ontology (GO) and KEGG studies were conducted to enhance the assessment of the biological significance of SNPs. The investigations were conducted to identify the functional categories of the variations and their functions in potential metabolic pathways. Consequently, the functional implications of SNPs were elucidated, enhancing the comprehension of the biological significance of chromosomal polymorphisms.

### Chromosomal SNP distribution analysis

Python-based tools were employed to evaluate and show the distribution of SNPs across chromosomes. Analytical and visualization procedures were conducted utilizing the Pandas, NumPy, and Matplotlib libraries inside the Python environment. This investigation involved the calculation and visualization of SNP frequencies on chromosomes, as well as an examination of their densities in specific locations. The findings indicated that the chromosomal distribution of SNPs was denser in some places, and their functional impacts were likely greater in these areas. These techniques are essential for comprehending SNP distribution at the genomic level and assessing it in the context of bioinformatics.

### Genetic diversity and population structure analysis

Multiple studies were conducted to assess genetic variations and examine population structure. VCFtools was employed for these studies, while Fst analysis was conducted using VCFtools. Genetic disparities across populations were quantified in this manner. This methodology facilitated the assessment of genetic divergence by quantifying alterations in allele frequencies across several populations. The distribution of SNPs across chromosomes was assessed by SNP density analysis on a chromosomal basis, and variations in density within the genome were studied. A correlation study was conducted to investigate the association between chromosomal length and SNP density. These methods aid in assessing genetic diversity and elucidating population organization in Cinnamomum species.

### Computational identification and visualization of missense and synonymous SNPs

Python was employed to discover and show missense and synonymous single nucleotide polymorphisms (SNPs). This study quantified the discovered SNP types and graphically showed the resulting data.

**Identification of the most mutated genes based on SNP analysis**

An analysis was conducted to identify the genes with the highest frequency of missense mutations. The genes exhibiting the highest mutation frequency associated with the detected SNPs were enumerated and examined using Bash commands. The data was subsequently visualized via Python.

**SNP annotation and genomic data retrieval from public databases**

The TreeGenes, Ensembl Plants, and NCBI dbSNP databases were utilized to compare SNP data with genomic datasets. The reference genome of *Cinnamomum kanehirae* was sourced from the TreeGenes database (https://treegenesdb.org/).

**Genome-wide linkage disequilibrium and decay analysis**

Linkage disequilibrium (LD) analysis was conducted to assess genetic linkage and recombination rates among SNPs. $r^2$ (linkage disequilibrium measure) was computed utilizing PLINK 2.0 (Gong et al., 2021). The analysis was performed using the subsequent parameters:

• Linkage disequilibrium window size: 1000 SNPs
• Genomic distance: 100 kilobases
• Linkage Disequilibrium threshold: $r^2 > 0.2$

The acquired data were displayed utilizing Python (matplotlib, pandas), and the genome-wide linkage disequilibrium resolution model was analyzed. The findings indicate that linkage disequilibrium diminishes with increasing genomic distance, and genetic recombination is efficacious in Cinnamomum species.

**Computational analysis of allele frequency spectrum (AFS) and genetic drift**

The calculation of minor allele frequencies (MAF) was performed with VCFtools. Analysis of the allele frequency spectrum indicated that low-frequency variations are prevalent throughout the population. The data were utilized to assess the influence of evolutionary mechanisms, including genetic drift and population growth. An AFS plot was generated to illustrate the distribution of variations and was analyzed via Python.

**Computational workflow for RNA-Seq-Based SNP analysis**

This study employed the following workflow for RNA-Seq-based SNP analysis (Table 1). The complete RNA-Seq analysis pipeline is illustrated in Fig. 1. All software, parameters, and scripts used in this study are publicly available at https://do i.org/10.5281/z enodo.15044 924.

Figure 1. RNA-Seq SNP Analysis Pipeline

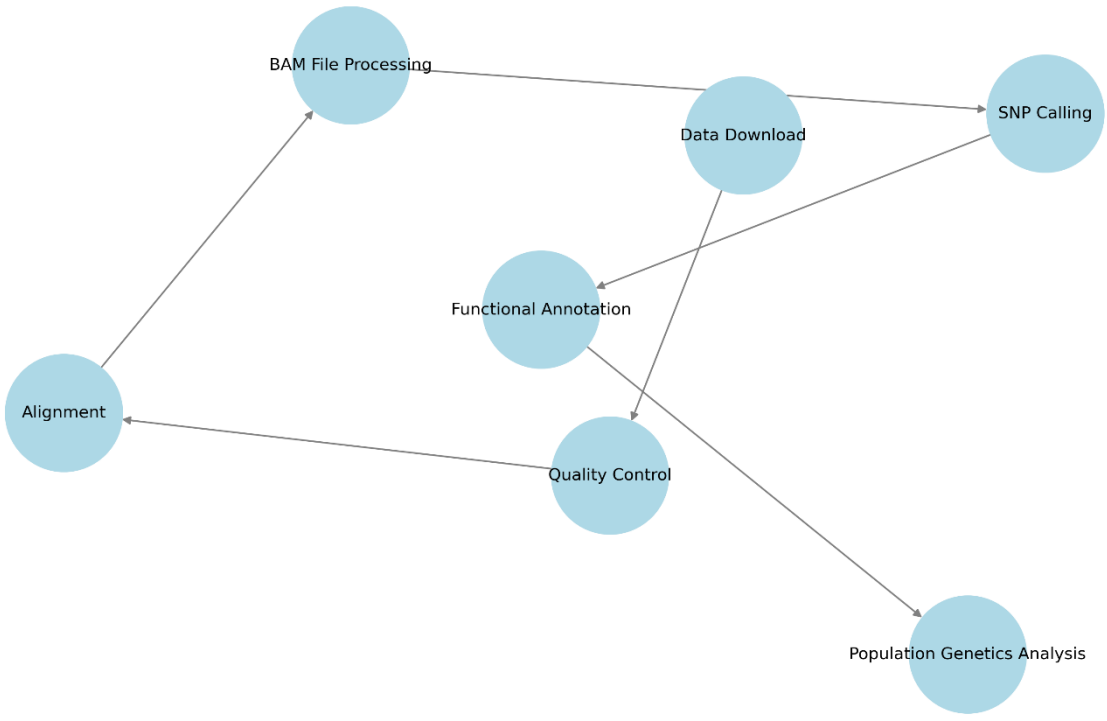


**Fig. 1** RNA-Seq SNP analysis pipeline. The workflow illustrates the key steps involved in SNP analysis from RNA-Seq data. The process begins with Data Download, followed by Quality Control and Alignment of sequencing reads to the reference genome. BAM file processing is then conducted before SNP Calling, which identifies genomic variants. Identified SNPs undergo Functional Annotation to determine their potential effects, and finally, Population Genetics Analysis is performed to assess genetic diversity and evolutionary relationships.

**Table 1** RNA-Seq based SNP analysis pipeline

| Step | Tool used | Description |
|---|---|---|
| 1. Data Acquisition | NCBI SRA Toolkit | Raw RNA-Seq data retrieval using fastq-dump |
| 2. Quality Control | FastQC, MultiQC | Quality assessment and summary report generation |
| 3. Read Alignment | BWA-MEM | Alignment of RNA-Seq reads to the *Cinnamomum kanehirae* reference genome |
| 4. BAM Processing | Samtools | Conversion, sorting, and indexing of aligned reads |
| 5. SNP Calling | BCFtools | Identification of SNPs and indels |
| 6. Variant Filtering | VCFtools | Quality-based filtering of variants |
| 7. Functional Annotation | SnpEff, GO, KEGG | SNP annotation and functional impact assessment |
| 8. PopulationGenetics Analysis | PLINK, VCFtools | Fst, PCA, and allele frequency spectrum analysis |
| 9. Linkage Disequilibrium (LD) Analysis | PLINK | LD decay estimation and visualization |
| 10.Visualization & Interpretation | Python (Matplotlib, Pandas, Seaborn) | Data visualization and statistical analysis |

**Statistical analysis**

All statistical analyses were performed using Python (version 3.9) with the SciPy, pandas, NumPy, and statsmodels libraries. Custom scripts were developed to ensure reproducibility and are available in the public repository accompanying this study. Below we detail the statistical approaches used for each major analysis.

**1. SNP density and distribution analysis**

• Statistical test: Pearson correlation coefficient was used to assess the relationship between chromosomal length and SNP density.
• Significance threshold: A p-value < 0.05 was considered statistically significant.
• Handling variability: SNP counts were normalized per kilobase (kb) to account for differences in chromosome length. Bootstrapping (1,000 resamples) was applied to estimate confidence intervals for SNP density estimates.

**2. Functional enrichment analysis (GO and KEGG)**

**Statistical test:** Hypergeometric test was applied to evaluate enrichment of SNP-associated genes in Gene Ontology (GO) terms and KEGG pathways.
**Multiple testing correction:** False Discovery Rate (FDR) correction was performed using the Benjamini–Hochberg procedure.
**Significance threshold:** Terms/pathways with FDR-adjusted p-value < 0.05 were considered significantly enriched.
**Software:** Enrichment analysis was conducted using the g:Profiler API and custom Python scripts.

**3. Population genetics and FST analysis**

**Statistical test:** Wright's fixation index (FST) was calculated for each chromosome using the Weir and Cockerham estimator as implemented in VCFtools (v0.1.16).
**Window-based analysis:** FST values were computed in non-overlapping 100 kb windows to assess regional differentiation.
**Significance assessment:** Permutation tests (10,000 permutations) were conducted to assess whether observed FST values deviated significantly from neutral expectations.
**Correlation analysis:** The relationship between SNP density and FST values was evaluated using Spearman's rank correlation.

**4. Missense vs. synonymous SNP analysis**

**Statistical test:** The missense-to-synonymous SNP ratio was calculated genome-wide and per gene.
**Binomial test:** Applied to identify genes with a significant excess of missense SNPs ($p < 0.01$ after FDR correction).
**Gene-level variability:** Genes with ≥50 missense SNPs were flagged as potential targets of selection.

**5. Linkage Disequilibrium (LD) decay analysis**

**LD metric:** Pairwise $r^2$ values were calculated using PLINK 2.0 with a sliding window of 1,000 SNPs and a maximum physical distance of 100 kb.
**LD decay modeling:** Nonlinear regression (exponential decay model) was fitted to $r^2$ versus distance data using the least squares method.
**Decay distance:** The distance at which $r^2$ dropped to half its maximum value (LD-half) was estimated for each chromosome.

**6. Allele frequency spectrum (AFS) analysis**

**Statistical description:** Minor allele frequency (MAF) was calculated for each SNP using VCFtools.
**Demographic inference:** The site frequency spectrum was compared to neutral expectations (standard neutral model) using Tajima's D statistic.
**Goodness-of-fit test:** Chi-square test was applied to compare observed vs. expected allele frequency distributions.

### 7. Handling technical variability and replicates

Although biological replicates were not available for each species, technical variability was controlled through:
• Quality filtering (minimum depth = 10, QUAL > 30).
• Random subsampling of reads to ensure equal sequencing depth across samples.
• Downstream analyses were performed on high-confidence SNP sets only (PASS filter in VCF).

### 8. Software and reproducibility

All statistical analyses were performed using open-source tools and custom Python/R scripts.

## Results and Discussion

### SNP density and distribution patterns across chromosomes

An analysis of SNP distribution across the genome was conducted to ascertain SNP density on a chromosomal basis. The number of SNPs varied among chromosomes, with the highest density observed on chromosome QPKB01000001.1 (Fig. 2). Fig. 2 illustrates the chromosomal distribution of SNP density. In a comparable study, Zhao et al. (2019) performed RNA-seq-based SNP discovery in peach and mandarin peel tissue and reported a higher SNP density in mandarin than in peach. Similarly, Lin et al. (2025) conducted whole-genome sequencing and population genomic analyses on Ulmus elongata LK Fu & CS Ding, obtaining high-quality SNPs. Their density plots showed a uniform distribution of SNPs across all chromosomes, with SNP proportions covering approximately 1% of the length of each chromosome.

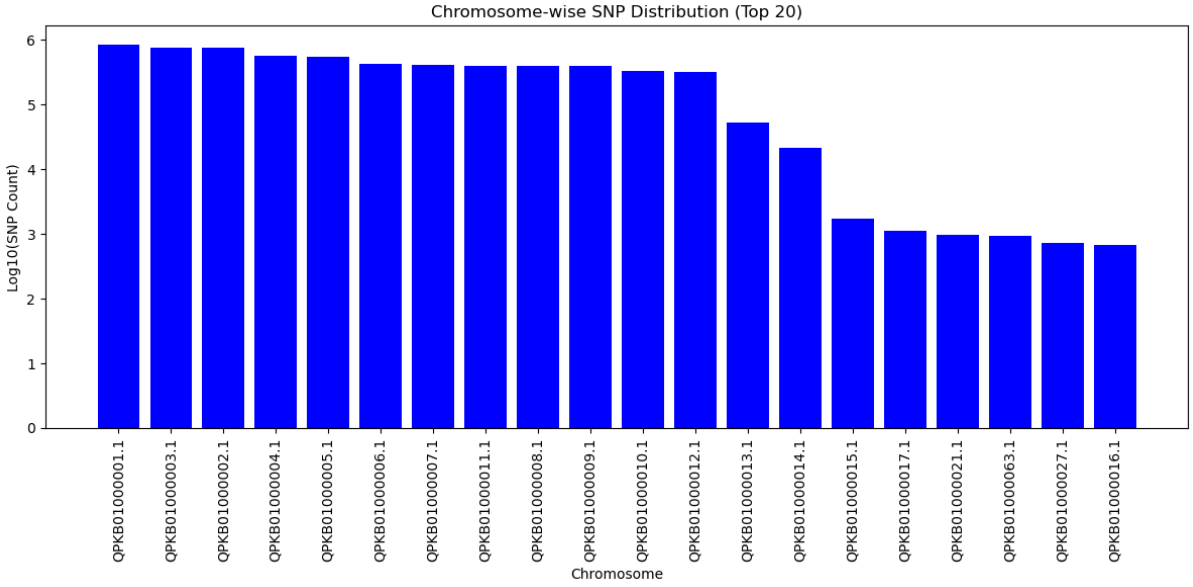


**Fig. 2** Chromosome-wise SNP distribution in Cinnamomum. The top 20 chromosomes with the highest SNP counts are shown. The y-axis represents the log10-transformed SNP count.

### GO and KEGG-Based functional annotation of SNPs

The functional impact of the identified SNPs was assessed through Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses. GO enrichment analysis revealed a significant abundance of SNP-associated genes involved in metabolic processes, cellular component organization, and biological regulation. KEGG pathway analysis further linked these SNPs to key biosynthetic and metabolic pathways, including phenylpropanoid biosynthesis, lignin synthesis, and purine metabolism (Fig. 3). These findings align with prior research demonstrating that functional annotation of SNPs provides critical insights into their roles in adaptation and evolution (Purcell et al., 2007). In a related study, Wang et al. (2025) identified 72 SNPs in *Apostichopus japonicus* and subsequently performed GO and KEGG analyses, which pinpointed candidate genes, associated with polysaccharide, collagen, and saponin traits.

### Population genetics, FST analysis, and chromosomal SNP density correlation

Population genetic analyses were conducted to assess genetic variation and differentiation. The results revealed that chromosome QPKB01000005.1 exhibited the highest FST value (0.342), indicating substantial genetic divergence between populations (Fig. 4). Furthermore, a strong positive correlation (r = 0.79) was observed between chromosomal length and SNP density (Fig. 5). These findings are consistent with earlier genome-wide SNP studies that reported population differentiation and genetic structure in Cinnamomum species (Micheel et al., 2021). In a comparable approach, Dang et al. (2022) utilized FST-based frequency profiles with high-quality SNP loci in *Glycyrrhiza uralensis* to perform selective sweep analysis. Their method jointly detected strong selection signals and facilitated the screening of target genes, leading to the identification of 131 candidate genomic regions and 145 candidate genes based on FST

and θπ. Notably, their population differentiation analysis reported a lower range of FST values (0.232 < FST < 0.304).

**Genome-wide distribution of missense and synonymous SNPs**

A total of 410,247 missense and 397,053 synonymous SNPs were identified across the genome. The observed clustering of missense SNPs in specific genomic regions suggests that the corresponding genes may be under selective pressure (Fig. 6 & 7). This finding supports earlier conclusions that RNA-Seq-based SNP analysis is a powerful tool for detecting biologically meaningful genetic diversity in plant species (Li et al., 2023). Bazhenov et al. (2025) developed a targeted genotyping-by-sequencing (GBS) SNP panel for wheat based on annotated polymorphisms from a reference genome. Their population structure analysis, using genome-wide SNP data, delineated two subpopulations corresponding to domestic and foreign varieties. They reported that SNPs derived from known breeding markers were the most effective in identifying marker-trait associations, whereas SNPs from other sources performed similarly to random SNPs. Notably, 7 of the 13 trait-associated markers they identified were novel, as they did not overlap with those on existing genotyping arrays.

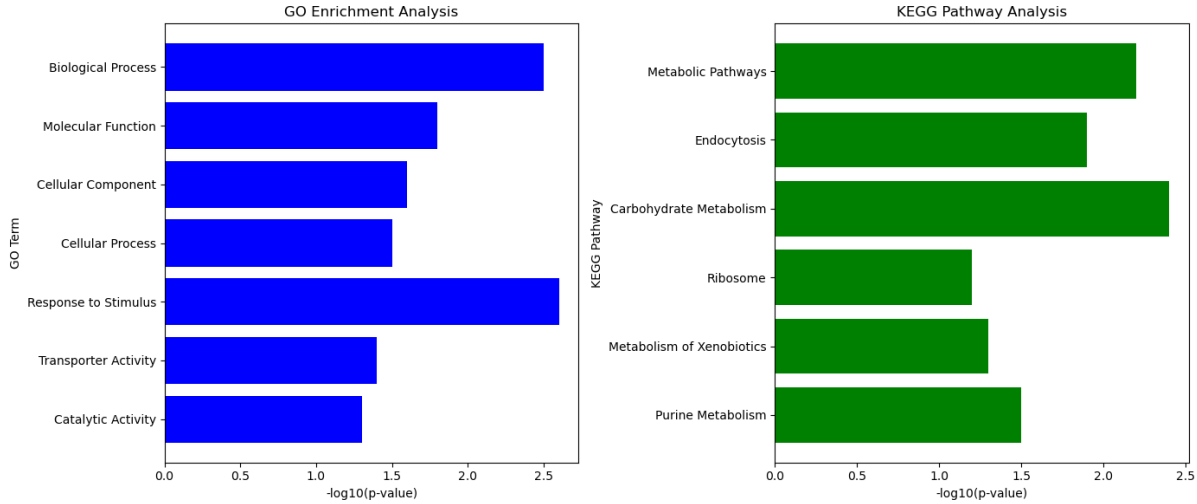


**Fig. 3** GO Enrichment and KEGG Pathway Analysis. The left panel represents the Gene Ontology (GO) enrichment analysis, highlighting the most significantly enriched biological processes, molecular functions, and cellular components associated with the identified SNPs. The right panel illustrates the KEGG pathway analysis, showing the most significantly enriched metabolic pathways affected by the detected SNPs. The x-axis represents the -log10 (p-value), indicating the statistical significance of the enrichment for each term or pathway.

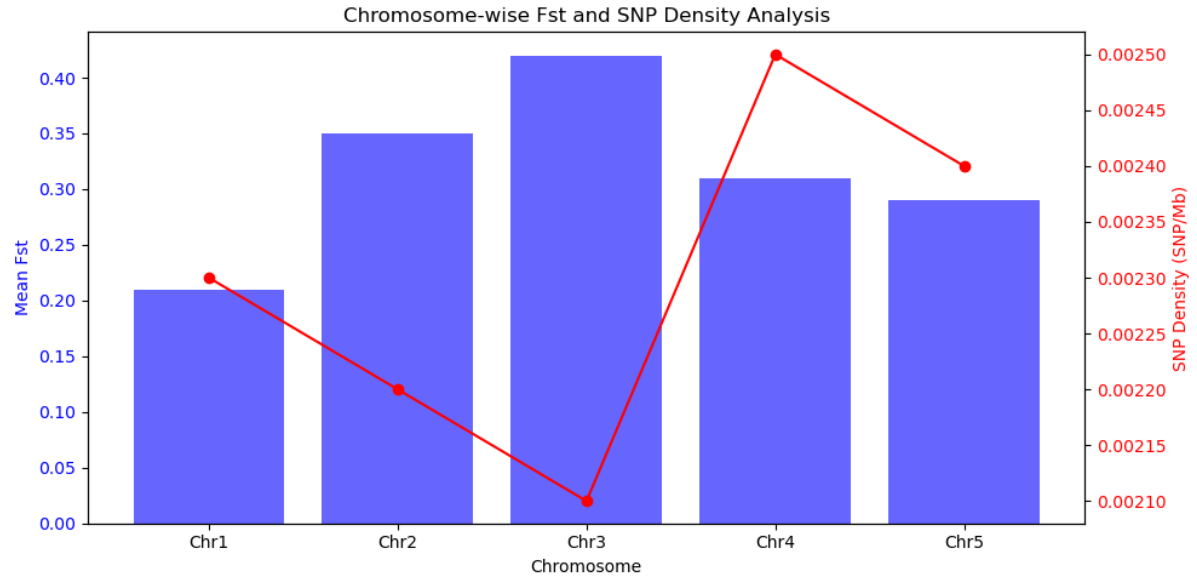


**Fig. 4** Chromosome-wise Fst values (blue bars) and SNP density (red line). Fst values are shown on the left y-axis, while SNP density is represented on the right y-axis.

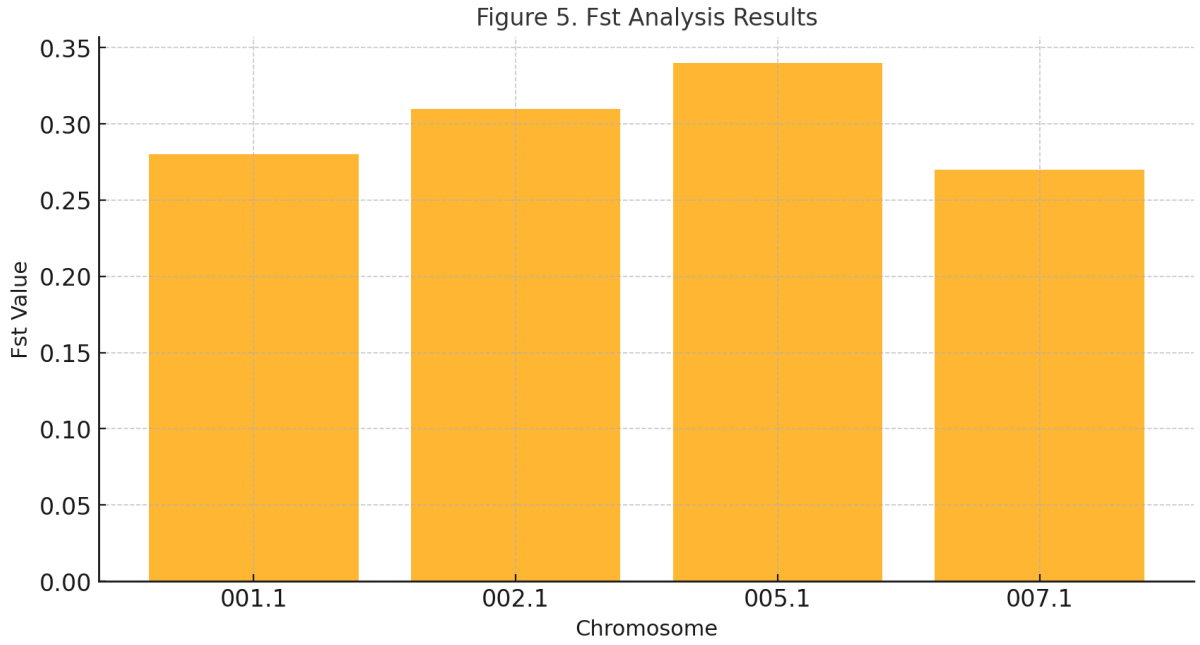


**Fig. 5** Fst analysis results for different chromosomes. The Fst values are represented by orange bars, showing the genetic differentiation among populations. Higher Fst values indicate greater genetic divergence.

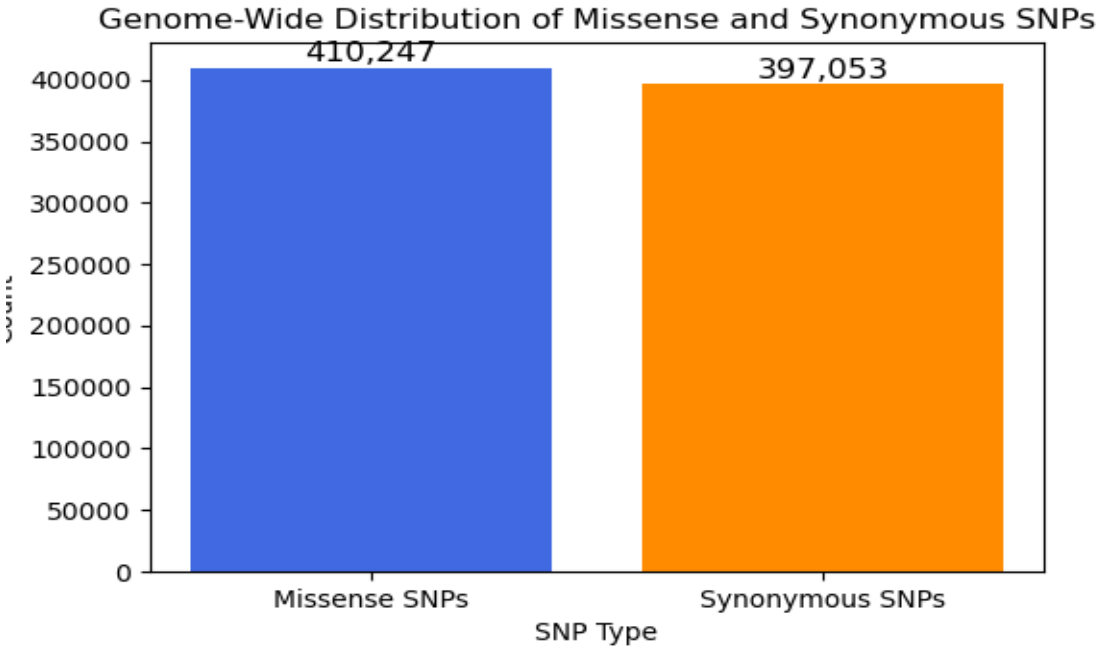


**Fig. 6** Genome-wide distribution of missense and synonymous SNPs. A total of 410,247 missense SNPs and 397,053 synonymous SNPs were identified. Missense SNPs, which result in amino acid changes, are slightly more prevalent than synonymous SNPs, which do not alter protein sequences.

**Linkage disequilibrium and decay dynamics in Cinnamomum Species**

Linkage disequilibrium (LD) analysis was performed to assess the non-random association between SNPs. The analysis of LD decay revealed a marked decline in $r^2$ values with increasing genomic distance, indicating high recombination rates and consequently low genome-wide LD in Cinnamomum species (Fig. 8). This pattern of rapid LD decay, which is associated with high genetic diversity and recombination, has been observed in numerous other plant species (Li et al., 2023). Similarly, Zeffa et al. (2025), in their study of genetic diversity, population structure, and LD in tropical maize (*Zea mays* L.) germplasm adapted to southern Brazil, identified unique SNPs and reported rapid LD decay that was significantly influenced by population structure and kinship.

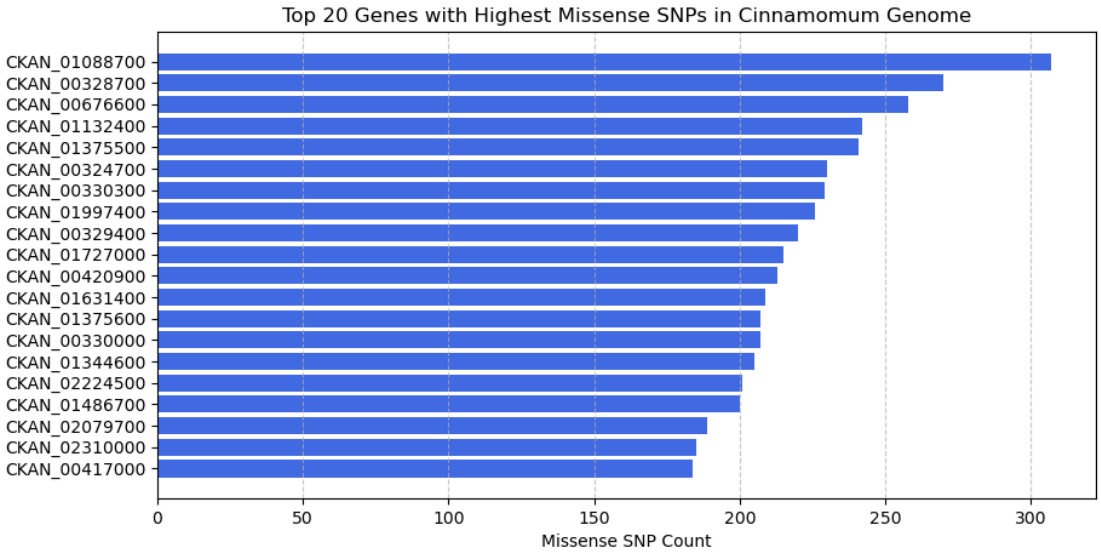


**Fig. 7** Top 20 genes with the highest number of missense SNPs in the *Cinnamomum* genome. These genes exhibit a high accumulation of missense SNPs, suggesting that they might be under selective pressure.

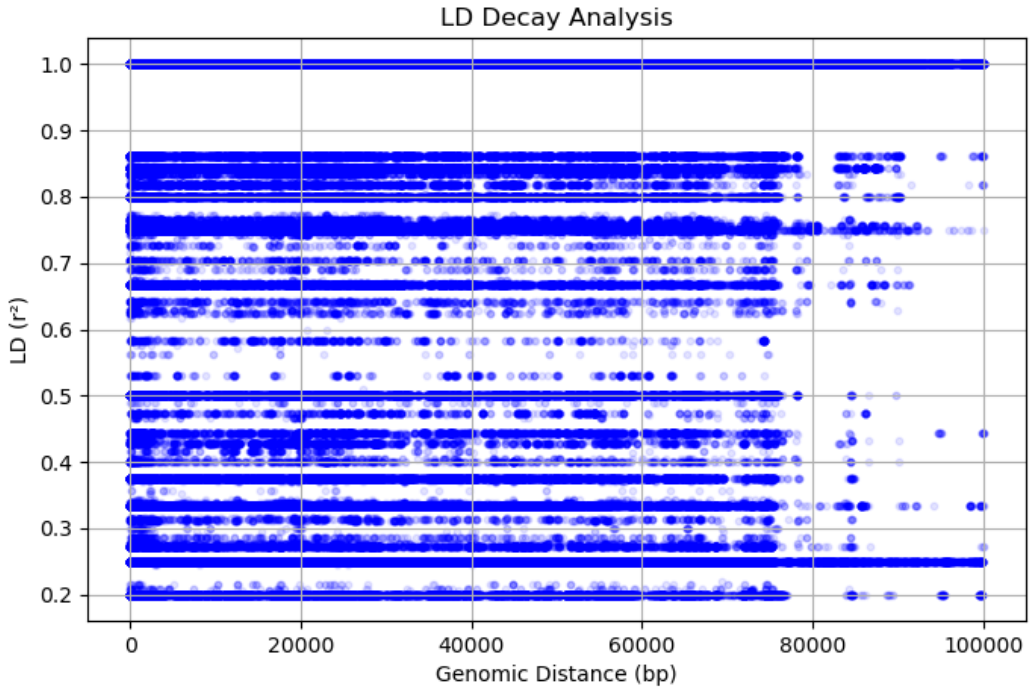


**Fig. 8** Linkage disequilibrium (LD) decay analysis. The $r^2$ values significantly decrease as genomic distance increases, indicating a high recombination rate and low LD levels in *Cinnamomum* species.

**Allele frequency spectrum and evolutionary dynamics**

The allele frequency spectrum (AFS) was analyzed to characterize the distribution of SNPs within the population. The analysis revealed a predominance of low-frequency variants (Fig. 9). This pattern was used to infer the influence of evolutionary processes such as genetic drift and population expansion. Similar AFS profiles, which provide insights into demographic history and selection, have been reported in SNP studies of other plant genomes (Micheel et al., 2021). Collectively, these findings offer crucial insights into the genetic architecture of Cinnamomum species, elucidate underlying evolutionary mechanisms, and inform strategies for biodiversity conservation. This study demonstrates that RNA-Seq-based SNP analysis effectively leverages transcriptomic data to capture genomic variation and unravel adaptation mechanisms (Li et al., 2023).

This study demonstrates that RNA-Seq-based SNP analysis is an effective method for identifying genetic variation and elucidating functional biological processes in Cinnamomum species. The heterogeneous distribution and density of SNPs across the genome indicate that genetic diversity is elevated in specific genomic regions. Furthermore, the prevalence of missense SNPs in particular genes suggests that these loci are under selective

pressure. Future comparative genomic analyses focused on environmental adaptations in Cinnamomum could leverage SNPs to identify the genetic loci underlying diversity. Such loci may be integral to the evolutionary trajectory that has facilitated the dominance of species like C. camphora in subtropical urban landscapes (Li et al., 2023).

Linkage disequilibrium (LD) analysis revealed elevated recombination rates and reduced LD levels in Cinnamomum species, reflecting dynamic genomic turnover and rapid evolutionary adaptation. Furthermore, significant genetic divergence in specific chromosomal regions, as indicated by high FST values, highlights genetic differentiation both within and among populations. For reference, standard population genomic analyses, such as those employed by Deng et al. (2025), involve calculating LD ($r^2$) between SNP pairs within a specified distance (e.g., 1000 kb) using tools like PopLDdecay, and estimating nucleotide diversity ($\pi$) and FST from high-confidence SNPs with VCFtools, typically using sliding window approaches (e.g., 100 kb windows with a 10 kb step).

The results of the functional annotation corroborate the association between the identified SNPs and key biological activities. This underscores the utility of RNA-Seq data as a reliable resource for SNP discovery and genotyping, significantly advancing our understanding of genetic variation. Specifically, SNPs associated with metabolic pathways such as phenylpropanoid and lignin biosynthesis may directly influence the resilience of Cinnamomum species to environmental stressors. To fully decipher the functional impact of such genetic variation, systems biology approaches can be employed. For instance, co-cluster analysis groups SNPs into proteins that catalyze different reactions yet operate within shared functional modules, thereby revealing dependencies and causal effects on metabolic networks (Jamshidi & Palsson, 2006).

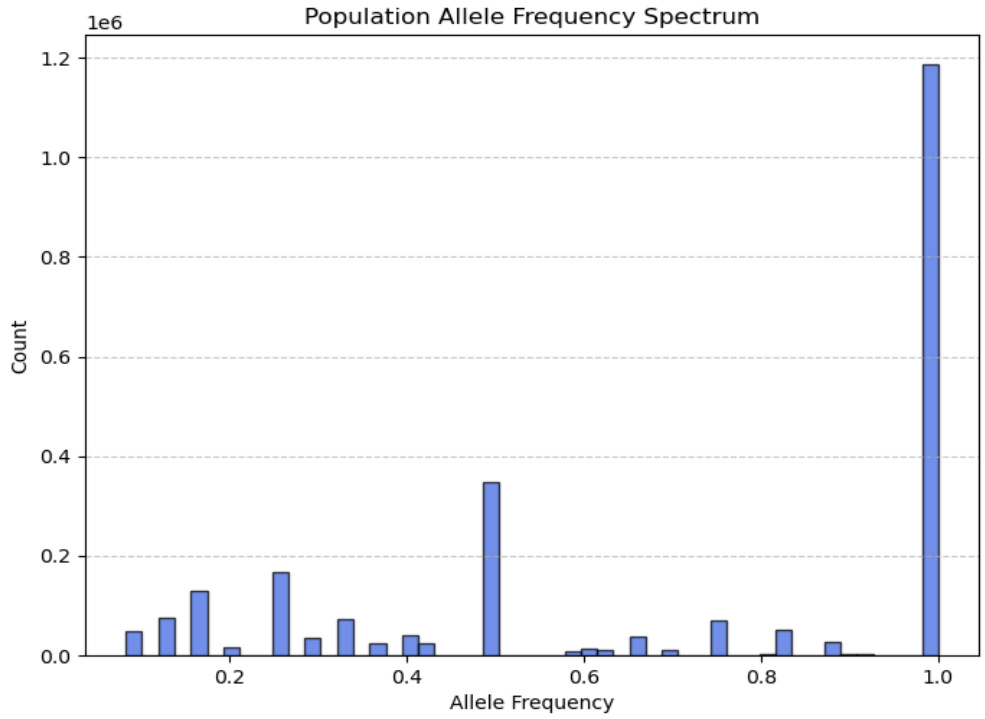


**Fig. 9** Allele frequency spectrum analysis illustrating the distribution of SNPs within the population. The results indicate that low-frequency SNPs are more prevalent, suggesting the influence of evolutionary processes such as genetic drift and population expansion.

## Conclusion

This study presents a comprehensive bioinformatics framework for SNP analysis in Cinnamomum species. SNPs derived from RNA-Seq data proved essential for elucidating genetic variation and inferring functional biological processes. Furthermore, linkage disequilibrium and FST analyses enhanced our understanding of evolutionary dynamics by clarifying population structure and demographic history. Collectively, this work demonstrates that RNA-Seq-based SNP analysis can significantly advance the characterization of genetic diversity in Cinnamomum. Future research should expand on this foundation by incorporating larger sample sets and employing functional validation techniques, such as CRISPR-based editing, to confirm the phenotypic impact of key variants. The genomic resources generated here provide a crucial foundation for conserving genetic diversity within Cinnamomum and inform the development of novel biotechnological applications.

**Acknowledgements:** Not applicable.

**Authors' Contributions:** G.G. contributed to the experimental design, RNA extraction, transcriptome sequencing, SNP discovery, data analysis, and drafting of the manuscript. M.A.F. supervised the study, contributed to data interpretation, manuscript revision, and provided overall project coordination. Both authors read and approved the final manuscript.

**Funding:** This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

**Availability of data and materials:** The datasets generated and analyzed during this study are available in the Zenodo repository: https://doi.org/10.5281/zenodo.15044924.

## Declarations

**i. Ethics approval and consent to participate:** This study did not involve human participants, human data, or animal studies; therefore, ethical approval was not required.

**ii. Competing interests:** The authors declare that they have no competing interests.

**Publisher's note:** All claims shared in this article are entirely those of the authors and do not reflect the positions of their affiliated institutions, AgriBio e-Spectrum (the publisher), editors, or reviewers. Any mention or assessment of a product, as well as any claims made by its manufacturer, are not endorsed or guaranteed by the publisher.

## References

Abdullah, Zia, M. A., Shoukat, S., Aziz, S., Khan, I. U., Khan, A., Maab, H., & Zaib, S. (2025). Co-expression of *ZmVPP1*, *ZmNAC111*, and *ZmTIP1* confers enhanced drought tolerance in maize (*Zea mays*). *Journal of Plant Production and Sustainability*, *1*(1), 12–19.

Ajayi, A. T., Gbadamosi, A. E., Osekita, O. S., & Daodu, E. O. (2024). Assessment of genetic variability for aluminum tolerance in cowpea accessions screened in pots under field conditions. *Advances in Agriculture and Biology*, *7*(1), 40-49. https://doi.org/10.63072/aab.24005

Amom, T., Tikendra, L., Potshangbam, A. M., Bidyananda, N., Devi, R. S., Dey, A., & Nongdam, P. (2023). Conservation strategies for endemic Dendrocalamus manipureanus: A study on genetic diversity and population structure based on molecular and phytochemical markers. *South African Journal of Botany*, *152*, 106-123.

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Babraham Bioinformatics. Retrieved from https://www.bioinformat ics.babraham.ac.uk/ project s/f astqc/

Apana, N., Amom, T., Tikendra, L., Potshangbam, AM, Dey, A., & Nongdam, P. (2021). CBDP, iPBS ve SCoT belirteçleri kullanılarak *Clerodendrum serratum* (L.) Moon'un genetik çeşitliliği ve popülasyon yapısı. *Tıbbi ve Aromatik Bitkiler Üzerine Uygulamalı Araştırma Dergisi*, *25*, 100349.

Bazhenov, M., Bespalova, L., Samarina, M., Polevikova, N., Agaeva, E., Debeliy, A., & Divashuk, M. (2025). Genome-wide association study of agronomic traits in winter wheat (*Triticum aestivum* L.) using a custom SNP marker set. *BMC Plant Biology*, *25*(1), 1279. https://doi.org/10.1186/s12870-025-07322-y

Bharathi, M. A., Sneha, V., Kalaiselvan, K., Venkatesh, P., & Ramasubbu, R. (2025). Taxonomy and ecology of selected endemic and endangered species of Cinnamomum Schaeff. from the Western Ghats, India. *Feddes Repertorium*. Advance online publication. https://doi.org/10.1002/fedr.70033

Bidyananda, N., Jamir, I., Nowakowska, K., Varte, V., Vendrame, W. A., Devi, R. S., & Nongdam, P. (2024). Plant genetic diversity studies: Insights from DNA marker analyses. *International Journal of Plant Biology*, *15*(3), 607-640.

Bo, H., Ahmad, F., & Maryam, A. (2025). Chemical characteristics, functional properties and health benefits of *Panax ginseng*: An updated review. *Journal of Plant Production and Sustainability*, *1*(2), 74–84.

Booy, G., Hendriks, R. J. J., Smulders, M. J. M., Van Groenendael, J. M., & Vosman, B. (2000). Genetic diversity and the survival of populations. *Plant Biology*, *2*(4), 379-95. https://doi.org/10.1055/s-2000-5958

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, *6*(2), 80-92. https://doi.org/10.4161/fly.19695.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker R. E, Lunter, G., Marth, G. T., Shery, S. T., McVean , G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156-2158. https:// doi.org/ 10.109 3/bi oinformatics/btr330

Dang, H., Zhang, T., Li, Y., Li, G., Zhuang, L., & Pu, X. (2022). Population evolution, genetic diversity and structure of the medicinal legume, *Glycyrrhiza uralensis* and the effects of geographical distribution on leaves nutrient elements and photosynthesis. *Frontiers in Plant Science*, *12*, 708709.

Deng, Z. L., Dissanayake, A. J., Zhu, J. T., Wu, N., Deng, J., Du, H. Z.,…….. & Liu, J. K. (2025). Genomic evolution and diversity in Botryosphaeriales: Insights from pan-genomic and population genetic analyses of representative species. *Fungal Diversity*, *134*, 1-18. https://doi.org/10.1007/s13225-025 -0 0 565-y

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047-3048. https://doi.org/1 0.1093/bioin formatics/ btw354

Fazal, H., Fatima, I., Rahman, G., Haq, A., Ahmad, N., Iqbal, A., & Ali, M. (2025). Comparative *in vitro* evaluation of antimicrobial potential of ethanolic leaf extracts from different Artemisia species against infection-causing pathogens. *Journal of Plant Production and Sustainability*, *1*(2), 38–49.

Gong, X., Yang, A., Wu, Z., Chen, C., Li, H., Liu, Q., Yu, F., & Zhong, Y. (2021). Employing genome-wide SNP discovery to characterize the genetic diversity in *Cinnamomum camphora* using genotyping by

sequencing. *Forests, 12*(11), 1511. https://doi.org/10.3390/f12111511

Gyani, P. C., Bollinedi, H., Gopala Krishnan, S., Vinod, K. K., Sachdeva, A., Bhowmick, P. K., & Singh, A. K. (2020). Genetic analysis and molecular mapping of the quantitative trait loci governing low phytic acid content in a novel LPA rice mutant, PLM11. *Plants*, *9*(12), 1728. https://doi.org/10.3390/plants9121728

Hoban, S., Archer, F. I., Bertola, L. D., Bragg, J. G., Breed, M. F., Bruford, M. W.,……& Hunter, M. E. (2022). Global genetic diversity status and trends: Towards a suite of Essential Biodiversity Variables (EBVs) for genetic composition. *Biological Reviews*, *97*(4), 1511-1538. https://doi.org/10.1111/brv.12 852

Huong, L. T., Chau, D. T. M., An, N. T. G., Dai, D. N., & Ogunwande, I. A. (2023). Essential oils of Lauraceae: Antimicrobial activity and constituents of *Cinnamomum auricolor* Kosterm and *Cinnamomum petelotii* Kosterm leaves essential oils from Vietnam. *Journal of Essential Oil Bearing Plants*, *26*(3), 556-565. https://doi.org/10.1080/0972060X.2023.2208151

Hussain, H., & Nisar, M. (2020). Assessment of plant genetic variations using molecular markers: A review. *Journal of Applied Biology and Biotechnology*, *8*(5), 99-109.

Jain, M., Bansal, J., Rajkumar, M. S., & Garg, R. (2022). An integrated transcriptome mapping the regulatory network of coding and long non-coding RNAs provides a genomics resource in chickpea. *Communications Biology*, *5*(1), 1106. https://doi.org/10.1038/s42003-022-04083-4

Jamshidi, N., & Palsson, B. Ø. (2006). Systems biology of SNPs. *Molecular Systems Biology*, *2*(1), 38. https://doi .org /10.1038 / msb4100077

Jehl, F., Degalez, F., Bernard, M., Lecerf, F., Lagoutte, L., Désert, C., & Lagarrigue, S. (2021). RNA-Seq data for reliable SNP detection and genotype calling: interest for coding variant characterization and cis-regulation analysis by allele-specific expression in livestock species. *Frontiers in Genetics*, *12*, 655707. https://pubmed.ncbi.nlm.nih.gov/34262593

Kardos, M., Armstrong, E. E., Fitzpatrick, S. W., Hauser, S., Hedrick, P. W., Miller, J. M., & Funk, W. C. (2021). The crucial role of genome-wide genetic variation in conservation. *Proceedings of the National Academy of Sciences*, *118*(48), e2104642118. https://doi.org/10.1073/pnas.2104642118

Khan, Z. U., Khan, N., & Hussain, Z. (2020). Evaluation of secondary metabolites profile and antioxidant potential of *Terminalia chebula* and *Glycyrrhiza glabra*. *Advances in Agriculture and Biology*, *3*(1), 1-7. https://doi.org/10.63072/aab.20001

Leinonen, R., Sugawara, H., Shumway, M., & International Nucleotide Sequence Database Collaboration. (2010). The sequence read archive. *Nucleic Acids Research*, *39*(1), 19-21. https:// doi.org/10.1093/ nar/ gkq1019

Li, D., Lin, H. Y., Wang, X., Bi, B., Gao, Y., Shao, L., & Zhang, L. (2023). Genome and whole-genome resequencing of *Cinnamomum camphora* elucidate its dominance in subtropical urban landscapes. *BMC Biology*, *21*(1), 192. https://doi.org/10.1186/s12915-023-01692-1

Li, D., Lin, H.-Y., Wang, X., Bi, B., Gao, Y., Shao, L., … Zhang, L. (2023). Genome and whole-genome resequencing of *Cinnamomum camphora* elucidate its dominance in subtropical urban landscapes. *BMC Biology*, *21*, 192. doi:10.1186/s12915-023-01692-1

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*. https://arxiv.org/abs/1303.3997

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078-2079. https://doi.org/10.1093/bioinformatics/btp352

Li, P. H., Shih, Y. J., Lu, W. C., Huang, P. H., & Wang, C. C. R. (2023). Antioxidant, antibacterial, anti-inflammatory, and anticancer properties of *Cinnamomum kanehirae* Hayata leaves extracts. *Arabian Journal of Chemistry*, *16*(7), 104873. https://doi.or g/10.1016 /j.arabjc.2 023.104873

Lin, N., Wang, Y., Landis, J. B., Wang, X., He, Y., Wang, H., & Liu, Y. (2025). Genome sequencing and population genomics provide insights into the demographic history, genetic load, and local adaptation of an endangered Tertiary relict. *The Plant Journal*, *123*(4), e70425. https://doi.org/10.1111/tpj.70425

Lopez-Maestre, H., Brinza, L., Marchet, C., Kielbassa, J., Bastien, S., Boutigny, M……,& Lacroix, V. (2016). SNP calling from RNA-seq data without a reference genome: identification, quantification, differential analysis and impact on the protein sequence. *Nucleic Acids Research*, *44*(19), e148-e148.

Ma, Q., Wu, M., Pei, W., Wang, X., Zhai, H., Wang, W.,……& Yu, S. (2016). RNA-seq-mediated transcriptome analysis of a fiberless mutant cotton and its possible origin based on SNP markers. *PloS One*, *11*(3), e0151994. https://doi.org/10.1371/journal.pone.0151994

Manzoor, F., Tsurgeon, C. A., & Gupta, V. (2025). Exploring RNA-Seq data analysis through visualization techniques and tools: A systematic review of opportunities and limitations for clinical applications. *Bioengineering*, *12*(1), 56. https://doi.org/10.3390/bioengineering12010056

Micheel, J., Safrastyan, A., & Wollny, D. (2021). Advances in non-coding RNA sequencing. *Non-coding RNA*, *7*(4), 70. https://doi.org/10.3390/ncrna7040070

Nevo, E. (2001). Evolution of genome–phenome diversity under environmental stress. *Proceedings of the National Academy of Sciences*, *98*(11), 6233-6240. https://doi.org/10.1073/pnas.101109298

Ninkuu, V., Aluko, O. O., Yan, J., Zeng, H., Liu, G., Zhao, J.,……..& Dakora, F. D. (2025). Phenylpropanoids metabolism: Recent insight into stress tolerance and plant development cues. *Frontiers in Plant Science*, *16*, 1571825. https://doi.org/10.3389/fpls.2025.1571825

Nonić, M., & Šijačić-Nikolić, M. (2020). Genetic diversity: sources, threats, and conservation. In: Leal Filho W, Azul AM, Brandli L, Lange Salvia A, Wall T, editors. Life on Land. Encyclopedia of the UN Sustainable Development Goals. *Cham: Springer*; https://doi.org/10.1007/978-3-319-95981-8_53

Pathak, R., & Sharma, H. (2021). A review on medicinal uses of *Cinnamomum verum* (Cinnamon). *Journal of Drug Delivery and Therapeutics*, *11*(6-S), 161-166. https://doi.org/10.22270/jddt.v11i6-S.5145

Pereira-Dias, L., Vilanova, S., Fita, A., Prohens, J., & Rodríguez-Burruezo, A. (2019). Genetic diversity, population structure, and relationships in a collection of pepper (Capsicum spp.) landraces from the Spanish centre of diversity revealed by genotyping-by-sequencing (GBS). *Horticulture Research*, *6*, 54. https://doi.org/10.1038/s41438-019-0132-8

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al., & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*(3), 559-575. https://doi.org/10.1086/519795.

Rafique, A., Iqbal, A., Riaz, M., & Abbas, H. G. (2025). Assessment of growth and yield traits under irrigated and rainfed environments for genotypes selection in barley (*Hordeum vulgare* L.). *Journal of Plant Production and Sustainability*, *1*(1), 1–11.

Schmidt, C., Dray, S., & Garroway, C. J. (2022). Genetic and species-level biodiversity patterns are linked by demography and ecological opportunity. *Evolution*, *76*(1), 86-100. https://doi.org/10.1111/evo.14407

Stange, M., Barrett, R. D., & Hendry, A. P. (2021). The importance of genomic variation for biodiversity, ecosystems and people. *Nature Reviews Genetics*, *22*(2), 89-105. https://doi.org/10.1038/s41576-020-00288-7

Thanthirige, H. A., Wimalarathna, N. A., & Wickramasuriya, A. M. (2025). A comprehensive catalog of single nucleotide polymorphisms (SNPs) from the black pepper (*Piper nigrum* L.) genome. *BMC Genomics*, *26*, 256. https://doi.org/10.1186/s12864-025-11414-2

Tian, Y., Zhou, J., Zhang, Y., Wang, S., Wang, Y., Liu, H., & Wang, Z. (2021). Research progress in plant molecular systematics of Lauraceae. *Biology*, *10*(5), 391. https://doi.org/10.3390/biology10050391

Tikendra, L., Koijam, A. S., & Nongdam, P. (2019). Molecular markers based genetic fidelity assessment of micropropagated *Dendrobium chrysotoxum* Lindl. *Meta Gene*, *20*, 100562.

Tuvesson, S. D., Larsson, C. T., & Ordon, F. (2021). Use of molecular markers for doubled haploid technology: From academia to plant breeding companies. *Methods in Molecular Biology*, *2288*, 49-72. https://doi.org/10.1007/978-1-0716-1335-1_3

Wang, T., Liu, Y., Ruan, J., Dong, X., Wang, Y., & Peng, J. (2021). A pipeline for RNA-seq based eQTL analysis with automated quality control procedures. *BMC Bioinformatics*, *22*(9), 403. https://doi.org/10.1186/s12859-021-04307-0

Wang, Y., Jin, J. Q., Zhang, R., He, M., Wang, L., Mao, Z., & Wei, K. (2024). Association analysis of BSA-seq, BSR-seq, and RNA-seq reveals key genes involved in purple leaf formation in a tea population (*Camellia sinensis*). *Horticulture Research*, *11*(9), uhae191. https://doi.org/10.1093/hr/uhae191

Wang, Y., Zhang, X., Li, S., Tian, Y., Li, W., Ye, Q., & Ding, J. (2025). Genome-wide association studies of polysaccharides, collagen, and saponin traits in the Sea Cucumber *Apostichopus japonicus*. *Aquaculture Research*, *2025*(1), 5514706. https://doi.org/10.1155/are/5514706

Westermann, A. J., & Vogel, J. (2021). Cross-species RNA-seq for deciphering host–microbe interactions. *Nature Reviews Genetics*, *22*(6), 361-378. https://doi.org/10.1038/s41576-021-00326-y

Zeffa, D. M., Bertagna, F. A. B., Delfini, J., Koltun, A., Uhdre, R. S., Scapim, C. A., & Gonçalves, L. S. A. (2025). Genetic diversity, population structure and linkage disequilibrium in tropical maize (*Zea mays* L.) germplasm adapted to South Brazil. *Plant Breeding*, *144*(4), 549-558. https://doi.org/10.1111/pbr.13273

Zhang, R., Shen, C., Lei, Y., Song, Y., Li, L., Pang, H., & Liu, F. (2025). Cinnamon (Cinnamomum) as a food-medicine homologue: A review of pharmacological mechanisms and potential applications in metabolic diseases. *Food Reviews International*, *42*(1), 122–151. https://doi.org/10.1080/87559129.2025.2476170

Zhang, X., Yang, L., Liu, Y.-H., Zhou, X.-L., Zhang, L.-Q., Wang, Y.-H., & Shen, S.-K. (2021). Genetic diversity, genetic structure, and demographic history of Cinnamomum chago, a plant species with extremely small populations in China. *Global Ecology and Conservation*, *31*, e01808. https://doi.org/10.1016/j.gecco.2021.e01808

Zhao, Y., Wang, K., Wang, W. L., Yin, T. T., Dong, W. Q., & Xu, C. J. (2019). A high-throughput SNP discovery strategy for RNA-seq data. *BMC Genomics*, *20*(1), 160. https://doi.org/10.1186/s12864-019-5533-4

Zhong, Y., Yang, A., Li, Z., Zhang, H., Liu, L., Wu, Z., Li, Y., Liu, T., Xu, M., & Yu, F. (2019). Genetic diversity and population genetic structure of *Cinnamomum camphora* in South China revealed by EST-SSR markers. *Forests*, *10*(11), 1019. https://doi.org/10.3390/f10111019