

RESEARCH ARTICLE

Open Access



Integrative genomic resequencing and transcriptome atlas identify InDel markers for identifying tea plants resistant to anthracnose

Min Li^{1,2}, Yuxuan Liu¹, Chong Sheng¹, Changyu Yan¹, Zhen Zeng¹, Zhusheng Liu³ and Yahui Huang^{1,2*}

Abstract

The development of insertion–deletion (InDel) markers provides a valuable tool for genetic and genomic studies in tea (*Camellia sinensis*) plants. In this study, 4,370,229 InDel markers were identified through genome resequencing of Lingtong Dancong (BY) and Shuixian 7 (SX7). Transcriptome based on RNA sequencing revealed differential expression of 1,125 genes involved in plant–pathogen interaction pathways. Based on genome resequencing of the two tea varieties with a shared genetic background and RNA-seq dataset for post-anthracnose infection, molecular markers for anthracnose resistance in tea plants were developed by identifying specific InDels within the disease resistance-associated interval. The InDel marker linked to anthracnose resistance was located at 123,508,930 bp on chromosome Chr10. Tea plants with DNA bands of 590-bp or heterozygous bands amplified by the primer were resistant to the disease, while those with 431-bp bands were susceptible. When tested across 206 different tea varieties, the InDel marker CsRc demonstrated an 88.59% genotype–phenotype match, confirming its applicability for detecting anthracnose resistance/susceptibility in tea plants. Additionally, the CsRc marker was found to be located within the serine/threonine protein kinase gene, suggesting that ATP plays a critical role in the metabolic processes of tea plants. Taken together, our findings unravel important InDel markers for efficiently distinguishing valuable tea genotypes with desirable resistance to anthracnose, which holds great potential for facilitating disease resistance breeding in the future.

Keywords *Camellia sinensis*, InDel, Genome resequencing, Transcriptome, Metabolome, *Colletotrichum*

Introduction

Tea, widely recognized for its health-promoting properties, exhibits significant neurorelaxant effects that help alleviate psychological stress (Gilbert et al. 2019), with growing evidence supporting the chemopreventive

potential of green tea against carcinogenesis (Eisenstein et al. 2019). The broad variety of tea cultivars (*Camellia sinensis*) provides a rich resource to cater to diverse consumer preferences (Drew et al. 2019). However, tea plants face considerable challenges from both biotic and abiotic stress factors during growth (Brody et al. 2019). Biotic stressors, particularly *Colletotrichum camelliae*, the causal agent of tea anthracnose, pose significant threats to tea cultivation (Orrock et al. 2020; Hassan et al. 2023). This pathogen predominantly infects leaves and woody branches, leading to severe growth retardation and a decline in the quality of tea products. To mitigate the adverse effects of tea anthracnose, identifying resistant

*Correspondence:

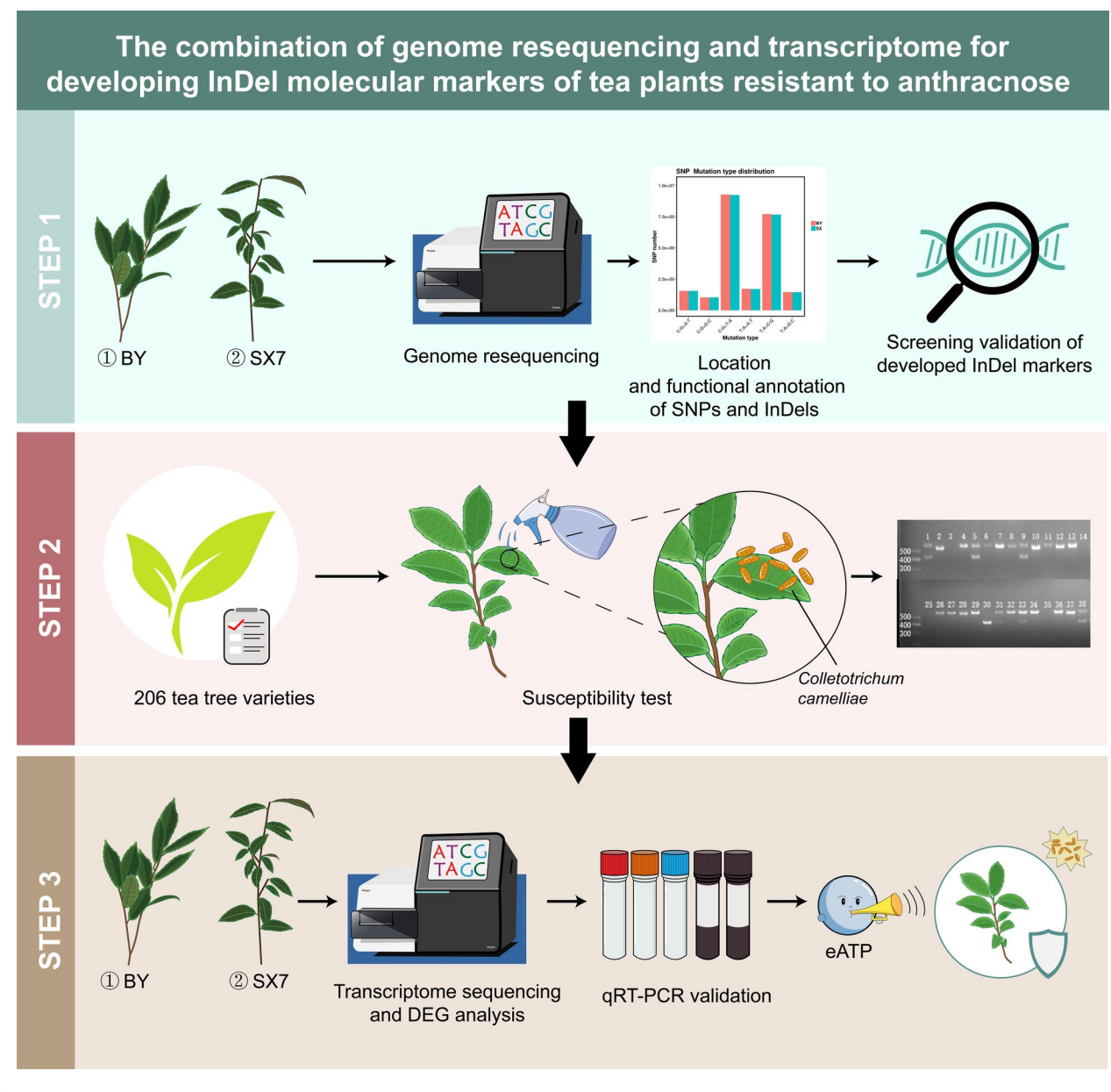
Yahui Huang
13501513191@163.com

¹ College of Horticulture, South China Agricultural University, Guang Zhou 510642, People's Republic of China

² Present Address: Guangdong Provincial Engineering Technology Research Center for Southern Specialty Tea, South China Agricultural University, Guang Zhou 510642, People's Republic of China

³ Guangxi Research Institute of Tea Science, Gui Lin 541010, People's Republic of China

Graphical Abstract



tea plant varieties and developing new resistant cultivars are critical.

While traditional breeding methods remain central to phytopathological programs, their scalability and temporal limitations hinder genetic progress. The application of molecular markers and other techniques for identifying and selecting disease-resistant varieties enhances breeding accuracy and efficiency (Khan et al. 2011). Molecular markers based on DNA polymorphisms have

transformed genetic breeding by enabling the decoding of genome-phenotype relationships at the nucleotide level. Numerous molecular markers have been successfully developed for genetic and genomic studies in tea plants (Mphangwe et al. 2013). Among these, single-nucleotide polymorphism (SNP) and insertion–deletion (InDel) markers have emerged as key tools in molecular breeding, offering substantial potential for research in plant genetics and breeding (Thakur et al. 2018).

In plant genetic improvement, the integration of InDel polymorphisms has significantly advanced disease resistance breeding (Liu et al. 2015). Researchers have extensively applied InDel technology, developing various molecular marker techniques based on InDel to enhance the efficiency and accuracy of gene identification (Lv et al. 2013). Foria et al. (2018) successfully converted a haplotype marker of the grape anti-downy mildew gene into an InDel marker, providing a detailed analysis of its origin. Hassan et al. (2018) cloned the tir-nbs-llr-type resistance gene related to sweet potato stem blight and developed the corresponding InDel marker to improve resistance in melons. With the continuous advancement of InDel technology, cleaved amplified polymorphic sequence (CAPS) and InDel markers have been employed to investigate the role of resistance genes in cucumbers against powdery mildew (Anarjan et al. 2021). Researchers have also pioneered the conversion of conventional CAPS markers into high-resolution PCR-based length polymorphism systems through optimized restriction enzyme digestion protocols. A notable example is the use of a conserved 41-bp deletion in the MR-1 melon (*Cucumis melo* L.) germplasm, which has enabled the development of codominant InDel markers (MePM-R1InDel) with greater discriminatory power (Choi et al. 2020). Additionally, PCR-based InDel marker systems have revolutionized resistance phenotyping across various pathosystems. For instance, in melon bacterial fruit blotch (BFB) research, Islam et al. (2020) developed a 12-InDel multiplex PCR panel. Moreover, SSR and InDel markers have been used to screen for rice genotype resistance to rice blast (Mishra et al., 2021). In onion, a combination of InDel polymorphisms and molecular marker-assisted backcross (MAB) technology led to the development of the DMR1 marker for resistance to downy mildew (Kim et al. 2016). Nvsvrot et al. (2020) employed next-generation sequencing-based genotyping methods to construct a genetic map of poplar, locating leaf rust resistance loci and identifying a 611-bp deletion of the *R* gene associated with resistance variation, thus providing valuable diagnostic markers for molecular breeding.

The combined application of SNP and InDel technologies is equally crucial. Lv et al. (2013) developed molecular markers for InDels and SNPs in Chinese cabbage, facilitating the breeding process for resistance to blackleg disease and creating an InDel genetic map for *Fusarium* wilt resistance, supporting marker-assisted selection. In tomato, newly developed SNP and InDel markers significantly improved genotyping accuracy for the *C. fulvum* gene 9 (*Cf-9*) allele in yellow stem blight resistance (Kim et al. 2017). Liu et al. (2019a, b) rapidly localized the *C. fulvum* gene 10 (*Cf-10*) resistance gene by adopting a combinatorial strategy (Kompetitive Allele-Specific PCR

[KASP] genotyping combined with SNP index, InDel index analysis, and local QTL mapping) Furthermore, a comparison between species-specific InDel markers in peppers and genotyping-by-sequencing (GBS)-derived SNP markers revealed that *C. baccatum* could be the source of exogenous introgression of the powdery mildew resistance gene into the "VK515R" (powdery mildew-resistant parental line) variety, advancing disease resistance breeding in peppers (Jo et al. 2017).

InDel markers are polymorphisms resulting from the insertion or deletion of nucleotides, and their widespread distribution, high density, low variation, strong polymorphism, and ease of detection make them increasingly valuable (Liu et al. 2013; Zhou et al. 2015). The identification of numerous InDel polymorphisms in genome-scale sequencing programs has facilitated their development as high-density markers (Mills et al. 2006). These markers have been extensively applied in the mapping, cloning, and polymorphism analysis of plant resistance genes (Jander et al. 2002; Hayashi et al. 2006; Rao et al. 2011), playing a pivotal role in marker-assisted breeding, germplasm resource analysis, and other areas. However, reports on the development and application of molecular markers for anthracnose resistance in tea plants are lacking.

In 1956, due to a scarcity of tea tree resources in Lingtou Village, thousands of Fenghuang Shuixian tea seedlings were introduced from Fenghuang Commune (now Fenghuang Town, Chaozhou). One of these seedlings underwent natural variation, becoming the progenitor of the Lingtou Dancong variety. This process demonstrates the direct inheritance of two key genes in the germplasm. The original mother plant of Lingtou Dancong was derived from the Fenghuang Shuixian population variety (Wang et al. 2022). In 1961, tea farmers in Lingtou Village, Pingxi Town, Raoping County, Guangdong Province, discovered a tea tree with extremely early budding and yellow-green buds and leaves in a Fenghuang Shuixian tea garden (now known as the "mother tree of Lingtou Dancong"). After three years of separate plucking and processing, the tea samples were found to have stable quality and a unique honey-like aroma, ultimately becoming an independent variety. In 1981, the Department of Agriculture of Guangdong Province officially named it the "Lingtou Dancong variety," noting that it was "selected from individual plants within the Fenghuang Shuixian population." Lingtou Dancong is a high-quality variant of Fenghuang Shuixian, resulting from specific geographic conditions and artificial selection. While both share the same genetic foundation, they differ significantly in aroma, stress resistance, and planting patterns (Qiu And Cao 2022).

Wang et al. (2016) demonstrate the strategic use of parental divergence in the study of tea plant resistance. Long Jing 43 (LJ43) and Zhong Cha 108 (ZC108) exhibit different responses to infection by *Colletotrichum* species. Pathogenicity assays revealed that the lesion area in ZC108 caused by *C. camelliae* was smaller than that in LJ43 under controlled inoculation conditions. The use of resistant-susceptible material pairs has proven valuable for marker development. Early in the laboratory study, a strain resistant to *C. camelliae* was isolated from a population of Fenghuang Shuixian tea, designated Shuixian 7 (SX7). Sensitivity testing showed that Lingtou Dancong is highly susceptible to *C. camelliae* (Fig. 1). This study conducted genome resequencing on both resistant and susceptible anthracnose materials, analyzing InDel variation sites through resequencing data. Additionally, combining transcriptome data from both materials exposed to anthracnose infection, InDel markers associated with anthracnose resistance were established based on the Kyoto encyclopedia of genes and genomes (KEGG) plant-pathogen interaction pathway. These InDel markers were validated through PCR. The selected markers were then used to verify 206 tea varieties. These findings generate valuable markers for genetic and genomic studies in tea plants, facilitating further advancements in breeding programs for anthracnose resistance.

Materials and methods

Plant materials and treatment

The disease-resistant cultivar SX7 and the disease-susceptible cultivar Lingtou Dancong (BY) were used, along with 206 tea cultivars, to validate the developed molecular marker primers. To assess the susceptibility of different tea varieties to anthracnose, tea leaves were sprayed with 1×10^6 CFU/mL *C. camelliae* spores and incubated at 28°C and 98% humidity for 5 days. The disease index (DI) was then recorded. Disease severity was classified according to Table S1, and DI was calculated based on the following formula: $= \sum (Si \times ni) / 6N \times 100$, where *S* represents the disease grade, *n* is the number of samples corresponding to each disease grade, *i* is the disease grade, 6 is the highest disease grade, and *N* is the total number of samples surveyed. The DI is classified into 7 levels, as detailed in Table S2 (Wu et al. 2023).

Genome resequencing

For genome resequencing, resistant cultivar SX7 and susceptible cultivar BY leaves were selected during their vegetative growth period. Genomic DNA was extracted using a cetyltrimethylammonium bromide (CTAB)-based protocol (Uzunova et al. 1995). The DNA concentration and quality were assessed using a NanoDrop2000 Spectrophotometer (Thermo Fisher Scientific, Shanghai,

China). DNA libraries (350 bp) were constructed for Illumina/BGI sequencing according to the manufacturer's guidelines. Sequencing was performed on an Illumina HiSeq X Ten/NovaSeq/BGI platform by a commercial service (Biomarker Technologies, Beijing, China).

SNP/INDEL calling and annotation

The quality of the paired-end raw sequencing reads was assessed with FastQC (v0.12.1), followed by stringent filtering to obtain clean reads. These reads were aligned to the reference genome of *C. sinensis* cultivar 'Shuchazao' (available at <https://eplant.njau.edu.cn/tea/index.html>; Zhao et al. 2019) using BWA-MEM (v0.7.17). Variant calling identified SNPs and InDels using GATK (v4.1.4.1) HaplotypeCaller, and functional annotation was performed. Base calling from raw image files generated FASTQ-formatted sequence reads. SNP annotation was conducted using snpEff (v3.6c) with the reference genome as the annotation source (Cingolani et al. 2012). Exonic InDels were categorized based on their potential to induce frameshift mutations. Circos (v0.69–9; Wei et al. 2018) was used to visualize the genomic distribution of identified variants through circular ideograms.

RNA isolation and transcriptome sequencing

Leaves were sampled from SX7 and BY at five days post inoculation (DPI) with *C. camelliae*, the causal agent of tea anthracnose, frozen in liquid nitrogen, and stored at –80°C for transcriptomic analysis. Total RNA extraction was performed using the Plant Quick RNA Isolation Kit (Huayueyang Biotech, Beijing) according to the manufacturer's protocols. The cleaned sequencing reads were aligned to the *C. sinensis* 'Shuchazao' reference genome (Zhao et al. 2019) using Hierarchical Indexing for Spliced Alignment of Transcripts (HISAT2, v2.2.1; Kim et al. 2019), which allows simultaneous mapping of genomic positions and sequence-specific features. Gene expression quantification was conducted using featureCounts (v2.0.3; Liao et al. 2014), based on the alignment coordinates and annotated gene boundaries in the reference genome.

Analysis of differentially expressed genes (DEGs) and functional annotation

Differential gene expression analysis between the sample groups was performed using DESeq2 (v1.38.3; Love et al. 2014), which applies the Benjamini–Hochberg procedure for false discovery rate (FDR) correction of *p*-values during multiple hypothesis testing. The analysis identified DEGs based on the thresholds of $|\log_2 \text{Fold Change}| \geq 1$ and $\text{FDR} < 0.05$ (Varet et al. 2016). Read quantification was achieved using featureCounts (Liao et al. 2014), with subsequent aggregation of counts across biological



Fig. 1 Resistance and susceptibility of two tea tree varieties to *Colletotrichum camelliae*. **a, b** Seedlings of the tea cultivars Lingtou Dancong (BY, **a**) and Shuixian 7 (SX7, **b**). **c, d** Disease progression on BY leaves (**c**) and SX7 leaves (**d**) 5 days after inoculation with *C. camelliae*

replicates. Functional annotation of DEGs was carried out using the Kyoto encyclopedia of genes and genomes (KEGG) database (Kanehisa et al. 2008) to identify enriched biological pathways. Gene Ontology (GO) categorization was performed in three ontological domains: biological process, cellular component, and molecular function. All available terms (fewer than 15 when insufficient) were visualized using GOplot, generating classification histograms according to established visualization standards (Ashburner et al. 2000).

Quantitative real-time polymerase chain reaction (qRT-PCR) assay

Total RNA was isolated from the leaf samples collected five days post-inoculation with *C. camelliae* using the Plant Quick RNA Isolation Kit (Huayueyang Biotech, Beijing, China), followed by cDNA synthesis using the PrimeScript™ RT Reagent Kit (Takara Bio, Kyoto, Japan), which eliminates genomic DNA. Quantitative reverse transcription PCR (qRT-PCR) was conducted following established protocols (Li et al. 2020) using *C. sinensis* Actin (*CsACTIN*) as an endogenous control. Relative gene expression levels were calculated using the $2^{-\Delta\Delta C_t}$ method (Livak And Schmittgen 2001), with three technical replicates per biological sample. Primer sequences are provided in Table S3.

Detection of InDel markers and primer design

Using the Shuchazao reference genome of *C. sinensis* (Zhao et al. 2019; accessible at <https://eplant.njau.edu.cn/tea/ahau.html>), genes involved in the plant-pathogen interaction pathway were identified through KEGG enrichment analysis of transcriptomic data. InDel loci (≥ 50 bp) derived from genome resequencing were prioritized as candidate molecular marker regions. Flanking sequences (± 300 bp) surrounding each InDel locus were extracted for primer design using Primer Premier 5.0 (PREMIER Biosoft, USA) under standardized parameters. The design criteria adhered to those used for InDel marker development (Liu et al. 2019a, b), emphasizing specificity and amplification efficiency.

DNA extraction

Leaf samples from *C. sinensis* cultivars were obtained from the Zengcheng Germplasm Resource Nursery at South China Agricultural University. Genomic DNA was extracted using the Ezup Spin Column Super Plant Genomic DNA Extraction Kit (Sangon Biotech, Shanghai) according to the manufacturer's protocol. DNA concentration was measured using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Shanghai, China) and standardized to 30 ng/ μ L for downstream

PCR applications, optimized for tea genomic studies (Liu et al. 2019a, b).

PCR amplification of InDel markers and polymorphism analysis

Genomic DNA from tea cultivars SX7 and BY was amplified using previously designed InDel-specific primer pairs. The primary screening was performed in 25 μ L reaction volumes containing 30 ng template DNA, using a Veriti™ 96-Well Thermal Cycler (Applied Biosystems, USA) under standard cycling conditions. Electrophoretic separation was carried out on 3% (w/v) agarose gels (Biowest) in 1×TAE buffer (Sangon Biotech) at 120 V for 45 min, with GoldRed nucleic acid stain (AG Bioengineering, China) for visualization. Poststaining optimization followed established silver staining protocols (Liu et al. 2016) to enhance band resolution. To assess marker discriminative power, primers showing consistent polymorphism and single-band amplification were validated across 206 diverse *C. sinensis* accessions. The final validated InDel marker primers, including sequences and amplification characteristics, are detailed in Table S4.

Statistical analysis

Relative gene transcript levels were analyzed using analysis of variance (ANOVA) via the Statistical Analysis System (SAS Program, SAS Institute Inc., Cary, NC, USA) to detect significant differences in gene expression. Tukey's test was applied to determine temporal variations in gene expression across different time points. The raw sequencing data were processed using the BMKCloud platform (www.biocloud.net). Integrated analysis of transcriptomic and metabolomic datasets was performed via MetWare Cloud (<https://cloud.metware.cn>), a publicly accessible analytical platform. Results were visualized through heatmap generation using TBtools software (Chen et al. 2020).

Results

Genome resequencing

To investigate molecular markers associated with anthracnose resistance in tea, whole-genome resequencing was conducted on two contrasting cultivars: the susceptible SX7 and the resistant BY. Post-sequencing data processing generated 122.96 Gbp of high-quality reads, achieving a 98.47% alignment rate against the 'Shuchazao' reference genome (20X coverage depth, 38.25% GC content). The mapped clean reads consisted of two types: paired-end and single-end reads. Following filtering for adapter sequences and low-quality reads, 410,370,472 and 411,659,116 clean reads were obtained from SX7 and BY, respectively (Table S5). No significant separation was

observed between AT and CG bases, and the curve was relatively flat, indicating normal sequencing results (Fig. S1a); the error rate distribution is shown in Fig. S1b. The coverage depth distribution across the chromosomes was relatively uniform, suggesting good sequencing randomness, with the chromosome coverage depth distribution shown in Fig. S1c. By detecting the starting and ending positions of the paired-end sequences on the reference genome, the actual size of the sequencing fragments was determined, as shown in Fig. S2a. The distribution of insert fragment lengths followed a normal distribution, indicating no abnormalities in the sequence data library construction. After alignment of the reads to the reference genome, nucleotide coverage across the genome was statistically analyzed. The distribution of nucleotide coverage depth was relatively uniform, further confirming sequencing randomness. The curves for nucleotide coverage depth and coverage degree for the samples are shown in Fig. S2b. Alignment rates, efficiency, sequencing depth, and coverage degree were statistically analyzed and compared with those of the reference genome using SAMtools flagstat/depth commands and other tools, as presented in Table S6.

Location and functional annotation of SNPs and InDels

The annotation of the 'Shuchazao' reference genome was utilized to explore the distribution of SNPs and InDels across distinct genomic regions. The results of SNP detection between the sample and reference genome are presented in Table S7. After filtering, a total of 33,454,128 SNP loci were identified. To validate the detected SNPs, read support counts for the SNPs and the cumulative distribution of distances between adjacent SNPs were analyzed (Fig. S3a). SNP mutations across the genome were classified into six categories. For instance, the T:A>C:G mutation includes T>C and A>G, as sequencing data can be mapped to both the positive and negative strands of the reference genome. When a T>C mutation occurs on the positive strand, an A>G mutation occurs at the same position on the negative strand, thus categorizing T>C and A>G as a single mutation type (Fig. S3b). The statistical results of SNPs across the samples are shown in Fig. S3c, with 8,657,071 common SNPs shared between SX7 and BY. Based on the classification of whole-genome SNP annotations for all samples, a pie chart is provided in Fig. S3d. The largest proportion of SNPs is found in the INTERGENIC region, accounting for 87.94% of the total. Using the mapping results of clean reads from the samples in the reference genome, small fragment insertions and deletions (small InDels) between the samples and the reference genome were detected. These InDels were identified *via* GATK. Small InDel variations, generally fewer than SNP variations, highlight the differences

between the samples and the reference genome. InDels in coding regions can cause frameshift mutations, leading to alterations in gene function. A total of 2,989,992 InDels and 3,015,906 InDels shared by the selected SX7 and BY strains, respectively, were identified (Table S8).

Statistical analysis of the InDel lengths in the coding sequence (CDS) region and the entire genome revealed the distribution of InDels across various length categories, with the ratios for each category calculated (Fig. 2a). Mononucleotide InDels were the most abundant, with InDels ranging from 1 to 20 bp in length being predominant. The number of InDels steadily decreased as the length of the InDels increased. The statistical results of InDels in the samples are shown in Fig. 2b, where 1,211,279 InDels were shared by SX7 and BY. Based on the classification of whole-genome InDel annotations for all samples, the INTERGENIC region had the highest proportion, accounting for 83.5% of the total number of InDels (Fig. 3a). The variant distribution for each type of InDel across the chromosomes was visualized in a Circos diagram (Fig. 3b), generated using Circos software (<http://circos.ca/>). Variations in the CDS region can result in changes to gene function. Genes with nonsynonymous mutation SNPs and InDels in the CDS region were identified by searching for variants between the reference genome and the samples, potentially revealing functional differences. SX7 and BY each had 11,312 and 11,393 InDels located in the CDS region, respectively (Table S9).

Transcriptome sequencing and DEG identification

Transcriptome reprogramming occurs in response to pathogen infection, involving multiple regulatory pathways. To explore genes associated with differences in susceptibility between anthracnose-resistant (SX7) and anthracnose-susceptible (BY) cultivars of *C. sinensis*, transcriptomic analyses were conducted. A total of six cDNA libraries were sequenced from SX7 and BY. DEGs were identified based on $|\log_2(\text{Fold Change})| \geq 1$ and $\text{P}_{\text{adj}} \leq 0.05$. The volcano plot visually summarizes the distribution of DEGs between the two sample groups (Fig. 4a). A total of 5,001 DEGs were identified, with 479 DEGs common to both susceptible and resistant varieties for further analysis. KEGG enrichment analysis revealed significant associations with several functional clusters, including the MAPK signaling pathway, phenylpropanoid biosynthesis, and flavonoid biosynthesis (Fig. 4b). The KEGG enrichment histogram indicated that plant-pathogen interactions constituted 14.84%, while metabolic pathways accounted for 44.98% (Fig. 5). The five KEGG pathways with the smallest q-values in each Level 1 KEGG pathway category were selected to generate the KEGG enrichment circle diagram, with Ko04626

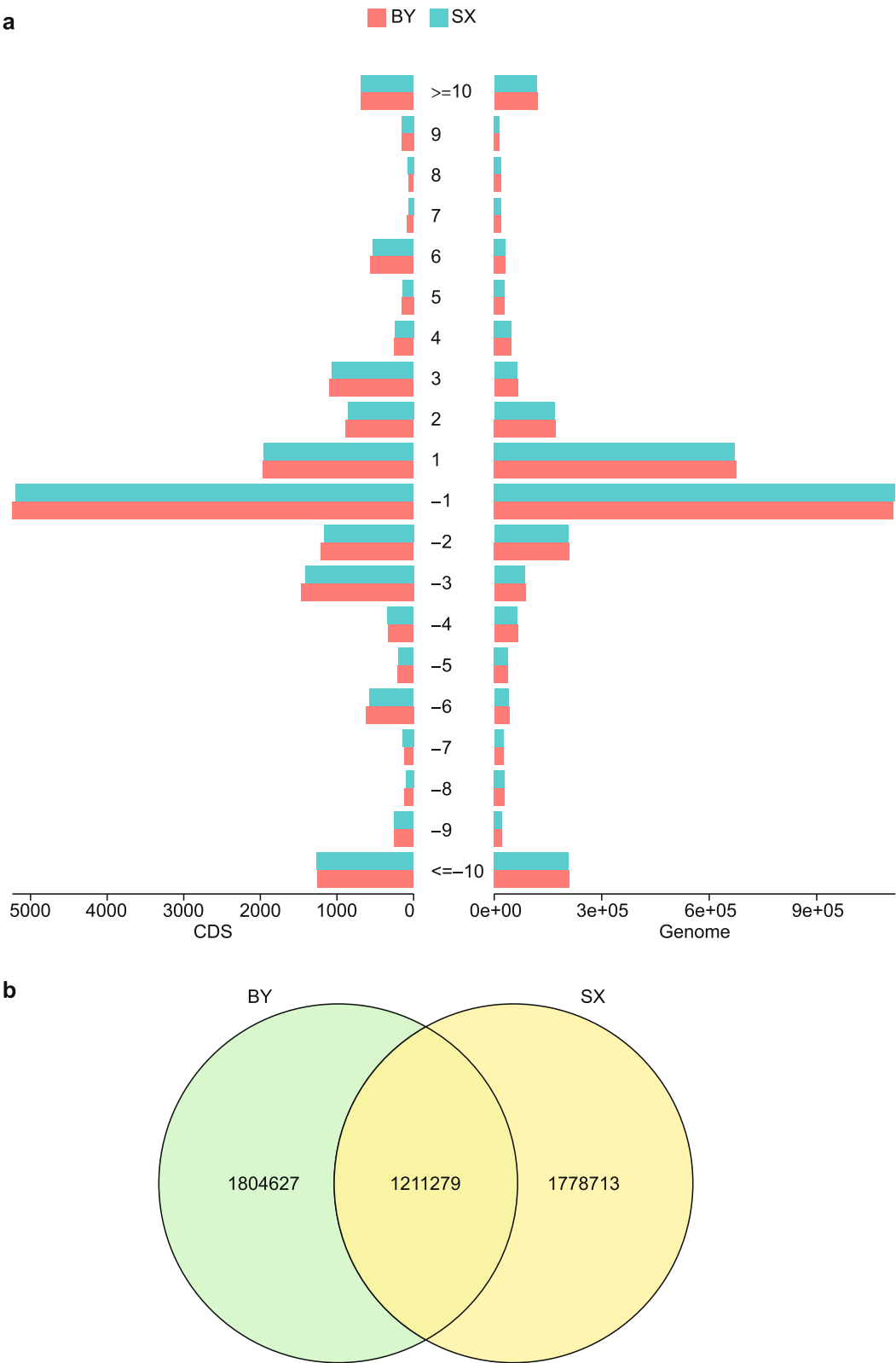
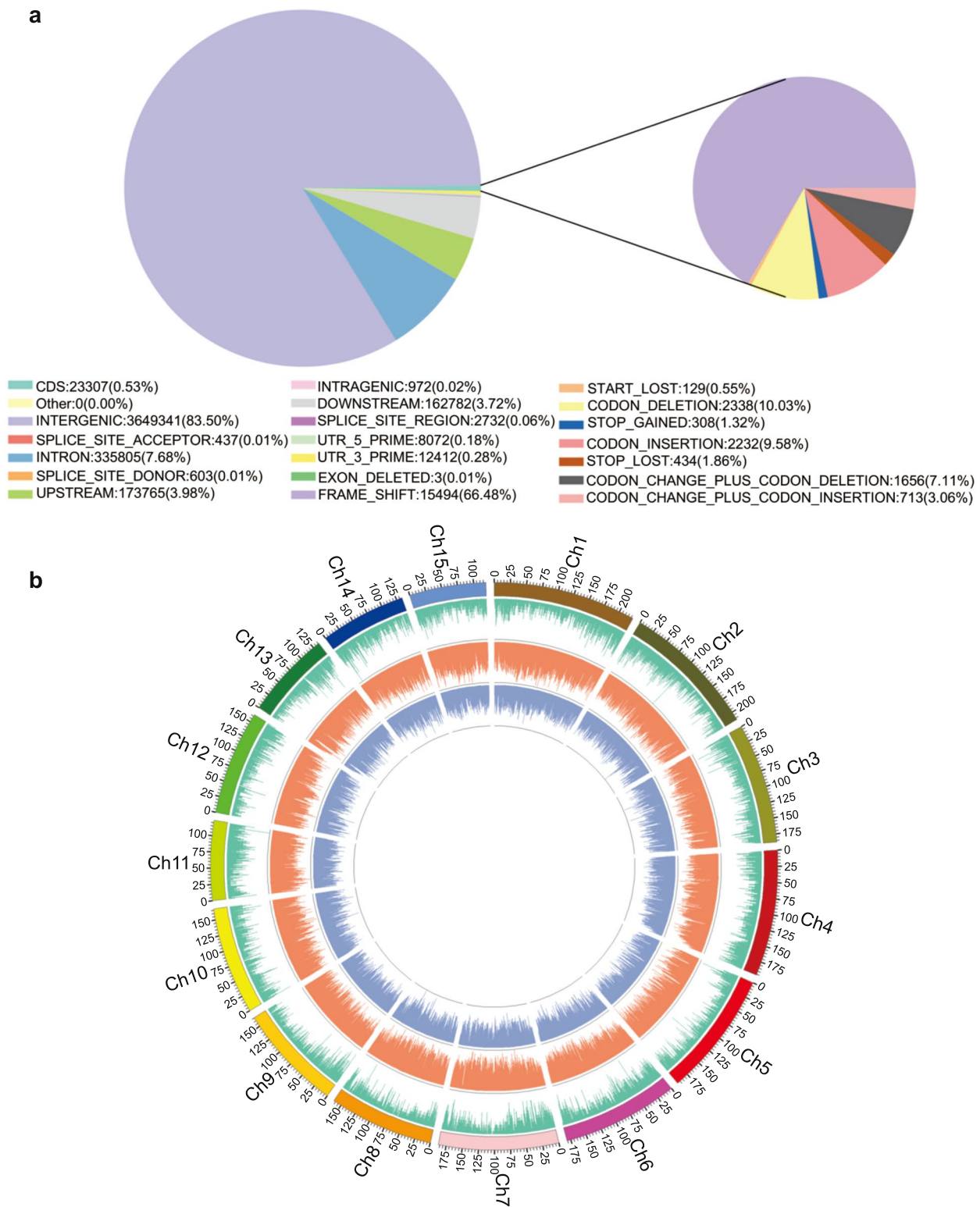


Fig. 2 Insertion–deletion (InDel) analysis *via* genome resequencing. **a** Distribution of InDel lengths across the whole genome and coding regions. The ordinate indicates the InDel length (within 10 bp), with values greater than 0 representing insertions and values less than 0 representing deletions; the abscissa represents the corresponding quantity. **b** Venn diagram showing InDel statistics among samples. BY, Lingtong Dancong; SX7, Shuixian 7; CDS, coding sequence



representing plant-pathogen interaction pathways (Fig. S4a). This study provides valuable molecular insights into the response to tea anthracnose infection, aiding the selection of resistant varieties through marker-assisted breeding.

qRT-PCR validation

The expression patterns of genes involved in plant-pathogen interaction pathways were validated using qRT-PCR. Six candidate DEGs from these pathways were selected for qRT-PCR analysis in the leaves of SX7 and BY (Fig. S4b). The results confirmed the reliability of the RNA-Seq data (Table S10), with expression patterns of the DEGs consistent with the transcriptome sequencing data, further verifying the accuracy of the transcriptomic findings.

Screening and validation of developed InDel markers

Candidate genes related to plant-pathogen interaction pathways were identified through KEGG enrichment analysis of the transcriptome data. Genomic regions containing InDel polymorphisms (> 50 bp variations between resistant and susceptible materials) were selected from these candidate genes (Table S11). Using the reference genome, 300-bp flanking sequences of target sites were extracted for primer design *via* Primer Premier 5. From an initial set of 139 designed InDel markers (Table S4), polymorphic primers were identified through amplification in SX7 and BY (Fig. S5). The CsRc primer set was then validated using DNA samples of 206 tea plants, showing a correlation between amplification patterns and anthracnose susceptibility phenotypes (Fig. S6). The band sizes amplified by the CsRc primers were used to classify resistance and susceptibility. The InDel marker linked to anthracnose resistance was located at 123,508,930 bp on chromosome Chr10. Tea plants with 590 bp or hybrid bands were resistant to disease, while those with 431 bp bands were susceptible. The molecular primers used were ACTCGGTCAAAGCAAGGCAT (CsRc-InDel-F) and GACACACCAGTTTGAAATGGCA (CsRc-InDel-R) (Table S4). The results indicated that the newly developed molecular markers have broad applicability in tea plant breeding. The susceptibility of 206 tea tree varieties from

different populations to *C. camelliae* was tested (Fig. 6; Table S12). DNA amplification was performed using the selected InDel marker primer CsRc, and the agarose gel electrophoresis results showed that 163 samples' detection results matched the resistance identification results. However, 21 samples showed inconsistencies, and 22 samples failed to produce bands (Fig. 7a). The marker primer demonstrated an accuracy of 88.59% in resistance identification (Fig. 7b). This molecular marker proves to be highly accurate (88.59%) for high-throughput screening of anthracnose-resistant germplasms, thus significantly accelerating breeding cycles by enabling early-stage selection.

Gene analysis of marked loci

The molecular marker locus of CsRc is located within the CSS0015304 gene. KEGG annotation places this gene in the plant-pathogen interaction pathway, while KOG annotation identifies it within signal transduction mechanisms, specifically as a serine/threonine protein kinase. GO enrichment analysis highlights ATP binding as a key function (Table S13). Comparative transcriptomic analysis demonstrated significant upregulation of ATP-dependent activity-related genes in SX7 relative to BY (Fig. 7c). This finding aligns with the conserved catalytic core structure of eukaryotic protein kinases (Hanks-type kinases), where a glycine-rich motif near the N-terminal lysine residue is essential for ATP binding (Hanks et al. 1995). Plant immunity is intricately linked to the quality of agricultural products; improving plant immunity not only enhances disease resistance but also boosts crop quality at various levels (Ma et al. 2024). ATP, a universal energy carrier, is present in all living cells and drives essential biochemical processes through its high-energy phosphoanhydride bonds (Chivasa et al. 2005). ATP released from plant cells into the extracellular space, known as extracellular ATP (eATP), serves as a damage-associated molecular pattern (DAMP) involved in plant immunity (Dark et al. 2011) (Fig. 8). Following plant damage, the concentration of eATP rises, binding to receptors on the plant cell surface and initiating downstream signaling events. This results

(See figure on next page.)

Fig. 4 Volcano plot of differentially expressed genes and Kyoto encyclopedia of genes and genomes (KEGG) enrichment scatter plot. **a** Volcano plot of differentially expressed genes. The abscissa represents the fold change in gene expression, and the ordinate represents the significance level of differential genes. Red dots represent upregulated differential genes, green dots represent downregulated differential genes, and gray dots represent non-differentially expressed genes. **b** KEGG enrichment scatter plot. The abscissa represents the Rich factor, where a larger Rich factor indicates a higher degree of enrichment. The ordinate represents the KEGG pathways. Larger dots indicate a greater number of differential genes enriched in the pathway, and redder dots indicate more significant enrichment

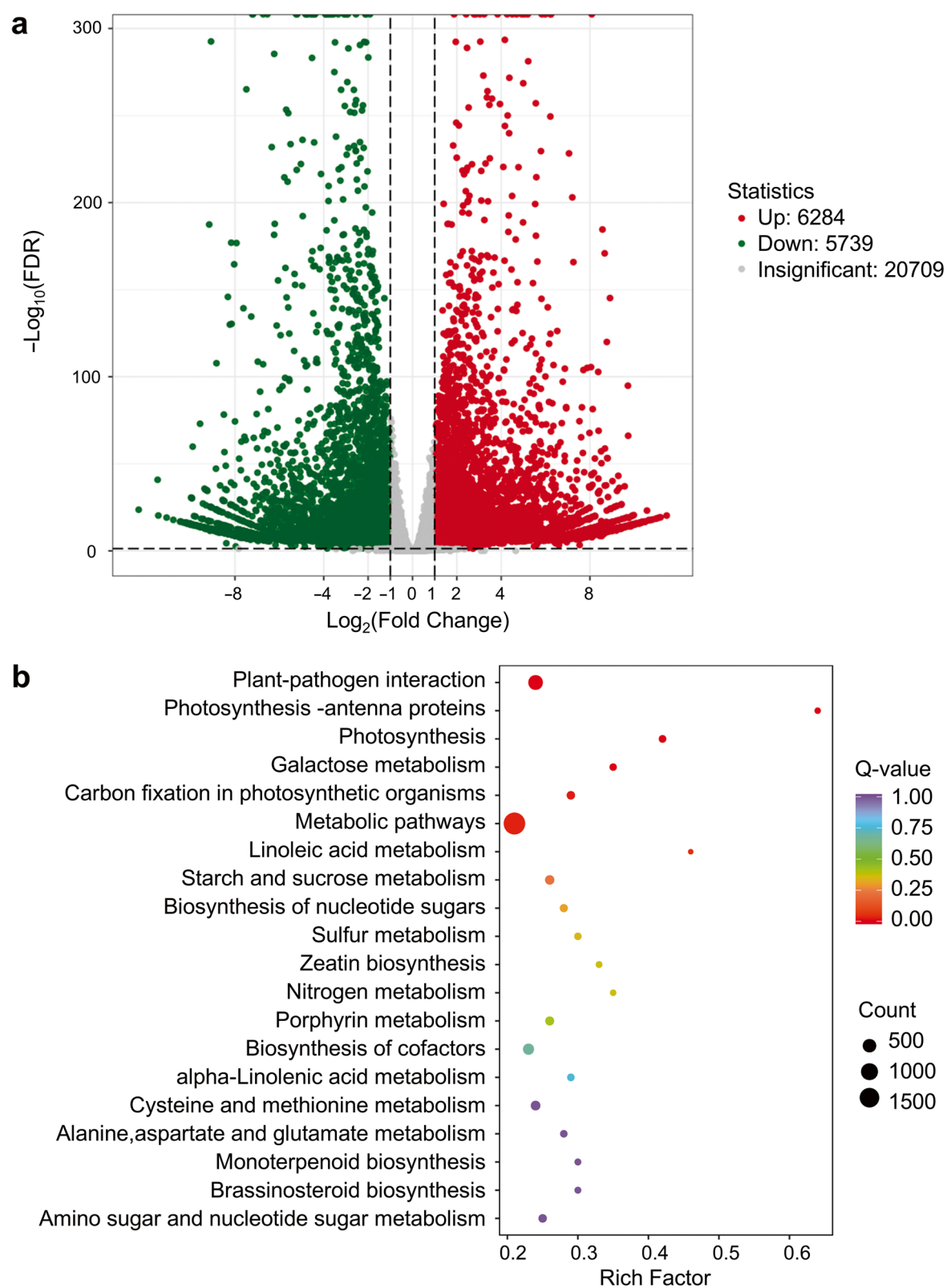


Fig. 4 (See legend on previous page.)

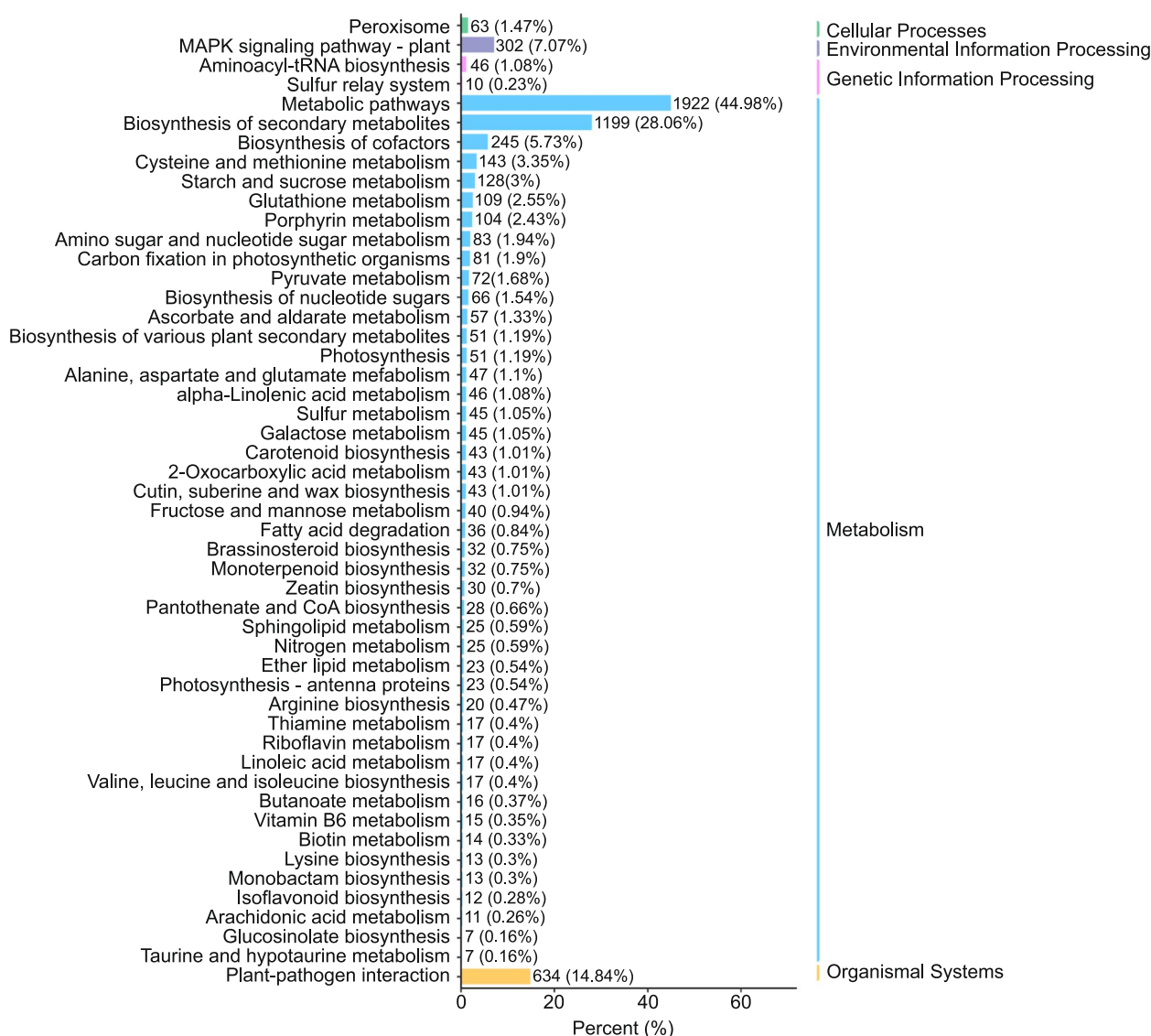


Fig. 5 Kyoto encyclopedia of genes and genomes (KEGG) enrichment graph. The graph displays the 50 most significantly enriched pathway entries. The abscissa represents the number of differential genes annotated to each pathway, and the ordinate represents the name of the KEGG pathway. The numbers within the graph indicate the number of differential genes annotated to the pathway, while the numbers in parentheses show the ratio of differential genes annotated to the pathway relative to the total number of differential genes. The label on the far right denotes the classification of the KEGG pathway

in the upregulation of defense-related genes and the production of defense substances, such as resistant proteins and phenolic compounds. Additionally, eATP indirectly regulates the nutritional quality of plants by influencing metabolic pathways, including the synthesis of proteins, vitamins, and minerals (Shan et al. 2022). Phenolic compounds, flavonoids, aromatic substances, and ascorbic acid, key secondary metabolites in plants, also play vital roles in the chemical defense against pathogens (Kumari et al. 2024).

Discussion

Molecular marker technology, which relies on nucleotide sequence variations between individuals, offers a direct and accurate reflection of genetic polymorphisms at the DNA level. This technology encompasses a variety of marker types, each playing a critical role in genetic analysis due to its unique characteristics. In recent years, molecular marker technology has been widely applied in the genetic analysis of several important crop species, including rice (Tian et al. 2010), potatoes (Barone et al. 2004), and tomatoes (Foolad et al. 2012). The

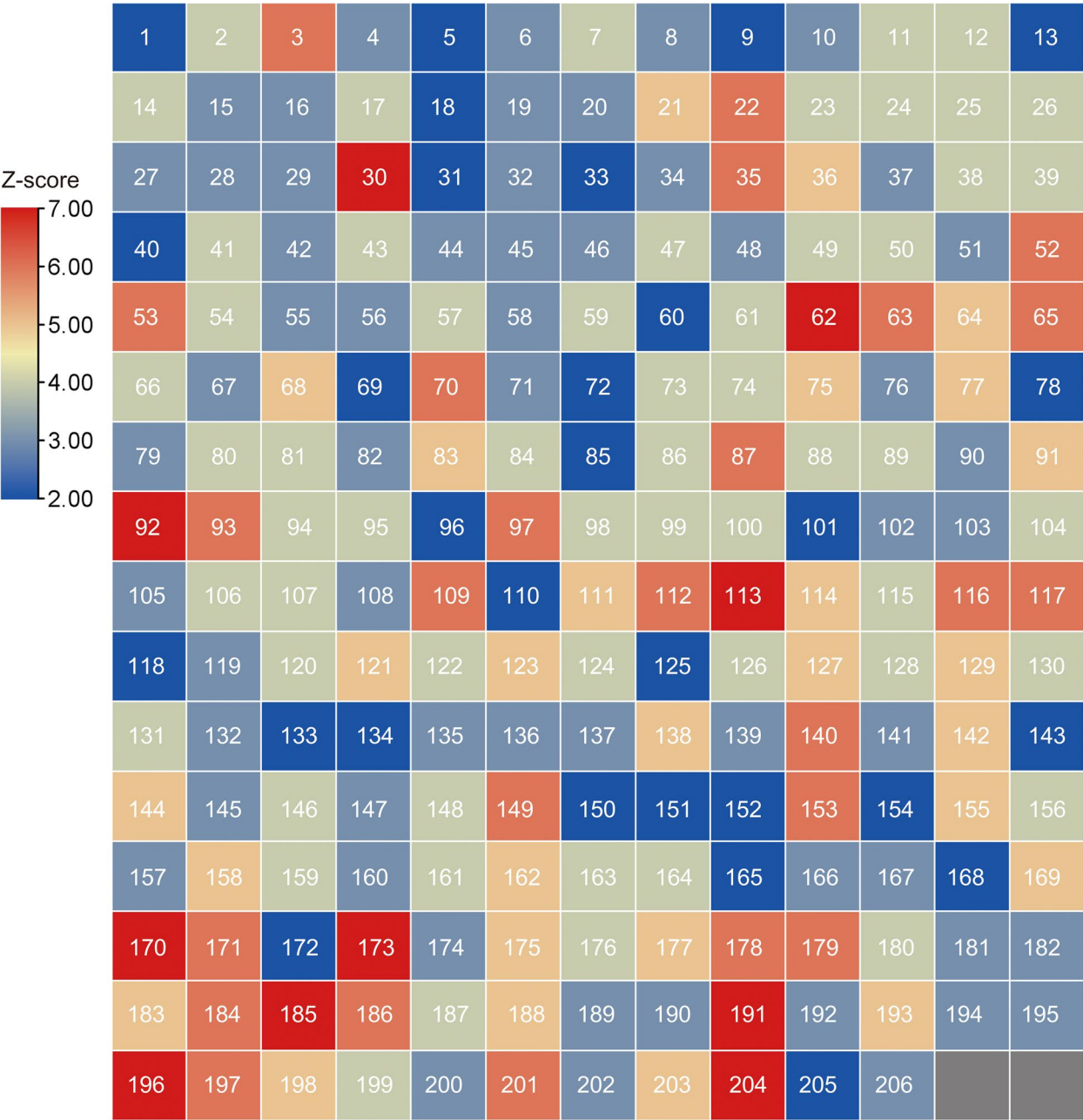


Fig. 6 Sensitivity of 206 tea plant resources to anthracnose. Statistical analysis was performed according to the disease severity index grading criteria in Table S2. Resources with a disease level ≤ 4 were classified as resistant, while those with a level ≥ 5 were classified as susceptible

application of this technology has significantly advanced plant genetics research, enabling researchers to analyze genetic variation with unprecedented precision, construct high-density genetic maps, assist in breeding selection, and explore the genetic basis of complex traits. Various molecular marker technologies have been extensively utilized in the analysis of key traits and quality in tea, with marker-assisted selection enhancing the

accuracy, efficiency, and quality of tea breeding. For example, nine specific bands generated by six random amplified polymorphic DNA (RAPD) primers have been successfully linked to six important traits—black tea quality, drought resistance, high-temperature tolerance, low-temperature tolerance, resistance to blue mold, and high yield—greatly improving the selection efficiency in tea breeding (Mphangwe et al. 2013). Research has

made notable progress in understanding the genetic mechanisms underlying tea traits, including growth and development, quality improvement, and cross-breeding. The function of auxin-related genes in leaf development has been investigated (Wang et al. 2020), as well as the molecular mechanisms of flowering induction and differentiation (Liu et al. 2020). The mechanism of selenium accumulation in tea has been systematically analyzed using transcriptomics (Zheng et al. 2023). Furthermore, exogenous melatonin, which influences polyphenol and amino acid metabolic pathways, such as those involving epigallocatechin gallate (EGCG) and theanine, has been shown to play a key role in optimizing tea quality under medium- and high-temperature conditions (Li et al. 2020).

InDel markers, as essential tools in genetic research, have played a pivotal role in numerous in-depth studies on tea plants. In the assessment of tea germplasm resources, a systematic evaluation of individual tea plants in the United States using 10 InDel markers addressed challenges related to genetic relationship identification and regional adaptability of germplasm resources. These findings offer a scientific foundation for the rational protection and utilization of tea germplasm resources (Clarke et al. 2023). An et al. (2020) employed InDel markers in combination with whole-genome resequencing technology to clearly define the genetic boundary between *C. sinensis* var. *sinensis* (Chinese species) and *C. sinensis* var. *assamica* (Assamica species), providing valuable insights into the evolution of tea varieties. Additionally, InDel markers are crucial in enhancing tea quality. The development of a 14-base-deletion InDel marker based on the novel F3'S'H (flavonoid 3', 5' hydroxylase gene) allele demonstrated a strong association between this gene and the high catechin index of wild tea plants, highlighting its potential for tea quality improvement (Jin et al. 2018). Furthermore, InDel markers have been applied to investigate the genetic regulatory mechanisms of specific traits in tea plants. In a study of the Shuchazao and Yunkang No. 10 varieties, InDel markers revealed significant differences in catechin/caffeine content and identified numerous SNP/

InDel variants in related biosynthetic genes. These findings provide novel insights into the molecular mechanisms of catechin/caffeine synthesis (Liu et al. 2019a, b).

ATP regulates cellular energy balance, and when a plant is attacked by pathogens or experiences other forms of damage, eATP is released. As a signaling molecule, eATP binds to membrane receptors, triggering the production of reactive oxygen species (ROS), opening calcium ion channels, and enhancing the expression of defense genes, thus promoting self-protection. As a key signaling molecule, eATP binds to plasma membrane receptors, stimulating calcium ions and other second messengers to initiate a range of intracellular signaling responses (Tanaka et al. 2010). Choi et al. (2014) identified the plant eATP receptor P2K1 (L-type lectin receptor kinase I.9, LecRK-I.9) through forward genetic screening of *Arabidopsis* mutants that were insensitive to ATP-mediated cellular calcium responses. In 2020, Pham et al. (2020) identified a second eATP receptor, P2K2, by screening all 116 members of the LecRK family, specifically LecRK-I.5. Upon binding with eATP, purinergic receptor kinase 1 (P2K1) and purinergic receptor kinase 2 (P2K2) induce autophosphorylation of their own tyrosine, threonine, or serine residues within the intracellular kinase domain, a critical step for receptor activation and signal transduction (Cho et al. 2022). When P2K1 undergoes autophosphorylation in the presence of P2K2, its conformation and activity change, enabling interaction with downstream signaling molecules. In *Arabidopsis* infected by *Pseudomonas syringae*, P2K1 enhances pathogen resistance by phosphorylating nicotinamide adenine dinucleotide phosphate (NADPH) oxidase, which increases ROS production and promotes stomatal closure (Chen et al. 2017). P2K1 also influences plant cell metabolism by interacting with mevalonate kinase (MVK), a key enzyme in the methylvalerate pathway (Cho et al. 2023). Additionally, P2K1 interacts with effector proteins of *Phytophthora* pathogens, modulating disease resistance by affecting cell wall adhesion to cell membranes (Kumar et al. 2020). Both P2K1 and P2K2 can interact with integrin-linked kinase 5 (ILK5), leading to the phosphorylation of MAP kinase kinase 5 (MKK5), which in

(See figure on next page.)

Fig. 7 Genotype–phenotype match statistics. **a** Comparison of the results of the sensitivity test for 206 tea plant resources to tea anthracnose with the results from agarose gel electrophoresis of Insertion–deletion (InDel) molecular markers. In this comparison, 22 samples showed no amplified bands, 163 samples were consistent with the sensitivity test results, and 21 samples showed discrepancies. **b** The identification accuracy rate of the InDel molecular marker CsRc was 88.59%. **c** Statistical classification of Gene Ontology (GO) annotations for differentially expressed genes in the transcriptome. Differentially expressed genes (DEGs) were categorized based on upregulation and downregulation for GO classification. The abscissa represents the secondary GO categories, while the ordinate represents the number of differential genes for each category. Yellow indicates upregulated genes, and blue indicates downregulated genes. BP, biological process; CC, cellular component; MF, molecular function

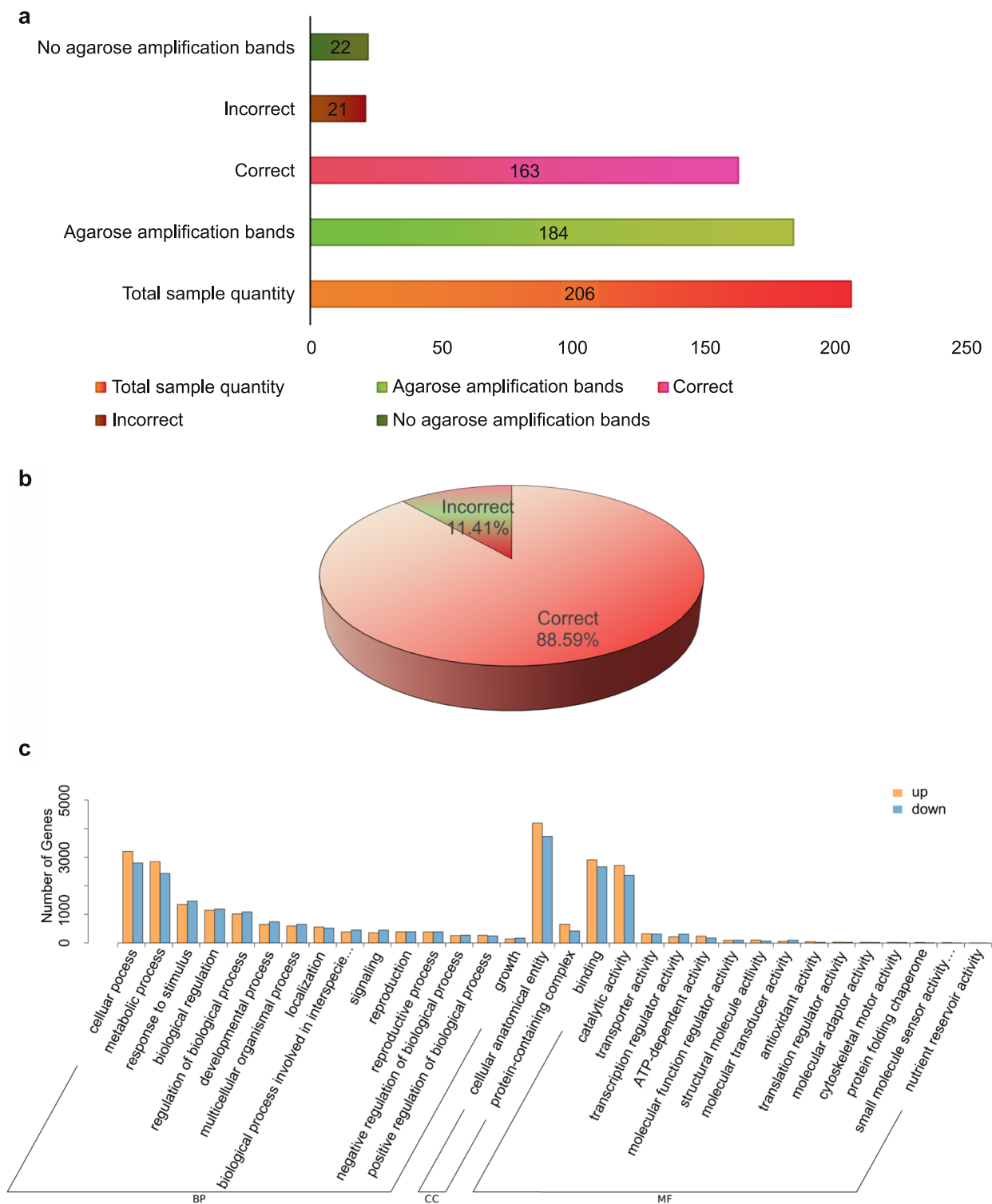


Fig. 7 (See legend on previous page.)

turn phosphorylates mitogen-activated protein kinases (MAPK3/MAPK6), amplifying the MAPK cascade and enhancing the plant’s defense response (Kim et al. 2023).

Conclusions
In the present study, whole-genome resequencing of anthracnose-resistant (SX7) and susceptible (BY) *C.*

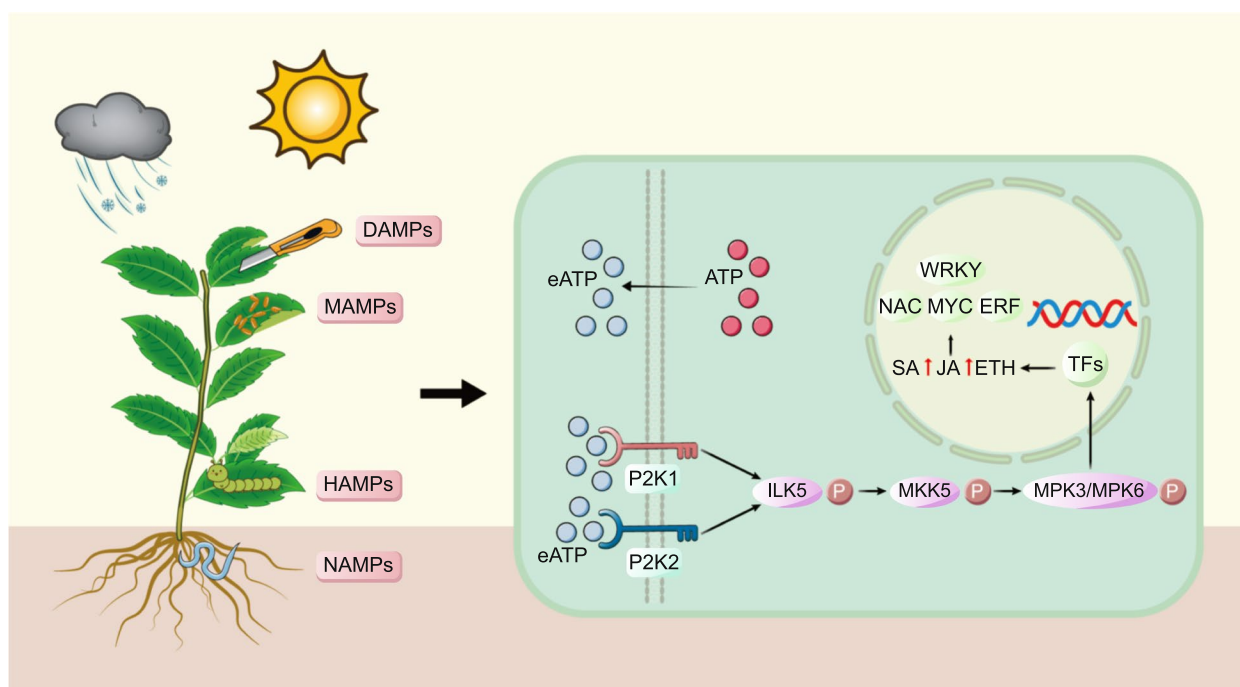


Fig. 8 Mechanistic diagram of extracellular ATP (eATP)-mediated plant immunity. When exposed to external stimuli—such as sunlight, rain, snow, insect herbivory, or mechanical wounding—plants convert intracellular ATP into eATP. eATP binds to the receptors P2K1 and P2K2, initiating a phosphorylation (P) cascade involving ILK5 and MKK5, which leads to the activation of MAPK3/6. Activated MAPK3/6 subsequently phosphorylate and modulate transcription factors (TFs) in the nucleus, including WRKY, NAC, MYC, and ERF families. These TFs regulate the expression of hormone-related genes associated with salicylic acid (SA), jasmonic acid (JA), and ethylene (ETH) signaling pathways, thereby elucidating the signal transduction mechanisms underlying plant responses to environmental stimuli. DAMPs, damage-associated molecular patterns; MAMPs, microbe-associated molecular patterns; HAMPs, herbivore-associated molecular patterns; NAMPs, nematode-associated molecular patterns; P2Ks, purinergic receptor kinases; ERF, ethylene response factor; ILK5, integrin-linked kinase 5; MKK5, MAP kinase kinase 5; MAPK3/MAPK6, mitogen-activated protein kinases

sinensis cultivars revealed 33,454,128 SNPs and 4,370,229 InDels. Complementary transcriptomic profiling was performed on leaves at 5 days post-anthrachnose inoculation. Integration of resequencing data (isogenic backgrounds) and RNA-seq signatures enabled targeted identification of resistance-associated InDels. A diagnostic InDel marker was mapped to Chr10:123,508,930 bp, with the following primer ACTCGGGTCAAAGCAAGGCAT (CsRc-InDel-F) and GACACACACCAGTTTGAAATGGCA (CsRc-InDel-R). Tea plants exhibiting a 590 bp or heterozygous band following PCR amplification were resistant, whereas those with a 431-bp band were susceptible. When tested across 206 different varieties, the assay results perfectly matched individual plant resistance, demonstrating its high applicability. This genotype can serve as an InDel marker for detecting anthracnose resistance and susceptibility in tea plants, facilitating the breeding programs by shortening breeding cycles and accelerating the development of anthracnose-resistant cultivars.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1007/s44281-025-00094-3>.

Supplementary Material 1: Figure S1. The proportional distributions of each base and base error rate distributions of Shuixian 7 (SX7) and Lingtou Dancong (BY) samples. a The proportional distributions of each base distributions of SX7 and BY samples. Different colors represent different base types, which indicates unidentified bases during sequencing. The first 150 bp corresponds to the base distribution of the first-end sequencing reads of the paired-end sequencing sequence, while the last 150 bp represents the base distribution of the opposite-end sequencing reads. Each cycle corresponds to a base in the sequence. For example, the first cycle indicates the distribution of A, T, G, C, and N at the first base of all sequencing reads in this project. b The proportional distributions of each base error rate distributions of SX7 and BY samples. The first 150 bp corresponds to the error rate distribution of the first-end sequencing reads, and the last 150 bp represents the error rate distribution of the second-end sequencing reads. c Chromosome coverage depth distribution map of SX7 and BY samples. The abscissa represents the position of the chromosome, and the ordinate represents the logarithmic ($\log_2$) value of the coverage depth at each position on the chromosome.

Supplementary Material 2: Figure S2. Distribution of insertion fragments and depth distribution in the SX7 and BY samples. a, b Distribution of insertion fragments (a) and depth distribution (b) in the SX7 and BY samples.

Supplementary Material 3: Figure S3. Single-nucleotide polymorphism (SNP) analysis in genome resequencing. a Quality distribution diagram: On the left, the cumulative graph of the number of SNP reads in support, and on the right, the cumulative graph of the distances between adjacent SNPs. b Distribution diagram of SNP mutation types: The ordinate represents the number of SNPs, and the abscissa represents the type of SNP mutation. c Venn diagram of SNP statistics among samples. d Pie chart of SNP annotation results.

Supplementary Material 4: Figure S4. Kyoto encyclopedia of genes and genomes (KEGG) enrichment circle graph and quantitative real-time PCR (qRT-PCR) analyses of gene expression. a The KEGG enrichment circle graph was drawn by selecting the five KEGG pathways with the smallest q value in each Level 1 KEGG pathway category. The first circle represents the Level 1 KEGG pathway category entries, with different colors representing various KEGG classifications. The second circle displays the number of genes in the background gene set belonging to each classification and their q-value. The longer and redder the bar, the more significant the enrichment. The third circle shows the proportion of upregulated and downregulated genes: light red represents upregulated genes, and light blue represents downregulated genes. Specific values are provided below. The fourth circle presents the RichFactor values for each classification, with background auxiliary lines representing 0.2 intervals. b qRT-PCR analysis of the expression of six plant-pathogen interaction genes in SX7 and BY inoculated with *C. camelliae*. Three independent biological replications were performed. All data were normalized to β -actin expression levels, and relative gene expression was calculated using the comparative $2^{-\Delta\Delta Ct}$ method. The bars represent the standard deviations among biological replicates.

Supplementary Material 5: Figure S5. The results of the 139 primers used to amplify the DNA of SX 7, BY, Zhongcha108, and Longjing43.

Supplementary Material 6: Figure S6. The 206 DNA samples extracted in each step were amplified via the CsRc primers.

Supplementary Material 7: Table S1. Classification criteria for the degree of resistance of tea leaves to anthracnose. Table S2. Grading of disease severity index. Table S3. The sequences of the qRT-PCR primers. Table S4. Primers of insertion-deletion (InDel) markers. Table S5. Evaluation and statistical computation of sample sequencing data. Table S6. Statistical analysis of the alignment results. Table S7. The statistical results of the SNPs detected. Table S8. Statistics of InDels in the whole genome and coding regions. Table S9. Classification and statistical computation of differential genes generated by various mutations. Table S10. RNA-Seq of plant-pathogen interaction genes. Table S11. Information on InDel mutation sites. Table S12. The determination of the susceptibility to tea anthracnose of 206 tea plant resources. Table S13. Locus-related information for the InDel molecular marker CsRc.

Acknowledgements

The methods guidance of Xiaoyu Qi is gratefully acknowledged.

Authors' contributions

ML and YHH conceived and designed research. ML conducted experiments. CS contributed new tea materials. ML and YXL analyzed the data. ML wrote the manuscript. CYY, ZZ, and ZSL modified the article. YHH provided funds. All authors read and approved the final manuscript.

Funding

This study was supported by the Technological Innovations in Sexual Hybridization Breeding of Tea Plants and the Creation of New Germplasm (2025B0202070007).

Data availability

All data generated or analysed during this study are included in this published article and its supplementary information files.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors have no relevant financial or no-financial interests to disclose. All authors read and approved the final manuscript.

Received: 28 September 2025 Revised: 14 November 2025 Accepted: 2 December 2025

Published online: 20 February 2026

References

- An Y, Mi X, Zhao S, Guo R, Xia X, Liu S, et al. Revealing distinctions in genetic diversity and adaptive evolution between two varieties of *Camellia sinensis* by whole-genome resequencing. *Front Plant Sci.* 2020;11:603819. <https://doi.org/10.3389/fpls.2020.603819>.
- Anarjan MB, Bae I, Lee S. Marker-assisted evaluation of two powdery mildew resistance candidate genes in Korean cucumber inbred lines. *Agronomy.* 2021;11:2191. <https://doi.org/10.3390/agronomy11112191>.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. *The Gene Ontology Consortium Nat Genet.* 2000;25:25–9. <https://doi.org/10.1038/75556>.
- Barone A. Molecular marker-assisted selection for potato breeding. *Am J Potato Res.* 2004;81:111–7. <https://doi.org/10.1007/BF02853608>.
- Brody H. Tea. *Nature.* 2019;566:S1. <https://doi.org/10.1038/d41586-019-00394-5>.
- Chen D, Cao Y, Li H, Kim D, Ahsan N, Thelen J, et al. Extracellular ATP elicits DORN1-mediated RBOHD phosphorylation to regulate stomatal aperture. *Nat Commun.* 2017;8:13. <https://doi.org/10.1038/s41467-017-02340-3>.
- Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant.* 2020;13:1194–202. <https://doi.org/10.1016/j.molp.2020.06.009>.
- Chivasa S, Ndimba BK, Simon WJ, Lindsey K, Slabas AR. Extracellular ATP functions as an endogenous external metabolite regulating plant cell viability. *Plant Cell.* 2005;17:3019–34. <https://doi.org/10.1105/tpc.105.036806>.
- Cho SH, Tóth K, Kim D, Vo PH, Lin CH, Handakumbura PP, et al. Activation of the plant mevalonate pathway by extracellular ATP. *Nat Commun.* 2022;13:450. <https://doi.org/10.1038/s41467-022-28150-w>.
- Cho SH, Nguyen CT, Pham AQ, Stacey G. Computational prediction and in vitro analysis of the potential ligand binding site within the extracellular ATP receptor, P2K2. *Plant Signal Behav.* 2023;18:6. <https://doi.org/10.1080/15592324.2023.2173146>.
- Choi J, Tanaka K, Cao Y, Qi Y, Qiu J, Liang Y, et al. Identification of a plant receptor for extracellular ATP. *Science.* 2014;343:290–4. <https://doi.org/10.1126/science.1243618>.
- Choi YR, Lee JY, Hwang S, Kim HU. PCR-based InDel marker associated with powdery mildew-resistant MR-1. *Agronomy.* 2020. <https://doi.org/10.3390/agronomy10091274>.
- Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. 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.* 2012;6:80–92. <https://doi.org/10.4161/fly.19695>.
- Clarke C, Richter BS, Rathinasabapathi B. Genetic and morphological characterization of United States tea (*Camellia sinensis*): insights into crop history, breeding strategies, and regional adaptability. *Front Plant Sci.* 2023;14:1149682. <https://doi.org/10.3389/fpls.2023.1149682>.
- Dark A, Demidchik V, Richards SL, Shabala S, Davies JM. Release of extracellular purines from plant roots and effect on ion fluxes. *Plant Signal Behav.* 2011;6:1855–7. <https://doi.org/10.4161/psb.6.11.17014>.

- Drew L. The growth of tea. *Nature*. 2019;566:S2–4. <https://doi.org/10.1038/d41586-019-00395-4>.
- Eisenstein M. Tea's value as a cancer therapy is steeped in uncertainty. *Nature*. 2019;566:S6–7. <https://doi.org/10.1038/d41586-019-00397-2>.
- Foolad MR, Panthee DR. Marker-assisted selection in tomato breeding. *Crit Rev Plant Sci*. 2012;31:93–123. <https://doi.org/10.1080/07352689.2011.616057>.
- Foria S, Magris G, Copetti D, Coleman C, Morgante M, Di Gaspero G. Indel markers for monitoring the introgression of downy mildew resistance from wild relatives into grape varieties. *Mol Breed*. 2018;38:124. <https://doi.org/10.1007/s11032-018-0880-4>.
- Gilbert N. The science of tea's mood-altering magic. *Nature*. 2019;566:S8–9. <https://doi.org/10.1038/d41586-019-00398-1>.
- Hanks SK, Hunter T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J*. 1995;9:576–96. <https://doi.org/10.1096/fasebj.9.8.7768349>.
- Hassan MZ, Rahim MA, Natarajan S, Robin AHK, Kim HT, Park JJ, et al. Gumy stem blight resistance in melon: inheritance pattern and development of molecular markers. *Int J Mol Sci*. 2018;19:2914. <https://doi.org/10.3390/ijms19102914>.
- Hassan O, Kim S, Kim K, Chang T. First report of leaf anthracnose caused by *Colletotrichum camelliae* on tea plants (*Camellia sinensis*) in South Korea. *Plant Dis*. 2023;107:2881. <https://doi.org/10.1094/PDIS-11-22-2622-PDN>.
- Hayashi K, Yoshida H, Ashikawa I. Development of PCR-based allele-specific and indel marker sets for nine rice blast resistance genes. *Theor Appl Genet*. 2006;113:251–60. <https://doi.org/10.1007/s00122-006-0290-6>.
- Islam MR, Hossain MR, Jesse DMI, Jung HJ, Kim HT, Park JJ, et al. Development of molecular marker linked with bacterial fruit blotch resistance in melon (*Cucumis melo* L.). *Genes*. 2020;11:220. <https://doi.org/10.3390/genes11020220>.
- Jander G, Norris SR, Rounsley SD, Bush DF, Levin IM, Last RL. Arabidopsis map-based cloning in the post-genome era. *Plant Physiol*. 2002;129:440–50. <https://doi.org/10.1104/pp.003533>.
- Jin JQ, Liu YF, Ma CL, Ma JQ, Hao WJ, Xu YX, et al. A novel f3'5'h allele with 14 bp deletion is associated with high catechin index trait of wild tea plants and has potential use in enhancing tea quality. *J Agric Food Chem*. 2018;66:10470–8. <https://doi.org/10.1021/acs.jafc.8b04504>.
- Jo J, Venkatesh J, Han K, Lee HY, Choi GJ, Lee HJ, et al. Molecular mapping of *PMR1*, a novel locus conferring resistance to powdery mildew in pepper (*Capsicum annuum*). *Front Plant Sci*. 2017;8:2090. <https://doi.org/10.3389/fpls.2017.02090>.
- Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res*. 2008;36:D480–4. <https://doi.org/10.1093/nar/gkm882>.
- Khan RS, Ntui VO, Chin DP, Nakamura I, Mii M. Production of marker-free disease-resistant potato using isopentenyl transferase gene as a positive selection marker. *Plant Cell Rep*. 2011;30:587–97. <https://doi.org/10.1007/s00299-010-0974-x>.
- Kim S, Kim CW, Choi MS, Kim S. Development of a simple PCR marker tagging the *Allium roylei* fragment harboring resistance to downy mildew (*Peronospora destructor*) in onion (*Allium cepa* L.). *Euphytica*. 2016;208:561–9. <https://doi.org/10.1007/s10681-015-1601-2>.
- Kim B, Hwang IS, Lee HJ, Oh CS. Combination of newly developed SNP and InDel markers for genotyping the *Cf-9* locus conferring disease resistance to leaf mold disease in the tomato. *Mol Breed*. 2017;37:59. <https://doi.org/10.1007/s11032-017-0663-3>.
- Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. 2019;37:907–15. <https://doi.org/10.1038/s41587-019-0201-4>.
- Kim D, Chen DQ, Ahsan N, Jorge GL, Thelen JJ, Stacey G. The raf-like MAPKKK integrin-linked kinase 5 regulates purinergic receptor-mediated innate immunity in Arabidopsis. *Plant Cell*. 2023;35:1572–92. <https://doi.org/10.1093/plcell/koad029>.
- Kumar S, Tripathi D, Okubara PA, Tanaka K. Purinoceptor P2K1/DORN1 enhances plant resistance against a soilborne fungal pathogen, *Rhizoctonia solani*. *Front Plant Sci*. 2020;11:572920. <https://doi.org/10.3389/fpls.2020.572920>.
- Kumari S, Nazir F, Maheshwari C, Kaur H, Gupta R, Siddique KHM, et al. Plant hormones and secondary metabolites under environmental stresses: enlightening defense molecules. *Plant Physiol Biochem*. 2024;206:108238. <https://doi.org/10.1016/j.plaphy.2023.108238>.
- Li X, Li MH, Deng WW, Ahmmed GJ, Wei JP, Yan P, et al. Exogenous melatonin improves tea quality under moderate high temperatures by increasing epigallocatechin-3-gallate and theanine biosynthesis in *Camellia sinensis* L. *J Plant Physiol*. 2020;253:153273. <https://doi.org/10.1016/j.jplph.2020.153273>.
- Liao Y, Smyth GK, Shi W. Featurecounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30:923–30. <https://doi.org/10.1093/bioinformatics/btt656>.
- Liu B, Wang Y, Zhai W, Deng J, Wang H, Cui Y, et al. Development of indel markers for *Brassica rapa* based on whole-genome re-sequencing. *Theor Appl Genet*. 2013;126:231–9. <https://doi.org/10.1007/s00122-012-1976-6>.
- Liu L, Dang PM, Chen CY. Development and utilization of indel markers to identify peanut (*Arachis hypogaea*) disease resistance. *Front Plant Sci*. 2015;6:988. <https://doi.org/10.3389/fpls.2015.00988>.
- Liu G, Xu JH, Zhang M, Li PF, Yao XF, Hou Q, et al. Exploiting illumina sequencing for the development of InDel markers in watermelon (*Citrullus lanatus*). *J Hortic Sci Biotechnol*. 2016;91:220–6. <https://doi.org/10.1080/14620316.2015.1133609>.
- Liu G, Zhao T, You X, Jiang J, Li J, Xu X. Molecular mapping of the *Cf-10* gene by combining SNP/InDel-index and linkage analysis in tomato (*Solanum lycopersicum*). *BMC Plant Biol*. 2019a;19:15. <https://doi.org/10.1186/s12870-018-1616-7>.
- Liu S, An Y, Tong W, Qin X, Samarina L, Guo R, et al. Characterization of genome-wide genetic variations between two varieties of tea plant (*Camellia sinensis*) and development of InDel markers for genetic research. *BMC Genomics*. 2019b;20:935. <https://doi.org/10.1186/s12864-019-6347-0>.
- Liu Y, Hao X, Lu Q, Zhang W, Zhang H, Wang L, et al. Genome-wide identification and expression analysis of flowering-related genes reveal putative floral induction and differentiation mechanisms in tea plant (*Camellia sinensis*). *Genomics*. 2020;112(3):2318–26. <https://doi.org/10.1016/j.ygeno.2020.01.003>.
- Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔCt} method. *Methods*. 2001;25:402–8. <https://doi.org/10.1006/meth.2001.1262>.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550. <https://doi.org/10.1186/s13059-014-0550-8>.
- Lv HH, Yang LM, Kang JG, Wang QB, Wang XW, Fang ZY, et al. Development of InDel markers linked to fusarium wilt resistance in cabbage. *Mol Breed*. 2013;32:961–7. <https://doi.org/10.1007/s11032-013-9925-x>.
- Ma S, An C, Lawson AW, Cao Y, Sun Y, Tan EYJ, et al. Oligomerization-mediated autoinhibition and cofactor binding of a plant NLR. *Nature*. 2024;632:869–76. <https://doi.org/10.1038/s41586-024-07668-7>.
- Mills RE, Luttig CT, Larkins CE, Beauchamp A, Tsui C, Pittard WS, et al. An initial map of insertion and deletion (INDEL) variation in the human genome. *Genome Res*. 2006;16:1182–90.
- Mishra A, Wickneswari R, Bhuiyan MAR, Jena KK, Shamsudin NA. Broad spectrum blast resistance alleles in newly developed Malaysian rice (*Oryza sativa* L.) genotypes. *Euphytica*. 2021;217:8. <https://doi.org/10.1007/s10681-020-02738-z>.
- Mphangwe NIK, Vorster J, Steyn JM, Nyirenda HE, Taylor NJ, Apostolides Z. Screening of tea (*Camellia sinensis*) for trait-associated molecular markers. *Appl Biochem Biotechnol*. 2013;171:437–49. <https://doi.org/10.1007/s12010-013-0370-4>.
- Nvsvrot T, Xia W, Xiao Z, Zhan C, Liu M, Yang X, et al. Combining QTL mapping with genome resequencing identifies an Indel in an *R* gene that is associated with variation in leaf rust disease resistance in poplar. *Phytopathology*. 2020;110:900–6. <https://doi.org/10.1094/PHYTO-10-19-0402-R>.
- Orrock JM, Rathinasabapathi B, Richter BS. Anthracnose in US tea: pathogen characterization and susceptibility among six tea accessions. *Plant Dis*. 2020;104:1055–9. <https://doi.org/10.1094/PDIS-07-19-1518-RE>.
- Pham AQ, Cho SH, Nguyen CT, Stacey G. *Arabidopsis* lectin receptor kinase P2K2 is a second plant receptor for extracellular ATP and contributes to innate immunity. *Plant Physiol*. 2020;183:1364–75. <https://doi.org/10.1104/pp.19.01265>.
- Qiu T, Cao FR. Leveraging the advantages of Lingtong single-variety tea industry to promote the rural revitalization of Lingtong Village. *Guangdong Tea Industry*. 2022;4:32–5 (Chinese).
- Rao IS, Srikanth B, Kishore VH, Suresh PB, Chaitanya U, Vemireddy LR, et al. Indel polymorphism in sugar translocation and transport genes associated with grain filling of rice (*Oryza sativa* L.). *Mol Breed*. 2011;28:683–91. <https://doi.org/10.1007/s11032-011-9618-2>.

- Shan Y, Zhang D, Luo Z, Li T, Qu H, Duan X, et al. Advances in chilling injury of postharvest fruit and vegetable: Extracellular ATP aspects. *Compr Rev Food Sci Food Saf*. 2022;21:4251–73. <https://doi.org/10.1111/1541-4337.13003>.
- Tanaka K, Gilroy S, Jones AM, Stacey G. Extracellular ATP signaling in plants. *Trends Cell Biol*. 2010;20:601–8. <https://doi.org/10.1016/j.tcb.2010.07.005>.
- Thakur O, Randhawa GS. Identification and characterization of SSR, SNP and InDel molecular markers from RNA-seq data of guar (*Cyamopsis tetragonoloba*, L. Taub.) roots. *BMC Genomics*. 2018;19(1):951. <https://doi.org/10.1186/s12864-018-5205-9>.
- Tian Z, Yan C, Qian Q, Yan S, Xie H, Wang F, et al. Development of gene-tagged molecular markers for starch synthesis-related genes in rice. *Chin Sci Bull*. 2010;55:3768–77. <https://doi.org/10.1007/s11434-010-4124-1>.
- Uzunova M, Ecke W, Weissleder K, Röbbelen G. Mapping the genome of rapeseed (*Brassica napus* L.). I. Construction of an RFLP linkage map and localization of QTLs for seed glucosinolate content. *Theor Appl Genet*. 1995;90:194–204. <https://doi.org/10.1007/BF00222202>.
- Varet H, Brillet-Guéguen L, Coppée JY, Dillies MA. SARTools: a DESeq2- and EdgeR-based R pipeline for comprehensive differential analysis of RNA-seq data. *PLoS ONE*. 2016;11:e0157022. <https://doi.org/10.1371/journal.pone.0157022>.
- Wang L, Wang Y, Cao H, Hao X, Zeng J, Yang Y, et al. Transcriptome analysis of an anthracnose-resistant tea plant cultivar reveals genes associated with resistance to *Colletotrichum camelliae*. *PLoS ONE*. 2016;11:e0148535. <https://doi.org/10.1371/journal.pone.0148535>.
- Wang Y, Liu H, Li H, Teng RM, Zhuang J. Genome-based identification and analysis of the genes involved in auxin biosynthesis and signal transduction during tea plant leaf development. *Sci Hortic*. 2020;261:109030. <https://doi.org/10.1016/j.scienta.2019.109030>.
- Wang M, Li J, Liu X, Liu C, Qian J, Yang J, et al. Characterization of key odorants in lingtong dancong oolong tea and their differences induced by environmental conditions from different altitudes. *Metabolites*. 2022;12:1063. <https://doi.org/10.3390/metabo12111063>.
- Wei C, Yang H, Wang S, Zhao J, Liu C, Gao L, et al. Draft genome sequence of *Camellia sinensis* var. *sinensis* provides insights into the evolution of the tea genome and tea quality. *Proc Natl Acad Sci U S A*. 2018;115:E4151–8. <https://doi.org/10.1073/pnas.1719622115>.
- Wu PF, Gong HE, Yao XH, Chang WX, Wang KL. Evaluating the resistance of *Camellia oleifera* clones to *Anthraco*se. *Forestry Science Research*. 2023;31:158–63. <https://doi.org/10.13275/j.cnki.lykxyj.2018.04.022>.
- Zhao Y, Wang K, Wang WL, Yin TT, Dong WQ, Xu CJ. A high-throughput SNP discovery strategy for RNA-seq data. *BMC Genomics*. 2019;20:160. <https://doi.org/10.1186/s12864-019-5533-4>.
- Zheng Q, Guo L, Huang J, Hao X, Li X, Li N, et al. Comparative transcriptomics provides novel insights into the mechanisms of selenium accumulation and transportation in tea cultivars (*Camellia sinensis* (L.) O. Kuntze). *Front Plant Sci*. 2023;14:1268537. <https://doi.org/10.3389/fpls.2023.1268537>.
- Zhou G, Zhang Q, Tan C, Zhang XQ, Li C. Development of genome-wide indel markers and their integration with SSR, DArT and SNP markers in single barley map. *BMC Genomics*. 2015;16:804. <https://doi.org/10.1186/s12864-015-2027-x>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.