Whole-genome resequencing analysis of the Medicinal Plant *Gardenia jasminoides*

Xinyu Xu[1,2], Bihua Chen[1*], Juan Zhang[1], Siren Lan[2], Shasha Wu[2]

[1]Fujian Academy of Forestry Sciences, Fuzhou Fujian 350012, China
[2]College of Landscape and Architecture, Fujian Agriculture and Forestry University, Fuzhou Fujian 350002, China

Xinyu Xu, Email: 18120872516@163.com. 1. Fujian Academy of Forestry Sciences, Fuzhou Fujian 350012, China; 2. College of Landscape and Architecture, Fujian Agriculture and Forestry University, Fuzhou Fujian 350002, China.
Bihua Chen, Email: chenbihua.happy@163.com. 1. Fujian Academy of Forestry Sciences, Fuzhou Fujian 350012, China.
Juan Zhang, Email: 262026619@qq.com. 1. Fujian Academy of Forestry Sciences, Fuzhou Fujian 350012, China.
Siren Lan, Email: lkzx@fafu.edu.cn. 2. College of Landscape and Architecture, Fujian Agriculture and Forestry University, Fuzhou Fujian 350002, China.
Shasha Wu, Email: shashawu1984@126.com. 2. College of Landscape and Architecture, Fujian Agriculture and Forestry University, Fuzhou Fujian 350002, China.

Corresponding Author:
Bihua Chen
Fujian Academy of Forestry Sciences, No.35 Shangchiqiao, Xindian Town, Jin'an District, Fuzhou City, Fujian Province, 350012, China
Email address: chenbihua.happy@163.com

Abstract

Background: *Gardenia jasminoides* is a species of Chinese medicinal plant, which has high medicinal and economic value and rich genetic diversity, but the study on its genetic diversity is far not enough.

Methods: In this study, one wild and one cultivated gardenia materials were resequenced using *Illumina* HiSeq sequencing platform and the data were evaluated to understand the genomic characteristics of *G. jasminoide*s.

Results: After data analysis, the results showed that clean data of 11.77G, Q30 reached 90.96%. The average comparison rate between the sample and reference genome was 96.08%, the average coverage depth was 15x, and the genome coverage was 85.93%. The SNPs of FD and YP1 were identified, and 3,087,176 and 3,241,416 SNPs were developed. In addition, SNP non-

synonymous mutation, InDel mutation, SV mutation and CNV mutation were also detected between the sample and the reference genome, and KEGG, GO and COG database annotations were made for genes with DNA level variation. The structural gene variation in the biosynthetic pathway of crocin and gardenia, the main medicinal substance of *G. jasminoides* was further explored, which provided basic data for molecular breeding and genetic diversity of *G. jasminoide*s in the future.

Keywords: *Gardenia jasminoides*, whole genome-resequencing, SNP, InDel, SV, CNV, Crocin, geniposide

Introduction

Gardenia jasminoides Ellis (2n=22) is an evergreen shrub of the genus *Gardenia* in the Rubiaceae family. It is distributed in many areas of China and is mainly used in the cultivation of medicinal plants and garden landscape. *G. jasminoides* is known for its special aroma and is an excellent greening tree (*Xiao et al., 2017*). Gardenia has very high medicinal value, and its fruits, leaves, flowers and roots can be used as medicine. It is an ideal Chinese medicinal plant (*Zhao et al., 2017; Lu et al., 2019; Zhang, Bian & Yao, 2022*). The chemical components of gardenia fruits are mainly iridoids, including gardenia and gardenia acid, among which gardenia is one of the most active components (*Long et al., 2013; Xu et al.,2016; Zhou et al., 2021*). In addition, gardenia also contains carotenoid compounds, such as crocin, which can regulate the neural center, improve memory and cognition (*Tian et al., 2020; Shen et al., 2022*), and phenolic acid compounds, mainly chlorogenic acid, which can lower blood pressure, antioxidant, free radical scavenging and regulate the body's immunity (*Liu et al., 2020*).With the continuous research, development and utilization, gardenia has become the first batch of edible and medicinal plants in China. Gardenia yellow can be used as natural food colorant and food coloring agent (*Wu et al., 2022*). In Thailand, *G. jasminoides* is used in religious rituals and hair decorations, and in other Asian countries, *G. jasminoides* is even used in cooking (*Chen et al., 2020a; Chen et al., 2020b*).

In recent years, whole genome resequencing has been conducted on many plants, such as soybean, rice and sorghum (*Long et al., 2022; Wang et al., 2022*). In order to detect various types of genetic variation and improve and domesticate genes better, whole genome sequencing on 292 pieces of pigeonpea were conducted and found several genome regions that might be targets for domestication and breeding. Then, through genome-wide association analysis, it was found that the genes related to flowering time control, seed development and pod cracking had sequence similarities with other plants. Meanwhile, some trait selection areas were consistent with geographical distribution, suggesting that different geographical environments led to the adaptive changes of some traits in pigeonpea (*Varshney et al., 2017*). On the other hand, some studies have provided a new perspective for the evolution of loquat genome and fruit domestication by resequencing wild loquat (*Jing et al., 2023*). In addition, resequencing analysis and KEGG enrichment analysis were used to identify differences in nitrite metabolism between 'Linhuang No.1' and 'Muzao' cultivars (*Li et al., 2021*). Genomes including heavy single

Formatted: Font: Italic

Formatted: Font: Not Italic

nucleotide polymorphisms (~~single nucleotide polymorphism,~~ SNPs) and insertions or deletions (~~insertion/deletion,~~ Indel) through high-throughput were sequenced to detect complex genetic variation~~,~~. ~~F~~for example, structure variation (SV) and copy number variation (CNV)~~, etc~~. (*Ganal, Altmann & Röder. 2009; Adedze et al., 2021; Ma et al., 2021*)~~.~~, A~~a~~ total of 520,260 SNP variations were identified after sequencing the whole genome of 18 flue-cured virginia (FCV) tobaccos (*Nicotiana tabacum*) ~~.FCVS.A total of~~ 4849 homozygotes and 28,584 polynucleotide polymorphisms were also detected, which are critical for the development of excellent tobacco breeding lines (*Thimmegowda et al., 2018*). InDel markers based on whole genome resequencing have attracted more and more attention due to their advantages of wide distribution, high density, stable variation, strong polymorphism and easy detection in the genome. For example, different InDel markers have been developed based on resequencing data in terms of watermelon peel traits and color (*Li et al., 2018*). InDel markers were used to identify genetic loci related to the regulation of peanut bacterial wilt resistance genes (*Zhang et al., 2022*). From the development of gardenia genome, it was found that the functional genes ~~in gardenia fruit~~ evolved through tandem gene replication ~~evolved~~ in gardenia fruit (*Xu et al., 2020*). At present, there are few reports on genetic diversity analysis of germplasm resources of gardenia, and due to the differences or incomplete species of the gardenia population selected by various researchers, there are abundant variation conditions among wild populations, the polymorphism between and within the species is not completely consistent~~., so there are some contradictions in the analysis results.~~ the research results can not really reflect the current situation of genetic resources of this species. At the same time, due to the lack of genomic data, there are few stable and reliable molecular markers, which greatly limits the progress of genetic breeding of gardenia.

In this study, we employed whole-genome resequencing technology to align the genomes of one ~~two~~ wild type~~s~~ and one cultivated variety of *G. jasminoides* to the reference genome of *G. jasminoides*. Based on this, we conducted differential analysis at both individual and population levels, aiming to explore the differences and structural variations in gene sequences. This analysis included InDel detection and annotation, SV detection and annotation, distribution of all variations on the genome, and analysis of genes involved in the variation of gardenoside and crocin. Functional enrichment analysis was performed to annotate the variant genes, laying the foundation for the genetic study of *G. jasminoides*.

Materials and Methods

2.1. Sample collection, DNA extraction and sequencing

The sample of *G. jasminoides* (FD) used for genome resequencing analysis is an improved variety cultivated in Fuding City, Fujian Province, it was selected by the fruit indexes of different cultivated varieties, and then obtained by asexual propagation. ~~and g~~Gardenia YP1 is a wild *G. jasminoides* in Yanping, Fujian Province. 20 fresh leaves of gardenia FD and 20 fresh leaves of Gardenia YP1 were collected at flowering stage, and genomic DNA was extracted from young leaves by cetyltrimethyl ammonium bromide (CTAB) (*Madhukar et al., 2020*). The

Formatted: Font: Times, 12 pt

Formatted: Font: Times, 12 pt, Italic

Formatted: Font: Times, 12 pt

Formatted: Font: Times, 12 pt, (Asian) Chinese (China)

Commented [TJ1]: is this the start of a new sentence? Uppercase 'The'

Why can't the result reflect the current situation? This is somewhat vague.

Commented [TJ2]: crocin production?

Commented [TJ3]: why lowercase 'g' for gardenia FD but uppercase for Gardenia YP1.

concentration and quality of total genomic DNA were determined by NanoDrop 2000 spectrophotometer (Thermo Fisher Science), and the DNA library of *Illumina*/BGI sequencing (350 bp) was constructed. After the library was constructed, the library was sequenced on the *Illumina* HiSeq X Ten/Nova Seq/BGI platform and read as 150 bp. ~~The original reading length is screened according to the following criteria: with > 10%: paired-end readings with bases; readings with more than 50% base mass fraction less than 20 (Phred-like score); and sequencing connectors.~~ Finally, high-quality sequences are obtained for follow-up analysis.

2.2. Tools for variation analysis

The detection of SNP and small InDel is mainly realized by GATK (v4.1) (*Liu et al, 2022*) software kit. Use SAMtools (v1.9) (*Danecek et al., 2021*) to filter redundant reads to ensure the accuracy of the test results. Then the local haplotype assembly (Haplotype Caller) algorithm of GATK is used to detect the mutation of SNP and InDel. Each sample first generates gVCF, and then carries on the population joint-genotype. In order to ensure the reliability of the mutation results, based on the subroutine vcfutils. pl (varFilter-w 5-W 10) in bcftools, ~~the~~ SNPs within ~~the~~ 5bp ~~near~~ of an INDEL and within 10bp if there is an ~~the~~ adjacent INDEL, ~~in the 10bp~~ are filtered out. The filtering parameters are QUAL < 30, QD < 2.0, MQ < 40, feeds FS > 60.0. Other variation filtering parameters are processed by the default values officially specified by GATK, and the final set of variation sites is obtained. SV mutation detection uses Manta (v1.6) (*Kosugi et al., 2019*) software to detect insertion (Insertion, INS), deletion (Deletion, DEL), inversion (Inversion, INV), chromosome translocation (Translocation, TRA) between sample and reference genome based on the relationship between Pair-end reads alignment and reference genome and actual Insert size. CNV detection uses FREEC (*Wei, Dugas & Sandmann, 2021*) to detect the depth distribution of reads on the reference genome by sample sequencing.

2.3. Alignment of reference genomes

The original reads (double-segment sequence) was analyzed by de-splicing and low-quality reads filtration analysis, and finally ~~C~~clean ~~R~~reads ~~was~~ were obtained for mutation gene detection. The ~~C~~clean ~~R~~reads of FD (GenBank BioSample: SAMN35881984) and YP1 (GenBank BioSample: SAMN35881985) of gardenia were compared with the reference genome. (https://www.ncbi.nlm.nih.gov/genome/?term=Gardenia_jasminoides, ie. GenBank: assembly ASM1310374v1) (*Xu et al., 2020*). TBtools (*Chen et al., 2020*) was used to search for the required variant gene information in the reference genome, and then compared and analyzed on NCBI.

2.4. Functional Annotation of the variant Gene of GO, COG, KEGG

Through the comparison of the reference genome, BLAST compared the mutant genes with the functional database such as GO, COG, KEGG (*Altschul et al., 1997; Kanehisa et al., 2004; Gene Ontology Consortium, 2015*) to get the annotations of these genes to analyze the gene function. The screening criteria for enrichment analysis are P-value≤10-5 and False discovery rate (FDR)~~FDR~~≤0.01, respectively.

Results

3.1. Analysis of the results compared with the reference genome

In this study, a total of ~~635.6~~534.9 Mb ~~was~~were ~~obtained~~mapped ~~according~~ to the reference genome (*Xu et al., 2020*). The sequencing data of the database were filtered after quality control. The average value of each sample was about 39 million ~~39.00 Mb C~~clean ~~R~~reads, 11.77 Gbp ~~C~~clean ~~D~~data, Q20, 96.~~68~~86%, 96.49% and Q30, 91.34%, 90.~~96~~58% (Table 1), and the average coverage depth of each sample was 10x. Compared with the reference genome, the percentage of ~~C~~clean ~~R~~reads mapped to the reference genome by FD ~~to all C~~clean ~~R~~reads ~~~~ is 95.57%. The comparison result is good (Table 2), and the ~~average~~ proportion of genome coverage corresponding to each depth of FD (1x, 5x, 10x) is 85.~~93~~02%, ~~69~~71.~~20~~13%, 4~~1~~6.0~~0~~2%, YP1 is 86.84%, 67.26%, 35.96% (Table S1). According to the coverage depth of each point of the chromosome, it was found that the coverage depth was evenly distributed on the chromosome (Fig. S1a, Fig. S1b), indicating that the genome was evenly covered and the sequencing was random.

3.2. Mutation detection and annotation

According to the mapping results of clean reads in the reference genome, the SNP, InDel, SV and CNV mutations of two samples FD and YP1 were detected. For SNP mutation, 3,087,176 and 3,241,416 high quality SNP sites were obtained respectively.~~The SNP, InDel, SV and CNV of the two samples FD and YP1 were detected and functionally annotated by the re-sequencing results. 3087176 and 3241416 high quality SNP sites were obtained respectively after detection. Then the SNP, InDel, SV, CNV of the two samples (FD, YP1) were detected and functionally annotated by the re-sequencing results, and 3,087,176 and 3,241,416 high quality SNP sites were obtained respectively.~~ Then the SNP was annotated, in which the sample FD conversion and subversion ratio was 1.87, the sample YP1 conversion and subversion ratio was 1.89, and the heterozygous SNP ratio FD was 49.64% and 67.07%. The heterozygosity of sample FD is lower than that of sample YP1 (Table S2). From the Venn map between samples, it was found that there were 1,737,580 SNP variants in FD and YP1, 1,503,846 SNP variants specific to YP1, and 1,349,606 in FD (Fig. ~~1~~S2). The SNP mutations of the whole genome can be divided into six types, among which the number of T:A>C:G in YP1 is more than that in FD (Fig. S~~2~~3). The results of SNP annotation showed that there was no significant difference in SNP variation between FD and YP1 (Fig. ~~2~~1a, Fig. ~~2~~1b). After that, the InDels~~insertions and deletions~~ of small fragments between the sample and the reference genome were detected. The numbers of InDels~~insertions and deletions~~ of FD CDS and genome fragments were 10,270 and 558,766, the YP1 is 10,752 and 540,581 ~~respectively~~. ~~T~~the number of InDels ~~insertions and deletions~~ in the coding region of YP1 was more than that of FD (Fig. S~~3~~4), and the number of homozygous Indel of FD in the whole genome and coding region was higher than that of YP1, and the number of YP1 heterozygous Indel in the whole genome and coding region was higher than that of FD (Table S3, Fig.~~3~~2a, Fig. ~~3~~2b). In addition, the genomic ~~structure variation (~~SV~~)~~ of the two samples was detected and annotated. There were 11,543 and 7,634 SV variants in FD and YP1. The number of insertion type variation, deletion type variation, inversion type variation, repetition type variation and chromosome translocation type variation in FD were higher than those in YP1. The annotation results of SV showed that the variation of FD in the gene was more

Commented [TJ4]: across two samples?

Commented [TJ5]: you're reporting data for two samples here in the text without stating that.

than that in the intergenic region, and the insertion and deletion in the intergenic region was much higher than that in the gene, while the reversal and repetition were much lower than those in the gene. In the annotation of CNV, 885 variant genes, 956 lost genes, 840 predicted single copy number, 885 variant genes, 780 deletion genes and 770 predicted single copy number were obtained by FD substitution type and YP1 substitution type, respectively. The number of variation obtained by FD and YP1 was the same, and the number of lost genes was more than YP1, and the number of single copy was also more than that of YP1 (Table S4).

Variation in the CDS region may cause changes in gene function. By looking for genes with non-synonymous mutations in SNP, InDel and SV in the CDS region between the reference genome and the sample, the genes were found that may have functional differences between the sample and the reference genome. Among the differential gene variations in FD, the type with the largest number of mutations was SNP, with a total of 23,338 variations. In YP1, the number of differential genes was also the most in SNP, with 24,631, in which the number of InDel mutants in YP1 was more than that in FD, while the number of SV variations was on the contrary, and SV variations in FD were more than YP1 (Table 3, Fig. ~~2~~1).

3.3. Functional annotation at the DNA level

Variations in the CDS region may cause changes in gene function, ~~T~~the variation genes of FD and YP1 were compared with GO, COG functional database, and these gene annotations were obtained to analyze gene function. In the GO annotation map of the variant genes of FD and YP1 (Fig. 43a, Fig. 43b), they are divided into three categories, namely biological process, cellular component and molecular function, in which the variant genes of FD and YP1 are expressed in 16 terms of biological process, of which FD annotated 3,049 GO items, 7,771 genes, YP1 annotated 3,061 items and 7,881 genes. The significantly rich items of FD and YP1 are biological process (GO:0008150), FD and YP1 have 1,752 gene mutations, cellular process (GO:0009987), FD and YP1 have 1,328 gene variation. The variant genes of FD and YP1 are expressed in 16 terms of cellular component, of which FD annotates 607 GO items, 7,599 genes, YP1 annotates 605 items and 7,788 genes. The most significant rich item of FD and YP1 is integral component of membrane (GO:0016021), which has 6,148 gene mutations. The mutant genes of FD and YP1 were expressed in 10 terms of molecular function, including 1,139 GO entries and 14,339 genes annotated by FD and 1,147 items and 18,585 genes by YP1. The most significant rich item of FD and YP1 was ATP binding (GO:0005524), which had 2,639 gene variations.

COG database can reflect the lineal homologous classification of gene products. COG functional classification map analysis showed that FD and YP1 were annotated in 24 categories (Fig. 43c, Fig. 43d), of which the richer categories were carbohydrate transport and metabolism, general function prediction only, secondly metabolites biosynthesis and transport and catabolism, and the least annotated category was RNA processing and modification.

3.4. Mining of variant genes of Crocin and Geniposide

Gardenia fruit will accumulate a large number of chemicals in the process of development, among which geniposide and Crocin are the main medicinal components. In order to better

Formatted: Font: Times, 12 pt

explore the related variant genes of Crocin and geniposide in cultivated species FD and wild species YP1, the structural genes in the biosynthesis pathway of Crocin and geniposide (Fig. S~~5~~5a, Fig. S~~5~~5b) were enriched by KEGG. The structural gene GjCCD4a (Gj9A597T69) of Crocin was annotated on Carotenoid biosynthesis (ko00906) (Table S5). In this pathway, FD annotated 91 genes, including 79 variant genes~~mutant genes~~, YP1 annotated 91 genes and 78 variant genes. The differential genes were annotated by Gj11A235T45, GjALDH2C3 (Gj9A24T70) in Phenylpropanoid biosynthesis (ko00940). FD annotated a total of 566 genes and 456 mutant genes. YP1 annotated 566 genes, 456 mutant genes, GjUGT94E13 (Gj9P1027T10) annotated on two KEGG pathways, Flavonoid biosynthesis (ko00941), Flavone and flavonol biosynthesis (ko00944), FD annotated 174 genes on Flavonoid biosynthesis, 134 genes were variant genes, YP1 annotated 174 genes, 136 genes were variant genes, FD annotated 45 genes on Flavone and flavonol biosynthesis, of which 37 genes were variant genes, YP1 annotated 40 genes, and 45 genes were variant genes.

The key structural genes GES (Gj7A350T74), G10H (Gj9A674T167), 10-HGO (Gj9P312T6) and IS (Gj9A1001T119) for geniposide synthesis were identified by TBtools comparison, in which GES (Gj7A350T74) was not annotated on the KEGG pathway, G10H (Gj9A674T167) was annotated on Monoterpenoid biosynthesis (ko00902), a total of 70 genes were annotated, FD annotated 59 mutant genes and YP1 annotated 58 mutant genes. 10-HGO (Gj9P312T6) annotates 96, 355219121 genes in Fatty acid degradation (ko00071), Carbon metabolism (ko00071), Glycolysis / Gluconeogenesis (ko00010), Tyrosine metabolism (ko00350), FD annotates 79 mutant genes on Fatty acid degradation (ko00071) pathway, YP1 annotates 80 variant genes, FD annotates 292 variant genes on Carbon metabolism (ko00071) pathway, YP1 annotates 291 variant genes on Glycolysis / Gluconeogenesis (ko00010) pathway. FD annotated 17_6variant genes and YP1 annotated 173_variant genes. In the Tyrosine metabolism (ko00350) pathway, FD annotated 97 variant genes, YP1 annotated 96 variant genes, and IS (Gj9A1001T119) did not annotate the synthetic pathway of KEGG.

Commented [TJ6]: enriched how? Do you mean identified?

Commented [TJ7]: ?? that's a lot of genes

~~**3.5. Mining of variant genes of Crocin and Geniposide**~~

~~Gardenia fruit will accumulate a large number of chemicals in the process of development, among which geniposide and Crocin are the main medicinal components. In order to better explore the related variant genes of Crocin and geniposide in cultivated species FD and wild species YP1, the structural genes in the biosynthesis pathway of Crocin and geniposide (Fig. 5 a) were enriched by KEGG. The structural gene GjCCD4a (Gj9A597T69) of Crocin was annotated on Carotenoid biosynthesis (ko00906). In this pathway, FD annotated 91 genes, including 79 mutant genes, YP1 annotated 91 genes and 78 variant genes. The differential genes were annotated by Gj11A235T45, GjALDH2C3 (Gj9A24T70) in Phenylpropanoid biosynthesis (ko00940). FD annotated a total of 566 genes and 456 mutant genes. YP1 annotated 566 genes, 456 mutant genes, GjUGT94E13 (Gj9P1027T10) annotated on two KEGG pathways, Flavonoid biosynthesis (ko00941), Flavone and flavonol biosynthesis (ko00944), FD annotated 174 genes on Flavonoid biosynthesis, 134 genes were variant genes, YP1 annotated 174 genes, 136 genes~~

~~were variant genes, FD annotated 45 genes on Flavone and flavonol biosynthesis, of which 37 genes were variant genes, YP1 annotated 40 genes, and 45 genes were variant genes.~~
~~The key structural genes GES (Gj7A350T74), G10H (Gj9A674T167), 10-HGO (Gj9P312T6) and IS (Gj9A1001T119) for geniposide synthesis were identified by TBtools comparison, in which GES (Gj7A350T74) was not annotated on the KEGG pathway, G10H (Gj9A674T167) was annotated on Monoterpenoid biosynthesis (ko00902), a total of 70 genes were annotated, FD annotated 59 mutant genes and YP1 annotated 58 genes. 10-HGO (Gj9P312T6) annotates 96, 355219121 genes in Fatty acid degradation (ko00071), Carbon metabolism (ko00071), Glycolysis / Gluconeogenesis (ko00010), Tyrosine metabolism (ko00350), FD annotates 79 mutant genes on Fatty acid degradation (ko00071) pathway, YP1 annotates 80 variant genes, FD annotates 292 variant genes on Carbon metabolism (ko00071) pathway, YP1 annotates 291 variant genes on Glycolysis / Gluconeogenesis (ko00010) pathway. FD annotated 176variant genes and YP1 annotated 173variant genes. In the Tyrosine metabolism (ko00350) pathway, FD was annotated 97 variant genes, YP1 was annotated 96 variant genes, and IS (Gj9A1001T119) did not annotate the synthetic pathway of KEGG.~~

Discussion

G. jasminoides is widely distributed in the tropics of Southeast Asia, but it has been cultivated in China for at least 1000 years. Gardenia population is rich in genetic diversity, and there is a close relationship between geographical distance and genetic distance between populations (Han et al., 2007; Yang et al., 2016). At present, it is generally believed that there is genetic variation among different populations of gardenia in China, and it has adapted to different living environments in the long-term process of cultivation and domestication (*Mei et al., 2015*). However, the differences and diversity between cultivated and wild populations have not been well explored.

In this study, a total of 39,406,575 clean ~~data~~reads were obtained from ~~two~~one wild and cultivated species, and the average depth of each sample was 1~~5~~0X. A large number of variations such as SNP, Indel, SV and CNV were detected. Compared with other mutations, SNP and InDel have more abundant genetic variation and can better represent the genetic information of the whole genome of gardenia. SNP and InDel variation analysis have been widely used in the fields of population genetic structure, genetic diversity analysis and functional gene mining of non-model organisms (*Chen et al., 2021; Giudice, Bazakos & Vassiliou, 2022*). 3,087,176 SNP variants were detected in cultivated species FD of gardenia, and 3,241,416 SNP variants were detected in YP1. The same was true of the number of SNP variations in wild species YP1 than in cultivated species FD1, InDel. It can be seen that wild species YP1 has higher genetic diversity, and in different *Brassica napus*, it is also found that winter rape has higher genetic diversity because the density of SNP and InDel is higher than that of spring rape (*Wu et al., 2019*). The most abundant type of SNP variation in sample YP1 and FD was the transformation type, and there was almost no difference between the conversion type of FD and the subversion type of SNP and YP1, especially the SNP ratio of heterozygous type of YP1 was much higher than that of FD, which indicated that YP1 had higher heterozygosity than FD and had higher homology

with gardenia material with reference genome, while the cultivated species FD had more homozygous SNP than wild species YP1. It is reported in different varieties of *Camellia sinensis* that the ratio of non-synonymous mutation to synonymous mutation of InDel variation is almost the same as that of long-term interspecific hybridization. It is inferred that to a large extent, these two samples and the reference genome of gardenia do not come from the same region (*An et al., 2020*)~~,~~ .. Ttransition (Ti) and transversion (Tv) are two types of SNP variants, and the proportion between them is related to the evolution of the species (*Mursyidin et al., 2022*),~~but~~ the proportion of Ti/Tv between the two samples is very close, which does not rule out the common origin and ancestor of cultivated species FD and wild species YP1. The number of the main mutation types of FD was also less than that of YP1, and the variations of the two samples were mainly concentrated in a large number of frameshift mutations in the intergenic region. Generally speaking, the variation of FD of cultivated species is relatively conservative, which has lost a lot of genetic diversity while having more stable quality, which is similar to the reported results of pear, cotton and wheat (*Li et al., 2021a; Li et al., 2021b; Hussain et al., 2022*).

This study also explored the variation genes of geniposide and crocetin, which are the main medicinal substances in gardenia fruit. The mining of variation genes related to medicinal substances is helpful to the improvement and application of gardenia genetic germplasm. The key genes in geniposide synthesis pathway are mainly annotated on five pathways, of which Tyrosine metabolism (ko00350) and Monoterpenoid biosynthesis (ko00902) have been shown to be closely related to geniposide synthesis in plants (*Pan et al., 2021*). Tyrosine acts as the precursor of many specialized metabolites, which have a variety of physiological functions. Some plant natural products derived from Tyrosine are also used in human medicine and nutrition, and in some Monoterpenoid plant products, they are used as effective drugs for a variety of diseases (*Schenck & Maeda, 2018; Nguyen et al., 2022*). Not only that Fatty acid degradation (ko00071), Carbon metabolism (ko00071), Glycolysis/Gluconeogenesis (ko00010) are also synthetic precursors of many metabolites (*Newman et al., 2008; Lin et al., 2019; Toleco et al., 2020*). The variant genes with differences in these pathways need to be further explored in the future. One of the important synthetic precursors of Crocin is carotenoid. It is reported that carotenoid biosynthesis pathway has a great influence on the synthesis of Crocin in gardenia (*Shen et al., 2022*), and 30 DEGs have been annotated on the Carotenoid biosynthesis KEGG pathway of gardenia (*Pan et al., 2021*). In addition to Carotenoid biosynthesis pathway, Flavonoid biosynthesis (ko00941) and Flavone and flavonol biosynthesis (ko00944) are important secondary metabolites that exist widely in plants, which not only play an important role in plant growth and development, but also have outstanding applications in food and medicine. The KEGG pathways annotated by CCD4a, ALDH2C3, UGT74F8, UGT94E13 of Crocin structural genes are all involved in regulating pigment metabolism. This is also the main reason for the color change of gardenia after ripening (*Sommano et al., 2020; Liu et al., 2021; Zhang et al., 2023*).

With the intensive artificial selection in pursuit of better quality and higher yield, the cultivated germplasm resources have lost a lot of genetic diversity. It has been reported that in order to

Formatted: Font: Times, 12 pt

Formatted: Font: Times, 12 pt

Formatted: Font: Times, 12 pt

Formatted: Font: Times, 12 pt, Italic, (Asian) Chinese (China)

accelerate the biological research and genetic improvement of tomato, pan-genomic comparison was carried out between cultivated and wild species of tomato to reveal ~~structural variation (~~SVs~~)~~ (*Li et al., 2023*). The number of samples selected in this study is small, and more wild and cultivated species of *G. jasminoides* have not been analyzed, so the genetic difference between samples is not significant. In the future, more germplasm resources are needed to explore the genetic diversity and origin of wild gardenia species, to establish a molecular marker system for the identification of different gardenia germplasm, and further functional characterization of the variation can better understand the genetic basis of the differences between cultivated species and their wild relatives.

Conclusions

The whole genome of one wild and one cultivated species of gardenia was resequenced, and a large number of SNPs, InDels, CNVs and SVs mutations were detected. SNP and InDel variation are important tools for studying genomic diversity and genome-based breeding, which have extensive application prospects and important practical value. This study have rapidly expanded our understanding of genetic variation of medicinal plants and provided rich resources for genetic research. KEGG significant enrichment analysis showed that the richness of metabolic pathways proved the possible role of differential genes in important biological and metabolic pathways., it was found that the related genes in geniposide and crocin biosynthesis pathway may have mutations, and the corresponding molecular markers can be developed in follow-up research to mine excellent genes for further verification. The extensive variation between wild and cultivated species provided in this study provides a good way to make further use of the genetic diversity of wild gardenia species for gene-based breeding, and also has a certain guiding significance for the evolution and classification of *G. jasminoides*.

Acknowledgements

The study was supported by the Key Laboratory of Timber Forest Breeding and Cultivation for Mountainous Areas in Southern China, China National Forestry and Grassland Bureau and the Key Laboratory of Forest Culture and Forest Product Processing Utilization of Fujian Province. Thanks to The BioMaker Technology Company Limited for sequencing service.

Funding

The research was funded by the Public Welfare Special Projects for Research Institutes from Fujian Department of Science and Technology, Project title: Study on the Breeding and Propagation Techniques of High Quality *G~~ardenia~~. jasminoides* (No. 2020R1009002); the research projects of Fujian Provincial Forestry Bureau, Project title: Study on the Evaluation and Selection of Germplasm of *Gardenia jasminoides* (Min Lin Ke Informer File [2020] 9); and the Fund for Special Commissioner of Science and Technology of Fujian Province in 2020.

References

Formatted: Font: Italic

Formatted: Font: Times, 12 pt, (Asian) Chinese (China)

Adedze YMN, Lu X, Xia Y, Sun Q, Nchongboh CG, Alam MA, Liu M, Yang X, Zhang W,Deng Z, Li W, Si L.2021. Agarose-resolvable InDel markers based on whole genome re-sequencing in cucumber. *Scientific Reports* **11(1)**:3872. DOI 10.1038/s41598-021-83313-x.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database sear chprograms. *Nucleic Acids Research* 25:3389-3402. DOI 10.1093/nar/25.17.3389.

An Y, Mi X, Zhao S, Guo R, Xia X, Liu S, Wei C. **2020.** Revealing Distinctions in Genetic Diversity and Adaptive Evolution Between Two Varieties of *Camellia sinensis* by Whole-Genome Resequencing. *Frontiers in Plant Science* **11:**603819. DOI 10.3389/fpls.2020.603819.

Chen C, Chen H,Zhang Y, Thomas HR, Frank MH, He Y, Xia Rui. **2020.**TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Molecular Plant* **13(8)**:1194-1202. DOI 10.1016/j.molp.2020.06.009.

Chen H, Zeng X, Yang J, Cai X, Shi Y, Zheng R, Wang Z, LiuJ,Yi X, Xiao S, Fu Q, Zou J, Wang C. **2021.** Whole-genome resequencing of *Osmanthus fragrans* provides insights into flower color evolution. *Horticulture Research* **8(1)**:98. DOI 10.1038/s41438-021-00531-0.

Chen L,Li M,Yang Z, Tao W, Wang P, Tian X, Li X,Wang W. **2020a**. *Gardenia jasminoides* Ellis: Ethnopharmacology, phytochemistry, and pharmacological and industrial applications of an important traditional Chinese medicine. *Journal of Ethnopharmacology* **257**:112829. DOI 10.1016/j.jep.2020.112829.

Chen Q, Xue G, Ni Q, Wang Y, Gao Q, Zhang Y, Xu G. **2020b**. Physicochemical and rheological characterization of pectin-rich polysaccharides from *Gardenia jasminoides* J. Ellis flower. *Food Science & Nutrition* **8(7)**:3335-3345. DOI 10.1002/fsn3.1612.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. **2021.** Twelve years of SAMtools and BCFtools. *GigaScience***10(2)**:giab008.DOI 10.1093/gigascience/giab008.

Ganal MW, Altmann T, Röder MS. **2009.**SNP identification in crop plants. *Current opinion in plant biology* **12(2)**: 211-217.DOI 10.1016/j.pbi.2008.12.009.

Gene Ontology Consortium. **2015.** Gene ontology consortium: going forward. *Nucleic Acids Research* **43**(D1): D1049-D1056.DOI10.1093/nar/gku1179.

Giudice LD, Bazakos C, Vassiliou MF. 2022. Study of genetic variation and its association with tensile strength among bamboo species through whole genome resequencing. *Frontiers Plant Science* 13:935751. DOI 10.3389/fpls.2022.935751.

Habyarimana E, Gorthy S, Baloch FS, Ercisli S, Chung G. **2022**. Whole-genome resequencing of *Sorghum bicolor* and *S. bicolor* × *S. halepense*lines provides new insights for improving plant agroecological characteristics. *Scientific Reports* **12 (1)**:5556. DOI 10.1038/s41598-022-09433-0.

Han J, Zhang W, Cao H, Chen S, Wang H. **2007.**Genetic diversity and biogeography of the traditional Chinese medicine, *Gardenia jasminoides*, based on AFLP markers. *Biochemical Systematics and Ecology* **35(3)**: 138-145. DOI 10.1016/j.bse.2006.05.021.

Hussain A, Farooq M, Naqvi RZ, Aslam MQ, Siddiqui HA, Amin I, Liu C, Liu X, Scheffler J, Asif M, Mansoor S. **2022**. Whole-Genome Resequencing Deciphers New Insight Into Genetic Diversity and Signatures of Resistance in Cultivated Cotton *Gossypium hirsutum*. *Molecular Biotechnology* **65(1)**:34-51. DOI 10.1007/s12033-022-00527-8

Jing D, Liu X, He Q, Dang J, Hu R, Xia Y, Wu D, Wang S, Zhang Y, Xia Q, Zhang C, Yu Y, Guo Q, Liang G. 2023. Genome assembly of wild loquat (*Eriobotrya japonica*) and resequencing provide new insights into the genomic evolution and fruit domestication in loquat. *Horticulture Research* **10 (2)**:uhac265. DOI 10.1093/hr/uhac265.

Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. **2004.** The KEGG resource fordeciphering the genome. *Nucleic Acids Research* **32**:D277-D280. https://doi.org/10.1093/nar/gkh063.

Kosugi S, Momozawa Y, Liu X, Terao C, Kubo M, Kamatani Y. **2019**. Comprehensive evaluation of structural variation detection algorithms for whole genome sequencing.*Genome Biology* **20(1)**:117. DOI 10.1186/s13059-019-1720-5.

Li B, Lu X, Dou J, Aslam A, Gao L, Zhao S, He N, Liu W. 2018. Construction of A High-Density Genetic Map and Mapping of Fruit Traits in Watermelon (*Citrullus lanatus* L.) Based on Whole-Genome Resequencing. *International Journal of Molecular Sciences* **19(10)**:3268. DOI 10.3390/ijms19103268.

Li N, He Q, Wang J, Wang B, Zhao J, Huang S, Yang T, Tang Y, Yang S, Aisimutuola P, Xu R, Hu J, Jia C, Ma K, Li Z, Jiang F, Gao J, Lan H, Zhou Y, Zhang X, Huang S, Fei Z, Wang H, Li H, Yu Q. **2023.** Super-pangenome analyses highlight genomic diversity and structural variation across wild and cultivated tomato species. *Nature Genetics* **55**:852-860. DOI 10.1038/s41588-023-01340-y

Li N, Song Y, Li J, Hao R, Feng X, Li L.**2021**.Resequencing and transcriptomic analysis reveal differences in nitrite reductase in jujube fruit (*Ziziphusjujuba* Mill.). *Plant Methods***17(1)**:75. DOI 10.1186/s13007-021-00776-9.

Li Y, Xiong H, Zhang J, Guo H, Zhou C, Xie Y, Zhao L, Gu J, Zhao S, Ding Y, Fang Z, Liu L. **2021a**. Genome-Wide and Exome-Capturing Sequencing of a Gamma-Ray-Induced Mutant Reveals Biased Variations in Common Wheat. *Frontiers in Plant Science* **12**:793496. DOI 10.3389/fpls.2021.793496.

Li Y, Zhang J, Wang S, Zhang Y, Yang M. **2021b**. The Distribution and Origins of *Pyrushopeiensis*- "Wild Plant With Tiny Population" Using Whole Genome Resequencing. *Frontiers in Plant Science* **12**:668796. DOI

10.3389/fpls.2021.668796.

Lin Y, Chang P, Hsu C, Hung M, Chien Y, Hwu W, Lai F, Lee N. 2022. Comparison of GATK and Deep Variant by trio sequencing. *Scientific Reports* **12 (1)**:1809. DOI 10.1038/s41598-022-05833-4

Lin Y, Li W, Zhang Y, Xia C, Liu Y, Wang C, Xu R, Zhang L. 2019. Identification of Genes/Proteins Related to Submergence Tolerance by Transcriptome and Proteome Analyses in Soybean. *Scientific Reports***9(1)**:14688. DOI 10.1038/s41598-019-50757-1

Liu W, Feng Y, Yu S, Fan Z, Li X, Li J, Yin H. 2021. The Flavonoid Biosynthesis Network in Plants. *International Journal of Molecular Sciences* **22(23)**:12824. DOI 10.3390/ijms222312824

Liu Z, Mohsin A, Wang Z, Zhu X, Zhuang Y, Cao L, Guo M, Yin Z. 2020. Enhanced Biosynthesis of Chlorogenic Acid and Its Derivatives in Methyl-Jasmonate-Treated *Gardenia jasminoides* Cells: A Study on Metabolic and Transcriptional Responses of Cells. *Frontiers in Bioengineering and Biotechnology* **8**:604957. DOI 10.3389/fbioe.2020.604957.

Long W, Li Y, Yuan Z, Luo Li, Luo L, Xu W, Cai Y, Xie H. 2022. Development of InDel markers for *Oryza sativa* ssp. *javanica* based on whole-genome resequencing. *PLoS One* **17(10)**:e0274418. DOI 10.1371/journal.pone.0274418.

Long Z, Zhang R, Zhao X, Meng X, Bi K, Chen X. 2013. Determination and pharmacokinetics of geniposidic acid in rat plasma after oral administration of *Gardenia jasminoides* fruit crude extract and Zhi-zi-chi decoction. *Biomedical Chromatography* **27(6)**:812-816. DOI 10.1002/bmc.2871.

Lu D, Zhang W, Jiang Y, Zhang Y, Pan D, Zhang D, Yao X, Yu Y. 2019. Two new triterpenoids from *Gardenia jasminoide*s fruits. *Nature Product Research* **33(19)**:2789-2794. DOI 10.1080/14786419.2018.1502764.

Ma Z, Zhang Y, Wu L, Zhang G, Sun Z, Li Z, Jiang Y, Ke H, Chen B, Liu Z, Gu Q, Wang Z, Wang G, Yang J, Wu J, Yan Y, Meng C, Li L, Li X, Mo S, Wu N, Ma L, Chen L, Zhang M, Si A, Yang Z, Wang N, Wu L, Zhang D, Cui Y, Cui J, Lv X, Yang Li, Shi R, Duan Y, Tian S, Wang X. 2021. High-quality genome assembly and resequencing of modern cotton cultivars provide resources for crop improvement. *Nature Genetics* **53(9)**:1385-1391. DOI 10.1038/s41588-021-00910-2.

Madhukar S, Radhakrishnan AV, Majhi AK, Raghunathan VA. 2020. Structure and stoichiometry of CTAB-DNA complexes. *The Journal of Chemical Physics* **153(22)**:224901. DOI 10.1063/5.0033193.

Mei Z, Khan MA, Yang L, Yang M, Fu J. 2015. Genetic characterization and authentication of *Gardenia jasminoides* in different regions of China by using improved RAPD analysis. *Indian Journal of Experimental Biology* **53(3)**:164-9. https://nopr.niscpr.res.in/handle/123456789/30745.

Field Code Changed

Mursyidin DH, Makruf MI, Badruzsaufari, Noor A. 2022. Molecular diversity of exotic durian (Durio spp.) germplasm: a case study of Kalimantan, Indonesia. J Genet Eng Biotechnol. **20 (1)**:39. doi:10.1186/s43141-022-00321-8

Newman KL, Chatterjee S, Ho KA, Lindow SE. 2008. Virulence of plant pathogenic bacteria attenuated by degradation of fatty acid cell-to-cell signaling factors.*Molecular Plant-microbe Interactions* **21(3)**:326-34. DOI 10.1094/MPMI-21-3-0326

Nguyen TM, McConnachie M, Nguyen TD, Dang TT. **2022**. Discovery and Characterization of Oxidative Enzymes Involved in Monoterpenoid Indole Alkaloid Biosynthesis. *Methods in MolecularBiology* **2505**:141-164. DOI 10.1007/978-1-0716-2349-7_11

Pan Y, Zhao X, Wang Y, Tan J, Chen D. 2021. Metabolomics integrated with transcriptomics reveals the distribution of iridoid and crocin metabolic flux in *Gardenia jasminoides* Ellis. *PLoS One* **16(9)**:e0256802. DOI 10.1371/journal.pone.0256802.

Schenck CA, Maeda HA.2018. Tyrosine biosynthesis, metabolism, and catabolism in plants. *Phytochemistry* **149**:82-102. DOI 10.1016/j.phytochem.2018.02.003

Shen T, Zheng Y, Liu Q, Chen C, Huang L, Deng S, Xu M, Yang C. **2022**. Integrated SMRT and *Illumina* Sequencing Provide New Insights into Crocin Biosynthesis of *Gardenia jasminoides*. *International Journal of Molecular Sciences* **23(11)**:6321. DOI 10.3390/ijms23116321.

Shen T, Zheng Y, Liu Q, Chen C, Huang L, Deng S, Xu M, Yang C. 2022. Integrated SMRT and *Illumina* Sequencing Provide New Insights into Crocin Biosynthesis of *Gardenia jasminoides*. *International Journal of Molecular Sciences* 23 (11):6231. DOI 10.3390/ijms23116321

Sommano SR, Suppakittpaisarn P, Sringarm K, Junmahasathien T, Ruksiriwanich W. **2020**. Recovery of Crocins From Floral Tissue of *Gardenia jasminoides* Ellis. *Frontiers in nutrition* **7**:106. DOI 10.3389/fnut.2020.00106

Thimmegowda GC, Ramadoss SK, Kaikala V, Rathinavelu R, Thamalampudi VR, Dhavala VNC, Saiprasad GVS. 2018. Whole genome resequencing of tobacco (*Nicotiana tabacum* L.) genotypes and high-throughput SNP discovery. *Molecular Breeding* 38: 1-10.DOI 10.1007/s11032-018-0876-0.

Tian Y, Pu X, Yu H, Ji A, Gao R, Hu Y, Xu Z, Wang H. **2020**. Genome-Wide Characterization and Analysis of bHLH Transcription Factors Related to Crocin Biosynthesis in *Gardenia jasminoides* Ellis (Rubiaceae). *BioMed Research International* **2020**:2903861. DOI 10.1155/2020/2903861.

Toleco MR, Naake T, Zhang Y, Heazlewood JL, Fernie AR. **2020**. Plant Mitochondrial Carriers: Molecular Gatekeepers That Help to Regulate Plant Central Carbon Metabolism. *Plants* **9(1)**:117. DOI 10.3390/plants9010117

Varshney RK, Saxena RK, Upadhyaya HD, Khan AW, Yu Y, Kim C, Rathore A,

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Indent: Left: 0 cm, Hanging: 2 ch

Formatted: Font: (Default) Times, 12 pt, Bold, (Asian) Chinese (China), Expanded by 0.6 pt

Formatted: Font: Bold

Kim D, Kim J, An S, Kumar V, Anuradha G, Yamini KN, Zhang W, Muniswamy S, Kim J-S, Penmetsa RV, von Wettberg E, Datta SK. **2017**. Whole-genome resequencing of 292 pigeonpea accessions identifies genomic regions associated with domestication and agronomic traits. *Nature Genetics* **49(7)**: 1082-1088.

Wang J, Hu Z, Liao X, Wang Z, Li W, Zhang P, Cheng H, Wang Q, Bhat JA, Wang H, Liu B, Zhang H, Huang F, Yu D. **2022**. Whole-genome resequencing reveals signature of local adaptation and divergence in wild soybean. *Evolutionary Applications* **15(11)**:1820-1833. DOI 10.1111/eva.13480.

Wei L, Dugas M, Sandmann S. **2021**. SimFFPE and FilterFFPE: improving structural variant calling in FFPE samples. *GigaScience* **10(9)**:2021, 1-12. DOI 10.1093/gigascience/giab065.

Wu D, Liang Z, Yan T, Xu Y, Xuan L, Tang J, Zhou G, Lohwasser U, Hua S, Wang H, Chen X, Wang Q, Zhu L, Maodzeka A, Hussain N, Li Z, Li X, Shamsi IH, Jilani G, Wu L, Zheng H, Zhang G, Chalhoub B, Shen L, Yu H, Jiang L. **2019**. Whole-Genome Resequencing of a Worldwide Collection of Rapeseed Accessions Reveals the Genetic Basis of Ecotype Divergence. *Molecular Plant* **12(1)**:30-43. DOI 10.1016/j.molp.2018.11.007.

Wu J, Wang X, Yu Y, Li J, Ma K, Zhang Y, Li H, Yin C, Zhang Y. **2022.**Stability evaluation of gardenia yellow pigment in presence of different phenolic compoun ds. *Food Chemistry* **373**(Pt A):131441. DOI 10.1016/j.foodchem.2021.131441.

Xiao W, Li S, Wang S, Ho C. **2017**. Chemistry and bioactivity of *Gardenia jasminoides*. *Journal of Food and Drug Analysis* **25(1)**:43-61. DOI 10.1016/j.jfda.2016.11.005.

Xu W, Yu J, Feng W, Su W. **2016.** Selective Extraction of Gardenia Yellow and Geniposide from *Gardenia jasminoides* by Mechanochemistry. *Molecules* **21(5)**: 540. DOI 10.3390/molecules21050540.

Xu Z, Pu X, Gao R, Demurtas OC, Fleck SJ, Richter M, He C, Ji A, Sun W, Kong J, Hu K, Ren F, Song J, Wang Z, Gao T, Xiong C, Yu H, Xin T, Albert VA, Giuliano G, Chen S, Song J. 2020. Tandem gene duplications drive divergent evolution of caffeine and crocin biosynthetic pathways in plants. *BMC Biology* **18 (1)**:63. DOI 10.1186/s12915-020-00795-3.

Yang C, Zhang T, Xu M, Zhu P, Deng SY.**2016**. Insights into biosynthetic genes involved in the secondary metabolism of *Gardenia jasminoides* Ellis using transcriptome sequencing. *Biochemical Systematics and Ecology* **67**:7-16.DOI 10.1016/j.bse.2016.05.011.

Zhang C, Xie W, Fu H, Chen Y, Chen H, Cai T, Qiang Yang, Zhuang Y, Zhong X, Chen K, Gao M, Liu F, Wan Y, Pandey MK, Varshney RK, Zhuang W. **2022**. Whole genome resequencing identifies candidate genes and allelic diagnostic markers for resistance to *Ralstonia solanacea* rum infection in cultivated peanut

(*Arachis hypogaea* L.). *Frontiers in Plant Science* **13**:1048168. DOI 10.3389/fpls.2022.1048168.

Zhang L, Cui D, Ma X, Han B, Han L. 2023. Comparative analysis of rice reveals insights into the mechanism of colored rice via widely targeted metabolomics. *Food Chemistry* **399**:133926. DOI 10.1016/j.foodchem.2022.133926

Zhang N, Bian Y, Yao L. 2022. Essential Oils of *Gardenia jasminoides* J. Ellis and *Gardenia jasminoides* f. *longicarpa* Z.W. Xie & M. Okada Flowers: Chemical Characterization and Assessment of Anti-Inflammatory Effects in Alveolar Macrophage. *Pharmaceutics***14(5)**:966. DOI 10.3390/pharmaceutics14050966.

Zhao D, Wang R, Meng J, Li Z, Wu Y, Tao J. 2017. Ameliorative effects of melatonin on dark-induced leaf senescence in gardenia (*Gardenia jasminoides* Ellis): leaf morphology, anatomy, physiology and transcriptome. *Scientific Reports* **7(1)**:10423.DOI 10.1038/s41598-017-10799-9.

Zhou L, Bao L, Wang Y, Chen M, Zhang Y, Geng Z, Zhao R, Sun J, Bao Y, Shi Y, Yao R, Guo S, Cui X. 2021. An Integrated Analysis Reveals Geniposide Extracted From *Gardenia jasminoides* J. Ellis Regulates Calcium Signaling Pathway Essential for Influenza A Virus Replication. *Frontiers in Pharmacology* **2**:755796. DOI 10.3389/fphar.2021.755796.