Brief Communication

Genomes of single- and double-petal jasmines (Jasminum sambac) provide insights into their divergence time and structural variations

Pengjie Wang¹,²,†, Jingping Fang³,†, Hongzheng Lin¹,², Wenwen Yang⁴, Jiaxin Yu⁵, Yaping Hong¹, Mengwei Jiang⁵, Mengya Gu¹, Qinchang Chen³, Yucheng Zheng¹, ZhenYang Liao⁵, Guixin Chen¹, Jiangfan Yang¹, Shan Jin¹,*, Xingtan Zhang²,§, and Naixing Ye¹,‡,*

¹College of Horticulture, Fujian Agriculture and Forestry University/Key Laboratory of Tea Science in Universities of Fujian Province, Fuzhou, China
²Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China
³College of Life Science, Fujian Normal University, Fuzhou, China
⁴Fuzhou City Green Food Development Center, Fuzhou, China
⁵College of Life Science, Fujian Agriculture and Forestry University, Fuzhou, China
⁶Center for Genomics and Biotechnology, Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, Key Laboratory of Genetics, Fujian Agriculture and Forestry University, Fuzhou, China

Keywords: Jasminum sambac, genome assembly, divergence time, structural variation.

Jasmine (Jasminum sambac) is a monoploid plant belonging to the genus Jasminum of the Oleaceae family. It is commercially cultivated worldwide for its sweet fragrance and is used in the fields of food, medicine, and industry. However, the cultivation area of jasmine is small due to its weak stress resistance and high yield. Notably, the SP flower shows a fresh and elegant aroma, which is related to its flower yield and resistance. The genome of jasmine (Jasminum sambac) is an important factor in determining its economic value.

Received 8 December 2021; revised 8 February 2022; accepted 30 March 2022.
*Correspondence (Tel +86-591-15306078980; fax +86-591-83789281; email ynxtea@126.com (NY), Tel +86-591-15985733087; fax +86-755-23251430; email zhangxingtan@caas.cn (XZ), Tel +86-591-15859000539; fax +86-591-83735681; jinshan0313@163.com (SJ))
†These authors contributed equally to this work.

86-591-83789281; fax 86-591-15985733087; email zhangxingtan@caas.cn (XZ), Tel +86-591-15859000539; fax +86-591-83735681; jinshan0313@163.com (SJ)

Plant Biotechnology Journal (2022) 20, pp. 1232–1234 doi: 10.1111/pbi.13820

We generated 15.35 Gb of HiFi PacBio subreads for SP jasmine and 12.11 Gb for DP jasmine on the PacBio Sequel II platform with an estimated coverage depth of at least 30-folds for each genome (Table S2, Material S1). The initial de novo assemblies using hifiasm yielded contig-level sequences of 653.28 and 535.89 Mb for SP and DP genomes, respectively (Table S3). To generate non-redundant genome sequences, the haplotigs and contig overlaps in the initial assemblies were removed by purge_dups program based on PacBio read depth and a K-mer counting strategy, which resulted in two purged assemblies of 502.27 and 486.97 Mb with improved contig N50 values of 7.34 and 11.82 Mb for SP and DP jasmine genomes, respectively (Figure 1d and e and Tables S3–S5), and observed that most redundant sequences have been removed by purge-dups program (Figure S2). These two genomes were further anchored into 13 super-scaffolds representing all chromosomes by incorporating PacBio HiFi subreads and physical mapping data from Hi-C (Figure S3 and Tables S6 and S7) using the ALLHIC pipeline. The final assemblies contain 44 contigs (>2 kb) for SP jasmine genome and 210 contigs (>2 kb) for DP jasmine genome. Genome assembly quality was also assessed based on long-terminal repeat (LTR) retrotransposons and LTR Assembly Index (LAI). LAI scores for SP and DP jasmine genomes were estimated to be ~28.16 and ~24.35, respectively (Figure 1e), which are much higher than that of the recently published MP jasmine genome using nanopore data (Xu et al., 2021). This finding indicates the better continuity and completeness of our jasmine monoploid genomes, which recovered more repetitive sequences. In addition, compared with the recently published MP jasmine genome, our two newly assembled jasmine genomes showed improvements in gene completeness according to BUSCO assessment results (Table S8). Genome annotation predicted 24,002 and 23,574 protein-coding genes with 96.6% BUSCO completeness in the two jasmine genomes, respectively (Figure 1e). Moreover, 307.62 and 298.26 Mb repetitive sequences (TE) were annotated in SP and DP Jasmine genomes, respectively, and both sequences accounted for 61.25% of the genome size (Table S9).
Genomes of single- and double-petal jasmines

(a) (b) (c) (d)

(e) (f) (g) (h) (i) (j) (k) (l) (m)

© 2022 The Authors. Plant Biotechnology Journal published by Society for Experimental Biology and The Association of Applied Biologists and John Wiley & Sons Ltd., 20, 1232–1234
The genomic collinearity between grape and the three jasmine types were analysed. We observed that the Ks peaks of SP and DP jasmines were on the left side of the grape peak (Figure 1f), and the syntenic blocks between grape and SP or DP jasmine showed a 1:3 syntenic relationship (Figures 1g and S4). This result suggests that SP and DP jasmines experienced an additional WGT event after the WGT-y event, which is consistent with MP jasmine (Xu et al., 2021). In addition, we found that most of the homologous gene pairs were caused by the subsequent WGT event (Ks=1.1–1.3, Figures 1g and S4), and the homologous gene pairs affected by the WGT-y event were lost and scattered in the genome. We further calculated the divergence time among the three types of jasmines. The base substitution rate of jasmine was calculated as \( r = 1 \times 10^{-8} \)–1.3 \( \times 10^{-8} \) with reference to the previous method (Badouin et al., 2017). Finally, the divergence time between SP and MP, DP and SP and MP jasmines were determined to be 0.03–2.4 Mya (Ks = 0.0006–0.049383), 0.035–5.76 Mya (Ks = 0.0007–0.115248), 0.015–5.28 Mya (Ks = 0.0003–0.105704), respectively, by calculating the Ks value. The divergence time of the three jasmines were very close; however, the divergence time between SP and DP jasmines may be more representative because of the large difference between the sequencing and assembly of MP jasmine with SP or DP jasmine (Figure 1h).

We identified 52 547 (134.99 Mb) and 26 773 (76.25 Mb) SVs, representing 26.88% and 15.66% of the genome, by comparing the raw HiFi reads of SP and DP jasmine with the MP jasmine genome, respectively (Tables S10 and S11). The SV-related genes of SP vs. MP were much more than those of DP vs. MP (Figure 1i). KEGG analysis showed that the SV-related genes of SP vs. MP were mainly involved in the energy metabolism pathways as the core (Figure 1j and l), which may contribute to the rich aroma of DP jasmine as a suitable raw material for jasmine tea.

In summary, these results greatly deepen our understanding of the evolution and differentiation of jasmine species and provide resources for the research on functional genomics of jasmines.

Funding
This research was funded by the Technology Research of Fuzhou Jasmine Tea Technology and Global Important Agricultural Cultural Heritage Joint Research Center (KH1501920), the Natural Science Foundation of Fujian Province, China (2019J05066) and the Natural Science Foundation of China (41906096).

Conflict of interest
The authors declare no conflict of interest.

Authors’ contributions
N.Y., X.Z., S.J., and P.W. conceived and designed the research. P.W., H.L., W.Y., and Y.H. collected the samples. P.W., J.F., J.Y., M.J., Y.H., Q.C., H.L., W.Y., M.G., Y.Z., Z.L. and G.C. performed the genome assembly and data analysis. P.W. and J.F. wrote the manuscript. N.Y., X.Z., S.J., J.Y., and P.W. revised the manuscript.

Data availability statement
All data are publicly available in the BIG Data Center (https://bigd.big.ac.cn/) under project number PRJCA006739.

References
Badouin, H., Gouzy, J., Grassa, C.J., Murat, F., Staton, S.E., Cotret, L., Lelandais-Brière, C. et al. (2017) The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature*, 546, 148–152.
Wang, P., Yu, J., Jin, S., Chen, S., Yue, C., Wang, W., Gao, S. et al. (2021) Genetic basis of high aroma and stress tolerance in the oolong tea cultivar genome. *Hortic. Res.*, 8, 107.
Xu, S., Ding, Y., Sun, J., Zhang, Z., Wu, Z., Yang, T., Shen, F. et al. (2021) A high-quality genome assembly of *Jasminum sambac* provides insight into floral trait formation and Oleaceae genome evolution. *Mol. Ecol. Resour.*, 22 (2), 724–739.

Supporting information
Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1–S4 Supplementary Figures.
Table S1–S11 Supplementary Datasets.
Materials and methods S1 Supplementary materials and methods.