Pangenomics and single-cell transcriptomics uncover the genetic basis of continuous bearing trait in grapevine

Yuting Liu[1, 2, †], Yang Dong[2, †], Xinyue Fang[2], Xu Wang[2], Jiacui Li[2], Tianhao Zhang[2], Qiming Long[2], Ying Su[2], Xiaoya Shi[2], Xiangnian Su[2], Yingchun Zhang[2], Ting Hou[1, 2], Mengrui Du[2], Yiran Ren[1], Peipei Wang[1], Xinglong Ji[1], Yunjun Chang[4], Sheng Yan[1, 2], Sifan Yang[1, 2], Yongfeng Zhou[2, 3], Yanling Peng[2, *], Xiangpeng Leng[1, *]

1 *College of Horticulture, Qingdao Agricultural University, Qingdao 266109, China*

2 *National Key Laboratory of Tropical Crop Breeding, Shenzhen Branch, Guangdong Laboratory of Lingnan Modern Agriculture, Key Laboratory of Synthetic Biology, Ministry of Agriculture and Rural Affairs, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China*

3 *National Key Laboratory of Tropical Crop Breeding, Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou, China*

4 *Shandong fresh table grape Research Institute, Qingdao, China*

† These authors contributed equally to this work.

** Corresponding authors:* pengyanling@caas.cn; lengxp2018@qau.edu.cn

Abstract

Grapevine (*Vitis vinifera*) is economically important for fresh consumption, winemaking and drying. The circadian systems of flowering and fruit development are crucial for viticulture and yield formation. However, the genetic basis of continuous flowering and bearing has been rarely elucidated. Here, we integrate pan-genomics, comparative genomics, single-cell transcriptomics, and bulk transcriptomics to investigate the continuous flowering and bearing trait (CFB) in the 'Julian' grape, which bears fruits at different development stages from flower to mature berries simultaneously. Pan-genomics and comparative genomics discovered 558 unique structural variations (SVs) and eight genes enriched in flowering pathways exclusively in Julian, based on the newly generated haplotype-resolved near telomere-to-telomere (T2T) genome of 'Julian' and 15 previously published genomes of grapevine, which bears fruits in the same developmental stage (i.e. seasonal flowering and bearing, SFB). Single-cell transcriptomes of flowering buds for CFB 'Julian' and SFB 'Muscat Hamburg' detected seven distinct cell types, which provide detailed cell-type-

specific gene expression profiles for both cultivars, with differential gene expression (DEG) insights highlighting growth, metabolism, and hormonal regulation pathways in 'Julian'. Integrating SVs and DEG data, we pinpoint 37 candidate genes potentially associated with CFB, including Auxin/IAA, Cytochrome P450, Vicilin-like antimicrobial peptides (*AMP*) and respiratory burst oxidase homolog protein B (*RBOH*) related genes. This study provides insights into the genetic basis of CFB in grapevine, facilitating grapevine breeding with continuous flowering and bearing.

Keywords: Viticulture, continuous flowering, multi-omics, structural variants, single cell RNA sequencing, grape breeding

Introduction

The grapevine is one of the most widely cultivated deciduous fruit trees, with significant economic value, primarily for fresh consumption, winemaking, and drying (Zhou et al., 2017). In the breeding process, the ability of the plant to form flowers is a key indicator of its potential for producing new varieties, and good bud differentiation is the basis for obtaining a considerable yield and quality of grapes. Furthermore, flower formation ability is a quantitatively inherited trait, which is influenced by environmental factors and multiple genes (Díaz-Riquelme et al., 2012; Shangguan et al., 2020). Most grape cultivars exhibit flowering and fruiting within a concentrated period, known as seasonal flowering and bearing (SFB), while the grape cultivar 'Julian' (JA) is notable for its continuous flowering and bearing (CFB), which offers significant advantages for agricultural production, including an extended harvest period, increased yield, and improved resource utilization throughout the growing season. Beyond grapevines, CFB is also found in other economically important plants, such as day-neutral or everbearing strawberries (Hytonen and Kurokura, 2020), indeterminate tomatoes (MacAlister et al., 2012), and certain varieties of roses (Yi et al., 2021), figs (Ikegami et al., 2013), and sweet pepper (Wubs et al., 2009). Although these traits are critical for agricultural production, most studies have focused on the physiological aspects of CFB in these crops, such as flowering timing and resource allocation (Wang et al., 2022, Ahmad et al., 2022). Previous studies on CFB have relied on single-omic approaches, such as identifying key regulators like *TERMINAL*

FLOWER 1 (*TFL1*) in roses and strawberries (Iwata et al., 2012). However, the dissection of genetic basis using single-omic is largely limited because its incomplete information. By contrast, integrated multi-omics, especially the integration of pan-genome, comparative genome, single-cell and bulk transcriptomes, provides a chance to uncover the genetic basis of CFB comprehensively and precisely.

Pan-genome can uncover rare or specific genes that may not be present in a single reference genome, and these genes can be crucial under specific conditions (Tong et al., 2024). A tomato pan-genome revealed rare alleles linked to fruit flavor, offering new possibilities for improving crop quality (Gao et al., 2019). In rice, pan-genome studies uncovered large deletions, such as those affecting OsWAK112d, which enhance disease resistance and environmental adaptation (Qin et al., 2021). Similarly, in sunflower, pan-genomes identified introgressed biotic resistance genes, enriching the genetic diversity of cultivated lines (Hübner et al., 2019). The application of the pan-genome in grapevine has enabled the identification of SVs/genes associated with 29 agronomic traits (Liu et al., 2024) and 7% of unique genes to each accession (Long et al., 2024). Similar to pan-genome, comparative genome studies have revealed the genetic basis of seedlessness in grapevine (Wang et al., 2024). Up to now, such an analysis had not been conducted on the dissection of the genetic basis for CFB. Nevertheless, these examples demonstrate the power of the pan-genome and comparative genome analysis to uncover critical variations associated with traits like CFB in grapevine in JA, providing valuable insights for breeding strategies.

The advantage of single-cell transcriptomics has enabled the characterization of the transcriptome in hundreds of thousands of individual cells, providing unique insights into gene expression, cellular heterogeneity, and dynamic cellular differentiation trajectories, which are often obscured in bulk transcriptomics (Liu et al., 2023, Long et al., 2021, Sunaga-Franze et al., 2021). This technology has been widely applied across various plant species, including model organisms and economically important crops such as Arabidopsis (Jean-Baptiste et al., 2019, Lee et al., 2023), rice (Liu et al., 2021), tomato (Tian et al., 2020), maize (Marand et al., 2021), and litchi (Yang et al., 2023). For instance, single-cell transcriptional profiles of etiolated, de-etiolating and light-grown *Arabidopsis thaliana* seedlings furnished information about the light signaling network at the cellular level and enhanced our comprehension of how light regulates plant development at the cellular and genome-wide levels (Han et al., 2023b). Similarly, the single-cell transcriptome of the stem tip meristematic

tissue cells in tomato generated high-resolution expression profiles and distinguished cell types (Tian et al., 2020). Furthermore, single-cell transcriptomes are able to identify genes associated with flower-forming transformation in lychee using apical buds at different developmental stages (Yang et al., 2023). Thus, the application of single-cell transcriptomes in grape flower bud is able to identify cell heterozygosity and detect lowly expressed genes at unprecedented resolution, and then help understand the contribution of genes to continuous flowering and fruiting of JA.

In this study, integrative analysis of multi-omics, including pan-genomics, comparative genomics, single-cell and bulk transcriptomics was used to elucidate the genetic basis underlying CFB in this JA grape. Firstly, we newly assembled the T2T genomes of JA and performed comparative genome analysis with 'Muscat Hamburg' (MH) and PN40024. Secondly, we constructed a graph-based pan-genome using the newly assembled genomes and 15 previously published chromosome-level grape genomes, which produce flowers and fruit at a concentrated time (i.e. seasonal flowering and bearing, SFB). The pan-genomic and comparative genomic studies enabled us to characterize the unique variations and genes present in JA grape that may be correlated with CFB. Thirdly, we constructed high-resolution single-cell gene expression profiles for JA and MH, with a total of 13 cell clusters, and 7 cell types identified. Differentially expressed genes (DEGs) in the two cultivars were compared across the different cell clusters. Finally, we integrated pan-genomics, comparative genomics, single-cell and bulk transcriptomics to screen a set of key candidate genes and SVs that may be involved in regulating the floral transition in JA. These findings from the present integrative genomic analysis will advance grapevine genomic breeding for CFB.

Results

Haplotype-resolved genome assembly and annotation of JA.

To investigate the genetic basis of CFB in the JA cultivar (Fig. 1B, C), we generated high-quality haplotype-resolved genome assemblies (JA_hap1 and JA_hap2) using ~26.23 Gb PacBio HiFi sequencing (~66× coverage) and ~29.59 Gb Hi-C sequencing (~38.49× coverage). K-mer analysis estimated a heterozygosity rate of 1.44% for the JA genome, comparable to seedless cultivars such as 'Thompson Seedless' (1.51%) and 'Black Monukka' (1.41%) (Wang et al., 2024) but significantly higher than wild grapes (0.54%, $P < 0.05$) (Zhang et al., 2024) (Supplementary Fig. 1A, B). The

assembled genomes, JA_hap1 (494.79 Mb) and JA_hap2 (494.13 Mb), were of high quality with BUSCO scores of 98.6% and 98.3%, and scaffold N50 values of 25.94 Mb and 26.38 Mb, respectively, and additional metrics including quality value (QV) (50.3426 for hap1, 50.0721 for hap2; mean 50.2053), k-mer completeness (81.3577% for hap1, 81.5623% for hap2; combined 99.1256%), LTR assembly index (LAI) values (19.21 for hap1, 19.20 for hap2, both meeting "Reference" classification), and a switch error rate of 1.19% (Table S2). Annotations revealed 19 centromeres and 35 telomeres (Table S2), and repeat masking covered ~333 Mb for each haplotype, consistent with typical *Vitis* genomes. GC content (35.07% and 34.98%) confirmed inter-haplotype consistency, while Hi-C data organized the assembly into 19 pseudo-chromosomes, validating structural integrity (Fig. 1D).

Comparative collinearity analysis with MH and PN40024 showed strong synteny, conserved genome sizes, and centromere positions, indicating evolutionary conservation (Fig. 1A). Between JA haplotypes, we identified 3,671 hemizygous genes (~10.21% of all genes), consistent with other highly heterozygous grapevine genomes (Peng et al., 2024). Gene ontology (GO) enrichment revealed these genes were involved in nitrogen metabolism, protein complex organization, and nucleic acid-related processes, suggesting roles in key cellular pathways (Fig. 1E).

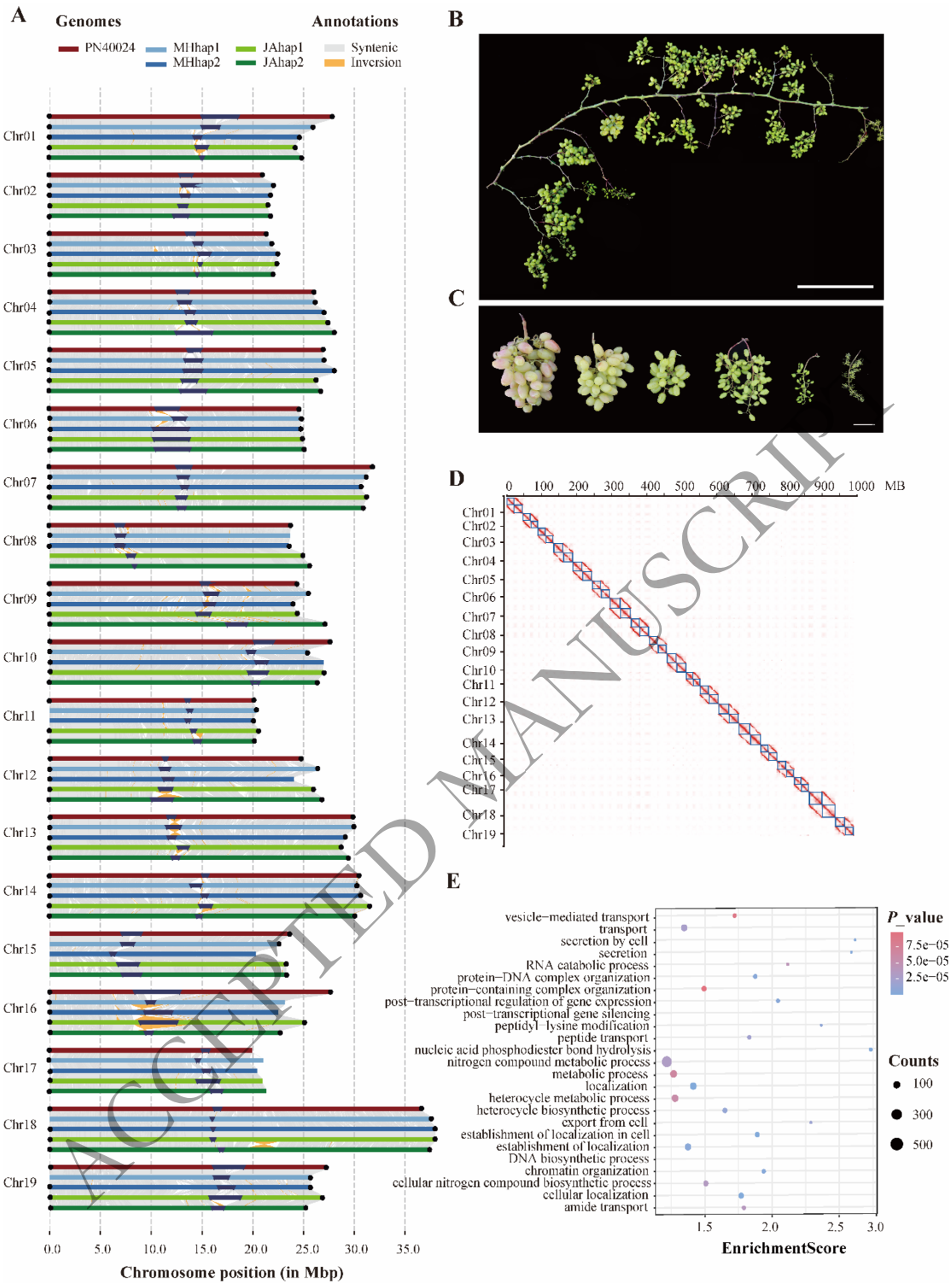

Fig 1 | Genome assembly and analysis of the 'Julian' grape cultivar. A, Synteny analysis comparing the haplotype-resolved 'Julian' genome (JA_hap1 and JA_hap2) with the reference genome PN40024 and 'Muscat Hamburg' (MH_hap1 and MH_hap2), displaying conserved genomic regions; **B**, **C** Morphological representation of JA phenotypes, including inflorescences and mature fruits at different node positions. Scale bars: 50 cm in (B) and 5 cm in (C); **D**, Hi-C

contact heatmap displaying the chromosomal interactions for the 19 pseudo-chromosomes of JA supporting the structural accuracy of the assembled genome; **E**, GO enrichment analysis for hemizygous genes in JA, showing enrichment in biological processes.

Pan-genomics and comparative genomics detected JA-specific SVs and genes

To investigate the genomic differences between the JA and other table grape cultivars, we constructed a pan-genome comprising the JA genome, 14 phased genome assemblies from other table grape cultivars, and the reference genome PN40024 (Shi et al., 2023). The newly sequenced JA genome was integrated into this analysis to enable a comparative exploration of sequence variations. Using the Pan Genome Graph Builder (PGGB), we constructed a pan-genome graph (Fig. 2A), which spanned 987.06 Mb—approximately double the size of a single haploid grape genome (Fig. 2B). The pan-genome approach provided a comprehensive framework to capture sequence variations that linear genome models often overlook, including structural variants (SVs) and presence/absence variations (PAVs), both critical for understanding the genetic diversity underlying complex traits like continuous flowering.

The analysis identified 19,574,412 variants across the pan-genome, including 16,748,352 single nucleotide polymorphisms (SNPs), 2,643,897 insertions/deletions (indels), and 1,383,324 multi-nucleotide polymorphisms (MNPs). These variants highlight the extensive genetic divergence among table grape cultivars, likely driven by artificial selection and domestication processes (Zhou et al., 2019). Pan-genome-based studies in other crops, such as rice and sunflower, have similarly revealed that domestication often introduces unique structural changes and genetic diversity essential for agronomic traits (Hübner et al., 2019, Qin et al., 2021). Such findings underscore the utility of the pan-genome approach in uncovering genetic variations linked to complex phenotypes. By applying a 50 bp threshold to define SVs, we identified 8,758 insertions and 2,059 deletions, with 558 classified as unique JA-specific SVs (Table S6). Notably, chromosomes 7, 13, and 14 exhibited a concentration of these structural variants (Fig. 2C). Further analysis of genes within 5 kb of these variants revealed significant enrichment in pathways related to the transition from vegetative to reproductive growth, photoperiod regulation, and long-day photoperiod response (Fig. 2D).

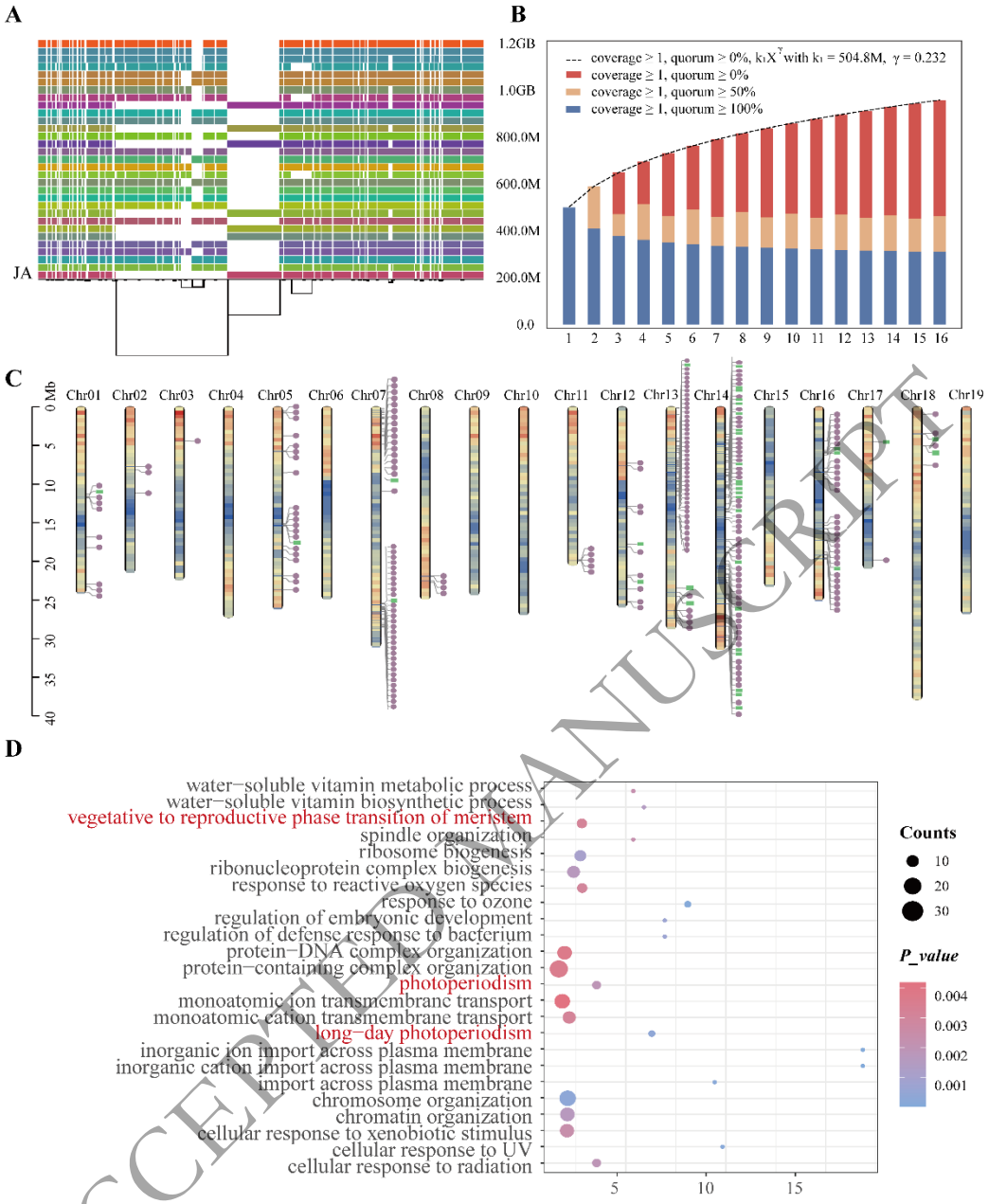

Fig 2 Pan-genomics identifies Julian-specific variations and genes in table grapes. A, Schematic representation of the graph pan-genome constructed from 16 grape genomes, including 31 haplotypes. The graph illustrates structural variations such as insertions and deletions across different haplotypes, with the two JA haplotypes shown at the bottom; **B**, Pan-genome growth curve depicting the increase in total pan-genome length as additional genomes are incorporated; **C**, Distribution of JA-specific structural variants (unique SVs) on chromosomes, showing the first 50 insertions per chromosome for JA_hap1 genome; **D**, GO enrichment analysis of genes located within 5 kb of JA unique SVs.

Single-cell transcriptome atlas of the grapevine bud

To investigate the molecular basis underlying continuous flowering in JA and compare it with MH, we conducted single-nucleus RNA sequencing (snRNA-seq) on both varieties. Gene expression analysis of grape buds from JA and MH revealed 13,482 high-quality single-nucleus transcriptomes from JA, covering 22,453 genes, with a median of 2,067 genes and 3,547 unique molecular identifiers (UMIs) per nucleus. For MH, 18,762 transcriptomes representing 24,462 genes were obtained (Table S8). Notably, approximately 77% of the genes detected by snRNA-seq overlapped with those identified in bulk RNA-seq, highlighting the complementarity of these methods.

Graph-based clustering analysis in Seurat identified 13 distinct cell clusters, visualized using UMAP (Fig. 3A). Mapping homologous genes from *Arabidopsis thaliana* allowed us to assign these clusters to major cell types in grapevine buds (Fig. 3B, Table S10), including vascular cells (xylem and phloem), epidermal, mesophyll, and meristematic cells. Xylem cells (9,524 cells) were most abundant, emphasizing their role in nutrient transport, followed by mesophyll cells (7,733 cells), essential for photosynthesis. Notably, the detection of 726 meristematic cells is a key component in dormancy and flowering regulation.

After identifying the top 50 marker genes for each cell cluster (Table S11), we performed GO enrichment analysis, which revealed distinct functional enrichments specific to each cell type. Xylem cells were enriched in pathways related to water transport and cell wall modification, reflecting their structural and resource distribution roles. Mesophyll cells were strongly enriched in photosynthetic pathways, highlighting their function in energy production. In contrast, meristematic cells exhibited enrichments in cell division and developmental regulation pathways (Fig. 3C, D), indicating their potential role in dormancy and flowering regulation. These findings demonstrate that the marker genes identified for each cluster are closely tied to the specialized functions of grapevine bud cells, advancing our understanding of grapevine development and flowering.

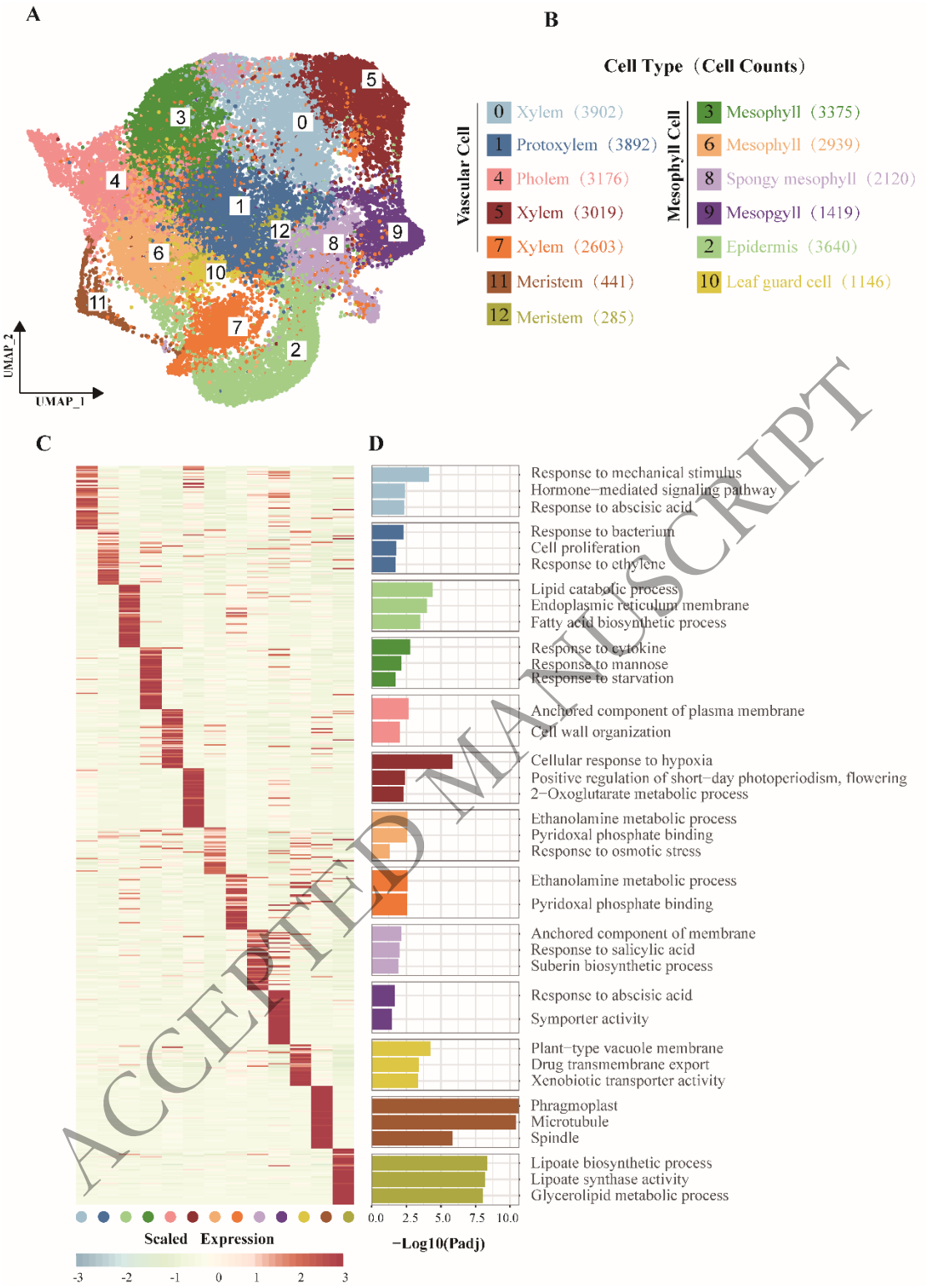

Fig 3 | Single-nucleus transcriptome atlas of grapevine buds. A, UMAP visualization of single-nucleus transcriptomes from JA and MH grape buds, revealing 13 distinct cell clusters; **B**, Annotated cell types based on marker gene expression, with the number of nuclei indicated in parentheses; **C**, Heatmap displaying the top 50 marker genes for each cell cluster, illustrating cell-type-specific gene expression; **D**, GO enrichment analysis of the marker genes, highlighting significantly enriched

biological processes associated with each cluster.

Differentially expressed genes (DEGs) in JA and MH

To further investigate the single-cell-level differences between JA and MH, we aligned their single-cell count matrices using the "anchor" method (Fig. 4A), enabling a direct comparison of cell composition and gene expression between the two cultivars. This alignment revealed significant differences in cell type distributions, particularly in the abundance of meristematic and mesophyll cells (Fig. 4B). Subsequently, DEG analysis was performed on the snRNA-seq data from both JA and MH, resulting in the identification of 5,807 DEGs (Fig. 4C, Table S12). Of these, 2,919 genes were upregulated in JA, while 2,888 genes were upregulated in MH. Among them (Table S12), key examples include Vitis18g01001, a *MED18* homolog regulating flowering time via *FLC* and *AG* in *Arabidopsi*s (Lai et al., 2014), significantly upregulated in Julian (5.2 TPM vs. 2.1 TPM in MH); and Vitis14g01538, an *AOC4* homolog critical for jasmonic acid biosynthesis, also upregulated in Julian, suggesting elevated jasmonic acid signaling (Stenzel et al., 2012). Additional DEGs involve auxin/cytokinin pathway genes like Vitis07g00043(homologous to *ATAUX2-11*), collectively supporting Julian's sustained flowering traits. To further explore the biological significance of these DEGs, we performed GO enrichment analysis, focusing on the upregulated genes in both JA and MH (Fig. 4D, E).

GO enrichment analysis of the upregulated genes in JA (Fig. 4D) showed significant enrichment in pathways related to growth, secondary metabolism, and hormonal regulation, including secondary cell wall biogenesis, phenylpropanoid biosynthesis, and auxin response. These pathways were associated with enhanced cell wall rigidity, tissue differentiation, and expansion, potentially supporting JA's CFB. In MH, upregulated genes (downregulated in JA) were enriched in pathways such as carboxylic acid catabolism, lipid metabolism, and interspecies interactions (Fig. 4E), which are linked to metabolic regulation and responses to environmental factors. These differences in enrichment profiles indicate distinct molecular priorities in JA and MH, with JA emphasizing growth-related processes and MH showing stronger activation of stress-associated pathways.

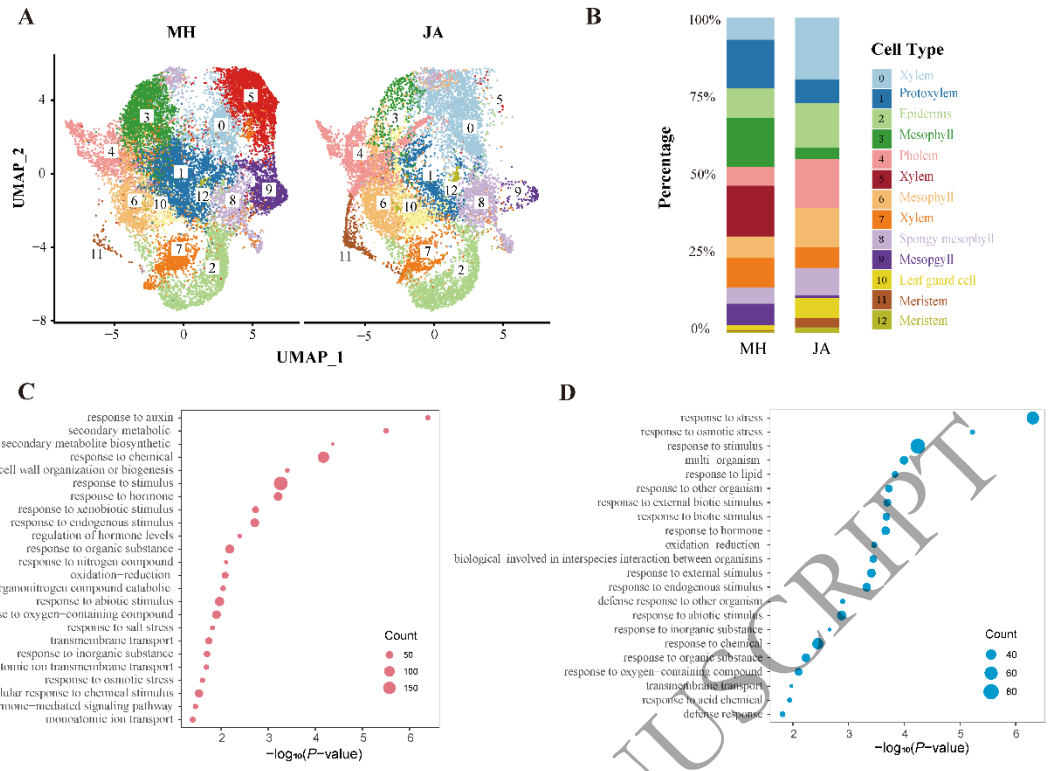

Fig. 4 | Differential expression and GO Enrichment in JA and MH grape buds. A, UMAP visualization of cell clusters from snRNA-seq data in JA and MH, displaying the distribution and clustering of cells in each grape variety; **B**, Comparative analysis of cell-type proportions between JA and MH, showing the relative abundance of different cell types in the grape bud samples; **C**, GO enrichment analysis of DEGs upregulated in JA for biological processes; **D**, GO enrichment analysis of DEGs upregulated in MH (downregulated in JA) for biological processes.

Integrative genomics of CFB in JA

To uncover the genetic basis of the CFB in JA, we integrated SVs, genomic hotspots, and DEGs from single-cell transcriptomic data. This analysis identified 757 genes within unique SV regions (Table S7), which overlapped with 5,807 DEGs, ultimately narrowing down to 37 candidate genes (Fig. 5A). Several candidate genes were implicated in pathways critical for rhythmic regulation, stress response, and cell wall modification, aligning with key traits observed in JA's CFB phenotype. Correlation analysis of the expression patterns between candidate genes, as shown in Fig. 5B and detailed in Supplementary Table S14, revealed several co-expression modules likely involved in shared regulatory pathways. For instance, Vitis07g00402 (pectinesterase) and Vitis07g00546 (GDSL esterase/lipase) displayed a strong correlation (0.964), suggesting a coordinated role in cell wall modification synchronized with circadian rhythms. This process may

facilitate daily restructuring of the bud architecture, supporting sustained growth and development. Similarly, Vitis07g02033 (zinc finger protein ZAT10) and Vitis13g00130 (LOB domain-containing protein 12-like), with a correlation of 0.928, are involved in stress response and differentiation. ZAT10, known for integrating circadian modulation and stress pathways, may regulate bud development under environmental rhythms, contributing to JA's adaptation to changing conditions. Vitis14g00733 (auxin-induced protein 22D) and Vitis14g00736 (pathogenesis-related protein PR-4), correlated at 0.678, are enriched in auxin signaling and defense pathways. Auxin's circadian regulation suggests a synchronized role in tissue differentiation and growth maintenance, directly supporting continuous flowering.

Circadian-regulated stress and defense mechanisms were also evident in unique genes. Vitis07g00128, encoding a Vicilin-like antimicrobial peptide (*AMP*), harbors an 11,768 bp insertion upstream (Fig 5C), potentially altering its expression. As a key player in defense and growth integration, its rhythmic expression aligns protective mechanisms with daily cycles, ensuring bud survival during vulnerable periods. Vitis14g00256 (respiratory burst oxidase homolog protein B, *RBOH*) features an intronic insertion and is highly expressed in buds and inflorescences during early development (Fig 5D). Its circadian-regulated expression suggests a role in coordinating early bud growth and floral development, critical for maintaining JA's CFB trait.

Further analysis identified gene pairs involved in lignin and phenolic biosynthesis, such as Vitis07g02162 and Vitis14g01562, which contribute to structural stability and bud hardening. These pathways are crucial for ensuring bud resilience and continuous flowering cycles. Additionally, photosynthesis-related genes such as Vitis14g00498 (BONZAI 3) and Vitis14g00582 (phytyl ester synthase 1), correlated at 0.954, highlight the integration of metabolic and rhythmic pathways in supporting energy demands during prolonged reproductive phases.

Our findings underscored the interplay of cell wall dynamics, hormonal regulation, and circadian control in JA unique CFB trait. The integration of SVs, DEGs, and functional annotations revealed a complex network of genetic regulation that enables JA to sustain flowering and fruiting cycles throughout the growing season.

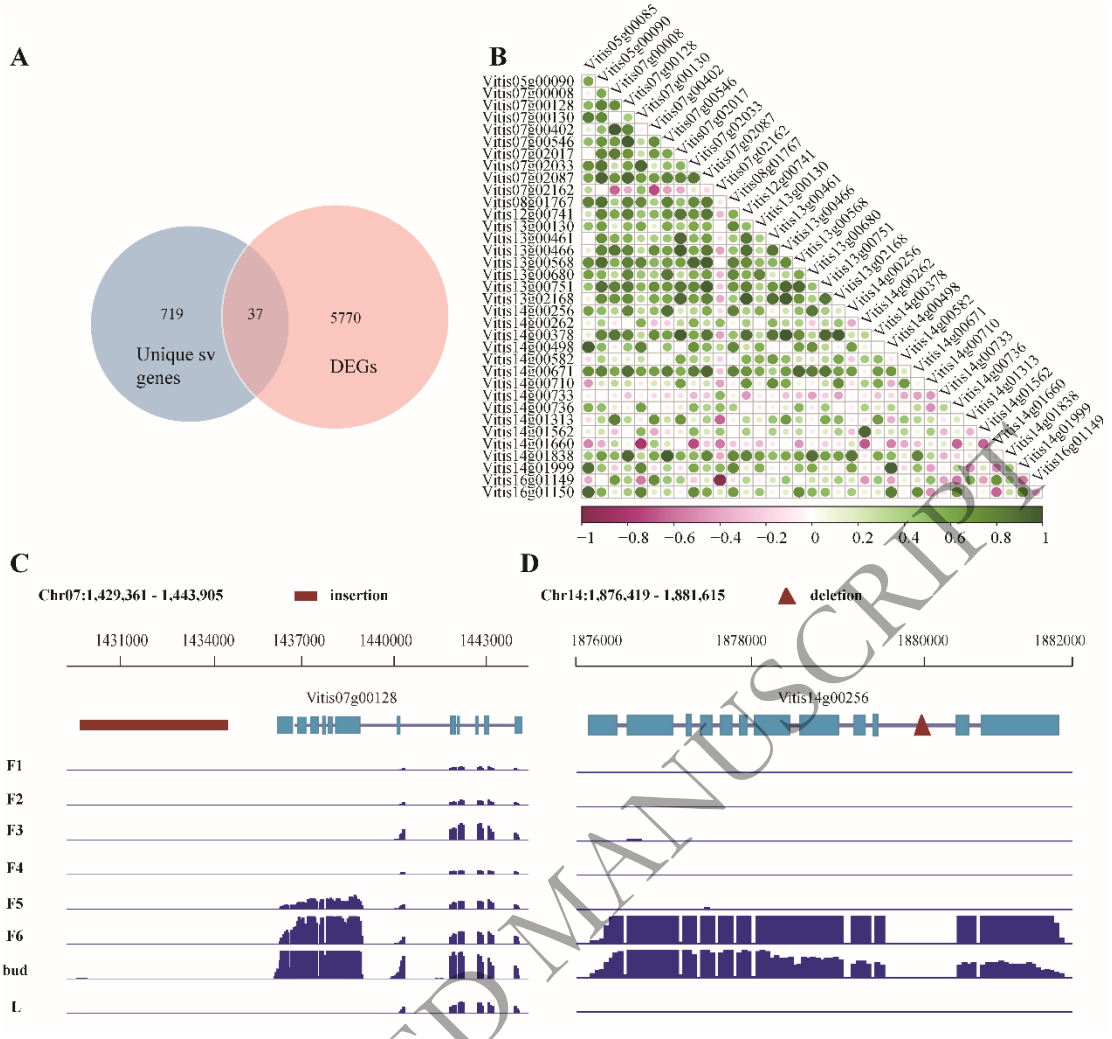

Fig. 5 | Analysis of candidate genes. A, Venn diagram showing the overlap of genes from JA unique SVs, and DEGs identified in single-cell transcriptomics; **B**, Correlation heatmap of candidate genes, with red indicating negative and green positive correlations, and circle size representing correlation strength; **C, D** Chromosomal location and expression patterns of candidate genes across fruit development stages (F1–F5), inflorescences (F6), leaves (L), and buds.

Discussion

Dissecting the genetic basis of CFB is crucial in plant biology and crop breeding, as these traits directly influence agricultural productivity and adaptability. However, the genetic basis of continuous flowering remains largely unexplored in crops. In our present study, we newly generated high-quality, haplotype-resolved genome assemblies for the grape cultivar JA, known for its CFB traits. By integrating multi-omics approaches, including pan-genomics, comparative genomics, single-cell and bulk transcriptomics, we identified unique large SVs and DEGs that distinguish JA

from control grapevines with concentrated flowering and fruiting periods. The study will help in understanding the genetic basis of continuous flowering trait in plants and provide a list of candidate unique SVs and genes for generating new grape varieties with continuous flowering.

Pan-genomic and comparative genome analysis capture unique SVs and genes

Pan-genome and comparative genome analysis can discover the unique SVs or genes associated with important agronomic traits at unprecedented resolution. For instance, a pan-genome analysis of wild and cultivated soybean detected lineage-specific genes and copy number variations associated with flowering time and biotic resistance (Li et al., 2014). Similarly, in grapevine, pan-genome and comparative analyses have been employed to uncover SVs or genes tied to essential agronomic traits (Liu et al., 2024, Long et al., 2024, Wang et al., 2024, Xiao et al., 2023). The pan-genome of a specific grapevine variety identified SVs and genes related to berry and bunch size (Liu et al., 2024). Another pan-genome grapevine study found that approximately 7% of genes were unique to each accession, contributing to traits such as aroma and disease resistance (Long et al., 2024). Furthermore, comparative genome analysis revealed large SVs that were crucial for the seedless trait (Wang et al., 2024). In our study, we focused on the pan-genome of JA, a grapevine with continuous flowering capability, compared to other grapevines that flower in concentrated periods. This analysis identified a number of SVs unique to JA. GO analysis of 664 genes within these JA-unique SV regions showed enrichment in pathways associated with photoperiod regulation and flowering, such as long-day photoperiodism. These SVs are likely to represent conserved mechanisms crucial for flowering regulation, forming the genetic foundation for breeding programs focused on flowering traits.

Importantly, our pan-genome analysis identified genes specific to JA, including those affecting key genes, such as Vitis07g02034 (a light signal transduction protein) and Vitis14g00398 (a transcription factor related to the Nuclear Factor Y complex). The *NF-Y complex* is known to regulate flowering by influencing key genes like FLOWERING LOCUS T (*FT*) and *SOC1*, as demonstrated in Arabidopsis (Cao et al., 2014). In addition, Vitis07g02159, encoding a SET domain-containing histone methyltransferase, and Vitis13g00556, a homolog of the regulator of nonsense transcripts, may contribute to the epigenetic regulation of flowering genes under specific light conditions, potentially allowing JA to maintain flowering activity across variable photoperiods. These structural variants provide a genomic basis for JA's extended flowering capability,

351 distinguishing it from other table grape varieties.

352 **Single-cell transcriptomics detect cell-type-specific gene expression**

353 Single-cell transcriptomic approaches have greatly advanced our ability to study shoot apical
354 meristems (SAMs) with unprecedented resolution, enabling the identification of cell types and
355 transcriptional dynamics that were previously masked in bulk-tissue analyses (Ogura et al., 2023).
356 For instance, in *Arabidopsis thaliana*, single-cell studies have highlighted the spatial organization
357 and dynamic transitions of gene expression during embryogenesis and seed germination (Han et al.,
358 2023a). Single-nucleus RNA sequencing also provided high-resolution insights into cell-type-
359 specific gene expression patterns in grapevine buds, a tissue rich in cellular heterogeneity and
360 comprising various specialized cell types such as xylem, phloem, and meristematic cells (Monteiro
361 et al., 2021). In our study, 13 distinct cell clusters were identified, each contributing to nutrient
362 transport, photosynthesis, or developmental regulation. Among these, meristematic cells play a
363 pivotal role in sustained growth and organogenesis, as they house pluripotent stem cells that
364 continuously generate new organs throughout a plant's lifespan. GO enrichment analysis revealed
365 that meristematic cells in JA were enriched in pathways associated with cell division, chromatin
366 remodeling, and hormonal regulation, suggesting their critical contribution to the continuous
367 flowering and bearing traits of JA.

368 Furthermore, to investigate the molecular basis of between JA (CFB) and MH (SFB), a comparative
369 analysis of single-cell transcriptomes from both cultivars were conducted to focus on DEGs across
370 various cell types, particularly in meristematic clusters. In JA, DEGs were significantly enriched in
371 pathways related to cell division, chromatin remodeling, and hormone-mediated signaling processes.
372 These findings suggested that JA's meristematic cells maintain a higher level of transcriptional and
373 epigenetic plasticity, which may support continuous organogenesis and floral initiation. For example,
374 the upregulation of histone modification-related genes in JA could facilitate the sustained activation
375 of flowering-related genes, enabling its prolonged reproductive activity. In contrast, MH displayed
376 enrichment in stress-response and dormancy-related pathways, indicating a developmental strategy
377 that prioritizes environmental resilience over continuous growth.

378 Auxin-responsive genes, which are critical for regulating cell differentiation and organ development,
379 were prominently upregulated in JA's meristematic cells compared to MH. This aligns with JA's
380 need for ongoing organogenesis to sustain its flowering phenotype. Similarly, DEGs in JA's

vascular tissues were enriched in pathways associated with nutrient transport and metabolic regulation, likely supporting the increased resource demands of continuous flowering and fruit production. Conversely, MH showed higher expression of genes involved in senescence and nutrient storage, reflecting its concentrated flowering strategy. These results highlight the distinct molecular landscapes of JA and MH, driven by differential transcriptional activity in key biological pathways. By linking DEGs to specific biological functions and cell types, our findings suggest that JA's CFB trait is supported by enhanced transcriptional reprogramming and hormonal regulation, particularly within its meristematic cells.

Integrative genomics dissect key genes associated with the circadian system

Integration of SVs, DEGs, and genomic hotspots led to the identification of 37 candidate genes potentially contributing to JA's continuous flowering and bearing phenotype. Noteworthy genes include Zinc finger protein ZAT10 and members of the Cytochrome P450 family, as well as auxin-responsive and NAC domain-containing proteins, which are central to stress response and developmental processes (Huysmans et al., 2018, Mittler et al., 2006, Mizutani and Ohta, 2010). These genes involvement in pathways such as GA-mediated cellulose synthesis, ethylene-regulated cell expansion, and ABA-controlled leaf senescence further underscored their potential regulatory roles in JA's growth rhythms and development.

The genetic framework of continuous flowering extends beyond grapevines, as evidenced by parallels in species like tomato, rose, and strawberry. The TERMINATING FLOWER (*TMF*) gene in tomato (MacAlister et al., 2012), for instance, coordinates the flowering process, while genes like *VIL1*, *FRI3*, and *CO-like 2* in roses are linked to hormone-mediated regulatory mechanisms (Dong et al., 2017). In strawberries, the SEASONAL FLOWERING LOCUS (*SFL*) enables perennial flowering irrespective of short-day conditions (Gaston et al., 2021). These cases illustrate the interplay between regulatory genes, hormone interactions, and environmental adaptability, providing a comparative framework for understanding continuous flowering across different plant species.

In conclusion, the integrated multi-omics approach employed in this research provides a robust framework for understanding the genetic underpinnings of continuous flowering in grapevines. Our findings not only enhance our knowledge of JA's unique phenotype but also offer practical implications for grape breeding strategies. By leveraging insights from structural variants and gene

expression patterns, this study paves the way for advancements in agricultural practices, aiming to optimize grape production, increase yield, and improve resource efficiency throughout the growing season. Future research should focus on functional validation of the identified candidate genes and exploration of their potential applications in breeding programs for other crops exhibiting continuous flowering traits.

Materials and Methods

Plant materials and genome sequencing

The JA grapes were cultivated in the vineyard of Qingdao Agricultural University. Young leaves were collected for genome sequencing. Additionally, buds from the same nodes of both JA and MH grapevines were harvested for RNA sequencing. For RNA-seq, samples were collected from JA at five developmental stages along the same branch, from mature berries to inflorescences, as well as from leaf samples, with three biological replicates for each stage. All collected samples were rapidly frozen in liquid nitrogen prior to sequencing. Genome sequencing of JA was performed using PacBio HiFi long-read sequencing and Hi-C technology, both on a third-generation sequencing platform.

Genome assembly and annotation

Before assembly, a preliminary genome assessment was conducted. A k-mer histogram was generated using Jellyfish (v2.3.0) (Marçais and Kingsford, 2011) with the parameters -C, -m 21, -s 1000000000 to compute k-mer frequencies from sequencing reads with a k-mer size of 21. The k-mer frequency distribution was subsequently visualized and analyzed using GenomeScope (v2.0) (Vurture et al., 2017) to infer genome characteristics, including genome size and heterozygosity rate. First, raw Hi-C data were subjected to quality control and filtering using fastp (v0.23.2) (Chen, 2023), which removed low-quality reads and adapter contamination (Table S1). Hifiasm (v0.15.5) (Cheng et al., 2021) was used to assemble and phase the diploid genome into preliminary haplotypes. Hi-C reads were aligned to contigs using Juicer (v2.0) (Durand et al., 2016b), enabling chromosome anchoring and size calculation. Additionally, RagTag (v2.1.0) (Alonge et al., 2022), with default parameters, aligned and oriented contigs to the Cabernet Sauvignon reference genome, forming scaffolds. The scaffolds were then refined at the chromosome level using 3D-DNA (v201008) (Dudchenko et al., 2017) to detect and correct potential mistakes. Manual curation of scaffolds was

performed in Juicebox (Durand et al., 2016a) to ensure assembly accuracy. The final scaffold-level assembly was completed using 3D-DNA, resulting in two phased haplotypes with minimal gaps. Next, genome quality metrics, including genome length, N50, and GC content, were calculated using seqkit (v2.2.0) (Shen et al., 2016). Genome completeness was assessed using BUSCO (v5.3.0) (Manni et al., 2021) with the embryophyta_odb10 database. Additionally, we evaluated additional assembly quality metrics using specialized tools: LAI (longest anchor interval) values were calculated by first identifying LTR sequences with LTR_FINDER (Xu and Wang, 2007) and then analyzing these sequences with LTR_retriever (Ou et al., 2018) to derive LAI scores; QV (quality value) and k-mer completeness values were computed using Merqury (v1.3.1) (Rhie et al., 2020) to assess sequence accuracy and continuity; and switch error rates, which reflect long-read alignment consistency, were determined using the calc_switchErr pipeline (Zhang et al., 2021) (https://github.com/tangerzhang/calc_switchErr).

The gene structure and repeat sequences (Table S4) of the JA genome were annotated using a genome-wide annotation pipeline (https://github.com/unavailable-2374/Genome-Wide-Annotation-Pipeline). Additionally, an annotation file was generated using Liftoff (v1.6.3) (Shumate and Salzberg, 2021) by mapping the PN40024 (V.5) reference genome to the JA genome.

To locate centromeres, we used trf (v4.09) (IBenson, 1999) to examine the length and distribution of tandem repeat monomers throughout the genome, discovering a continuous and frequent occurrence of 107 bp monomers in centromeric regions (Table S3). The centromeric regions were then defined by integrating TRF findings with TE annotations. For telomeric repeat identification, we employed the Telomere Identification Toolkit (v0.2.0) (Shen et al., 2016) to detect instances of the 6 bp motif (TTAGGG) within the genomic sequence.

Synteny analysis and identification of hemizygous genes

The entire genomes of JA, MH and PN40024 were aligned using Nucmer (v4.0.0) (Marçais et al., 2018) with the parameters --maxmatch, -c 100, -b 500, and -l 50 to generate pairwise alignments. Alignment outputs in delta format were filtered with Delta-filter using the parameters -m, -i 90, and -l 100 to retain alignments with high identity and alignment length. Filtered alignments were converted to coordinate format using show-coords -THrd for further downstream analysis. Structural variations (SVs), including inversions, translocations, and duplications, were identified and analyzed using Syri (v1.6.4) (Goel et al., 2019), which processed filtered delta files and

coordinates with the parameter -F T to detect both collinear regions and non-collinear events across the genomes (Table S5). The results were visualized using Plotsr (v1.0.1) (Goel and Schneeberger, 2022).

To identify hemizygous genes, we first mapped HiFi reads to the JA haplotype genome using Minimap2 (v2.24) (Li, 2021) and sorted the resulting BAM files with samtools (v1.15.1) (Danecek et al., 2021). Structural variants were detected using Sniffles (v2.0.6) (Smolka et al., 2024) with default parameters. Variants were filtered based on the following criteria: DEL type, END present, precise breakpoints, length ≥ 50 bp, quality ≥ 60, and support ≥ 4 reads (Peng et al., 2024). Variants meeting these criteria were further analyzed, and those overlapping entirely with gene regions were classified as hemizygous genes.

Pan-genome construction and variant calling

To construct a pan-genome graph, fifteen haplotype-assembled grape genomes, including the haploid PN40024 and both haplotypes from fourteen table grape varieties, were used. These genomes were specifically chosen to represent cultivated table grapes characterized by high levels of diploid heterozygosity. To ensure consistency and reliability in the pan-genome construction, all selected genomes were generated using third-generation sequencing technologies and assembled with a unified methodology, which collectively contributed to their high-quality assembly outcomes. To ensure consistency across datasets, all haplotypes were uniformly named using the format Name#Haplotype#Chromosome (e.g., JA#1#chr01 for chromosome 1 of haplotype 1 in JA). All sequences were merged into a single FASTA file for input. The initial graph construction was performed with PGGB (v0.6.0) (Garrison et al., 2023) using the parameters -s 5000, -l 25000, -p 90, -c 1, -K 19, -F 0.001, -g 30, -k 23, -f 0, and -B 10M, generating partitioned chromosome-specific files based on genomic community structure. The final GFA file was obtained by running commands extracted from PGGB's log file. For structural variation (SV) calling, vg deconstruct was used with JA as the reference genome to produce a VCF file. We first extracted rows from the merged VCF file where all genotype values were non-zero and not missing (.|.). From these, rows with all genotype calls consistently 1|1 were further identified and defined as unique-SV, representing loci with uniform homozygous alternate allele calls across genomes. After graph construction, Panacus (v0.0.2) (Parmigiani et al., 2024) was used to calculate pan-genome growth, pan-genome size, and core genome size. Following SV calling, bcftools (v1.13) (Danecek et al., 2021) view was employed

to analyze the VCF file, extracting and statistically analyzing SNPs, indels, and MNPs. Additionally, a custom Python script was developed to determine the type of unique SVs: if the reference length was greater than the query length, the SV was classified as an insertion (Supplementary Figure 2B); otherwise, it was classified as a deletion (Supplementary Figure 2A).

10×snRNA-seq library construction and processing

Nuclei were isolated using a previously published method (Ichino et al., 2022), with approximately 20,000 nuclei loaded into the 10× Genomics Chromium Controller to generate single-nucleus RNA-seq libraries. The nuclei were then encapsulated with reagents and barcoded gel beads in oil droplets using a microfluidic chip to form Gel Beads in Emulsions (GEMs), where RNA was released and reverse transcribed. The resulting cDNA was recovered, amplified, and subjected to quality control, followed by library construction. Sequencing was performed on the Illumina NovaSeq 6000 platform with 150 bp paired-end reads, generating approximately 100 GB of clean data per sample, across two biological replicates, totaling four samples and 400 GB of clean data.

The snRNA-seq data were processed using 10x Genomics Cell Ranger (v7.1.0) (Zheng et al., 2017) to generate gene-by-cell matrices from demultiplexed FASTQ files. First, a custom reference genome for JA nuclear transcriptome was built using cellranger mkref function, specifying the genome sequence and annotation files. This custom reference was then used to align reads from each biological replicate of JA and MH using the cellranger count function. For each sample, alignment was performed with the '--include-introns' flag enabled (Table S8). The resulting outputs for two biological replicates per cultivar were aggregated using the cellranger aggr function, with the *--normalize=mapped* parameter to normalize read depths across replicate samples before calculating cell-type proportions. (Table S9) (Supplementary Figure 3A, B). The aggregated gene-by-cell matrices were subjected to downstream analysis using Seurat (v5.1.0) (Hao et al., 2024). Quality control criteria were applied to retain nuclei with detectable gene counts between 200 and 7,500 and fewer than 15,000 unique molecular identifier (UMI) counts (Supplementary Figure 3C, D). The resulting high-quality feature matrix was used for further clustering, dimensionality reduction, and visualization of cell populations. By integrating data across biological replicates, we ensured robust and reproducible analyses of gene expression patterns.

Cell clustering, marker gene identification and differential expression analysis

To integrate the single-cell transcriptome data from the JA and MH samples, we first identified

highly variable genes across both datasets using the 'SelectIntegrationFeatures' function in Seurat. The top 3,000 genes with the most significant variability were selected as features for integration. These genes were then used to identify common features between the datasets through the 'FindIntegrationAnchors' method, which calculates integration anchors. The integration process was then carried out using 'IntegrateData (normalization.method = "LogNormalize")' ensuring that data from both samples could be combined effectively while preserving biological signals and minimizing technical variability.

After integration, dimensionality reduction and clustering were performed to identify distinct cell populations. To determine the optimal number of principal components (PCs) for downstream analysis, a quantitative approach was employed. First, the proportion of variance explained by each PC was calculated, and the cumulative variance was computed to assess the contribution of the PCs collectively. Two criteria were applied to identify the optimal PC cutoff: (1) the cumulative variance explained exceeded 90%, and (2) the individual variance explained by a PC was less than 5%. Additionally, the differences in variance explained between consecutive PCs were analyzed to capture the point where the variance drop stabilized. To validate this selection, an elbow plot was generated, displaying the variance explained by each PC. A vertical line was used to mark the selected cutoff point, corresponding to the optimal number of PCs. Based on this evaluation, 13 PCs were selected for downstream analysis (Supplementary Figure 4A). Dimensionality reduction was performed using Uniform Manifold Approximation and Projection (UMAP) with the selected PCs, enabling visualization of cell populations. Clustering was achieved using the FindNeighbors function with dimensions set to 1:13, followed by the FindClusters function at a resolution parameter of 0.4, resulting in distinct cell clusters. Marker genes for each cluster were identified using the FindAllMarkers function with the parameters only.pos = TRUE, min.pct = 0.25, and logfc.threshold = 0.25. These settings ensured that only positively enriched marker genes, expressed in at least 25% of cells within a cluster, were reported. The top 10, 20, and 50 marker genes for each cluster were ranked by average log fold change (avg_log2FC) and exported as CSV files for further analysis (Tables S11 and S12). Heatmaps of the top 10, 20, and 50 marker genes were generated to visualize their expression patterns across clusters, highlighting the distinct transcriptional profiles of each population.

To annotate cell types, cluster-specific marker genes were compared with previously reported

marker genes from *Arabidopsis thaliana* through orthologous gene comparisons. Protein sequences for A. thaliana were downloaded from the TAIR database (Lamesch et al., 2011) and aligned to grapevine protein sequences (Table S10). Marker gene identification was facilitated using the scPlant (Cao et al., 2023). Annotation of clusters was further supported by GO enrichment analysis performed on the top 50 marker genes from each cluster (Table S11). For each cluster, two representative marker genes were selected, and their correlation was visualized using custom R scripts (Supplementary Figures 5, 6, 7)

Cluster-specific marker genes were subjected to differential expression analysis between JA and MH samples. Metadata in the Seurat object was updated to include sample group information, categorizing cells into JA or MH groups based on their orig.ident labels. For each cluster, a subset of the Seurat object corresponding to that cluster was created, and the FindMarkers function was used to identify DEGs between JA and MH samples within each cluster. The parameters min.pct and logfc.threshold were kept at their default values of 0.1 and 0.25, respectively, ensuring the identification of biologically meaningful DEGs. The DEG results for all clusters were consolidated into a single data frame and annotated with cluster identifiers. This comprehensive DEG dataset was exported as a CSV file for downstream analyses (Table S12). Heatmaps of DEGs were generated to visualize expression differences between JA and MH across all clusters.

RNA-Seq analysis and gene expression correlation

RNA-Seq data from JA grapevine buds at developmental stages F1–F5 were aligned to the JA haplotype 1 genome reference using HISAT2 (v2.2.1) (Kim et al., 2019) with paired-end mode enabled (-q -x), specifying the genome index and the paired read files (-1 and -2). The alignment outputs in SAM format were converted to BAM, sorted, and indexed using samtools (v1.10) (Danecek et al., 2021) with the view -bS, sort, and index commands. Gene-level counts were then generated using featureCounts (v2.0.4) (Liao et al., 2014) with the parameters -a for the GTF annotation file, -p for paired-end reads, -g gene_id to group counts by gene ID, and -t exon to target exon-level counts. For expression quantification, raw counts were normalized to Transcripts Per Million (TPM) values (Table S13). For correlation analysis, TPM data for candidate genes were filtered to exclude rows with zero values, and Spearman's rank correlation coefficient was calculated to assess co-expression patterns among genes (Table S14). Correlation matrices were generated and visualized with the R package using the parameters order="hclust" and

method="color", applying hierarchical clustering to reveal gene expression relationships, and output for further analysis.

Acknowledgements

We thank members of the Zhou lab at Agricultural Genomics Institute at Shenzhen (AGIS) for discussions and comments on the project. This work was supported by the Key Research and Development Plan of Shandong Province (2024LZGC033 and 2023TZXD015), the National Natural Science Foundation of China (NSFC) (32372646), the Qingdao Natural Science Foundation (24-4-4-zrjj-135-jch), the Project of National Key Laboratory for Tropical Crop Breeding (SKLTCBYWF202507), the Science Fund Program for Distinguished Young Scholars of the National Natural Science Foundation of China (Overseas) to Yongfeng Zhou and Hainan Province Key Research and Development Project (ZDYF2024XDNY156).

Author contributions

X.L. and Y.P. conceived and designed the study; Y.C. and X.L. collected the plant materials; Y.L., Y.D. and X.F. assembled and annotated the haplotype-resolved T2T genomes; Y.L., X.W., J.L., T.Z. and Q.L. jointly constructed the pan-genome; Y.L., Y.D., Y.S., X.S., X.N.S.,Y.C.Z.,Y.R., P.W., X.J. and T.H. jointly conducted single-cell transcriptome analysis; M.D., S.Y. and S.F.Y. jointly conducted bulk transcriptomics analysis; Y.L. and Y.D. wrote the manuscript; All authors provided critical feedback and approved the final manuscript.

Data availability

The raw sequencing data, comprising PacBio HiFi long-reads, Illumina Hi-C reads, RNA-seq reads, and snRNA-seq reads, is accessible on NCBI under BioProject ID PRJNA192950 and on the National Genomics Data Center (NGDC) under BioProject ID PRJCA033121. The genome assembly and their annotations have been deposited into in Zenodo:10.5281/zenodo.14263262.

Competing interests

The authors declare no competing interests.

Supplementary Data

Extended Data Fig. 1 | Pre-assembly assessment and quality evaluation of the haplotype-resolved genome assembly. A-B, Pre-assembly evaluation results using k-mer analysis (k-mer length of 21) to assess genome size and heterozygosity, showing a heterozygosity rate of 1.44%, indicating high genome heterozygosity. Panel A is a magnified view of Panel B, providing a closer examination of heterozygosity. C, BUSCO assessment of genome completeness for the assembled haplotypes using the embryophyta_odb10 database, with both JA haplotypes exhibiting over 98% completeness, reflecting high assembly quality.

Extended Data Fig. 2 | Comparative genomics of the pan-genome to identify JA-specific variations. A, Identification of a unique 70 bp deletion on chromosome 16 of JA haplotype 1 (ZLA#1#) through pan-genome analysis. B, Detection of a 281 bp JA-specific insertion on chromosome 14 of JA haplotype 1.

Extended Data Fig. 3 | Pre-processing of single-nucleus transcriptome data. A, t-SNE projection of JA samples, showing data distribution across two replicates, with each color representing a different replicate. B, t-SNE projection of MH samples, displaying the data distribution of its two replicates, with colors indicating each replicate. C, Distribution of feature and count before data filtering, with different colors representing JA and MH samples. D, Distribution of feature and count after data filtering, showing the cleaned data set.

Extended Data Fig. 4 | Elbow plot for selecting the optimal number of principal components (PCs).

The plot shows the variance explained by each PC, with a red line indicating the cutoff at 13 PCs, selected based on a cumulative variance exceeding 90% and an individual variance below 5%.

Extended Data Fig. 5–7 | UMAP distribution of marker genes across clusters in single-nucleus transcriptome data. Each panel displays the spatial distribution of two selected marker genes within individual clusters on UMAP plots. Extended Data Fig. 5 (A–D) illustrates clusters 0–3, Extended Data Fig. 6 (A–E) covers clusters 4–8, and Extended Data Fig. 7 (A–D) represents clusters 9–12. Red and green indicate the expression of individual marker genes, while yellow highlights regions of high co-expression within each cluster.

Supplementary Table 1. Quality Control Summary for Hi-C Sequencing Data. A comparison

of sequencing metrics before and after filtering, including total reads, base quality (Q20 and Q30), GC content, and filtering statistics.

Supplementary Table S2. Summary of JA Genome Assembly. Assembly metrics for the two haplotypes of JA, including total sequence length, chromosome number, N50 statistics, maximum scaffold length, gap count, annotation of centromeres and telomeres, BUSCO completeness, gene count, masked bases, and GC content.

Supplementary Table S3. Centromere Identification in JA and MH Varieties. Coordinates of centromeres identified in chromosomes of JA and MH haplotypes, including start and end positions, centromere length, and repeat unit size for each chromosome.

Supplementary Table S4. Detailed Annotation of Repeated Sequences in JA Genome. Summary of repeated sequence annotations in both haplotypes of JA, including total length, GC content, and classifications of retroelements, LINEs, LTR elements, DNA transposons, and other repeat types. Metrics include the number of elements, occupied length, and percentage of the genome for each category.

Supplementary Table S5. Synteny and Variant Analysis Between Haplotypes. Comparative analysis of structural variant types (breakpoints, duplications, gaps, inversions, and jumps) across different haplotypes, including comparisons between JA and MH using PN40024 as the reference genome, as well as within and between haplotypes of JA and MH. Metrics include counts and percentage for each variant type in each comparison.

Supplementary Table S6. Unique Structural Variations Identified in the JA Genome.

Supplementary Table S7. Genes Located within 5 kb Upstream or Downstream of JA Unique Structural Variations.

Supplementary Table S8. Quality Control and Mapping Statistics for Single-Nucleus RNA Sequencing. Summary of sequencing and mapping quality metrics for single-nucleus RNA sequencing across four samples (JA1, JA2, MH1, and MH2), including sequencing read counts, barcode and UMI validation, sequencing saturation, Q30 base quality scores, genome mapping rates, and cell-specific statistics. Metrics include the number of estimated cells, mean reads per cell, median UMI counts, median genes per cell, and total genes detected.

Supplementary Table S9. Aggregated Statistics for JA and MH Replicates. Summary of aggregated sequencing and cell statistics for JA and MH replicates, including pre- and post-

normalization total reads, mean reads per cell, fraction of reads retained per replicate, confidently mapped reads, estimated cell counts, and median UMI and gene counts per cell.

Supplementary Table S10. Marker Genes Used for Cluster Annotation and Cell Type Identification. List of marker genes used for annotating cell clusters. Pct.1 and Pct.2 indicate the percentage of cells expressing each gene in the first and second clusters, respectively. *P_val*_adj represents the adjusted p-value for multiple testing correction, indicating the statistical significance of differential gene expression in each cluster.

Supplementary Table S11. Top 50 Marker Genes Based on Differential Expression Analysis.

Supplementary Table S12. Differentially Expressed Genes (DEGs) in JA and MH Single-Nucleus RNA Sequencing. List of differentially expressed genes (DEGs) identified from snRNA-se data comparing JA and MH. Pct.1 represents the percentage of cells expressing each gene in JA, and Pct.2 represents the percentage of cells expressing each gene in MH.

Supplementary Table S13. Correlation Matrix of Gene Expression Levels for Candidate Genes. Pairwise correlation coefficients among 36 candidate genes with non-zero expression across developmental stages (F1-F5), leaf, and bud tissues in JA.

Reference

AHMAD, S., PENG, D., ZHOU, Y. & ZHAO, K. 2022. The genetic and hormonal inducers of continuous flowering in orchids: an emerging view. *Cells* **11**, 4.

ALONGE, M., LEBEIGLE, L., KIRSCHE, M., JENIKE, K., OU, S., AGANEZOV, S., WANG, X., LIPPMAN, Z.B., SCHATZ, M.C. & SOYK, S. 2022. Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol.* **23**, 258.

CAO, S., HE, Z., CHEN, R., LUO, Y., FU, L.Y., ZHOU, X., HE, C., YAN, W., ZHANG, C.Y. & CHEN, D. 2023. scPlant: a versatile framework for single-cell transcriptomic data analysis in plants. Plant Commun. **4**, e107.

CAO, S., KUMIMOTO, R.W., GNESUTTA, N., CALOGERO, A.M., MANTOVANI, R. & HOLT, B.F. 2014. A distal CCAAT/NUCLEAR FACTOR Y complex promotes chromatin looping at the FLOWERING LOCUS T promoter and regulates the timing of flowering in Arabidopsis. *Plant Cell* **26**, 1009–1017.

CHEN, S. 2023. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *Imeta.* **2**, e107.

CHENG, H., CONCEPCION, G.T., FENG, X., ZHANG, H. & LI, H. 2021. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175.

DANECEK, P., BONFIELD, J.K., LIDDLE, J., MARSHALL, J., OHAN, V., POLLARD, M.O., WHITWHAM, A., KEANE, T., MCCARTHY, S.A., DAVIES, R.M. & LI, H. 2021. Twelve

years of SAMtools and BCFtools. *Gigascience* **10**, giab008.
DíAZ-RIQUELME, J., GRIMPLET, J., MARTíNEZ-ZAPATER, J. M. & CARMONA, M. J. 2012. Transcriptome variation along bud development in grapevine (Vitis vinifera L.). *BMC Plant Biol*, **12**, 181.
DONG, X., JIANG, X., KUANG, G., WANG, Q., ZHONG, M., JIN, D. & HU, J. 2017. Genetic control of flowering time in woody plants: roses as an emerging model. *Plant Divers.* **39**, 104–110.
DUDCHENKO, O., BATRA, S.S., OMER, A.D., NYQUIST, S.K., HOEGER, M., DURAND, N.C., SHAMIM, M.S., MACHOL, I., LANDER, E.S., AIDEN, A.P. & AIDEN, E.L. 2017. De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. *Science* **356**, 92–95.
DURAND, N.C., ROBINSON, J.T., SHAMIM, M.S., MACHOL, I., MESIROV, J.P., LANDER, E.S. & AIDEN, E.L. 2016a. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst.* **3**, 99–101.
DURAND, N.C., SHAMIM, M.S., MACHOL, I., RAO, S.S., HUNTLEY, M.H., LANDER, E.S. & AIDEN, E.L. 2016b. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell Syst. 3, 95–98.
GAO, L., GONDA, I., SUN, H., MA, Q., BAO, K., TIEMAN, D.M., BURZYNSKI-CHANG, E.A., FISH, T.L., STROMBERG, K.A., SACKS, G.L., THANNHAUSER, T.W., FOOLAD, M.R., DIEZ, M.J., BLANCA, J., CANIZARES, J., XU, Y., VAN DER KNAAP, E., HUANG, S., KLEE, H.J., GIOVANNONI, J.J. & FEI, Z. 2019. The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat. Genet.* **51**, 1044–1051.
Garrison, E., Guarracino, A., Heumos, S., Villani, F., Bao, Z., Tattini, L., Hagmann, J., Vorbrugg, S., Marco-Sola, S., Kubica, C., Ashbrook, D.G., Thorell, K., Rusholme-Pilcher, R.L., Liti, G., Rudbeck, E., Nahnsen, S., Yang, Z., Moses, M.N., Nobrega, F.L., Wu, Y., Chen, H., De Ligt, J., Sudmant, P.H., Soranzo, N., Colonna, V., Williams, R.W. and Prins, P. (2024). Building pangenome graphs. *Nat. Methods*, **21**, 2008–2012.
GASTON, A., POTIER, A., ALONSO, M., SABBADINI, S., DELMAS, F., TENREIRA, T., COCHETEL, N., LABADIE, M., PRéVOST, P., FOLTA, K. M., MEZZETTI, B., HERNOULD, M., ROTHAN, C. & DENOYES, B. 2021. The FveFT2 florigen/FveTFL1 antflorigen balance is critical for the control of seasonal flowering in strawberry while FveFT3 modulates axillary meristem fate and yield. *New Phytol,* **232,** 372-387.
GOEL, M. & SCHNEEBERGER, K. 2022. plotsr: visualizing structural similarities and rearrangements between multiple genomes. *Bioinformatics,* **38,** 2922-2926.
GOEL, M., SUN, H., JIAO, W.-B. & SCHNEEBERGER, K. 2019. SyRI: finding genomic rearrangements and local sequence differences from whole-genome assemblies. *Genome Biol.,* **20,** 277.
HAN, X., ZHANG, Y., LOU, Z., LI, J., WANG, Z., GAO, C., LIU, Y., REN, Z., LIU, W., LI, B., PAN, W., ZHANG, H., SANG, Q., WAN, M., HE, H. & DENG, X. W. 2023a. Time series single-cell transcriptional atlases reveal cell fate differentiation driven by light in Arabidopsis seedlings. *Nat. Plants,* **9**, 2095-2109.
HAO, Y., STUART, T., KOWALSKI, M. H., CHOUDHARY, S., HOFFMAN, P., HARTMAN, A., SRIVASTAVA, A., MOLLA, G., MADAD, S., FERNANDEZ-GRANDA, C. & SATIJA, R. 2024. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.,* **42**, 293-304.

HüBNER, S., BERCOVICH, N., TODESCO, M., MANDEL, J. R., ODENHEIMER, J., ZIEGLER, E., LEE, J. S., BAUTE, G. J., OWENS, G. L., GRASSA, C. J., EBERT, D. P., OSTEVIK, K. L., MOYERS, B. T., YAKIMOWSKI, S., MASALIA, R. R., GAO, L., ĆALIĆ, I., BOWERS, J. E., KANE, N. C., SWANEVELDER, D. Z. H., KUBACH, T., MUñOS, S., LANGLADE, N. B., BURKE, J. M. & RIESEBERG, L. H. 2019. Sunflower pan-genome analysis shows that hybridization altered gene content and disease resistance. *Nat. Plants,* **5**, 54-62.

HUYSMANS, M., BUONO, R. A., SKORZINSKI, N., RADIO, M. C., DE WINTER, F., PARIZOT, B., MERTENS, J., KARIMI, M., FENDRYCH, M. & NOWACK, M. K. 2018. NAC Transcription Factors ANAC087 and ANAC046 Control Distinct Aspects of Programmed Cell Death in the Arabidopsis Columella and Lateral Root Cap. *Plant Cell,* **30**, 2197-2213.

HYTONEN, T. & KUROKURA, T. 2020. Control of Flowering and Runnering in Strawberry. *Hortic. J.,* **89**, 96-107.

ICHINO, L., PICARD, C. L., YUN, J., CHOTAI, M., WANG, S., LIN, E. K., PAPAREDDY, R. K., XUE, Y. & JACOBSEN, S. E. 2022. Single-nucleus RNA-seq reveals that MBD5, MBD6, and SILENZIO maintain silencing in the vegetative cell of developing pollen. *Cell Rep*, **41**, 111699.

IKEGAMI, H., NOGATA, H., INOUE, Y., HIMENO, S., YAKUSHIJI, H., HIRATA, C., HIRASHIMA, K., MORI, M., AWAMURA, M. & NAKAHARA, T. 2013. Expression of FcFT1, a FLOWERING LOCUS T-like gene, is regulated by light and associated with inflorescence differentiation in fig (Ficus caricaL.). *BMC Plant Biol.,* **13**, 216.

IWATA, H., GASTON, A., REMAY, A., THOUROUDE, T., JEAUFFRE, J., KAWAMURA, K., OYANT, L. H.-S., ARAKI, T., DENOYES, B. & FOUCHER, F. 2012. The TFL1 homologue KSN is a regulator of continuous flowering in rose and strawberry. *Plant J.,* **69**, 116-125.

JEAN-BAPTISTE, K., MCFALINE-FIGUEROA, J. L., ALEXANDRE, C. M., DORRITY, M. W., SAUNDERS, L., BUBB, K. L., TRAPNELL, C., FIELDS, S., QUEITSCH, C. & CUPERUS, J. T. 2019. Dynamics of Gene Expression in Single Root Cells of Arabidopsis thaliana. *Plant Cell,* **31**, 993-1011.

KIM, D., PAGGI, J. M., PARK, C., BENNETT, C. & SALZBERG, S. L. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.,* **37**, 907-915.

LAI, Z., SCHLUTTENHOFER, C. M., BHIDE, K., SHREVE, J., THIMMAPURAM, J., LEE, S. Y., YUN, D.-J. & MENGISTE, T. 2014. MED18 interaction with distinct transcription factors regulates multiple plant functions. *Nat. Commu.*, **5**, 3064.

LAMESCH, P., BERARDINI, T. Z., LI, D., SWARBRECK, D., WILKS, C., SASIDHARAN, R., MULLER, R., DREHER, K., ALEXANDER, D. L., GARCIA-HERNANDEZ, M., KARTHIKEYAN, A. S., LEE, C. H., NELSON, W. D., PLOETZ, L., SINGH, S., WENSEL, A. & HUALA, E. 2011. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.,* **40**, D1202-D1210.

LBENSON, G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.,* **27**, 573-580.

LEE, T. A., NOBORI, T., ILLOUZ-ELIAZ, N., XU, J., JOW, B., NERY, J. R. & ECKER, J. R. 2023. A Single-Nucleus Atlas of Seed-to-Seed Development in Arabidopsis. *bioRxiv*, 2023.03.23.533992.

LI, H. 2021. New strategies to improve minimap2 alignment accuracy. *Bioinformatics,* **37**, 4572-4574.

LI, Y.-H., ZHOU, G., MA, J., JIANG, W., JIN, L.-G., ZHANG, Z., GUO, Y., ZHANG, J., SUI, Y., ZHENG, L., ZHANG, S.-S., ZUO, Q., SHI, X.-H., LI, Y.-F., ZHANG, W.-K., HU, Y., KONG,

G., HONG, H.-L., TAN, B., SONG, J., LIU, Z.-X., WANG, Y., RUAN, H., YEUNG, C. K. L., LIU, J., WANG, H., ZHANG, L.-J., GUAN, R.-X., WANG, K.-J., LI, W.-B., CHEN, S.-Y., CHANG, R.-Z., JIANG, Z., JACKSON, S. A., LI, R. & QIU, L.-J. 2014. De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. *Nat. Biotechnol.,* **32**, 1045-1052.

LIAO, Y., SMYTH, G. K. & SHI, W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics,* **30**, 923-30.

LIU, Q., LIANG, Z., FENG, D., JIANG, S., WANG, Y., DU, Z., LI, R., HU, G., ZHANG, P., MA, Y., LOHMANN, J. U. & GU, X. 2021. Transcriptional landscape of rice roots at the single-cell resolution. *Mol. Plant,* **14**, 384-394.

LIU, Z., KONG, X., LONG, Y., LIU, S., ZHANG, H., JIA, J., CUI, W., ZHANG, Z., SONG, X., QIU, L., ZHAI, J. & YAN, Z. 2023. Integrated single-nucleus and spatial transcriptomics captures transitional states in soybean nodule maturation. *Nat. Plants,* **9**, 515-524.

LIU, Z., WANG, N., SU, Y., LONG, Q., PENG, Y., SHANGGUAN, L., ZHANG, F., CAO, S., WANG, X., GE, M., XUE, H., MA, Z., LIU, W., XU, X., LI, C., CAO, X., AHMAD, B., SU, X., LIU, Y., HUANG, G., DU, M., LIU, Z., GAN, Y., SUN, L., FAN, X., ZHANG, C., ZHONG, H., LENG, X., REN, Y., DONG, T., PEI, D., WU, X., JIN, Z., WANG, Y., LIU, C., CHEN, J., GAUT, B., HUANG, S., FANG, J., XIAO, H. & ZHOU, Y. 2024. Grapevine pangenome facilitates trait genetics and genomic breeding. *Nat Genet*., **56**, 2804-2814

LONG, Q., CAO, S., HUANG, G., WANG, X., LIU, Z., LIU, W., WANG, Y., XIAO, H., PENG, Y. & ZHOU, Y. 2024. Population comparative genomics discovers gene gain and loss during grapevine domestication. *Plant Physiol.,* **195**, 1401-1413.

LONG, Y., LIU, Z., JIA, J., MO, W., FANG, L., LU, D., LIU, B., ZHANG, H., CHEN, W. & ZHAI, J. 2021. FlsnRNA-seq: protoplasting-free full-length single-nucleus RNA profiling in plants. *Genome Biol.,* **22**, 66.

MACALISTER, C. A., PARK, S. J., JIANG, K., MARCEL, F., BENDAHMANE, A., IZKOVICH, Y., ESHED, Y. & LIPPMAN, Z. B. 2012. Synchronization of the flowering transition by the tomato TERMINATING FLOWER gene. *Nat. Genet.,* **44**, 1393-1398.

MANNI, M., BERKELEY, M. R., SEPPEY, M., SIMãO, F. A. & ZDOBNOV, E. M. 2021. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Mol. Biol. Evol.,* **38**, 4647-4654.

MARAND, A. P., CHEN, Z., GALLAVOTTI, A. & SCHMITZ, R. J. 2021. A cis-regulatory atlas in maize at single-cell resolution. *Cell,* **184**, 3041-3055.e21.

MARçAIS, G., DELCHER, A. L., PHILLIPPY, A. M., COSTON, R., SALZBERG, S. L. & ZIMIN, A. 2018. MUMmer4: A fast and versatile genome alignment system. *PLoS Comput. Biol.,* **14**, e1005944.

MARçAIS, G. & KINGSFORD, C. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics,* **27**, 764-770.

MITTLER, R., KIM, Y., SONG, L., COUTU, J., COUTU, A., CIFTCI-YILMAZ, S., LEE, H., STEVENSON, B. & ZHU, J. K. 2006. Gain- and loss-of-function mutations in Zat10 enhance the tolerance of plants to abiotic stress. *FEBS Lett.,* **580**, 6537-42.

MIZUTANI, M. & OHTA, D. 2010. Diversification of P450 genes during land plant evolution. *Annu. Rev. Plant Biol.,* **61**, 291-315.

MONTEIRO, A. I., MALHEIRO, A. C. & BACELAR, E. A. 2021. Morphology, Physiology and Analysis Techniques of Grapevine Bud Fruitfulness: A Review. **11**, 127.

MYLES, S., BOYKO, A. R., OWENS, C. L., BROWN, P. J., GRASSI, F., ARADHYA, M. K., PRINS, B., REYNOLDS, A., CHIA, J. M., WARE, D., BUSTAMANTE, C. D. & BUCKLER, E. S. 2011. Genetic structure and domestication history of the grape. *Proc Natl Acad Sci U S A,* **108**, 3530-5.

OU, S., CHEN, J. & JIANG, N. 2018. Assessing genome assembly quality using the LTR Assembly Index (LAI). *Nucleic Acids Research*, **46**, e126-e126.

OU, S. & JIANG, N. 2018. LTR_retriever: A Highly Accurate and Sensitive Program for Identification of Long Terminal Repeat Retrotransposons. *Plant Physiol*, **176**, 1410-1422.

OGURA, N., SASAGAWA, Y., ITO, T., TAMESHIGE, T., KAWAI, S., SANO, M., DOLL, Y., IWASE, A., KAWAMURA, A., SUZUKI, T., NIKAIDO, I., SUGIMOTO, K. & IKEUCHI, M. 2023. WUSCHEL-RELATED HOMEOBOX 13 suppresses de novo shoot regeneration via cell fate control of pluripotent callus. *Sci. Adv.,* **9**, eadg6983.

PARMIGIANI, L., GARRISON, E., STOYE, J., MARSCHALL, T. & DOERR, D. 2024. Panacus: fast and exact pangenome growth and core size estimation. *Bioinformatics,* **40**, btae720.

PENG, Y., WANG, Y., LIU, Y., FANG, X., CHENG, L., LONG, Q., XU, Q., WANG, N., ZHANG, F., LIU, Z., XIAO, H., HU, W., CHEN, S., HUANG, S., GAUT, B. S. & ZHOU, Y. 2024. The genomic and epigenomic dynamics of hemizygous genes across crops with contrasting mating systems. *Proc Natl Acad Sci U S A*., **122**, e2422487122.

QIN, P., LU, H., DU, H., WANG, H., CHEN, W., CHEN, Z., HE, Q., OU, S., ZHANG, H., LI, X., LI, X., LI, Y., LIAO, Y., GAO, Q., TU, B., YUAN, H., MA, B., WANG, Y., QIAN, Y., FAN, S., LI, W., WANG, J., HE, M., YIN, J., LI, T., JIANG, N., CHEN, X., LIANG, C. & LI, S. 2021. Pan-genome analysis of 33 genetically diverse rice accessions reveals hidden genomic variations. *Cell,* **184**, 3542-3558.e16.

RHIE, A., WALENZ, B. P., KOREN, S. & PHILLIPPY, A. M. 2020. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol*, **21**, 245.

SHANGGUAN, L., CHEN, M., FANG, X., XIE, Z., GONG, P., HUANG, Y., WANG, Z. & FANG, J. 2020. Comparative transcriptome analysis provides insight into regulation pathways and temporal and spatial expression characteristics of grapevine (Vitis vinifera) dormant buds in different nodes. *BMC Plant Biol*, **20**, 390.

SHEN, W., LE, S., LI, Y. & HU, F. 2016. SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PLOS ONE,* **11**, e0163962.

SHI, X., CAO, S., WANG, X., HUANG, S., WANG, Y., LIU, Z., LIU, W., LENG, X., PENG, Y., WANG, N., WANG, Y., MA, Z., XU, X., ZHANG, F., XUE, H., ZHONG, H., WANG, Y., ZHANG, K., VELT, A., AVIA, K., HOLTGRäWE, D., GRIMPLET, J., MATUS, J. T., WARE, D., WU, X., WANG, H., LIU, C., FANG, Y., RUSTENHOLZ, C., CHENG, Z., XIAO, H. & ZHOU, Y. 2023. The complete reference genome for grapevine (Vitis vinifera L.) genetics and breeding. *Hortic. Res.,* **10**, uhad061.

SHUMATE, A. & SALZBERG, S. L. 2021. Liftoff: accurate mapping of gene annotations. *Bioinformatics,* **37**, 1639-1643.

SMOLKA, M., PAULIN, L. F., GROCHOWSKI, C. M., HORNER, D. W., MAHMOUD, M., BEHERA, S., KALEF-EZRA, E., GANDHI, M., HONG, K., PEHLIVAN, D., SCHOLZ, S. W., CARVALHO, C. M. B., PROUKAKIS, C. & SEDLAZECK, F. J. 2024. Detection of mosaic

and population-level structural variants with Sniffles2. *Nat. Biotechnology,* **42**, 1571-1580.
STENZEL, I., OTTO, M., DELKER, C., KIRMSE, N., SCHMIDT, D., MIERSCH, O., HAUSE, B. & WASTERNACK, C. 2012. ALLENE OXIDE CYCLASE (AOC) gene family members of Arabidopsis thaliana: tissue- and organ-specific promoter activities and in vivo heteromerization. *J Exp Bot.*, **63**, 6125-38.
SUNAGA-FRANZE, D. Y., MUINO, J. M., BRAEUNING, C., XU, X., ZONG, M., SMACZNIAK, C., YAN, W., FISCHER, C., VIDAL, R., KLIEM, M., KAUFMANN, K. & SAUER, S. 2021. Single-nuclei RNA-sequencing of plant tissues. *bioRxiv*, 2020.11.14.382812.
TIAN, C., DU, Q., XU, M., DU, F. & JIAO, Y. 2020. Single-nucleus RNA-seq resolves spatiotemporal developmental trajectories in the tomato shoot apex. *bioRxiv*, 2020.09.20.305029.
TONG, C., JIA, Y., HU, H., ZENG, Z., CHAPMAN, B. & LI, C. 2024. Pangenome and pantranscriptome as the new reference for gene family characterisation – a case study of basic helix-loop-helix (bHLH) genes in barley. *Plant Commun.,* **6**, 101190.
VURTURE, G. W., SEDLAZECK, F. J., NATTESTAD, M., UNDERWOOD, C. J., FANG, H., GURTOWSKI, J. & SCHATZ, M. C. 2017. GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics,* **33**, 2202-2204.
WANG, Q., GAO, G., CHEN, X., LIU, X., DONG, B., WANG, Y., ZHONG, S., DENG, J., FANG, Q. & ZHAO, H. 2022. Genetic studies on continuous flowering in woody plant Osmanthus fragrans. *Front Plant Sci,* **13**, 1049479.
WANG, X., LIU, Z., ZHANG, F., XIAO, H., CAO, S., XUE, H., LIU, W., SU, Y., LIU, Z., ZHONG, H., ZHANG, F., AHMAD, B., LONG, Q., ZHANG, Y., LIU, Y., GAN, Y., HOU, T., JIN, Z., WU, X., LIU, G., WANG, Y., PENG, Y. & ZHOU, Y. 2024. Integrative genomics reveals the polygenic basis of seedlessness in grapevine. *Curr. Biol.,* **34**, 3763-3777.
WUBS, A. M., MA, Y., HEMERIK, L. & HEUVELINK, E. 2009. Fruit Set and Yield Patterns in Six Capsicum Cultivars. *HortScience horts.*, **44**, 1296-1301.
XIAO, H., LIU, Z., WANG, N., LONG, Q., CAO, S., HUANG, G., LIU, W., PENG, Y., RIAZ, S., WALKER, A. M., GAUT, B. S. & ZHOU, Y. 2023. Adaptive and maladaptive introgression in grapevine domestication. *Proc Natl Acad Sci U S A.,* **120**, e2222041120.
XU, Z. & WANG, H. 2007. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.*, **35**, W265-8.
YANG, M. C., WU, Z. C., CHEN, R. Y., ABBAS, F., HU, G. B., HUANG, X. M., GUAN, W. S., XU, Y. S. & WANG, H. C. 2023. Single-nucleus RNA sequencing and mRNA hybridization indicate key bud events and LcFT1 and LcTFL1-2 mRNA transportability during floral transition in litchi. *J. Exp. Bot.,* **74**, 3613-3629.
YI, X., GAO, H., YANG, Y., YANG, S., LUO, L., YU, C., WANG, J., CHENG, T., ZHANG, Q. & PAN, H. 2021. Differentially Expressed Genes Related to Flowering Transition between Once- and Continuous-Flowering Roses. *Biomolecules,* **12**.
ZHANG, T., PENG, W., XIAO, H., CAO, S., CHEN, Z., SU, X., LUO, Y., LIU, Z., PENG, Y., YANG, X., JIANG, G. F., XU, X., MA, Z. & ZHOU, Y. 2024. Population genomics highlights structural variations in local adaptation to saline coastal environments in woolly grape. *J. Integr. Plant Biol.,* **66**, 1408-1426.
ZHANG, X., CHEN, S., SHI, L., GONG, D., ZHANG, S., ZHAO, Q., ZHAN, D., VASSEUR, L., WANG, Y., YU, J., LIAO, Z., XU, X., QI, R., WANG, W., MA, Y., WANG, P., YE, N., MA, D., SHI, Y., WANG, H., MA, X., KONG, X., LIN, J., WEI, L., MA, Y., LI, R., HU, G., HE, H., ZHANG,

L., MING, R., WANG, G., TANG, H. & YOU, M. 2021. Haplotype-resolved genome assembly provides insights into evolutionary history of the tea plant Camellia sinensis. *Nat. Genet.*, **53**, 1250-1259.

ZHENG, G. X. Y., TERRY, J. M., BELGRADER, P., RYVKIN, P., BENT, Z. W., WILSON, R., ZIRALDO, S. B., WHEELER, T. D., MCDERMOTT, G. P., ZHU, J., GREGORY, M. T., SHUGA, J., MONTESCLAROS, L., UNDERWOOD, J. G., MASQUELIER, D. A., NISHIMURA, S. Y., SCHNALL-LEVIN, M., WYATT, P. W., HINDSON, C. M., BHARADWAJ, R., WONG, A., NESS, K. D., BEPPU, L. W., DEEG, H. J., MCFARLAND, C., LOEB, K. R., VALENTE, W. J., ERICSON, N. G., STEVENS, E. A., RADICH, J. P., MIKKELSEN, T. S., HINDSON, B. J. & BIELAS, J. H. 2017. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.,* **8**, 14049.

ZHOU, Y., MASSONNET, M., SANJAK, J. S., CANTU, D. & GAUT, B. S. 2017. Evolutionary genomics of grape (Vitis vinifera ssp. vinifera) domestication. *Proc Natl Acad Sci U S A.,* **114**, 11715-11720.

ZHOU, Y., MINIO, A., MASSONNET, M., SOLARES, E., LV, Y., BERIDZE, T., CANTU, D. & GAUT, B. S. 2019. The population genetics of structural variants in grapevine domestication. *Nat. Plants,* **5**, 965-979.