

Integrative genomic and transcriptomic analyses of a bud sport mutant ‘Jinzaowu’ with the phenotype of large berries in grapevines

Jianquan Huang^{1,*}, Guan Zhang^{2,3,*}, Yanhao Li^{1,4}, Mingjie Lyu², He Zhang¹, Na Zhang¹ and Rui Chen²

¹The Research Institute of Forestry and Pomology, Tianjin Academy of Agricultural Sciences, Tianjin, China

²Institute of Crop Germplasm and Biotechnology, Tianjin Academy of Agricultural Sciences, Tianjin, China

³College of Biotechnology and Food Science, Tianjin University of Commerce, Tianjin, China

⁴College of Horticulture and Gardening, Tianjin Agricultural University, Tianjin, China

* These authors contributed equally to this work.

ABSTRACT

Background: Bud sport mutation occurs frequently in fruit plants and acts as an important approach for grapevine improvement and breeding. ‘Jinzaowu’ is a bud sport of the elite cultivar ‘Himord Seedless’ with obviously enlarged organs and berries. To date, the molecular mechanisms underlying berry enlargement caused by bud sport in grapevines remain unclear.

Methods: Whole genome resequencing (WGRS) was performed for two pairs of bud sports and their maternal plants with similar phenotype to identify SNPs, InDels and structural variations (SVs) as well as related genes. Furthermore, transcriptomic sequencing at different developmental stages and weighted gene co-expression network analysis (WGCNA) for ‘Jinzaowu’ and its maternal plant ‘Himord Seedless’ were carried out to identify the differentially expressed genes (DEGs), which were subsequently analyzed for Gene Ontology (GO) and function annotation.

Results: In two pairs of enlarged berry bud sports, a total of 1,334 SNPs, 272 InDels and 74 SVs, corresponding to 1,022 target genes related to symbiotic microorganisms, cell death and other processes were identified. Meanwhile, 1,149 DEGs associated with cell wall modification, stress-response and cell killing might be responsible for the phenotypic variation were also determined. As a result, 42 DEGs between ‘Himord Seedless’ and ‘Jinzaowu’ harboring genetic variations were further investigated, including pectin esterase, cellulase A, cytochromes P450 (CYP), UDP-glycosyltransferase (UGT), zinc finger protein, auxin response factor (ARF), NAC transcription factor (TF), protein kinase, *etc.* These candidate genes offer important clues for a better understanding of developmental regulations of berry enlargement in grapevine.

Conclusion: Our results provide candidate genes and valuable information for dissecting the underlying mechanisms of berry development and contribute to future improvement of grapevine cultivars.

Submitted 20 June 2022

Accepted 1 December 2022

Published 3 January 2023

Corresponding authors

Na Zhang, 2371254838@qq.com

Rui Chen, chenrui_aglab@126.com

Academic editor

Othmane Merah

Additional Information and
Declarations can be found on
page 17

DOI 10.7717/peerj.14617

© Copyright

2023 Huang et al.

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Agricultural Science, Bioinformatics, Genomics, Plant Science

Keywords Grapevine, Bud sport, Berry enlargement, Genetic variation, RNA-seq, WGRS

INTRODUCTION

Grape (*Vitis vinifera* L.) is one of the most favorite fruits worldwide with great economic and nutritional value (Yang & Xiao, 2013). Different from routine improvement strategies in crops, grapevine is highly heterozygous and exhibit serious genetic segregation, which hinders the development process of excellent traits. So far, various breeding techniques had been applied to improve the genetic background of grapes and achieve reliable and stable inheritance of desirable traits (Yamada & Sato, 2016). Out of these approaches, bud sport is an important strategy for grapevine improvement and breeding. Bud sport mutation is a phenomenon usually happened in fruit plants in which stable somatic cells are mutated in the meristematic tissue of the bud and exhibit novel traits due to endogenous and exogenous factors. In this process, favorable phenotypes caused by mutation (e.g., shorter ripening cycle, larger berries, and different skin color) will be selected by breeders that are significantly different from the maternal plants with desired and superior characteristics (Foster & Aranzana, 2018). To date, numerous bud sport mutants have been observed and successfully utilized in grapevine breeding, such as early-ripening cultivars ‘Fengzao’, ‘Tiangong Moyu’, and ‘Nantaihutezao’ (Guo & Zhang, 2015; Wei et al., 2020; Leng et al., 2021), as well as berry skin color mutants ‘Pinot Grigio’, ‘Pinot Blanc’ and ‘Benitaka’ (Azuma et al., 2009; Pan et al., 2012; Negri et al., 2015). However, the phenotype with large berries induced by bud sport has not been reported and the corresponding regulatory mechanisms remain poorly understood.

Berry size is a polygenic quantitative trait in grapevine. Similar to other fruit plants, grapevine follows a double sigmoid curve corresponding to the fruit enlargement, véraison and maturity stages. Few days after flowering, berry development occurs through cell division and expansion. During the enlargement stage, the berry grows slowly and ends with the véraison stage, where it softens and the color of the skin changes. In the véraison stage, berry growth restarts but exclusively through cell enlargement. In the maturity stage, all the tissues that make up the berries are formed. Previous studies in Arabidopsis (Krizek, 1999; Wang et al., 2016), apple (Malladi & Hirst, 2010), tomato (Su et al., 2014) and peach (Scorzal et al., 1991) indicated that cell division and cell expansion play crucial roles in controlling fruit size.

Several key genes and regulators had been demonstrated to be responsible for berry enlargement which involved in cell wall modification, phytohormone production, water and sugar transport processes. “Pectin Lyase (PL)”, “wall-associated receptor kinase (WAK)” and “Pectin Esterase (PE)” are related with the accumulation of pectin, that can directly bind pectin polymers to reduce stiffness and stimulate cell wall development (Tucker et al., 2018). Ethylene was previously thought to have limited effects on grape growth and development, that is why grapevines were classified as non-climacteric berries (Chervin et al., 2004). Recent evidence suggested that ethylene coupled with increased ACO transcription could produce a modest peak during berry ripening and promote larger berry diameter, higher sugar accumulation (De La Torre-Ruiz et al., 2016), and higher

anthocyanin synthesis (*El-kereamy et al., 2003*). UGT, as sugar donors, are primarily responsible for the formation of secondary metabolites such as antibiotics, flavanols, and steroids through glycosylation modifications that affect plant growth and development (*Kiedrowski et al., 1992; Mitchell, Hall & Barber, 1994; Breton, Fournel-Gigleux & Palcic, 2012*). Cytochrome P450 (CYP) is a highly conserved enzyme family that plays significant roles in plant development and growth by increasing cell size, responding to biotic and abiotic stresses, and stimulating phytohormone signal transduction (*Ma et al., 2015*). *PaCYP78A9* was inferred to be involved in the regulation of organ size in sweet cherry (*Qi et al., 2017*). CYPs, such as *EOD3/CYP78A6* and *CYP78A9* in Arabidopsis (*Fang et al., 2012*); *CYP78A5* in tomato (*Chakrabarti et al., 2013*), *GmCYP78A10* in soybean (*Wang et al., 2015*), *CYP78A13* (*Xu et al., 2015*) and *GW10* in rice (*Zhan et al., 2021*), have all been identified as seed size regulators.

Transcription factors (TFs) are also important regulators in controlling berry size which are typically transcribed in response to hormonal and environmental stimuli. NAC transcription factor has been proven to be functional in fruit development and maturation (*Migicovsky et al., 2021*). For example, the participation of *VvNAC26* determinates the berry ultimate size in grapevine (*Tello et al., 2015*). Overexpression of *VvCEB1* and *VvSAUR41* genes in grapevine embryos increase fruit diameter by stimulating cell expansion (*Nicolas et al., 2013; Li et al., 2021*). *VvANT1*, an ortholog of Arabidopsis ANT, could regulate cell proliferation and organ expansion (*Chialva et al., 2016*). In addition, *Martins et al., (2018)* have demonstrated that calcium-related genes have a significant regulatory effect on the cell wall biogenesis by influencing the expression of pectin methyl esterase (PME) and polygalacturonase (PG).

In this study, ‘Jinzao Wuhe’ is a very rare bud mutation discovered and developed by our laboratory due to its remarkably enlarged berry size. Herein, we carried out WGRS and genetic variation analysis for ‘Jinzao Wuhe’ and its maternal cultivar to explore the candidate genes in response to berry development, as well as the potential mechanisms underlying bud sport in grapevines. Furthermore, transcriptome sequencing and WGCNA analyses were performed at four developmental time-points, leading to the identification of causal genes and regulators. These results would shed light on the molecular mechanisms of bud sport mutation and offer useful information for future grapevine improvement.

MATERIALS AND METHODS

Plant materials

‘Himord Seedless’ is an early maturing cultivar with seedless and small berries, ripening before the onset of the rainy season in Tianjin, China. ‘Jinzao Wuhe’ is a bud sport derived from ‘Himord Seedless’ discovered at Tianjin Academy of Agricultural Sciences (Tianjin, China) orchard in 2014. The 3-year-old vines ‘Himord Seedless’ and ‘Thompson Seedless’ and their bud sports ‘Jinzao Wuhe’ and ‘Dawuhebai’ were planted in the grape garden of Tianjin Academy of Agricultural Sciences, Tianjin, China, grown on Y-shaped trellises with 1.5 m (column) × 3 m (row) interval. Management guidelines include techniques for irrigation, fertilization, pruning, weeding, disease prevention, pest control and precautions during rainy season. Compared with its maternal plant ‘Himord Seedless’, ‘Jinzao Wuhe’

exhibits larger leaves, inflorescences, berries, thicker stems, roots, and greater vigorous growth compared to its maternal plant. Especially, the berry size of ‘Jinzao Wuhe’ is almost twice as large as its maternal plant, which is an excellent germplasm resource for grapevine breeding. In addition, ‘Dawuhebai’, another bud sport derived from table grape ‘Thompson Seedless’, also possesses enlarged berries that would be helpful for investigating the regulatory mechanisms underlying berry enlargement (Baiano & Terracone, 2012).

The longitudinal diameter, transverse diameter, volume and weight of each ten berries for each replicate were measured individually for every 7 days started from the 10 days post-anthesis (DPA). Three biological replicates were set for each treatment. For each grapevine cultivar, 2-week-old leaves without pathogens and insect infestations were harvested for genomics DNA isolation, library construction and WGRS. Inflorescences and berries of ‘Himord Seedless’ and ‘Jinzao Wuhe’ that were healthy and vigorous were sampled at the full-flowering stage (stage 1: 1 DPA), berry expansion stage (stage 2: 14 DPA), véraison stage (stage 3: 35 DPA) and maturity stage (stage 4: 56 DPA) for RNA extraction and transcriptome sequencing, respectively. Three biological replicates were prepared for each sampling time point. These leaf, inflorescence and berry samples were immediately snap-frozen by immersion in liquid nitrogen and stored at -80°C until use.

DNA and RNA extraction and sequencing

Genomic DNA was extracted from grapevine leaves by using the modified protocol of cetyl trimethyl ammonium bromide (CTAB). The concentration and purity of the extracted genomic DNA was measured by NanoDrop-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and the quality and integrity of isolated genomic DNA was assessed by 1% agarose gel electrophoresis. Paired-end (PE) Deep sequencing libraries were prepared according to the standard protocol (Illumina) with the Paired-end (PE) strategy. Finally, the qualified libraries were sequenced using the Illumina HiSeq 2500 platform in Novogene Biotechnology Company (Tianjin, China). Raw data was deposited in the NCBI SRA database under the BioProject accession [PRJNA795039](#) and [PRJNA872374](#) (Table S1).

Total RNA was isolated from inflorescences and berries of ‘Himord Seedless’ and ‘Jinzao Wuhe’ at 1, 14, 35, and 56 DPA using TRIzol reagent according to the manufacturer’s instructions. The quality of RNA was evaluated by NanoDrop-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA), and RNA integrity was determined by 1% denaturing gel electrophoresis. Three replicates at different stages were set for each sample. RNA-seq libraries were then prepared with the Illumina TruSeq RNA Sample Prep Kit and sequenced by using the Illumina HiSeq X Ten System in Novogene Biotechnology Company (Tianjin, China). All raw RNA-seq data were deposited in the NCBI SRA database under BioProject accession [PRJNA795246](#) (Table S1).

SNP, InDel and SV calling and annotation

Illumina raw reads were filtered and trimmed for low quality bases and adapter sequences using Fastp (default parameters), and then mapped onto the reference genome of *V. vinifera* Pinot Noir PN40024 (12X, GCF_000003745.3) using BWA (version 0.7.17) (Li

& Durbin, 2009) with default parameters. Marking duplicated sequences caused by PCR reactions using the Genome Analysis Toolkit (GATK4) (McKenna et al., 2010) with the MarkDuplicates module (default parameters). Both SNPs and InDels were identified for each sample using the GATK4 pipeline with the HaplotypeCaller module according to the recommended variation detection workflow (gatk -java-options HaplotypeCaller-R \$genome_file-ERC GVCF-I \$bam-O \$out). The criteria for variant filtration were as follows: (i) biallelic SNPs and InDels were retained; (ii) sequencing depth of higher than 5 was selected; (iii) variant allele frequencies of less than 40% were excluded; (iv) missing genotypes were deleted; (v) InDels shorter than or equal to 50 bp were retained.

Structural variations (SVs) were detected in all samples using LUMPY (version 0.2.13) (Layer et al., 2014) and DELLY (version 0.8.7) (Rausch et al., 2012). Splitters and discordant reads were extracted by using Lumpyexpress (lumpyexpress-B \$sample.bam-S \$sample.splitters.bam-D \$sample.discordants.bam-o lumpyout_single_\$tag.vcf) for each individual sample. Bam files were sorted and indexed using Samtools (version 1.12) (Li et al., 2009) and SVs were genotyped using SVTyper software. For DELLY (version 0.8.7), SVs were presented with the recommended instructions in BCF format (delly call-t ALL-g \$genome \$sample.bam-o \$sample.bcf), which were converted to VCF format using BCFtools (bcftools view \$sample.bcf > \$sample.vcf) (Danecek & McCarthy, 2017). Moreover, SURVIVOR (SURVIVOR merge) was used to merge SVs generated from both LUMPY and DELLY pipelines. As for filtration, SVs with a length of 50 bp or more and the quality flag PASS, as well as the precise breakpoints, were selected as the final dataset for further analysis.

The intersection of SNP, InDel, and SV datasets of the two pairs of bud mutation and their maternal were selected separately. Then, the effect of all selected variants located in chromosomes was annotated using SnpEff (java-jar snpEff.jar eff-csvStats \$sample.csv-s \$sample.html-c snpEff.config-v grape \$sample.vcf.gz >\$sample_snpEff.gz) (Cingolani et al., 2012), which classified them into HIGH, LOW, MODERATE, and MODIFIER categories according to the predicted results for influencing gene function. Important and non-synonymous SNPs, InDels and SVs with high impact were selected as target variants that might be responsible for phenotype variation.

Transcriptome analysis

Tophat-Cufflinks pipeline was used to generate gene expression profiles and identify DEGs (Trapnell, Pachter & Salzberg, 2009). First of all, raw reads were preprocessed using Fastp (default parameters) (Chen et al., 2018) to remove adaptors and low quality bases. Tophat2 (default parameters) was carried out for genomic mapping. Gene expression levels were estimated as fragments per kilobase of transcript per million (FPKM) by Cufflink (default parameters) based on the alignments. DEGs between different developmental stages were identified using Cuffdiff with regarding a false discovery rate (FDR) ≤ 0.05 and \log_2 Fold-Change ≥ 1 as thresholds. Meanwhile, HiSAT2 (Kim, Langmead & Salzberg, 2015), feature counts (Liao, Smyth & Shi, 2014) and DESeq2 (Love, Huber & Anders, 2014, p. 2) workflows were also performed to achieve a more comprehensive result of DEGs analysis. The statistical power calculated using RNASeqPower (Bioconductor—RNASeqPower) was

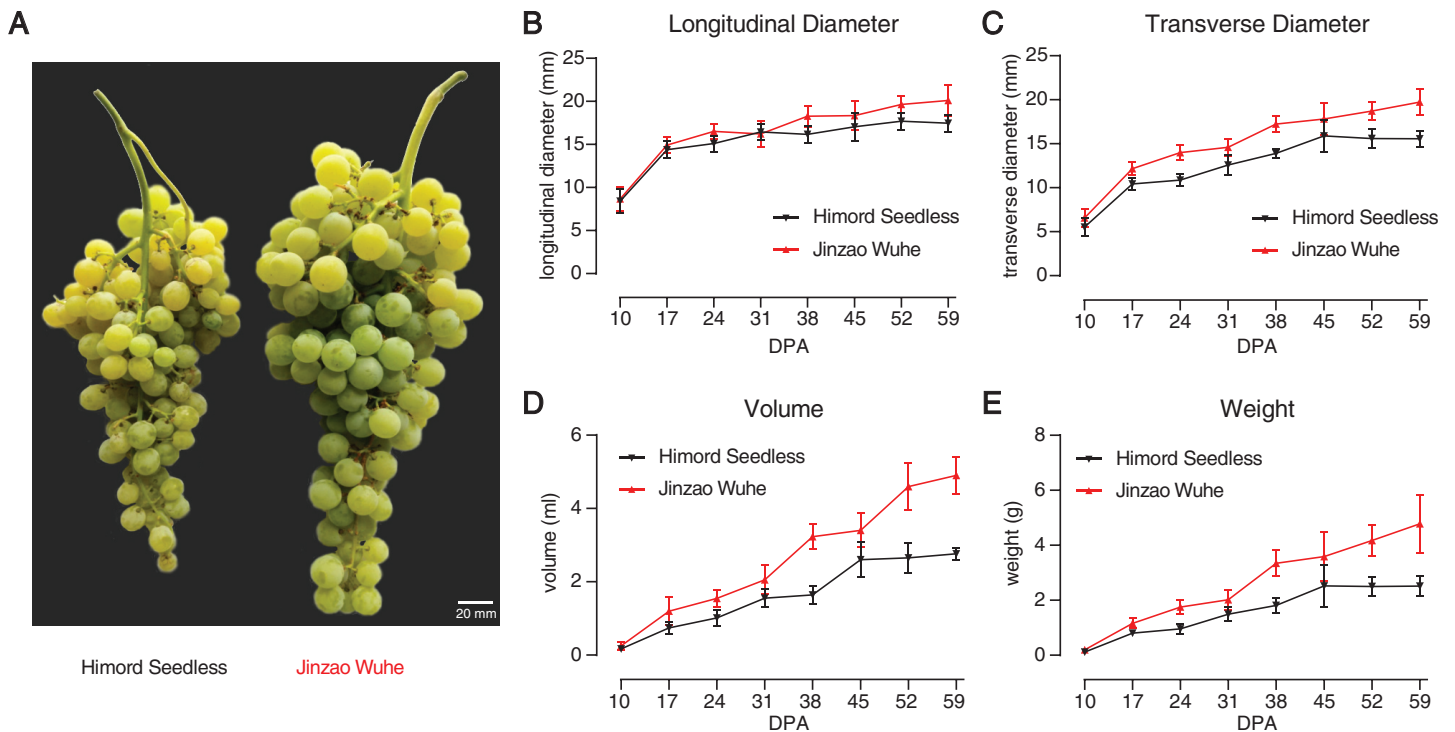


Figure 1 The comparison of ‘Himord Seedless’ and its bud sport ‘Jinzao Wuhe’. Comparison of longitudinal diameter (B), transverse diameter (C), volume (D) and weight (E) of ‘Himord Seedless’ (black line) and ‘Jinzao Wuhe’ (red line) at different developmental stages. Standard deviations (SDs) were calculated with 30 biological replicates for each timepoint. The error bars represented 1 SD.

Full-size DOI: [10.7717/peerj.14617/fig-1](https://doi.org/10.7717/peerj.14617/fig-1)

0.96. We visualized the correlation among the three biological replicates for each sample and gene expression level using the R package ‘scatterplot3d’ and ‘pheatmap’, respectively.

WGCNA network analysis

WGCNA was carried out in R program with the WGCNA package (Langfelder & Horvath, 2008). Following the WGCNA tutorial, the FPKM values of DEGs were chosen for further investigation. The soft-thresholding power (β) was determined using the “pickSoftThreshold” function. Hierarchical cluster tree displaying co-expression modules was constructed as follows: minimodule size = 50, deep split = 2; merge cut height = 0.35. To identify the relationship between co-expression modules and phenotypes, the module-trait relationships were subsequently determined by Pearson correlation.

RESULTS

Phenotypical comparison between ‘Himord Seedless’ and its bud sport mutant ‘Jinzao Wuhe’

Physiological traits were measured, where ‘Jinzao Wuhe’ showed a double sigmoid pattern curve from fruiting (10 DPA) to maturity (59 DPA) with higher berry longitudinal diameter, transverse diameter, volume and weight than ‘Himord Seedless’ (Figs. 1B–1E; Table S2). Two discriminable stages for berry expansion could be observed during berry

Table 1 The statistics of whole-genome resequencing data.

Sample	Raw reads	Clean reads	Q20	Q30	Mapping rates
Thompson seedless	54,003,664	53,434,316	95.52%	88.93%	93.57%
Dawuhebai	75,630,160	75,128,802	93.57%	90.50%	95.87%
Himord seedless	91,982,400	91,164,032	96.30%	90.52%	91.91%
Jinzaowu	83,784,034	83,013,230	93.70%	90.19%	95.39%

growth. At the end of maturity (59 DPA), ‘Jinzaowu’ was about twice as large as the ‘Himord Seedless’ in terms of volume and weight (Fig. 1A).

SNP, InDel and SV calling

To identify the genome-wide genetic variations between two maternal grapevines (WTs: ‘Thompson Seedless’ and ‘Himord Seedless’) and their bud mutants (MUTs: ‘Dawuhebai’ and ‘Jinzaowu’), WGRS was performed, producing a total of 305,400,258 raw reads (~45.8 Gb). After removing the low-quality reads and bases, a total of 302,740,380 clean reads (~45.5 Gb) were retrieved, accounting for more than 99.34% of the raw data. The Q20 and Q30 Phred Scores of all samples were greater than 93.57% and 88.93%, respectively (Table 1). The genomic mapping rates for clean reads ranged from 91.91% to 95.87% (Table 1).

For ‘Thompson Seedless’, ‘Dawuhebai’, ‘Himord Seedless’ and ‘Jinzaowu’, 3,763,122, 3,947,775, 5,512,478 and 5,499,289 SNPs were detected, respectively. As a result, MUTs are richer in SNPs than WTs and the number of SNPs of ‘Jinzaowu’ is higher than ‘Dawuhebai’ (Fig. 2A). We also discovered 2,658,536, 3,663,339, 2,531,663 and 3,670,669 transitions (Ts) and 1,289,239, 1,835,950, 1,231,459 and 1,841,809 transversions (Tv) in ‘Thompson Seedless’, ‘Dawuhebai’, ‘Himord Seedless’, and ‘Jinzaowu’, respectively, with genome-wide transitions to transversions ratio (Ts/Tv) were 2.06, 2.00, 2.06, and 1.99 (Table S3). SnpEff annotation revealed that the SNPs were mainly located in the upstream and downstream of genes, followed by intergenic regions, intronic regions, and exon regions, while they were scarce in splice site regions, UTR3’ and UTR5’ regions. In total, the heterozygous rates in ‘Himord Seedless’ and ‘Jinzaowu’ were apparently higher than those in ‘Thompson Seedless’ and ‘Dawuhebai’ (Table S3).

‘Thompson Seedless’, ‘Dawuhebai’, ‘Himord Seedless’ and ‘Jinzaowu’ have 698,381 (insertion: 382,040; deletion: 316,341), 752,785 (insertion: 341,594; deletion: 411,191), 1,113,041 (insertion: 516,558; deletion: 596,483) and 1,108,155 (insertion: 513,699; deletion: 594,456) InDels, respectively. It could be inferred that ‘Himord Seedless’ and ‘Jinzaowu’ have more distant genetic relationships with the reference genome (Table S4). Besides, the length distribution of InDels showed that mononucleotide insertions or deletions accounted for the majority of InDels (Fig. 2B). The frequency of InDels located in the upstream and downstream of genes was higher, averaging at 29.30% and 30.26%. Approximately 15.00%, 11.00%, and 1.00% of InDels harbored in intergenic, intronic, and exon regions, respectively. And they were almost undetectable for the splice

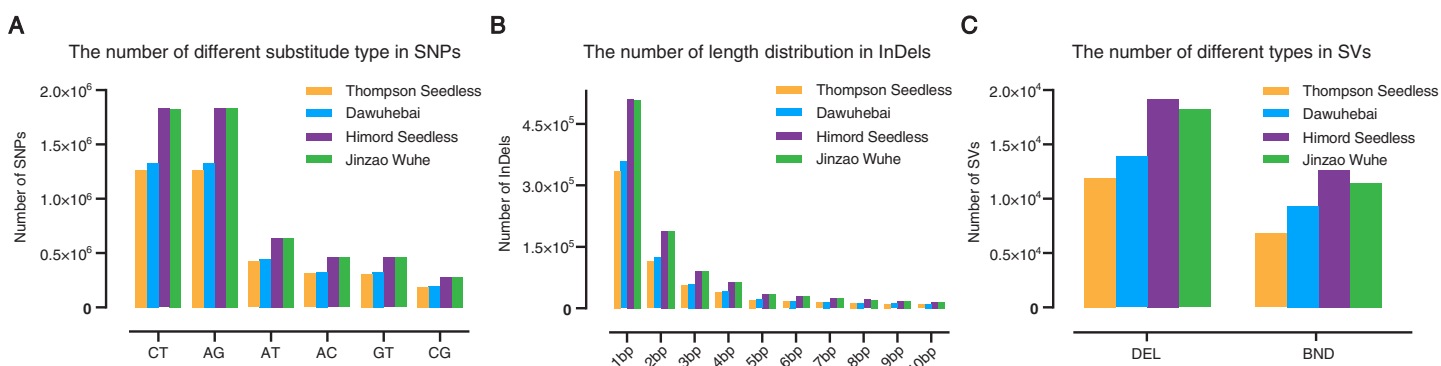


Figure 2 Annotation and distribution of (A) SNPs, (B) InDels and (C) SVs of ‘Thompson Seedless’ and its bud sports ‘Dawuhebai’, ‘Himord Seedless’ and its bud sports ‘Jinzao Wuhe’.

Full-size DOI: 10.7717/peerj.14617/fig-2

site, UTR3’ and UTR5’ regions (all less than 1.00%). In addition, homozygous and heterozygous InDels in four grapevine samples were also calculated, with the heterozygous rate of 79.32% and 79.19% in ‘Himord Seedless’ and ‘Jinzao Wuhe’, respectively, which were higher than 68.02% and 70.50% in ‘Thompson Seedless’ and ‘Dawuhebai’, respectively (Table S4).

SVs play an important functional role in genetic and morphological divergence, mainly including deletions (DEL), translocation (BND), inversions (INV), duplications (DUP) and insertions (INS). Due to the limitations of Illumina sequencing platform that usually producing short reads less than 200 bp, the results of DUP, INV, and INS were excluded for further investigation. Only DEL and BND were remained for subsequent analysis. In ‘Thompson Seedless’, ‘Dawuhebai’, ‘Himord Seedless’ and ‘Jinzao Wuhe’, 18,717, 23,237, 31,807 and 29,715 SVs were identified, respectively. The number of SVs in MUTs was markedly higher than in WTs. The distribution of SVs of four grape lines showed that the proportion of deletions was more than 53.73%, whereas other SVs accounted for less than 46.27%. (Fig. 2C). In summary, the distribution and proportions of SVs among these four grapevine lines were highly similar to those of SNPs and InDels (Table S5).

Identification of candidate genes harboring SNPs, InDels and SVs

To identify potential causal genes that are affected by DNA variations, comparative analysis between WTs and their MUTs were performed firstly. For ‘Thompson Seedless’ and ‘Dawuhebai’, 867,546, 184,194 and 21,817 different SNPs, InDels and SVs, as well as 3,535,491, 633,486 and 10,066 common SNPs, InDels and SVs were identified.

For ‘Himord Seedless’ and ‘Jinzao Wuhe’, 706,821, 216,022 and 30,521 different SNPs, InDels and SVs, as well as 5,152,473, 1,002,587 and 15,497 common SNPs, InDels and SVs were identified, respectively (Figs. 3A–3C). Then, we compared the genetic variants between WTs and MUTs individually. Between ‘Thompson Seedless’ and ‘Himord Seedless’, 2,682,997 SNPs, 456,929 InDels and 8,250 SVs were regarded as common genetic variants. Between ‘Dawuhebai’ and ‘Jinzao Wuhe’, 2,775,884 common SNPs and 481,550 common InDels and 9,000 common SVs were identified (Figs. 3A–3C).

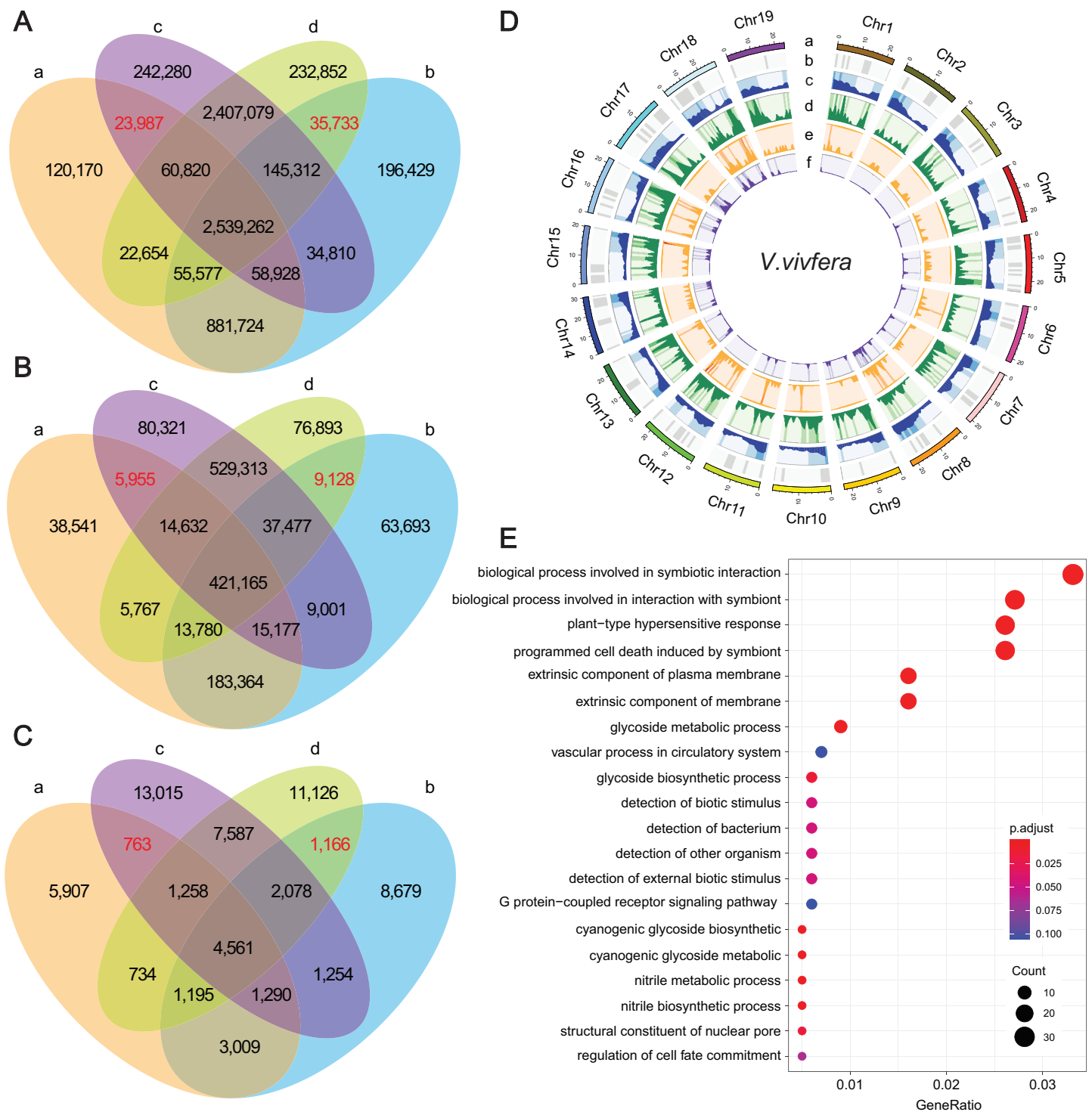


Figure 3 The statistics of SNPs, InDels and SVs and corresponding genes. (A–C) Venn diagrams represent the numbers of unique and shared SNPs, InDels and SVs, respectively. (a) ‘Thompson Seedless’; (b) ‘Dawuhebai’; (c) ‘Himord Seedless’; (d) ‘Jinzaowu’ (Wu). (D) The circos plots of SNPs, InDels and SVs distribution on the nineteen *V. vivifera* chromosomes (a) Chromosome; (b) GC content; (c) Gene density; (d) SNP density; (e) InDel density; (f) SV density. (E) The GO annotation of potential causal genes harbored with determining SNPs, InDels and SVs.

Full-size [DOI: 10.7717/peerj.14617/fig-3](https://doi.org/10.7717/peerj.14617/fig-3)

These common variants derived from two independent bud sport mutation events could offer crucial clues for explaining the phenotypic variation in MUT berries that growing larger. Finally, 35,733 SNPs, 9,128 InDels and 1,166 SVs that were absent in WTs but co-existed in MUTs, as well as 23,987 SNPs, 5,955 InDels and 763 SVs that were absent in MUTs but co-existed in WTs were deemed as target variants that caused by bud sport mutation (Figs. 3A–3C). According to the results of snpEff analysis, 61 SNPs, 272 InDels and 74 SVs were predicted to have HIGH impact, whereas 1,273 SNP entries (missense variants) to have MODERATE impact (Table S6). These variants were distributed unevenly across 19 chromosomes, with SNPs and InDels predominating on Chr18 and Chr12, and SVs predominating on Chr 18 and Chr5 (Fig. 3D).

As a result, a total of 1,022 unique genes were finally identified based on these determining genetic variations, in which 891, 171 and 68 genes were associated with SNPs, InDels and SVs, respectively (Table S6). GO enrichment analysis of these genes showed 20 representative GO terms, including “biological process involved in symbiotic interaction”, “biological process involved in interaction with symbiont”, “plant-type hypersensitive response”, “programmed cell wall”, “glycoside metabolic process”, “detection of biotic stimulus”, “detection of external biotic stimulus”, “structural constituent of nuclear pore”, “regulation of cell fate commitment” *etc.* (Fig. 3E).

Identification of the DEGs between ‘Himord Seedless’ and ‘Jinzaowu’

In order to obtain evidence at the transcriptomic level, RNA-seq was carried out and DEGs were investigated at four stages during berry development for ‘Himord Seedless’ and its mutant ‘Jinzaowu’. There was a good correlation among the three biological replicates for each sample (Fig. 4A). According to the analysis results of Cuffdiff, 928 (309 up-regulated and 619 down-regulated), 841 (358 up-regulated and 483 down-regulated), 1,627 (846 up-regulated and 781 down-regulated) and 727 (548 up-regulated and 179 down-regulated) DEGs were identified at stages 1–4, respectively (Figs. 4B, 4C, 4E and 4F; Table S7). After the redundancy removal, a total of 3,321 DEGs were identified that differentially expressed at least one time point (Fig. 4D). To achieve a more reliable result, another pipeline for RNA-seq and DEG analysis was performed using the same datasets. The results of DESeq2 generated 995 (346 up-regulated and 649 down-regulated), 1,080 (425 up-regulated and 655 down-regulated), 2,055 (1,125 up-regulated and 930 down-regulated) and 999 (733 up-regulated and 266 down-regulated) DEGs at stages 1–4, respectively (Table S8). As a result, 2,984 DEGs were finally confirmed by coupling two separate bioinformatic pipelines.

The GO enrichment analysis of these DEGs identified 20 representative GO terms at four berry development stages. The terms were related to cell wall modification, such as “plant-type cell wall modification”, “cell wall modification”, “callose deposition in cell wall”, “cell wall thickening”, “pectin esterase activity” in all stages, which could provide a reasonable explanation between cell wall biogenesis and berry enlargement. Several GO terms were associated with stress-response, such as “response to heat”, “response to toxic substance”, “response to antibiotic”, “response to reactive oxygen”, “response to auxin”,

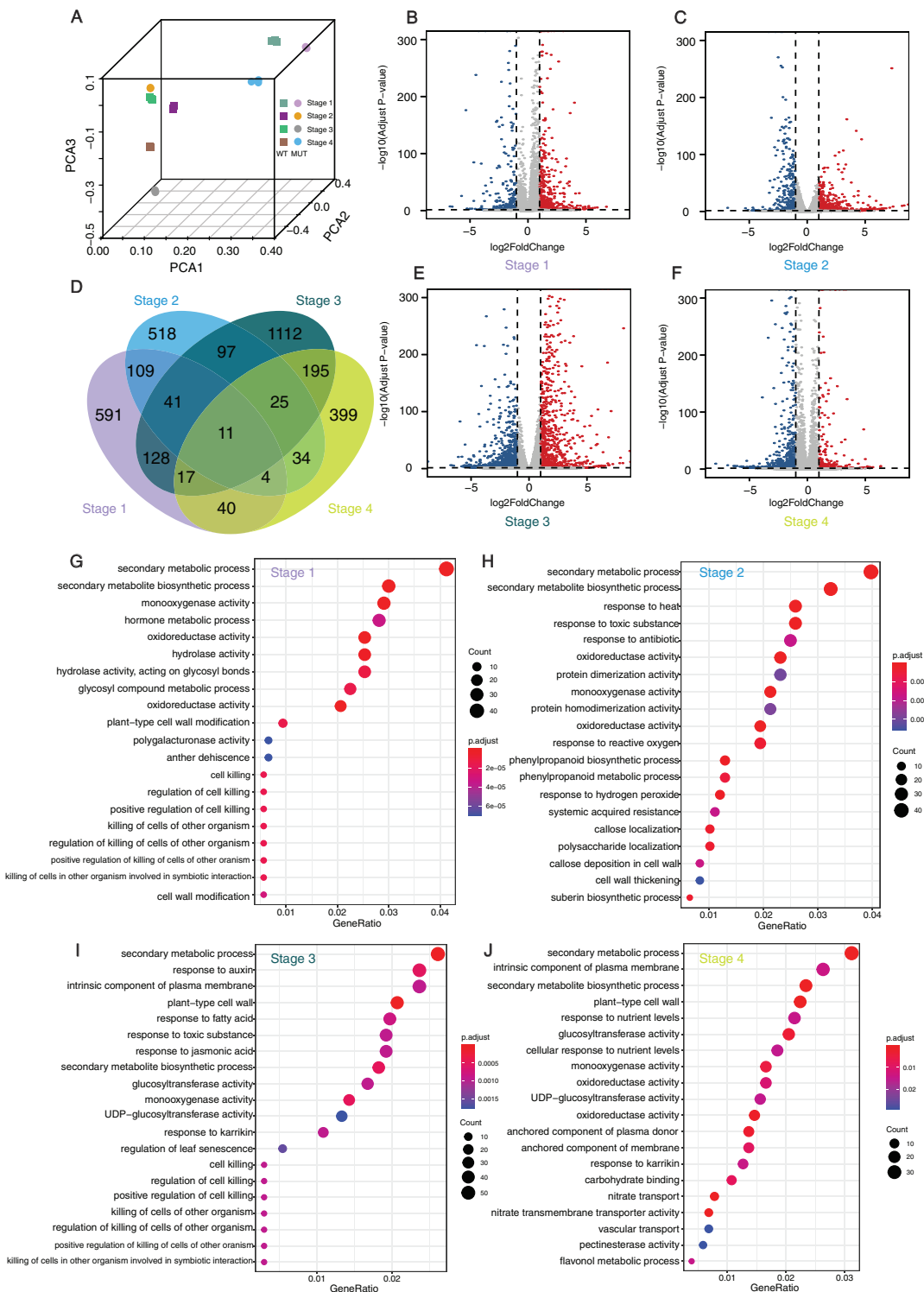


Figure 4 Analyses and statistics of DEGs and GO enrichment. Analyses and statistics of DEGs and GO enrichment. (A) Principal component analysis (PCA) of transcriptome data. (B, C, E, and F) Volcano plots for expressed genes at different developmental stages. Red dots represented the upregulated genes while blue dots represented the down-regulated genes. (D) Venn diagram of DEGs at four different stages. (G–J) GO enrichment results of DEGs at four development stages. WT: ‘Himord Seedless’; MUT: ‘Jinzaowu’. Stage 1–4: full-flowering stage, fruit enlargement stage (14 days post-anthesis), véraison stage (35 days post-anthesis) and maturity stage (56 days post-anthesis). The ‘Himord Seedless’ in the stage 1–4 were the control group and ‘Jinzaowu’ in the stage 1–4 were the experiment group.

Full-size [DOI: 10.7717/peerj.14617/fig-4](https://doi.org/10.7717/peerj.14617/fig-4)

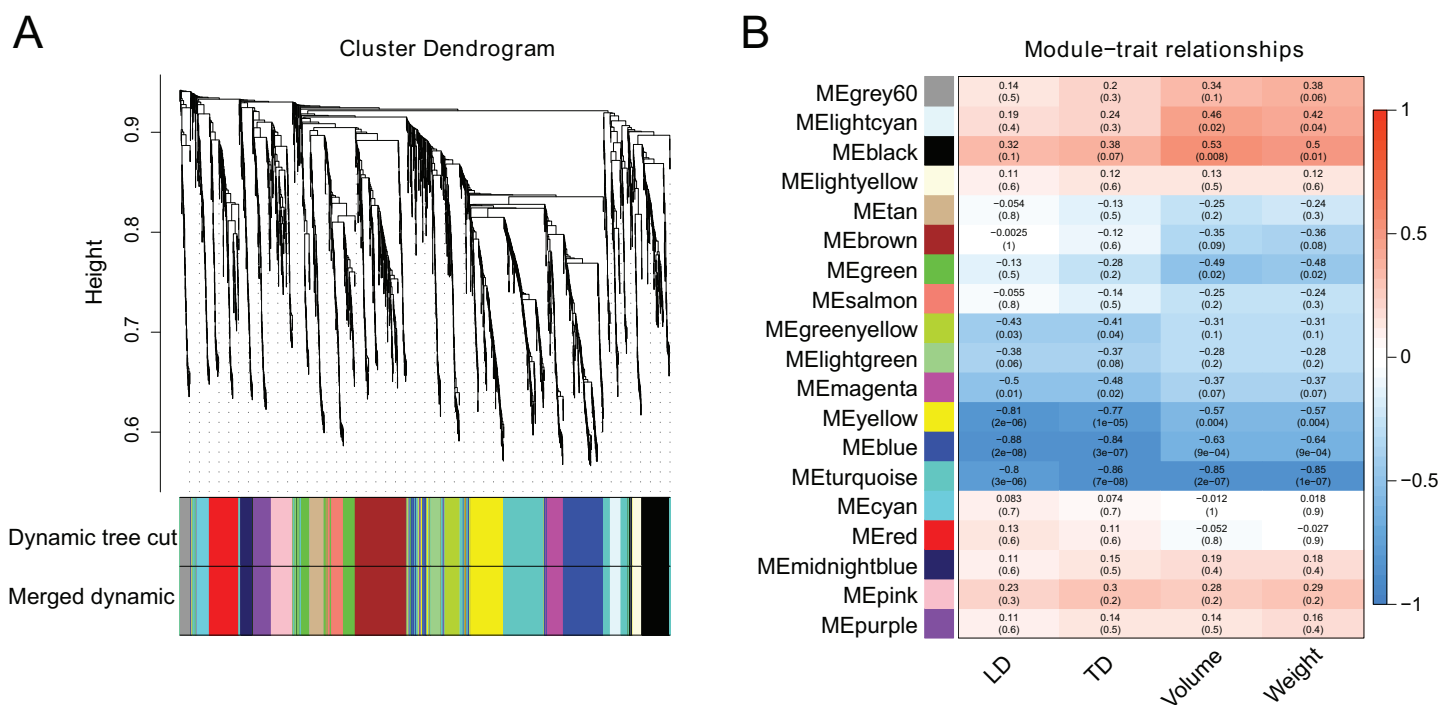


Figure 5 Weighted gene co-expression network analysis of berry development transcriptomes. (A) Dendrogram plot of co-expression modules. This plot shows 18 modules and different colors represent different modules. (B) Module-trait relationship and corresponding p values. LD: longitudinal diameter; TD: transverse diameter. [Full-size !\[\]\(5f471a71b78d7676bc356df190b88ab4_img.jpg\) DOI: 10.7717/peerj.14617/fig-5](https://doi.org/10.7717/peerj.14617/fig-5)

“response to fatty acid”, “response to toxic substance”, “response to jasmonic acid” and “response to karrikin”. For DEGs in stage 1 and 3, significantly enriched GO terms were “cell killing”, “regulation of cell killing”, “positive regulation of cell killing”, “killing of cells of other organism” and “regulation of killing of cells of other organism”. At the stage 2, the GO terms “phenylpropanoid biosynthetic process”, “phenylpropanoid metabolic process” and “suberin biosynthetic process” were significantly enriched. These detected pathways and related genes offered novel clues and contributed to our better understanding of the mechanisms taking place behind the phenomenon of bud sport mutation which leading to berry enlargement in ‘Jinzaowu’. (Figs. 4G–4J).

Identification of genes related to berry enlargement using WGCNA analysis

WGCNA analysis was performed utilizing the FPKM values of 2,984 DEGs to investigate co-expression patterns and determine the hub genes from DEGs. When the soft threshold β was set to 2, the gene network initially corresponded to a scale-free network (Fig. S1). Based on their co-expression patterns, all genes were further grouped into 18 modules (Fig. 5A). Berry phenotype data was used for correlation analysis including transverse diameter, longitudinal diameter, volume and weight. As a result, the ‘blue’ and ‘brown’ modules ($p < 0.01$), corresponding to 816 and 333 genes, were significantly clustered and linked to berry enlargement (Fig. 5B; Table S9).

Integrated results of DEGs and WGCNA for bud sport grapevine ‘Jinzaowu’

Based on the previous results from WGRS, RNA-seq and WGCNA, key genes were gradually determined by means of DEGs, specified DNA variations and co-expression pattern with berry phenotypic data. Finally, 42 key genes carrying high or moderate SNPs, InDels and SVs were discovered, which closely related to cell wall modifications (*LOC100258861*, *LOC100258876*), cytochromes P450 (*LOC100244431*, *LOC100253660*, *LOC100266067*, *LOC100254242*), UGT (*LOC100242998*, *LOC100260618*), zinc finger protein (*LOC100260826*), RING-type E3 ubiquitin transferase (*LOC100250551*), ARFs (*LOC100265555*), NAC domain-containing protein (*LOC100251801*), ABC transporter C family member (*LOC100259590*), LRR receptor-like serine threonine-protein kinase (*LOC100243949*, *LOC100247557*, *LOC109124293*), protein kinase (*LOC100853763*, *LOC100258962*, *LOC100267572*, *LOC100266874*, *LOC100258928*, *LOC104880223*, *LOC100244602*, *LOC100246192*), mitochondrial transcription (*LOC100242048*, *LOC100245527*, *LOC100853897*, *LOC100264034*) and other functions (Fig. 6, Table S10).

DISCUSSION

In grapevine, very few studies have been focused on the berry enlargement through the identifications of genetic variation, quantitative trait loci (QTL) and candidate genes (Doligez *et al.*, 2002; Mejía *et al.*, 2007). So far, there are very limited materials available for exclusively dissecting the regulatory mechanisms of berry development, especially in bud sport mutation. Doligez *et al.* (2002) detected QTLs related to berry weight using F1 progenies from a cross between two partially seedless genotypes. Muñoz-Espinoza *et al.* (2020) discovered a number of informative and transferable SNP and InDel markers, as well as 68 candidate genes linked with berry size using ‘Ruby’ and ‘Sultanina’ cultivars. Wang *et al.* (2022) underlined candidate genes for berry related traits using ‘Moldova’ (*V. labruscana* × *V. vinifera*) and ‘Ruiduxiangyu’ cultivars. These varieties are seedless with enlarged berries. In this study, the seedless cultivars ‘Jinzaowu’ and ‘Dawuhebai’ are ideal objects with enlarged berries due to bud sport mutation that derived from ‘Himord Seedless’ and ‘Thompson Seedless’ cultivars. Especially in ‘Jinzaowu’, the berries were enlarged nearly twice bigger than its maternal plant, making it an excellent breeding material for studying the regulatory mechanisms underlying berry enlargement.

WGRS is a powerful strategy for genomic variation detection, which could provide direct evidence for phenotypic alterations in plants mainly including SNPs and InDels (Horton *et al.*, 2012). Moreover, structural variations (SVs) could be also identified by using Illumina paired end reads and offer clues to the molecular basis of bud sport mutations. In apples, for instance, abundant SNPs and InDels were discovered in various bud mutants and other variants (Lee *et al.*, 2016; Wang *et al.*, 2020). In grapevines, WGRS of ‘Summer Black’ and its bud mutant ‘SBBM’ identified 635 genes carrying differential SNPs and InDels associated with the early-ripening trait (Xu *et al.*, 2016). ‘Kyoho’, ‘Zana’ and their bud mutants ‘Fengzao’ and ‘90-1’ were analyzed separately by Pei *et al.* (2021) to identify variations that could explain the altered phenotype between bud sport mutations and their maternal plants in terms of early maturing trait. Herein, WGRS of ‘Himord

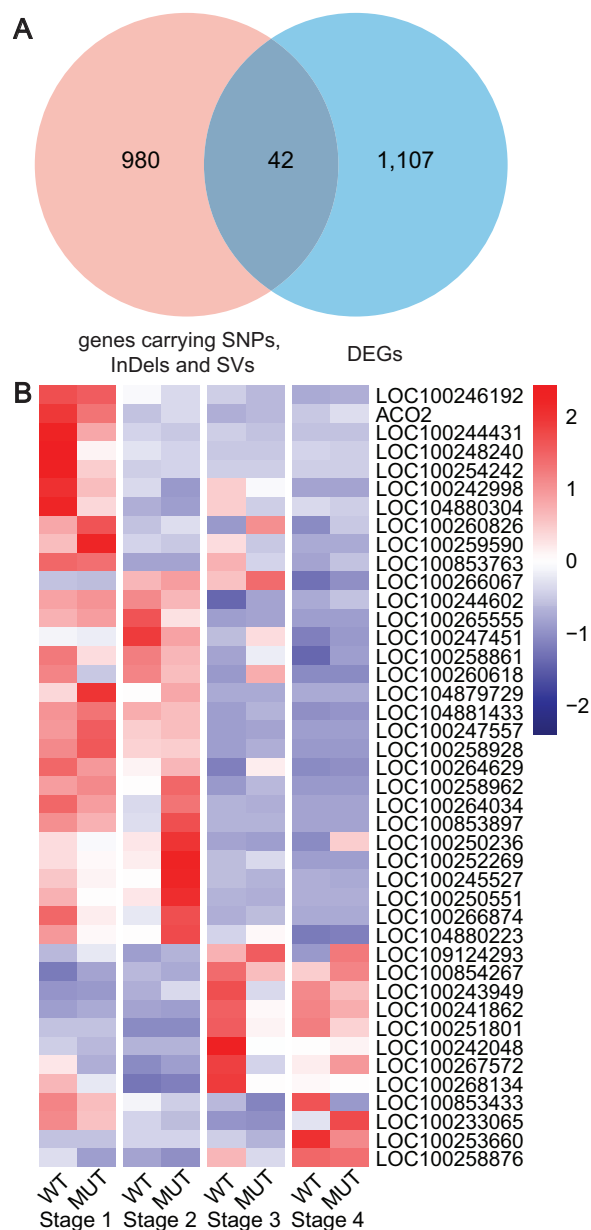


Figure 6 Analyses and statistics of DEGs carrying specified SNPs, InDels and SVs. (A) The intersection between DEGs and genes containing SNPs, InDels and SVs. (B) Expression profiles of DEGs harboring specified SNPs, InDels and SV. [Full-size !\[\]\(1679558f37f6db0dd8360a2a7e913e90_img.jpg\) DOI: 10.7717/peerj.14617/fig-6](https://doi.org/10.7717/peerj.14617/fig-6)

Seedless’ and ‘Thompson Seedless’ and their bud mutants ‘Jinzaowu’ and ‘Dawuhebai’ resulted in the identification of 1,022 genes carrying differential genetic variants. Combined with the GO enrichment results, in this case, symbiotic microorganisms and plant-type hypersensitive response might be the reason or consequence for berry enlargement caused by bud mutation in grapevines (Marrano *et al.*, 2018).

Transcriptome sequencing and WGCNA are also practical methods for characterizing and quantifying the gene expression and exploring the potential genes and mechanisms of agronomic traits. RNA-Seq of bud mutants ‘Tiangong Moyu’, ‘Fengzao’ and their maternal

‘Summer Black’, ‘Kyoho’, respectively, at different stages of berry development led to uncovering potential causal genes regulating ripening period (Wei et al., 2020). With the aid of examining gene expression profiles, Adetunji et al. (2019) discovered regulatory genes that either delay or promote berry development at the véraison stage of ‘Pinot Noir’ and ‘Pinot Noir Precoce’ grapes. Muñoz-Espinoza et al. (2020) combined genetic variation detection and transcriptomic approaches to find functional genes involved in berry size and prematurity and develop molecular markers based on SNPs, InDels and SVs, demonstrating the practicability and applicability of RNA-seq. Herein, we discovered 1,149 DEGs associated with berry size based on WGCNA and GO functional categories between ‘Jinzaowu’ and ‘Himord Seedless’. These results indicated that cell wall modification, secondary metabolic processes, stress responses and cell death regulation might play significant roles during berry enlargement (Sweetman et al., 2012).

The combination of high-throughput sequencing and QTL mapping is a practical strategy for screening candidate genes for important agronomic traits in plants. In a very recent study, 772 candidate genes associated with berry size in table grapes have been identified by QTL approach (Wang et al., 2022). We compared the results with 1,022 genes harboring SNPs, InDels and SVs and 2,984 DEGs identified in this study by WGRS and RNA-seq approaches. As a result, 81 genes were found in the reported QTL regions including ‘RING-type E3 ubiquitin transferase’ (LOC100266002, LOC100250551), ‘Probable xyloglucan endotransglucosylase/hydrolase protein 23’ (LOC100266747, LOC100251303, LOC100246166, LOC100263324, LOC100241056), ‘Expansin’ (LOC100261426, LOC100253046), ‘Auxin-induced protein PCNT115’ (LOC100252909), ‘Auxin efflux carrier component’ (LOC100256460), ‘beta-1,3-galactosyltransferase’ (LOC100246152), ‘Ethylene-responsive transcription factor’ (LOC100259767), ‘LRR receptor-like serine/threonine-protein kinase’ (LOC100241613, LOC100244275, LOC100244999, LOC100262226), ‘Heat shock protein’ (LOC100264870, LOC100252916, LOC104879973), ‘Cytochrome P450’ (LOC100251316), ‘MATE efflux family protein’ (LOC100242781, LOC100263305) and several genes involved in ‘Cell wall modifications’ (LOC100267301, LOC100252982, LOC100242036, LOC100252799) (Table S11). These genes offered imported clues for understanding the mechanisms of berry development which had been reported to be functionally related to berry size in grapes or other plants (Xia et al., 2013; Doligez et al., 2013; de Jong et al., 2015; Muñoz-Espinoza et al., 2016, 2020).

Pectin and cellulose are key substances involved in the mechanical strength of the primary cell wall and are important to the physical properties of plant cell. The degradation of pectin during ripening seems to be responsible for tissue softening, as reported in several fruit plants including kiwifruit (*Actinidia deliciosa*), lemon (*Citrus limon*), apple (*Malus pumila* Mill) and grapevines (*Vitis vinifera* L.). Moreover, previous studies demonstrated that the cellulose-hemicellulose networks and related genes could regulate cell wall modifications and cellulase degradation, leading to decreased rigidity in grape berry during ripening (Fasoli et al., 2016). In this study, PE and glycosyl hydrolase may catalyze the hydrolysis of pectin, resulting in soft berries during the expansion stage of berry development. In rice, UGT has been demonstrated to regulate grain size by modulating cell proliferation and expansion (Dong et al., 2020). The expression of UGT

had been shown to gradually decrease with fruit ripening, regulating the early stages of berry development. In our results, the expression levels of two ‘UDP-glycosyltransferase-related genes’ (*LOC100266067* and *LOC100254242*) significantly varied between MUT and WT. This is consistent with the reported expression trend, suggesting their potential roles in regulating berry size through cell proliferation or differentiation. In addition, ‘arabinogalactan protein (AGP)’ is highly complex cell wall glycoprotein that plays a role in cell expansion and division (*Leszczuk et al., 2020*). As a result, AGP (*LOC100252982*) was detected to have higher expression in MUT than in WT, with specific accumulation in the second stage of berry development, implying its regulatory role during berry expansion. However, these genes need to be further investigated and functionally validated.

The CYP superfamily genes were also reported to be positively associated with fruit size (*Qi et al., 2017*). Four CYP genes (*LOC100244431*, *LOC100254242*, *LOC100258062*, *LOC100266067*) were identified in this study, of which *LOC100258062* and *LOC100266067* were upregulated at the berry expansion stage in MUT, implying their functional roles in enlarging berry size in MUT. Similar phenomena were also discovered in Arabidopsis that overexpression of *AtCYP78A9* resulted in enlarged fruit organs (*Ito & Meyerowitz, 2000*). In addition, auxin was regulated by NAC, which thereby triggered the expression of ARFs (*Tello et al., 2015; Li et al., 2022*). Ethylene coupled with increased ACO transcription could produce a modest peak during berry ripening and promote larger berry diameter. Simultaneously, zinc finger protein (*Zhao et al., 2021*), ABC transporter C family member, RING-type E3 ubiquitin transferase (*Xia et al., 2013*), multi-antimicrobial extrusion (MATE) (*Suzuki et al., 2015*), mitochondrial transcription, leucine-rich repeat receptor-like protein kinase (*Wang et al., 2019*), F-box kelch-repeat protein (*Li et al., 2011*) and protein kinase (*Ye et al., 2017*) resulted in the enlargement of berries. This result is in agreement with genes related to fruit enlargement between pears, peaches, and strawberries (*Pei et al., 2020*). However, limited results supported simultaneously by evidence at genomic and transcriptomic levels suggested that bud sport mutation might be more complex than previously assumed.

CONCLUSIONS

In conclusion, combined with WGRS, RNA-seq approaches and integrated bioinformatic analysis, key genes including pectin esterase, cellulase A, CYP, UGT, zinc finger protein, ARF, NAC TF, protein kinase were identified in ‘Himord Seedless’ and its bud sport mutant ‘Jinzaowu’. These results could partially explain its phenotypic changes, especially for berry enlargement in bud sport ‘Jinzaowu’. However, DEGs during berry development and DNA variations between bud sport and its maternal cultivar could not agree well with each other, suggesting the complexity of mechanisms underlying bud sport mutation. Our findings offer new clues for a better understanding of bud sport in grapevines and provide useful information for table grape improvement through molecular breeding.

ACKNOWLEDGEMENTS

We would like to thank Dr. Ming Li at Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences, for helping to provide materials of ‘Thompson Seedless’ and ‘Dawuhebai’.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This research was funded by the China Agriculture Research System of MOF and MARA, grant number CARS-29, and the Innovative Research and experimental Projects for young researchers of Tianjin Academy of Agricultural Science, grant number 20211002.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

China agriculture research system of MOF and MARA: CARS-29.

Tianjin Academy of Agricultural Science: 20211002.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Jianquan Huang performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Guan Zhang performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Yanhao Li analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Mingjie Lyu conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- He Zhang conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Na Zhang conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Rui Chen conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The whole-genome resequencing data set generated and analyzed in the current study are available at NCBI BioProject: [PRJNA795039](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA795039) and [PRJNA872374](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA872374).

All raw RNA-seq data set generated and analyzed in the current study is available at NCBI BioProjects: [PRJNA795246](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA795246), [PRJNA872374](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA872374).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.14617#supplemental-information>.

REFERENCES

- Adetunji MO, Lamont SJ, Abasht B, Schmidt CJ. 2019.** Variant analysis pipeline for accurate detection of genomic variants from transcriptome sequencing data. *PLOS ONE* **14**(9):e0216838 DOI [10.1371/journal.pone.0216838](https://doi.org/10.1371/journal.pone.0216838).
- Azuma A, Kobayashi S, Goto-Yamamoto N, Shiraishi M, Mitani N, Yakushiji H, Koshita Y. 2009.** Color recovery in berries of grape (*Vitis vinifera* L.) “Benitaka”, a bud sport of Italia, is caused by a novel allele at the VvmybA1 locus. *Plant Science: An International Journal of Experimental Plant Biology* **176**(4):470–478 DOI [10.1016/j.plantsci.2008.12.015](https://doi.org/10.1016/j.plantsci.2008.12.015).
- Baiano A, Terracone C. 2012.** Effects of bud load on quality of Beogradska besemena and Thompson seedless table grapes and cultivar differentiation based on chemometrics of analytical indices. *Journal of the Science of Food and Agriculture* **92**(3):645–653 DOI [10.1002/jsfa.4625](https://doi.org/10.1002/jsfa.4625).
- Breton C, Fournel-Gigleux S, Palcic MM. 2012.** Recent structures, evolution and mechanisms of glycosyltransferases. *Current Opinion in Structural Biology* **22**(5):540–549 DOI [10.1016/j.sbi.2012.06.007](https://doi.org/10.1016/j.sbi.2012.06.007).
- Chakrabarti M, Zhang N, Sauvage C, Muños S, Blanca J, Cañizares J, Diez MJ, Schneider R, Mazourek M, McClead J, Causse M, van der Knaap E. 2013.** A cytochrome P450 regulates a domestication trait in cultivated tomato. *Proceedings of the National Academy of Sciences of the United States of America* **110**(42):17125–17130 DOI [10.1073/pnas.1307313110](https://doi.org/10.1073/pnas.1307313110).
- Chen S, Zhou Y, Chen Y, Gu J. 2018.** fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics (Oxford, England)* **34**(17):i884–i890 DOI [10.1093/bioinformatics/bty560](https://doi.org/10.1093/bioinformatics/bty560).
- Chervin C, El-Kereamy A, Roustan J-P, Latché A, Lamon J, Bouzayen M. 2004.** Ethylene seems required for the berry development and ripening in grape, a non-climacteric fruit. *Plant Science* **167**(6):1301–1305 DOI [10.1016/j.plantsci.2004.06.026](https://doi.org/10.1016/j.plantsci.2004.06.026).
- Chialva C, Eichler E, Grissi C, Muñoz C, Gomez-Talquenca S, Martínez-Zapater JM, Lijavetzky D. 2016.** Expression of grapevine AINTEGUMENTA-like genes is associated with variation in ovary and berry size. *Plant Molecular Biology* **91**(1–2):67–80 DOI [10.1007/s11103-016-0443-1](https://doi.org/10.1007/s11103-016-0443-1).
- Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012.** A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **6**(2):80–92 DOI [10.4161/fly.19695](https://doi.org/10.4161/fly.19695).
- Danecek P, McCarthy SA. 2017.** BCFTools/csqr:haplotype-aware variant consequences. *Bioinformatics (Oxford, England)* **33**(13):2037–2039 DOI [10.1093/bioinformatics/btx100](https://doi.org/10.1093/bioinformatics/btx100).
- de Jong M, Wolters-Arts M, Schimmel BCJ, Stultiens CLM, de Groot PFM, Powers SJ, Tikunov YM, Bovy AG, Mariani C, Vriezen WH, Rieu I. 2015.** *Solanum lycopersicum* AUXIN RESPONSE FACTOR 9 regulates cell division activity during early tomato fruit development. *Journal of Experimental Botany* **66**(11):3405–3416 DOI [10.1093/jxb/erv152](https://doi.org/10.1093/jxb/erv152).
- De La Torre-Ruiz N, Ruiz-Valdiviezo VM, Rincón-Molina CI, Rodríguez-Mendiola M, Arias-Castro C, Gutiérrez-Miceli FA, Palomeque-Dominguez H, Rincón-Rosales R. 2016.**

- Effect of plant growth-promoting bacteria on the growth and fructan production of *Agave americana* L. *Brazilian Journal of Microbiology: Publication of the Brazilian Society for Microbiology* 47(3):587–596 DOI 10.1016/j.bjm.2016.04.010.
- Doligez A, Bertrand Y, Farnos M, Grolier M, Romieu C, Esnault F, Dias S, Berger G, François P, Pons T, Ortigosa P, Roux C, Houel C, Laucou V, Bacilieri R, Péros J-P, This P. 2013.** New stable QTLs for berry weight do not colocalize with QTLs for seed traits in cultivated grapevine (*Vitis vinifera* L.). *BMC Plant Biology* 13(1):217 DOI 10.1186/1471-2229-13-217.
- Doligez A, Bouquet A, Danglot Y, Lahogue F, Riaz S, Meredith P, Edwards J, This P. 2002.** Genetic mapping of grapevine (*Vitis vinifera* L.) applied to the detection of QTLs for seedlessness and berry weight. *TAG. Theoretical and Applied Genetics*. 105(5):780–795 DOI 10.1007/s00122-002-0951-z.
- Dong N-Q, Sun Y, Guo T, Shi C-L, Zhang Y-M, Kan Y, Xiang Y-H, Zhang H, Yang Y-B, Li Y-C, Zhao H-Y, Yu H-X, Lu Z-Q, Wang Y, Ye W-W, Shan J-X, Lin H-X. 2020.** UDP-glucosyltransferase regulates grain size and abiotic stress tolerance associated with metabolic flux redirection in rice. *Nature Communications* 11(1):2629 DOI 10.1038/s41467-020-16403-5.
- El-kereamy A, Chervin C, Roustan J-P, Cheynier V, Souquet J-M, Moutounet M, Raynal J, Ford C, Latché A, Pech J-C, Bouzayen M. 2003.** Exogenous ethylene stimulates the long-term expression of genes related to anthocyanin biosynthesis in grape berries. *Physiologia Plantarum* 119(2):175–182 DOI 10.1034/j.1399-3054.2003.00165.x.
- Fang W, Wang Z, Cui R, Li J, Li Y. 2012.** Maternal control of seed size by EOD3/CYP78A6 in *Arabidopsis thaliana*. *The Plant Journal: For Cell and Molecular Biology* 70(6):929–939 DOI 10.1111/j.1365-313X.2012.04907.x.
- Fasoli M, Dell’Anna R, Dal Santo S, Balestrini R, Sanson A, Pezzotti M, Monti F, Zenoni S. 2016.** Pectins, hemicelluloses and celluloses show specific dynamics in the internal and external surfaces of grape berry skin during ripening. *Plant and Cell Physiology* 57(6):1332–1349 DOI 10.1093/pcp/pcw080.
- Foster TM, Aranzana MJ. 2018.** Attention sports fans! The far-reaching contributions of bud sport mutants to horticulture and plant biology. *Horticulture Research* 5(1):44 DOI 10.1038/s41438-018-0062-x.
- Guo D-L, Zhang G-H. 2015.** A new early-ripening grape cultivar—‘Fengzao’. *Acta Horticulturae* 1082(1082):153–156 DOI 10.17660/ActaHortic.2015.1082.20.
- Horton MW, Hancock AM, Huang YS, Toomajian C, Atwell S, Auton A, Mulyati NW, Platt A, Sperone FG, Vilhjálmsson BJ, Nordborg M, Borevitz JO, Bergelson J. 2012.** Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. *Nature Genetics* 44(2):212–216 DOI 10.1038/ng.1042.
- Ito T, Meyerowitz EM. 2000.** Overexpression of a gene encoding a cytochrome P450, CYP78A9, induces large and seedless fruit in *Arabidopsis*. *The Plant Cell* 12(9):1541–1550 DOI 10.1105/tpc.12.9.1541.
- Kiedrowski S, Kawalleck P, Hahlbrock K, Somssich IE, Dangl JL. 1992.** Rapid activation of a novel plant defense gene is strictly dependent on the *Arabidopsis* RPM1 disease resistance locus. *The EMBO Journal* 11(13):4677–4684 DOI 10.1002/j.1460-2075.1992.tb05572.x.
- Kim D, Langmead B, Salzberg SL. 2015.** HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* 12(4):357–360 DOI 10.1038/nmeth.3317.
- Krizek BA. 1999.** Ectopic expression of AINTEGUMENTA in *Arabidopsis* plants results in increased growth of floral organs. *Developmental Genetics* 25:224–236 DOI 10.1002/(ISSN)1520-6408(1999)25:3<224::AID-DVG5>3.0.CO;2-Y.

- Langfelder P, Horvath S. 2008.** WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**(1):559 DOI [10.1186/1471-2105-9-559](https://doi.org/10.1186/1471-2105-9-559).
- Layer RM, Chiang C, Quinlan AR, Hall IM. 2014.** LUMPY: a probabilistic framework for structural variant discovery. *Genome Biology* **15**(6):R84 DOI [10.1186/gb-2014-15-6-r84](https://doi.org/10.1186/gb-2014-15-6-r84).
- Lee HS, Kim GH, Kwon SI, Kim JH, Kwon YS, Choi C. 2016.** Analysis of “Fuji” apple somatic variants from next-generation sequencing. *Genetics and Molecular Research: GMR* **15**(3):gmr8185 DOI [10.4238/gmr.15038185](https://doi.org/10.4238/gmr.15038185).
- Leng F, Ye Y, Zhu X, Zhang Y, Zhang Z, Shi J, Shen N, Jia H, Wang L. 2021.** Comparative transcriptomic analysis between ‘Summer Black’ and its bud sport ‘Nantaihutezao’ during developmental stages. *Planta* **253**(1):23 DOI [10.1007/s00425-020-03543-7](https://doi.org/10.1007/s00425-020-03543-7).
- Leszczuk A, Kalaitzis P, Blazakis KN, Zdunek A. 2020.** The role of arabinogalactan proteins (AGPs) in fruit ripening—a review. *Horticulture Research* **7**(1):1–12 DOI [10.1038/s41438-020-00397-8](https://doi.org/10.1038/s41438-020-00397-8).
- Li M, Chen R, Gu H, Cheng D, Guo X, Shi C, Li L, Xu G, Gu S, Wu Z, Chen J. 2021.** Grape small auxin upregulated RNA (SAUR) 041 is a candidate regulator of berry size in grape. *International Journal of Molecular Sciences* **22**(21):11818 DOI [10.3390/ijms222111818](https://doi.org/10.3390/ijms222111818).
- Li H, Durbin R. 2009.** Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* **25**(14):1754–1760 DOI [10.1093/bioinformatics/btp324](https://doi.org/10.1093/bioinformatics/btp324).
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009.** The sequence alignment/map format and SAMtools. *Bioinformatics (Oxford, England)* **25**(16):2078–2079 DOI [10.1093/bioinformatics/btp352](https://doi.org/10.1093/bioinformatics/btp352).
- Li M, Tang D, Wang K, Wu X, Lu L, Yu H, Gu M, Yan C, Cheng Z. 2011.** Mutations in the F-box gene LARGER PANICLE improve the panicle architecture and enhance the grain yield in rice. *Plant Biotechnology Journal* **9**(9):1002–1013 DOI [10.1111/j.1467-7652.2011.00610.x](https://doi.org/10.1111/j.1467-7652.2011.00610.x).
- Li X, Zhang X, Shi T, Chen M, Jia C, Wang J, Hou Z, Han J, Bian S. 2022.** Identification of ARF family in blueberry and its potential involvement of fruit development and pH stress response. *BMC Genomics* **23**(1):329 DOI [10.1186/s12864-022-08556-y](https://doi.org/10.1186/s12864-022-08556-y).
- Liao Y, Smyth GK, Shi W. 2014.** featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics (Oxford, England)* **30**(7):923–930 DOI [10.1093/bioinformatics/btt656](https://doi.org/10.1093/bioinformatics/btt656).
- Love MI, Huber W, Anders S. 2014.** Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**(12):550 DOI [10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8).
- Ma M, Wang Q, Li Z, Cheng H, Li Z, Liu X, Song W, Appels R, Zhao H. 2015.** Expression of TaCYP78A3, a gene encoding cytochrome P450 CYP78A3 protein in wheat (*Triticum aestivum* L.), affects seed size. *The Plant Journal: For Cell and Molecular Biology* **83**(2):312–325 DOI [10.1111/tpj.12896](https://doi.org/10.1111/tpj.12896).
- Malladi A, Hirst PM. 2010.** Increase in fruit size of a spontaneous mutant of “Gala” apple (*Malus x domestica* Borkh.) is facilitated by altered cell production and enhanced cell size. *Journal of Experimental Botany* **61**(11):3003–3013 DOI [10.1093/jxb/erq134](https://doi.org/10.1093/jxb/erq134).
- Marrano A, Micheletti D, Lorenzi S, Neale D, Grando MS. 2018.** Genomic signatures of different adaptations to environmental stimuli between wild and cultivated *Vitis vinifera* L. *Horticulture Research* **5**(1):34 DOI [10.1038/s41438-018-0041-2](https://doi.org/10.1038/s41438-018-0041-2).
- Martins V, Garcia A, Costa C, Sottomayor M, Gerós H. 2018.** Calcium- and hormone-driven regulation of secondary metabolism and cell wall enzymes in grape berry cells. *Journal of Plant Physiology* **231**(3):57–67 DOI [10.1016/j.jplph.2018.08.011](https://doi.org/10.1016/j.jplph.2018.08.011).

- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Research* 20(9):1297–1303 DOI 10.1101/gr.107524.110.
- Mejía N, Gebauer M, Muñoz L, Hewstone N, Muñoz C, Hinrichsen P. 2007. Identification of QTLs for seedlessness, berry size, and ripening date in a seedless x seedless table grape progeny. *American Journal of Enology and Viticulture* 58(4):499–507 DOI 10.5344/ajev.2007.58.4.499.
- Migicovsky Z, Yeats TH, Watts S, Song J, Forney CF, Burgher-MacLellan K, Somers DJ, Gong Y, Zhang Z, Vrebalov J, van Velzen R, Giovannoni JG, Rose JKC, Myles S. 2021. Apple ripening is controlled by a NAC transcription factor. *Frontiers in Genetics* 12:671300 DOI 10.3389/fgene.2021.671300.
- Mitchell HJ, Hall JL, Barber MS. 1994. Elicitor-induced cinnamyl alcohol dehydrogenase activity in lignifying wheat (*Triticum aestivum* L.) leaves. *Plant Physiology* 104(2):551–556 DOI 10.1104/pp.104.2.551.
- Muñoz-Espinoza C, Di Genova A, Correa J, Silva R, Maass A, González-Agüero M, Orellana A, Hinrichsen P. 2016. Transcriptome profiling of grapevine seedless segregants during berry development reveals candidate genes associated with berry weight. *BMC Plant Biology* 16(1):104 DOI 10.1186/s12870-016-0789-1.
- Muñoz-Espinoza C, Di Genova A, Sánchez A, Correa J, Espinoza A, Meneses C, Maass A, Orellana A, Hinrichsen P. 2020. Identification of SNPs and InDels associated with berry size in table grapes integrating genetic and transcriptomic approaches. *BMC Plant Biology* 20(1):365 DOI 10.1186/s12870-020-02564-4.
- Negri AS, Prinsi B, Failla O, Scienza A, Espen L. 2015. Proteomic and metabolic traits of grape exocarp to explain different anthocyanin concentrations of the cultivars. *Frontiers in Plant Science* 6:603 DOI 10.3389/fpls.2015.00603.
- Nicolas P, Lecourieux D, Gomès E, Delrot S, Lecourieux F. 2013. The grape berry-specific basic helix-loop-helix transcription factor VvCEB1 affects cell size. *Journal of Experimental Botany* 64(4):991–1003 DOI 10.1093/jxb/ers374.
- Pan Q-H, Chen F, Zhu B-Q, Ma L-Y, Li L, Li J-M. 2012. Molecular cloning and expression of gene encoding aromatic amino acid decarboxylase in “Vidal blanc” grape berries. *Molecular Biology Reports* 39(4):4319–4325 DOI 10.1007/s11033-011-1219-y.
- Pei M-S, Cao S-H, Wu L, Wang G-M, Xie Z-H, Gu C, Zhang S-L. 2020. Comparative transcriptome analyses of fruit development among pears, peaches, and strawberries provide new insights into single sigmoid patterns. *BMC Plant Biology* 20(1):108 DOI 10.1186/s12870-020-2317-6.
- Pei M-S, Liu H-N, Wei T-L, Yu Y-H, Guo D-L. 2021. Detection and characterization of genome-wide genetic variation associated with the early-ripening phenotype of grape mutants. *Scientia Horticulturae* 285(2):110195 DOI 10.1016/j.scienta.2021.110195.
- Qi X, Liu C, Song L, Li Y, Li M. 2017. PaCYP78A9, a cytochrome P450, regulates fruit size in sweet cherry (*Prunus avium* L.). *Frontiers in Plant Science* 8:20115 DOI 10.3389/fpls.2017.02076.
- Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO. 2012. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics (Oxford, England)* 28(18):i333–i339 DOI 10.1093/bioinformatics/bts378.
- Scorzal R, May LG, Purnell B, Upchurch B. 1991. Differences in number and area of mesocarp cells between small- and large-fruited peach cultivars. *Journal of the American Society for Horticultural Science* 116(5):861–864 DOI 10.21273/JASHS.116.5.861.

- Su L, Bassa C, Audran C, Mila I, Cheniclet C, Chevalier C, Bouzayen M, Roustan J-P, Chervin C. 2014. The auxin SI-IAA17 transcriptional repressor controls fruit size via the regulation of endoreduplication-related cell expansion. *Plant & Cell Physiology* 55(11):1969–1976 DOI 10.1093/pcp/pcu124.
- Suzuki M, Sato Y, Wu S, Kang B-H, McCarty DR. 2015. Conserved functions of the MATE transporter BIG EMBRYO1 in regulation of lateral organ size and initiation rate. *The Plant Cell* 27(8):2288–2300 DOI 10.1105/tpc.15.00290.
- Sweetman C, Wong DC, Ford CM, Drew DP. 2012. Transcriptome analysis at four developmental stages of grape berry (*Vitis vinifera* cv. Shiraz) provides insights into regulated and coordinated gene expression. *BMC Genomics* 13(1):691 DOI 10.1186/1471-2164-13-691.
- Tello J, Torres-Pérez R, Grimplet J, Carbonell-Bejerano P, Martínez-Zapater JM, Ibáñez J. 2015. Polymorphisms and minihaplotypes in the VvNAC26 gene associate with berry size variation in grapevine. *BMC Plant Biology* 15(1):253 DOI 10.1186/s12870-015-0622-2.
- Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics (Oxford, England)* 25(9):1105–1111 DOI 10.1093/bioinformatics/btp120.
- Tucker MR, Lou H, Aubert MK, Wilkinson LG, Little A, Houston K, Pinto SC, Shirley NJ. 2018. Exploring the role of cell wall-related genes and polysaccharides during plant development. *Plants (Basel, Switzerland)* 7(2):E42 DOI 10.3390/plants7020042.
- Wang X, Li Y, Zhang H, Sun G, Zhang W, Qiu L. 2015. Evolution and association analysis of GmCYP78A10 gene with seed size/weight and pod number in soybean. *Molecular Biology Reports* 42(2):489–496 DOI 10.1007/s11033-014-3792-3.
- Wang X, Shen F, Gao Y, Wang K, Chen R, Luo J, Yang L, Zhang X, Qiu C, Li W, Wu T, Xu X, Wang Y, Cong P, Han Z, Zhang X. 2020. Application of genome-wide insertion/deletion markers on genetic structure analysis and identity signature of Malus accessions. *BMC Plant Biology* 20(1):540 DOI 10.1186/s12870-020-02744-2.
- Wang Q, Sun G, Ren X, Du B, Cheng Y, Wang Y, Li C, Sun D. 2019. Dissecting the genetic basis of grain size and weight in barley (*Hordeum vulgare* L.) by QTL and comparative genetic analyses. *Frontiers in Plant Science* 10:237 DOI 10.3389/fpls.2019.00469.
- Wang Q, Xue X, Li Y, Dong Y, Zhang L, Zhou Q, Deng F, Ma Z, Qiao D, Hu C, Ren Y. 2016. A maize ADP-ribosylation factor ZmArf2 increases organ and seed size by promoting cell expansion in Arabidopsis. *Physiologia Plantarum* 156(1):97–107 DOI 10.1111/ppl.12359.
- Wang H, Yan A, Wang X, Zhang G, Liu Z, Xu H, Sun L. 2022. Identification of QTLs and candidate genes controlling berry size in table grape by integrating QTL and transcriptomic analysis. *Scientia Horticulturae* 305(3):111403 DOI 10.1016/j.scienta.2022.111403.
- Wei L, Cao Y, Cheng J, Xiang J, Shen B, Wu J. 2020. Comparative transcriptome analyses of a table grape ‘Summer Black’ and its early-ripening mutant ‘Tiangong Moyu’ identify candidate genes potentially involved in berry development and ripening. *Journal of Plant Interactions* 15(1):213–222 DOI 10.1080/17429145.2020.1760367.
- Xia T, Li N, Dumenil J, Li J, Kamenski A, Bevan MW, Gao F, Li Y. 2013. The ubiquitin receptor DA1 interacts with the E3 ubiquitin ligase DA2 to regulate seed and organ size in Arabidopsis. *The Plant Cell* 25(9):3347–3359 DOI 10.1105/tpc.113.115063.
- Xu F, Fang J, Ou S, Gao S, Zhang F, Du L, Xiao Y, Wang H, Sun X, Chu J, Wang G, Chu C. 2015. Variations in CYP78A13 coding region influence grain size and yield in rice. *Plant, Cell & Environment* 38(4):800–811 DOI 10.1111/pce.12452.
- Xu Y, Gao Z, Tao J, Jiang W, Zhang S, Wang Q, Qu S. 2016. Genome-wide detection of SNP and SV variations to reveal early ripening-related genes in grape. *PLOS ONE* 11(2):e0147749 DOI 10.1371/journal.pone.0147749.

- Yamada M, Sato A. 2016.** Advances in table grape breeding in Japan. *Breeding Science* **66**(1):34–45
DOI [10.1270/jsbbs.66.34](https://doi.org/10.1270/jsbbs.66.34).
- Yang J, Xiao Y-Y. 2013.** Grape phytochemicals and associated health benefits. *Critical Reviews in Food Science and Nutrition* **53**(11):1202–1225 DOI [10.1080/10408398.2012.692408](https://doi.org/10.1080/10408398.2012.692408).
- Ye Y, Ding Y, Jiang Q, Wang F, Sun J, Zhu C. 2017.** The role of receptor-like protein kinases (RLKs) in abiotic stress response in plants. *Plant Cell Reports* **36**(2):235–242
DOI [10.1007/s00299-016-2084-x](https://doi.org/10.1007/s00299-016-2084-x).
- Zhan P, Wei X, Xiao Z, Wang X, Ma S, Lin S, Li F, Bu S, Liu Z, Zhu H, Liu G, Zhang G, Wang S. 2021.** GW10, a member of P450 subfamily regulates grain size and grain number in rice. *TAG. Theoretical and Applied Genetics. Theoretische und Angewandte Genetik* **134**(12):3941–3950
DOI [10.1007/s00122-021-03939-3](https://doi.org/10.1007/s00122-021-03939-3).
- Zhao F, Zhang J, Weng L, Li M, Wang Q, Xiao H. 2021.** Fruit size control by a zinc finger protein regulating pericarp cell size in tomato. *Molecular Horticulture* **1**(1):6
DOI [10.1186/s43897-021-00009-6](https://doi.org/10.1186/s43897-021-00009-6).