Genome-wide perspective reveals the potential mechanism of maturation differentiation in cabbage (*Brassica oleracea* var. *capitata*)

Guoli Zhang[1] [iD], Yuting Sun[2], Na Lyu[1], Wenfang Dong[3], Xuechao Liu[3], Shamsiah Abdullah[4] and Qiang Li[1*]

[1] *Faculty of Life Science, Tangshan Normal University, Tangshan Key Laboratory of Cruciferous Vegetable Genetics and Breeding, Tangshan 063000, China*
[2] *Heilongjiang Institute of Construction Technology, Harbin 150025, China*
[3] *Tangshan Fengnan District agricultural service center, Tangshan 063000, China*
[4] *Faculty of Plantation and Agrotechnology, Universiti Teknologi MARA, Jasin Campus, Merlimau 77300, Malaysia*
* Correspondence: liqiang@tstc.edu.cn (Li Q)

Abstract

Sustainable agricultural production is an effective way to address global climate change and the rapid growth of population, among which the multi-maturing crops play a significant role, but its underlying genetic background has rarely been studied. We sequenced the genomes of early, middle, and late-maturing cabbage, and analyzed them comparatively. Phylogenetic and genetic structure analysis showed that the early- and late-maturing varieties had a certain genetic distance, and the early- and late-maturing individuals were likely to have been acquired by the mid-maturing individuals through continuous breeding. Selective-sweep and gene function enrichment analysis revealed candidate genes in early and late-maturing resources were significantly enriched in metabolic pathways, and candidate genes in early-maturing resources were associated with cell wall modification and pectinesterase activity. The present study provides initial insights into the mechanism of maturation differentiation in cabbage, and laid an important scientific basis for the future of higher yield and wider human needs of cabbage.

Introduction

A rapidly growing population and the impact of global climate change on crop production are urging the improvement of agricultural yields and quality in a sustainable manner. Multi-maturing of crops has the characteristics of high efficiency, intensiveness, and high yield. Reasonable arrangement of maturities is conducive to creating a harmonious and sustainable agro-ecological environment. Although the maturity of crop varieties is often affected by temperature, precipitation, soil, sowing date, cultivation technology conditions, and other factors, it is mainly determined by the heritability of the varieties[1]. Heredity, variation, and complex genetic regulation are the important basis of crop evolution and variety breeding.

Leafy vegetable crops play an important role in the human diet, and are important sources of vitamins, carotenoids, flavonoids, and other health-promoting substances. The *Brassica* genus contains a diverse range of oilseed and leafy vegetable crops important for human nutrition, and *Brassica oleracea* is one of the most representative leafy vegetable crops, widely cultivated in the world, with an extreme morphological diversity and crop form[2]. Compared to other *Brassica* species, *B. oleracea* vegetables have the characteristics of rapid diversification and domestication of morphological types. The construction of the *B. oleracea* pan-genome revealed the hidden driving force of the rapid domestication of *B. oleracea* varieties, and obtained several key genes for important traits[3]. However, systematic characterization of gene function is lacking[4].

Cabbage (*B. oleracea* var. *capitata* L.) is a herbaceous plant native to the coasts from the Mediterranean to the North Sea, and is a widely cultivated vegetable worldwide. Rich in nutrients such as vitamin C, carotenoids, cellulose, folic acid, and amino acids, it has a good dietary health care effect[5,6], and important economic value[7]. The importance of cabbage in tropical and subtropical regions has increased considerably in recent decades. Maturity is one of the important horticultural characteristics of cabbage, which is the fundamental trait system formed gradually in the reform of long-term tillage system. Cabbage can be divided into three categories of early-, middle-, and late-maturing, belonging to the green body vernalization type. Seedlings need to grow to a certain size and be exposed to a certain low temperature for a certain period before they can undergo vernalization[8–11]. The maturation time of cabbage varieties refers to the period from transplanting to harvest (early-maturing: t < 60 d, transplanted in April; mid-maturing: 60 d < t < 90 d, transplanted in May; late-maturing: t > 90 d, transplanted in June). The typical temperate monsoon climate and large latitude span have led to different breeding needs for cabbage. The maturation time of some extremely early-maturing varieties can be as short as 45 d, such as Chungan 45 and Zhonggan 12. Following a long period of direct selection, the adaptability of the three different varieties of cabbage has improved and has been widely cultivated in different areas at different times. Evaluating cabbage varieties for maturity will help farmers, breeders, and seed companies select and develop varieties best suited to the local environment and market.

With the publication of the *B. oleracea* genome, an increasing number of questions in the fields of evolution, domestication, crop quantitative genetics, molecular genetics, and crop breeding are being addressed[7,12–14]. To examine the mechanism of maturation differentiation of cabbage in more detail, the whole genome of cabbage varieties with different maturation times were resequenced and investigated the genetic changes during maturation differentiation by phylogenetic tree construction, genetic structure analysis, selective-sweep, and gene function enrichment. This study

provided initial insights into the mechanism of maturation differentiation of cabbage, laying an important theoretical basis for the breeding and identification of dominant species and new varieties, improvement of agronomic traits, conservation of germplasm resources, and marker-assisted breeding of cabbage, and also provided an important scientific basis for the future of higher yield and wider human needs of cabbage.

Materials and methods

Plant materials and DNA extraction

A total of 91 cabbage samples from 14 countries were collected for this study. All materials were seeded in a laboratory tray and grown under suitable conditions. Twenty-two samples were early maturing resources, 46 samples were mid-maturing resources, and 23 samples were late-maturing resources (Supplementary Table S1). After phenotypic observation and evaluation, the differences among cabbage varieties with different maturation times were mainly reflected in the shape of the leaf ball, including round head type, flat head type, and pointed head type. Total genomic DNA was extracted from 3–5 tender leaves at the seedling stage using the CTAB method[15]. RNase treatment was performed to remove RNA contaminants. Genomic DNA quality was measured using agarose gel electrophoresis (1%), and the concentration was quantified using Qubit 3.0 Fluorometer. More than 1.5 μg of DNA samples were used to build the library.

Database building and sequencing

The DNA samples processed with a Covaris crusher were randomly broken into 350 bp fragments. Libraries were constructed using the TruSeq Library Construction Kit. Initial quantification was performed using a Qubit 3.0, and libraries were diluted to 1 ng/μl. The insert size of the libraries was then determined using an Agilent 2100. Quantitative PCR (qPCR) was used to accurately quantify the effective concentration of the libraries to ensure their quality (effective concentration > 2 nM). The whole genome resequencing was performed on Illumina X-TEN platform at Novogene Corporation (Beijing, China), with a paired-end sequencing length of 150 bp. Sequencing and base calling were performed according to the standard Illumina protocols.

Resequencing data processing and SNP calling

The raw data generated by sequencing were rigorously filtered via SOAPnuke 1.5.6 software to produce high-quality clean data[16]. The specific filtering conditions were as follows: (1) filter out reads containing adapter sequences; (2) filter out reads containing the number of N greater than 10%; (3) filter out reads with more than 50% low-quality bases (Phred score ≤ 5). The genome sequence of *B. oleracea* JZS 2.0 (http://brassicadb.cn/#/Download) was downloaded as the reference genome. Burrows Wheeler Alignment (BWA) software was used to align the high-quality clean data with the reference genome[17], and duplicate reads in the alignment results were removed using SAMtools[18]. The results of sample comparison showed that their similarity with the reference genome met the requirements of resequencing analysis, and the coverage depth and genome coverage were excellent. SNPs were detected using SAMtools, with polymorphic loci identified via a Bayesian model. SNPs detected were filtered to obtain high-quality SNPs for subsequent analysis. The filtering conditions were as follows: (1) the coverage depth is 2; (2) the missing rate was set at 0.2; (3) the minimum allele frequency was 0.05.

Population genetics analysis

To analyze the phylogenetic relationships among cabbages with different maturities and evaluate the evolutionary trends of leaf ball shape, we constructed a neighbor-joining tree with TreeBeST[19], with bootstrap values set to 1,000. FigTree (http://tree.bio.ed.ac.uk/software/figtree/) was used to visualize the phylogenetic trees. The genetic structure was inferred using PLINK[20]. The number of subgroups (K values) was set to 2–8, and the optimal K values were determined by the cross-validation (CV) error method.

Genome scan for selection

To identify genomic regions that may have been subject to selection during maturation differentiation, three comparisons, were conducted namely, pairwise comparisons of samples with different maturation times: early-maturing vs mid-maturing, mid-maturing vs late-maturing, and early-maturing vs late-maturing. The present scan used a 0.05-cM sliding window with 100-bp steps across the whole genome. The population differentiation statistic (F_{ST}) was calculated using a 50 kb window with a 20 kb step for the three contrasts, and the sequence diversity statistics (π) for cabbage varieties with different maturation times separately using VCFtools[21]. The π log-ratio was calculated as $\ln(\pi_{Early})$ - $\ln(\pi_{Mid})$, $\ln(\pi_{Mid})$ - $\ln(\pi_{Late})$ and $\ln(\pi_{Early})$ - $\ln(\pi_{Late})$, respectively, where π_{Early}, π_{Mid} and π_{Late} are the nucleotide diversity values for the early-maturing, mid-maturing, and late-maturing cabbage samples, respectively. Plots were generated using the ggplot2 package (https://github.com/tidyverse/ggplot2; accessed on 4 March 2023) in R. Putative selection targets with the top 5% and bottom 5% of π-ratios, the top 5% of F_{ST} values, were extracted as high-confidence outliers. All candidate genes were used for subsequent analysis and discussion.

To test whether the candidate selective-sweep regions had an excess of singleton polymorphisms, Tajima's *D* values were calculated for samples with different maturation times, respectively, using the same sliding window approach. Regions under selective sweeps had significantly lower values of Tajima's *D*.

Gene function enrichment analysis

Functional classification into Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) categories was performed for candidate genes selected from different maturation gene pools in each set of comparative analyses[22]. KEGG pathway enrichment analysis was performed using the KOBAS program[23], and GO annotation was performed using the GOseq R package[24]. The hypergeometric test was used to calculate *p*-value, and Benjamini-Hochberg (BH) correction was applied[25]. Terms with a corrected *p*-value < 0.05 were considered significantly enriched. The 20 most significantly enriched KEGG pathways were displayed in a scatter plot. In the GO analysis, the 30 GO terms with the most significant enrichment were shown in the bar chart, which directly reflected the distribution of candidate genes enriched in these GO terms. In addition, the Directed Acyclic Graph (DAG) for each ontology was drawn, and associated GO terms were clearly displayed through branch inclusion relationships.

Results

Genetic relationship

A total of 3.43 million high-quality single-nucleotide polymorphisms (SNPs) were detected, and the phylogenetic tree

constructed based on this dataset showed genome-wide relationships and divergence among cabbage resources with different maturation times. They did not cluster into an independent clade, respectively, but are embedded within each other (Fig. 1a). Early-maturing resources were clustered in a clade, but embedded with some mid-maturing individuals. Similarly, late-maturing resources were clustered in a clade, but there were also individuals of other maturities embedded within them. Mid-maturing resources were not clustered into a monophyletic clade, but were distributed evenly throughout the phylogenetic tree. The genetic differences between early-maturing and late-maturing individuals are clearly reflected in the tree, revealing that climatic differences across seasons have shaped distinct genetic backgrounds, and the diversity of variety resources has been enriched. The close pedigree relationship between mid-maturing individuals and early/late-maturing individuals was shown. The results of genetic structure analysis were consistent with this (Fig. 1b). When K = 2, early- and late-maturing individuals were obviously divided into two groups, although each groups contained mid-maturing individuals. When K = 3, some late-maturing individuals showed significant genetic differentiation from mid-maturing individuals. The results of K = 4 showed that some early-maturing individuals had a certain genetic distance from the mid-maturing individuals.

Selective-sweep for maturation differentiation

Genomic regions that have been subject to selection as inferred from low/high π log-ratios and significant genetic differentiation were detected (Fig. 2). More specifically, in the analysis of early-maturing vs mid-maturing, at a significance level of $p > 0.005$ (Z test, with π log-ratio ≥ -0.31 or ≤ 1.91 and $F_{ST} \geq 0.17$), a total of 433 potential selective-sweep regions overlapping with 818 candidate genes were identified in early-maturing gene pools, and 182 potential selective-sweep regions overlapping with 271 candidate genes were identified in mid-maturing gene pool; similarly, in mid-maturing vs late-maturing (π log-ratio ≥ -0.64 or ≤ 1.17 and $F_{ST} \geq 0.19$), 312 potential selective-sweep regions, and 557 candidate genes were identified in mid-maturing gene pools, 491 potential selective-sweep regions covering 802 candidate genes were identified in late-maturing gene pool; in early-maturing vs late-maturing (π log-ratio ≥ -2.07 or ≤ 1.04 and $F_{ST} \geq 0.38$), 181 and 316 potential selective-sweep regions, 456 and 470 candidate genes were identified in early-maturing and late-maturing gene pools, respectively (Table 1; Supplementary Tables S2–S7). The role of these regions was confirmed by significantly lower values of Tajima's D (Supplementary Fig. S1).

Secondary metabolism was reinforced during maturation differentiation

In the comparative analysis, the candidate genes from resources with different maturities were functionally enriched, and the enriched metabolic pathways showed anisotropy (Fig. 3). In early-maturing resources, candidate genes were significantly enriched in metabolic pathways, and more genes were included (Fig. 3a, e). Although these pathways were significant (e.g., phenylalanine metabolism, pentose, and glucuronate interconversions and biosynthesis of amino acids), the number of enriched genes was small. For mid-maturing, candidate genes were significantly enriched in ubiquitin-mediated proteolysis, spliceosome, glutathione metabolism, cyanoamino acid metabolism, and arachidonic acid metabolism (Fig. 3b, c). Relatively more genes were enriched in the biosynthesis of secondary metabolites and RNA transport pathways, but these enrichments were not significant. For late-maturing, candidate genes were mainly enriched in metabolic pathways, although not highly significant (Fig. 3d, f).

GO term enrichment analysis suggested that candidate genes in early-maturing resources were associated with cell wall modification and pectinesterase activity ($p < 0.05$, Fig. 4a, Supplementary Fig. S2). Candidate genes in mid-maturing were overexpressed in multiple terms, e.g. DNA restriction-modification system, clearance of foreign intracellular DNA, Type II site-specific deoxyribonuclease activity, G-protein-coupled acetylcholine receptor signaling pathway, gamma-glutamyltransferase activity ($p < 0.05$, Fig. 4b, c, Supplementary Fig. S3). Candidate genes of late-maturing resources showed no significant enrichment (Fig. 4d, f).

Discussion

Phylogenetic relationships among different maturing resources

Phylogenetic analysis and genetic structure analysis showed that the cabbage varieties with different maturation times did not cluster into a single genetic cluster, but were embedded with each other and shared some genetic background. The phylogenetic tree showed that early- and late-maturing individuals were clustered together in one clade, respectively, but both had intercalations individuals of mid-mature individuals. Similarly, in genetic structure analysis, when K = 2, early- and late-maturing individuals were clearly identified, and when K = 3 and 4, some early and late-maturing individuals had a certain genetic distance from mid-maturing individuals. The results revealed that the genetic
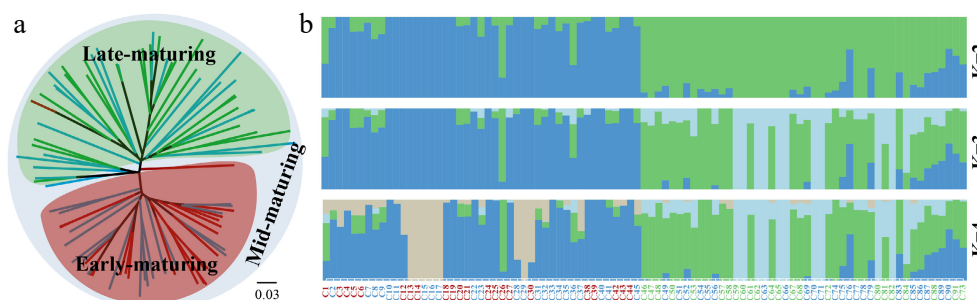

Fig. 1 Phylogenetic and genetic structure analyses of cabbage varieties with different maturation times. (a) A neighbour-joining phylogenetic tree constructed using whole-genome SNPs data. The red lines represent early-maturing of cabbages; green lines represent late-maturing of cabbages; blue lines represent mid-maturing of cabbages. The clades clustered by each maturity resource were identified with corresponding colors. (b) Genetic structure plots with K = 2–4. The y axis quantifies the proportion of the individual's genome from inferred ancestral populations, and x axis shows the different individuals. The color of the sample code represents the same meaning as in (a).
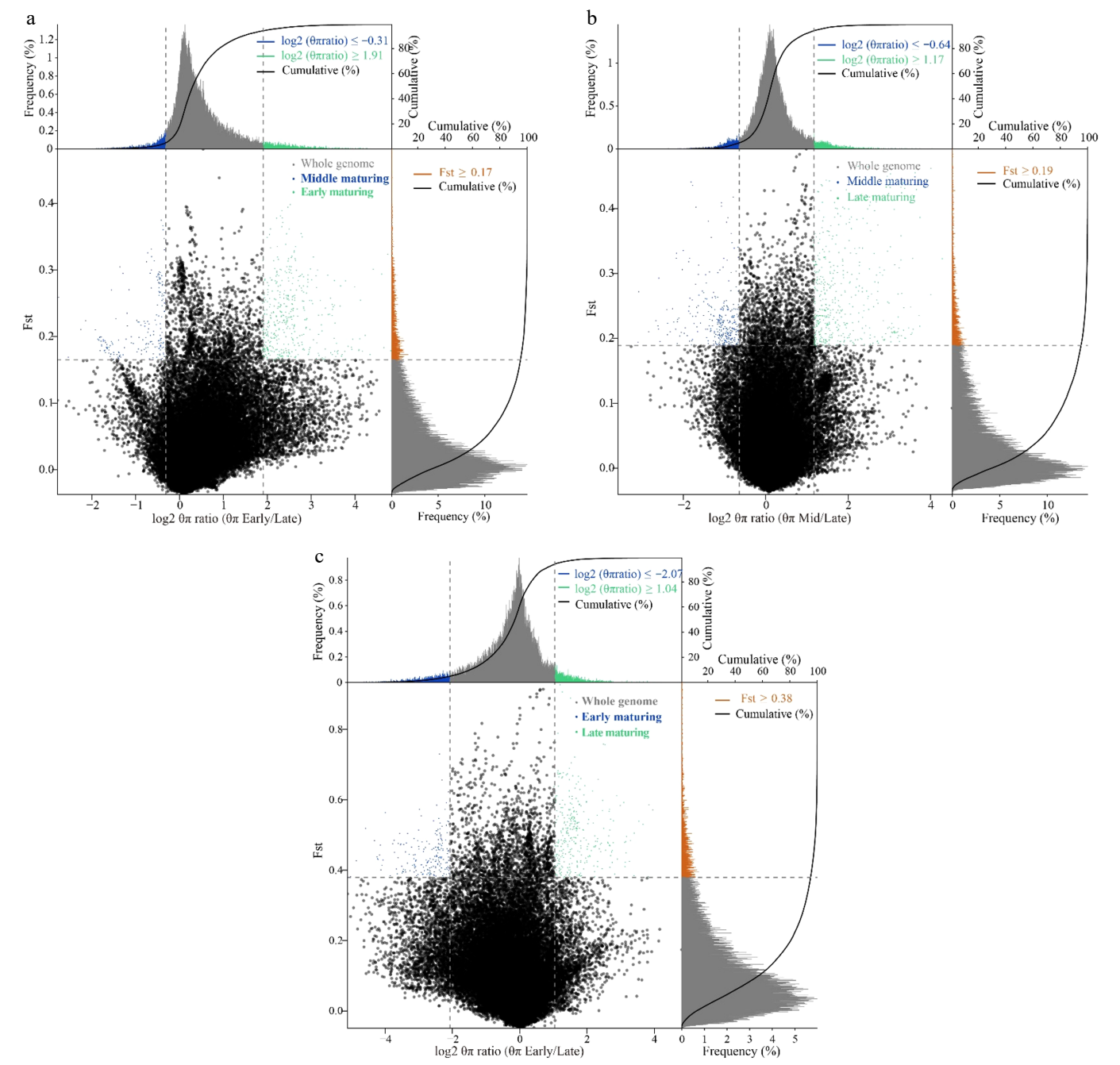
Fig. 2 Genomic regions with selection sweep signals in different mature cabbages. (a) Distribution of ln ratio ($\theta\pi$, early/middle) and F_{ST}. Green dots represent windows fulfilling the selected regions requirement in early-maturing cabbages; blue dots represent selected regions in mid-maturing. (b) Distribution of ln ratio ($\theta\pi$, middle/late) and F_{ST}. Blue dots represent selected regions in mid-maturing; green dots represent selected regions in late-maturing. (c) Distribution of ln ratio ($\theta\pi$, early/late) and F_{ST}. Blue dots represent selected regions in early-maturing; green dots represent selected regions in late-maturing.

Table 1. The numbers of regions and candidate genes with selection sweep signals in the comparative analysis of selective-sweep of cabbage varieties with different maturation times.

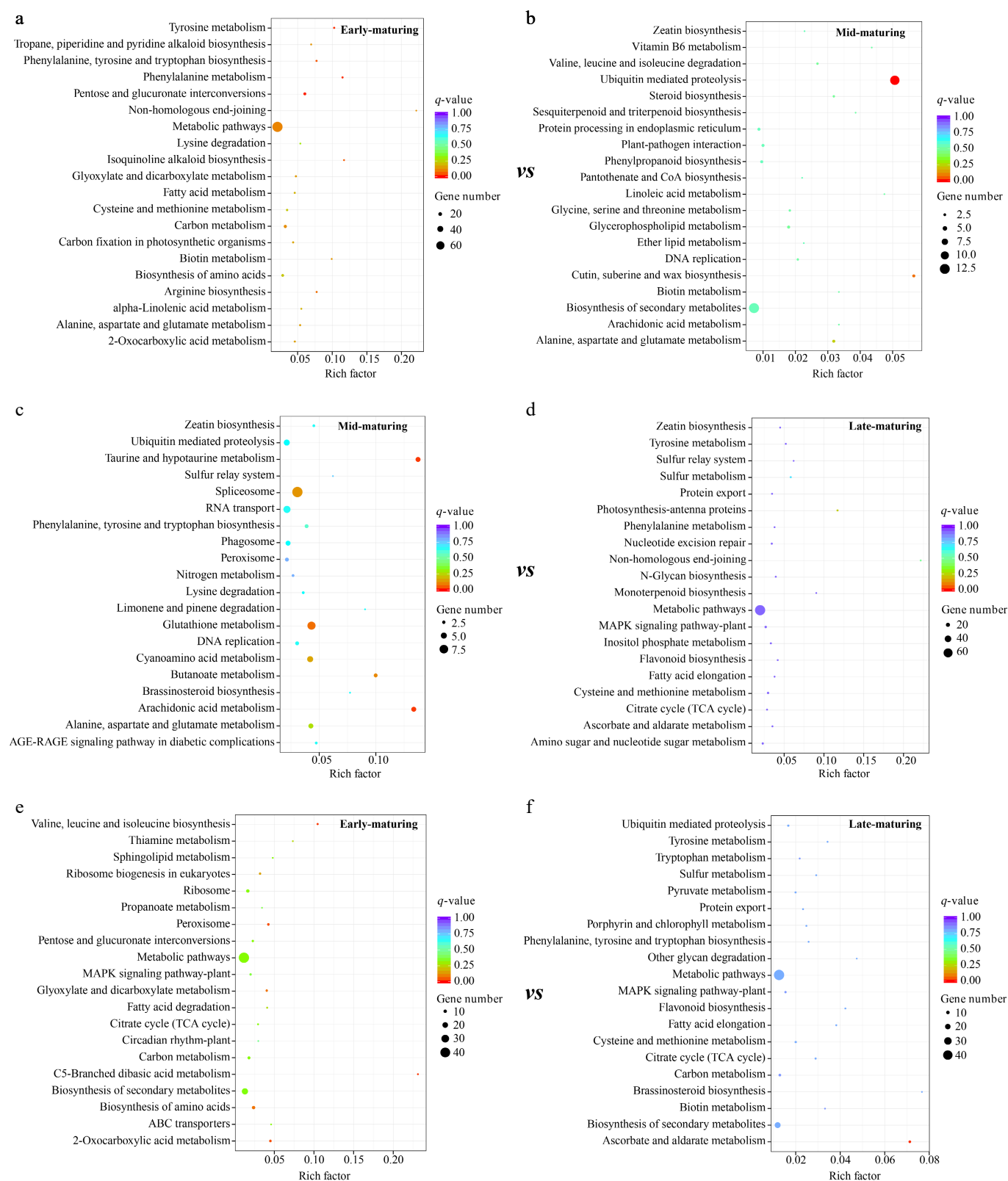| Maturation	With selection sweep signals	Analysis among cabbages with different maturation times		
		Early vs Mid	Mid vs Late	Early vs Late
Early-maturing	Regions	433	–	181
	Candidate genes	818	–	456
Mid-maturing	Regions	182	312	–
	Candidate genes	271	557	–
Late-maturing	Regions	–	491	316
	Candidate genes	–	802	470

Fig. 3 KEGG pathway enrichment of the candidate genes in different mature cabbages. Enrichment result of candidate genes located in the (a) early-maturing and (b) mid-maturing resources respectively in early-maturing vs mid-maturing analysis; (c) mid-maturing and (d) late-maturing resources respectively in mid-maturing vs late-maturing analysis; (e) early-maturing and (f) late-maturing resources respectively in early-maturing vs late-maturing analysis.
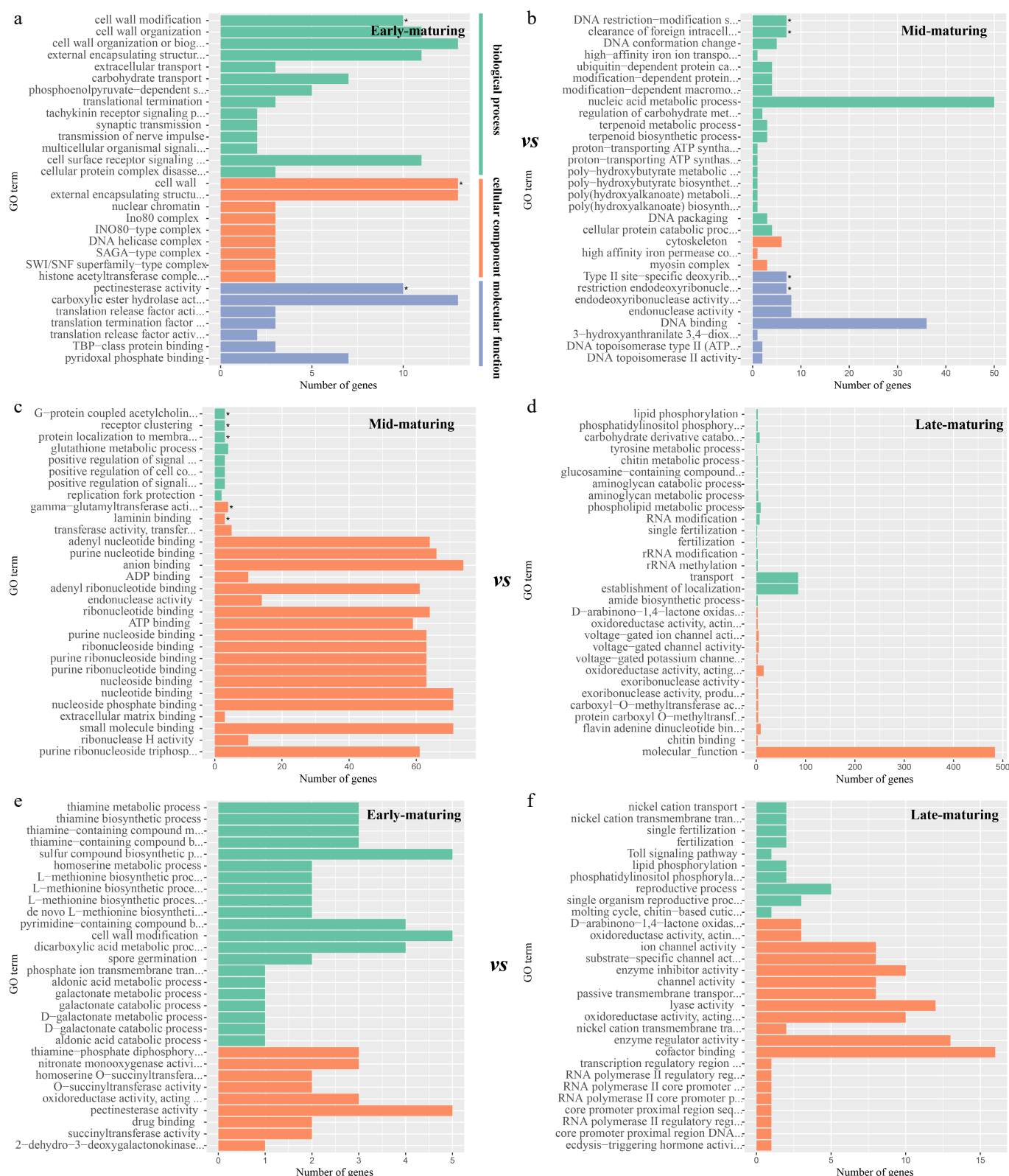
Fig. 4 GO annotation of the candidate genes in different mature cabbages. Annotation result of candidate genes located in the (a) early-maturing, and (b) mid-maturing resources respectively in early-maturing vs mid-maturing analysis. Result of candidate genes located in the (c) mid-maturing, and (d) late-maturing resources respectively in mid-maturing vs late-maturing analysis. Result of candidate genes located in the (e) early-maturing, and (f) late-maturing resources respectively in early-maturing vs late-maturing analysis.

divergence between mid-maturing resources, and the other two were relatively low, while the divergence was obvious between early-maturing and late-maturing resources. It is reasonable to assume that the mid-maturing individuals represent a relatively primitive group, and the early- and late-maturing individuals are likely to have been acquired by the mid-maturing individuals

through continuous breeding. Species from *Brassica*'s 'U's triangle' model are widely used in studies on polyploidization and genome hybridization[26]. As an important component of this model, cabbage has its evolutionary process worthy of in-depth exploration—which will provide a more reliable basis for enriching its genetic background. The initial inferences of this study offer a new perspective for subsequent research. The initial differentiation of *B. oleracea* was influenced by artificial selection on leaves or stems[27], with cabbage originating from the 'leafy head' domestication lineage[28]; thus, leaf ball shape is the most important phenotypic trait of cabbage. The leaf ball shape of cabbage with different maturities was also differentiated to some extent. The leaf balls of early-maturing cabbage are generally pointed and round heads, those of late-maturing cabbage are generally flat heads, and those of mid-maturing cabbage have all three types of balls. These observations further strengthen the findings of this study and indicate that, along with the evolution of maturities, the leaf ball shape of cabbage has also undergone convergent evolution. Early-maturing cabbages are more inclined to have pointed and round head traits, while late-maturing varieties tend to have flat heads. Gene flows between cabbages of different maturities were also detected. These gene flows are likely the result of human agricultural activities and ongoing variety improvement processes, which have further shaped the genetic diversity and structure of cabbage.

The ripening time of cabbage is closely related to temperature, light intensity and duration, water, and nutrients[29]. The results showed that there were obvious genetic differences between early and late-maturing varieties, with mid-maturing varieties as the reference. Mid-maturing varieties are suitable for planting in most areas, while the planting area of early and late-maturing varieties is relatively limited, suggesting that environmental factors played an important role in their phenotypic and molecular changes. The breeding of new varieties aims to further meet the needs of production and daily life. Early breeding is a more purposeful method, prioritizing ripening while integrating other favorable traits. Early maturity is an important index in vegetable crop breeding. The selection of early-maturing cabbage varieties enables two harvests in one growing season, frees up space for early planting of other crops. Owing to their shorter growth cycle, early-maturing varieties can be marketed earlier, filling market gaps and satisfying consumers' demand for fresh vegetables. As a result, they can command higher prices and increase the economic benefits of the growers[30,31], while the breeding of late-maturing varieties plays an important role in the annual supply of vegetables. These varieties have strong cold resistance and excellent storage resistance, allowing them to be harvested well into the deep winter. They thus play a significant role in the long-term vegetable supply and storage chain[32]. More importantly, the combination of early- and late-maturing varieties can stagger sowing or harvesting periods, facilitate the efficient arrangement of agricultural machinery and labor, and avoid missing farming deadlines. The cultivation of cabbage for multiple maturities has made it a staple of people's diet, like wheat and rice[33]. Overall, the relative pedigree and genetic structure analysis of different maturities sheds light on the domestication of cabbage, providing valuable insights into the evolutionary history and genetic variation of this important vegetable species. The results of this study have implications for future molecular breeding and genetic improvement of cabbage and contribute to our broader understanding of crop domestication processes.

Genetic mechanism of maturation differentiation

The continuous development of sequencing technology has prompted many researchers to study the genetic mechanisms of various fruit and vegetable crops based on genomic data. The ripening of fruits and vegetables is a crucial agronomic trait, but most of the current studies focus on the changes of physicochemical characteristics during ripening[34–37], or the differences in gene expression at different ripening stages[38–40], and rarely involve the genetic differences between resources with different maturities. The maturity of cabbage is one of the important indicators for selecting varieties with a suitable maturity stage in different ecological regions. A deep understanding of the genetic mechanisms underlying maturity is helpful to accelerate the breeding process, and rapid selection of more new varieties with different maturity stages is of great significance to revitalize the cabbage industry.

Whole-genome resequencing was performed to identify the key genetic differences among cabbage varieties with different maturities and reveal the important factors influencing these differences. In KEGG enrichment analysis, candidate genes were mainly enriched in metabolic pathways in both early and late-maturing resources (Fig. 3), indicating that related genes promote early or late ripening of different cabbage varieties by influencing metabolic pathways, a phenomenon also observed in other crop of the *Brassica* family, such as *Brassica rapa*, critical genes involved in the terpenoid biosynthesis pathway and terpene synthase (TPS) family genes have been identified through integrated genomic and transcriptomic analyses[41]. Additionally, for other vegetables like tomatoes[42]. GO term enrichment analysis on the candidate genes in early-maturing resources suggested that genes associated with cell wall modification and pectinesterase activity were involved (Fig. 4). Cell wall substances (alcohol-insoluble solids, AIS) mainly include cellulose, hemicellulose, and pectin, among which pectin plays the most important role in fruit and vegetable ripening. Cell wall modification is a core process of cabbage ripening. The firmness of cabbage heads is determined by the degree of cross-linking of pectin in the middle layer of the cell wall[43]. Early-maturing resources exhibit rapid head softening, mainly due to rapid cell wall disintegration, which accelerates pectin degradation and subsequently causes further head softening. In contrast, late-maturing resources maintain a higher degree of pectin esterification for a longer period, resulting in stronger cell wall stability and thus delaying the softening process[44]. Previous studies showed that the structure of the fruit undergoes obvious changes during the ripening process, and these structural changes were considered to be the main cause of fruit ripening[45,46]. During the ripening of fruits and vegetables, the structure and composition of the cell wall undergo significant changes[47]. Pectin is degraded first into soluble pectin, the intercellular layer disappears, and cells separate from each other[48]. The present results revealed that changes in cell walls play a key role in the maturation of cabbage. Vegetable ripening is a highly programmed and complex biological process. This process involves the spatiotemporally specific expression of many ripening-related genes, which leads to a series of physiological and biochemical changes in color, texture, nutrition, flavor, and aroma, among others. In a sense, vegetable ripening is a finely regulated transcriptional reprogramming process[49]. These findings have been well demonstrated in the model species *Arabidopsis thaliana*, which belongs to the same family (Brassicaceae) as cabbage. A number of functional genes have been identified, including those involved in glucosinolate biosynthesis, flowering regulation, auxin signaling, and anthocyanin biosynthesis[26]. The maturity of cabbage is reflected in its heading, and the size, texture, and flavor of the head directly affect its quality. A large number of hormones and enzymes are specifically upregulated in the heading phase of cabbage[50]. As a key enzyme involved in decomposing cell walls, pectinesterase is

active during cabbage ripening and may be related to the final taste and texture.

After continuous domestication and improvement, cabbage has exhibited different types of differentiation through selective breeding to meet production and livelihood needs. Different breeding selection processes have different effects on the metabolome, and the maturity differentiation of cabbage is mainly reflected in metabolic pathways. Although our ancestors knew nothing about the metabolome, this did not hinder their selection and modification of the cabbage metabolome, because, in a sense, what we consume from cabbage is an important part of its metabolome, and the heading metabolome comprehensively reflects its nutrients and flavor substances. Therefore, based on this study, the key genes and specific metabolites in metabolic pathways that affect the maturity differentiation process of cabbage can be further explored in combination with metabolomics analysis. Furthermore, with the integration and optimization of genomic data platforms, it will be feasible to construct a cabbage biological network based on multi-omics data[51].

Author contributions

The authors confirm contribution to the paper as follows: study design and manuscript writing: Zhang G, Li Q; data collection and organization: Zhang G, Sun Y, performing research, data analysis: Zhang G, Lyu N, Dong W, Liu X; manuscript revision: Sun Y, Abdullah S; supplying guidance: Abdullah S. All authors reviewed the results and approved the final version of the manuscript.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Conflict of interest

The authors declare that they have no conflict of interest.

References

[1] Quiros CF, Farnham MW. 2011. The genetics of *Brassica oleracea*. In *Genetics and Genomics of the Brassicaceae*, eds. Schmidt R, Bancroft I. New York, USA: Springer. pp. 261–289 doi: 10.1007/978-1-4419-7118-0_9

[2] Warwick SI, Francis A, Al-Shehbaz IA. 2006. Brassicaceae: species checklist and database on CD-Rom. *Plant Systematics and Evolution* 259:249–258

[3] Li X, Wang Y, Cai C, Ji J, Han F, et al. 2024. Large-scale gene expression alterations introduced by structural variation drive morphotype diversification in *Brassica oleracea*. *Nature Genetics* 56:517–529

[4] Ma W, Zhang P, Zhao J, Hong Y. 2023. Chinese cabbage: an emerging model for functional genomics in leafy vegetable crops. *Trends in Plant Science* 28:515–518

[5] Kopsell DA, Kopsell DE. 2006. Accumulation and bioavailability of dietary carotenoids in vegetable crops. *Trends in Plant Science* 11:499–507

[6] Liang Y, Li Y, Sun AD, Lu HY, Liu XJ. 2019. Effects of different cultivation methods on nutritional constituents of cabbage. *Quality And Safety Of Agro-Products* 2019:73–77 (in Chinese)

[7] Liu S, Liu Y, Yang X, Tong C, Edwards D, et al. 2014. The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nature Communications* 5:3930

[8] Qin WB, Dai ZL, Pan YP. 2010. Effects of sowing data on yield, quality and resistance of overwintering cabbage hanyu No. 21. *Acta Agriculturae Jiangxi* 22:67–69 (in Chinese)

[9] Pan YF, Dai ZL, Pan YF, Pan YP. 2012. Suitable planting time of high quality Brass-resistant spring cabbage variety Runyu. *Jiangsu Agricultural Sciences* 2012:142–143 (in Chinese)

[10] Pan YF, Dai ZL, Mao ZL, Wu GP. 2010. Effect of seedling age on yield and harvest time of cabbage. *Acta Agriculturae Jiangxi* 22:64–65 (in Chinese)

[11] He J, Xiong XQ, Lu YL. 2011. Effect of density and seeding time on yield of rutabaga. *Prataculture Animal Husbandry* 2011:7–10 (in Chinese)

[12] Sheng X, Cai S, Shen Y, Yu H, Wang J, et al. 2024. QTL analysis and fine mapping of a major QTL and identification of candidate genes controlling curd setting height in cauliflower. *Vegetable Research* 4:e008

[13] Zhang B, Wu Y, Li S, Yang L, Zhuang M, et al. 2024. Two large inversions seriously suppress recombination and are essential for key genotype fixation in cabbage (*Brassica oleracea* L. var. *capitata*). *Horticulture Research* 11:uhae030

[14] Meng C, Liu X, Wu F, Ma X, Ma L, et al. 2025. DNA methylation, metabolome, and transcriptome analysis reveal epigenomic differences between diploid Chinese cabbage seeds and tetraploid Chinese cabbage seeds. *Genomics* 117:111088

[15] Yi S, Jin WT, Yuan YN, Fang YH. 2018. An optimized CTAB method for genomic DNA extraction from freshly-picked pinnae of fern, *Adiantum capillus-veneris* L. *Bio-protocol* 8:e2906

[16] Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, et al. 2018. SOAPnuke: a mapreduce acceleration supported software for integrated quality control and preprocessing of high-throughput sequencing data. *Gigaence* 7:1–6

[17] Li H, Durbin R. 2009. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25:1754–1760

[18] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079

[19] Vilella AJ, Severin J, Ureta-Vidal A, Heng L, Durbin R, et al. 2009. EnsemblCompara GeneTrees: complete, duplication-aware phylogenetic trees in vertebrates. *Genome research* 19:327–335

[20] Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, et al. 2015. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience* 4:7

[21] Danecek P, Auton A, Abecasis G, Albers CA, Banks E, et al. 2011. The variant call format and VCFtools. *Bioinformatics* 27:2156–2158

[22] Song X, Hu J, Wu T, Yang Q, Feng X, et al. 2021. Comparative analysis of long noncoding RNAs in angiosperms and characterization of long noncoding RNAs in response to heat stress in Chinese cabbage. *Horticulture Research* 8:48

[23] Bu D, Luo H, Huo P, Wang Z, Zhang S, et al. 2021. KOBAS−i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Research* 49:W317–W325

[24] oung MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-SEQ: accounting for selection bias. *Genome biology* 11:R14

[25] Ferreira JA, Zwinderman AH. 2006. On the benjamini-hochberg method. *The Annals of Statistics* 34:1827−1849

[26] Liu Z, Li N, Yu T, Wang Z, Wang J, et al. 2022. The brassicaceae genome resource (TBGR): a comprehensive genome platform for brassicaceae plants. *Plant Physiology* 190:226−237

[27] Guo N, Wang S, Wang T, Duan M, Zong M, et al. 2024. A graph-based pan-genome of *Brassica oleracea* provides new insights into its domestication and morphotype diversification. *Plant Communications* 5:100791

[28] Cai C, Bucher J, Bakker FT, Bonnema G. 2022. Evidence for two domestication lineages supporting a middle-eastern origin for *Brassica oleracea* crops from diversified kale populations. *Horticulture Research* 9:uhac033

[29] Xing S, Liu L, Guan H, Zhang X, Sun X, et al. 2021. Dynamic prediction model of ripening degree of Chinese spicy cabbage under fluctuation temperatures. *Food Packaging and Shelf Life* 30:100744

[30] Qiao J, Shi Y, Wang L. 2015. Research progress on early maturity breeding of fruit vegetables. *Chinese Vegetable* 1:14 (in Chinese)

[31] Yang LM, Fang ZY. 2022. Researches on cabbage genetics and breeding in china for 60 years. *Acta Horticulturae Sinica* 49:2075−2098 (in Chinese)

[32] Qin WB, Shan X, Zhang ZC, Dai ZL. 2023. Breeding of new cabbage variety ruigan 35 and its industrialization application prospect. *Hans Journal of Agricultural Sciences* 13:1−8 (in Chinese)

[33] Yan SB, Hu LJ, Ge CQ, Yu B, Huang KM. 2024. Anniversary efficiency cultivation techniques of stem lettuce-fresh corn-cabbage. *China Academic* 4:162−164 (in Chinese)

[34] Liang R, Su Y, Qin X, Gao Z, Fu Z, et al. 2022. Comparative transcriptomic analysis of two *Cucumis melo* var. *saccharinus* germplasms differing in fruit physical and chemical characteristics. *BMC Plant Biology* 22:193

[35] Dabesor PA, Sanni DM, Kolawole AO, Enujiugha VN, Lawal OT, et al. 2022. Changes in physicochemical properties and enzymes associated with ripening of snake tomato (*Trichosanthes Cucumerina* L.) fruit. *Biocatalysis and Agricultural Biotechnology* 40:102313

[36] Saladié M, Cañizares J, Phillips MA, Rodriguez-Concepcion M, Larrigaudière C, et al. 2015. Comparative transcriptional profiling analysis of developing melon (*Cucumis melo* L.) fruit from climacteric and non-climacteric varieties. *BMC Genomics* 16:440

[37] Zhu Q, Gao P, Liu S, Zhu Z, Amanullah S, et al. 2017. Comparative transcriptome analysis of two contrasting watermelon genotypes during fruit development and ripening. *BMC Genomics* 18:3

[38] Guo S, Sun H, Zhang H, Liu J, Ren Y, et al. 2015. Comparative transcriptome analysis of cultivated and wild watermelon during fruit development. *PLoS One* 10:e0130267

[39] Yue Q, Xie Y, Yang X, Zhang Y, Li Z, et al. 2024. An InDel variant in the promoter of the NAC transcription factor *MdNAC18.1* plays a major role in apple fruit ripening. *The Plant Cell* 37:koaf007

[40] Niu Q, Xu Y, Huang H, Li L, Tang D, et al. 2025. Two transcription factors play critical roles in mediating epigenetic regulation of fruit ripening in tomato. *Proceedings of the National Academy of Sciences of the United States of America* 122:e2422798122

[41] Liu Z, Fu Y, Wang H, Zhang Y, Han J, et al. 2023. The high-quality sequencing of the *Brassica rapa* 'XiangQingCai' genome and exploration of genome evolution and genes related to volatile aroma. *Horticulture Research* 10:uhad187

[42] Zuo J, Grierson D, Courtney LT, Wang Y, Gao L, et al. 2020. Relationships between genome methylation, levels of non-coding RNAs, mRNAs and metabolites in ripening tomato fruit. *The Plant Journal* 103:980−994

[43] Cui S, Yao B, Gao M, Sun X, Gou D, et al. 2017. Effects of pectin structure and crosslinking method on the properties of crosslinked pectin nanofibers. *Carbohydrate Polymers* 157:766−774

[44] Lyu M, Iftikhar J, Guo R, Wu B, Cao J. 2020. Patterns of expansion and expression divergence of the polygalacturonase gene family in *Brassica oleracea*. *International Journal of Molecular Sciences* 21:5706

[45] Mohammad M, Ding P. 2019. Physico-textural and cellular structure changes of *Carissa congesta* fruit during growth and development. *Scientia Horticulturae* 246:380−389

[46] Jusoh NAM, Ding P. 2023. Morphometrics and structural changes of fig (*Ficus carica* L. var. Ipoh Blue Giant) syconium during growth, maturation and on-tree ripening. *Scientia Horticulturae* 321:112347

[47] Li R, Wang Y, Li W, Shao Y. 2023. Comparative analyses of ripening, texture properties and cell wall composition in three tropical fruits treated with 1-methylcyclopropene during cold storage. *Horticulturae* 9:126

[48] Proctor A, Peng LC. 1989. Pectin transitions during blueberry fruit development and ripening. *Journal of Food Science* 54:385−387

[49] Deng L, Yang T, Li Q, Chang Z, Sun C, et al. 2023. Tomato MED25 regulates fruit ripening by interacting with EIN3-like transcription factors. *The Plant Cell* 35:1038−1057

[50] Zhang K, Yang Y, Wu J, Liang J, Chen S, et al. 2022. A cluster of transcripts identifies a transition stage initiating leafy head growth in heading morphotypes of *Brassica*. *The Plant Journal* 110:688−706

[51] Liu Z, Zhang C, He J, Li C, Fu Y, et al. 2024. plantGIR: a genomic database of plants. *Horticulture Research* 11:uhae342