scientific reports

OPEN

Comparative transcriptome analysis of midgut of *Bombyx mori* under a sucrose-supplemented diet

Zhenbin Hu[1,2,3], Lijie Zhan[1,2,3], Xinyi Chen[1,2], Jia Li[1,2], Meirong Zhang[1,2] & Pingzhen Xu[1,2]✉

As a monophagous insect, the silkworm possesses a restricted feeding capacity. To investigate the impact of varying sucrose concentrations on silkworm growth and development, transcriptomic analyses were conducted on the silkworm midgut. A total of 909 differentially expressed genes (DEGs) were identified, with down-regulated genes predominantly associated with starch and sucrose metabolism, fructose and mannose metabolism, galactose metabolism, and amino sugar and nucleotide sugar metabolism. These alterations suggest that excessive sucrose consumption may suppress the carbohydrate metabolism pathway in silkworm. This study offers a comprehensive analysis of transcriptomic studies on high sucrose intake in silkworm, providing insights into the effects of varying sucrose concentrations on these organisms.

Keywords Silkworm, Sucrose, Midgut, Metabolism, Transcriptome

The silkworm (*Bombyx mori*), as a species within the order Lepidoptera, is a widely used model organism in entomological research. This insect has undergone long-term domestication, primarily for silk production, and its breeding history is also quite extensive[1]. Silkworms mainly feed on mulberry leaves, which provide the necessary nutrients and moisture as a single food source[2]. Silkworms, throughout their entire life cycle, only ingest food from the external environment during the larval stage to meet the needs of growth and development. The nutrients required for the life activities of the pupal, adult, and egg stages are entirely dependent on the nutritional reserves accumulated during the larval stage[3,4]. The maximum increase in the body size of silkworm occurs in the last fifth instar. During the silkworm larval development day 3 of the fifth instar (I5D3) is the boundary[5], the growth rate is not constant throughout and is dependent on this time, the larva grows quickly and the growth rate is higher before this time, afterward, the growth rate declines[6]. Prior to the boundary time, the fat body, as the main energy storage tissue, stores a large amount of nutrients, causing a rapid increase in larval weight. Afterward, silkworm enters a voracious feeding period, with a sharp increase in mulberry leaf consumption, begins to synthesize mass silk proteins in the silk gland.

The mulberry leaves are rich in various nutritional components, the components of mulberry leaves that can be used by silkworm mainly include protein, carbohydrates, lipids, inorganic substances, moisture and vitamins[7–9]. The digestion and absorption of nutrients in silkworms are inseparable from the digestive tract, of which the midgut is the thickest part, accounting for approximately 78% of the total length of the digestive tract. Cylindrical cells are the main components of the midgut epithelium, especially well-developed in the posterior part of the midgut, capable of secreting digestive fluid and absorbing nutrients. Therefore, the majority of the enzymatic digestion and the absorption of nutrients largely take place in the midgut.

Glucose is the basic form of energy acquisition, but also an important substance of body tissue, involved in many physiological processes. For silkworms and many insects, sucrose is essential to obtain enough nutrients for growth and reproduction[10]. Insects regulate their energy metabolism and the conversion of energy substrates through hormonal control and various signaling pathways, allowing them to adapt to environmental fluctuations and optimize their growth and survival strategies across different physiological states[10–12]. Carbohydrate metabolism, in particular, plays a pivotal role in maintaining sugar balance in the hemolymph, modulating chitin synthesis in the cuticle, and influencing other physiological processes[13]. In summary, carbohydrate metabolism is a critical aspect of insect physiology, influencing energy homeostasis, structural integrity, and overall survival.

[1]Jiangsu Key Laboratory of Sericultural and Animal Biotechnology, School of Biotechnology, Jiangsu University of Science and Technology, Zhenjiang 212100, Jiangsu, China. [2]Key Laboratory of Silkworm and Mulberry Genetic Improvement, Ministry of Agriculture and Rural Affairs, Sericultural Scientific Research Center, Chinese Academy of Agricultural Sciences, Zhenjiang 212100, China. [3]Zhenbin Hu and Lijie Zhan contributed equally to this work. ✉email: xpz198249@just.edu.cn

By gaining a deeper understanding of the mechanism underlying glucose conversion and metabolic regulation in insect, we can develop more effective strategies for silkworm rearing, potentially leading to increased silk production and more sustainable agricultural practices. In this study, we investigated the effects of adding sucrose with four different concentrations (6%, 12%, 18% and 24%), respectively, from the 5th instar molted silkworm (I5D1) to wandering (spinning the cocoon). From I5D1 to I5D4, the diet supplemented with sucrose had positive effects on the development of silkworm, but after I5D5, that had negative effects. Among them, the growth situation of the 12% treatment group was relatively the better. Therefore, we conducted a transcriptomic comprehensive analysis of gene expression of midgut in I5D3 silkworm under condition of sucrose 12% treatment. The results revealed various alterations in carbohydrate metabolism and glycolytic pathways induced by changes in sucrose consumption. The findings indicate that the appropriate sucrose concentration and the optimal feeding time will positively impact the development of silkworms. This research not only can provide valuable understanding regarding the impact of sucrose on the midgut of silkworm but also holds significant guiding implications for enhancing silkworm breeding efficiency.

Materials and methods
Experimental insect
The silkworm 871 strain (monovoltine race) was provided by the Sericulture Research Institute of Chinese Academy of Agricultural Sciences, Zhenjiang, China. The rearing temperature should be maintained between 25 and 26 °C, and the humidity should be controlled between 70 and 85%. Ensure appropriate lighting and ventilation. Control the stocking density to avoid overcrowding. Provide fresh, clean, and uncontaminated mulberry leaves as the feed.

Sucrose treatment and RNA preparation
In the preliminary stage of the formal feeding trial, we conducted a pre-experiment to select the appropriate concentration levels. Initially, we used five concentrations, 6%, 12%, 18%, 24% and 30%, as evaluation criteria and found that the silkworms grew slowly and appeared mass death in the 30% concentration groups. Consequently, in the formal experiment, we decided to refine the concentration gradients to 6%, 12%, 18% and 24% to allow for a more reasonable assessment of the experimental results.

Select well-developed and similar-sized female silkworm at the 1st of the 5th instar (molted silkworm, I5D1) as the experimental subjects. Establish a Control Group (or C-Group) and four treatment groups with sucrose supplementation (or S-Group), each with three replicates consisting of 150 silkworm per replicate. The C-Group was fed mulberry leaves washed with deionized water, while the S-Groups were fed mulberry leaves soaked in 6%, 12%, 18% and 24% sucrose solutions. In handling the mulberry leaves, we firstly placed the calculated amount of sucrose into a certain volume of sterile water and heated it until fully dissolved. After cooling, additional sterile water was added to reach the desired volume. All leaves were selected to ensure a similar state, minimizing significant differences among them. During processing, the petioles and stems of the mulberry leaves were removed, and the leaves were cut into equal sizes. During the mulberry leaf immersion process, each leaf is individually soaked in a sucrose solution for 15 min to ensure that the mulberry leaves are fully exposed to the sucrose solution. Finally, we lay out the mulberry leaves taken from the solution flatly on a sterilized wire mesh for natural drying until that the mulberry leaves were moistened with sucrose solution without any dripping. Then the leaves were placed in the silkworm box for consumption by the silkworms. During feeding, the density of silkworms was also controlled to ensure that each silkworm had access to a sufficient amount of leaf.

Upon analysis of the feeding results, the growth situation of the 12% treatment group was relatively the better among all treat groups. Therefore, we conducted RNA-Seq of midgut in I5D3 female silkworm under condition of sucrose 12% treatment. The midguts of 10 female larvae per sample were collected from the sucrose 12% group and the control group, respectively. The total RNA of the sample was extracted by EASYspin Plus rapid tissue/cell RNA extraction kit of Beijing Aide-Lai Biotechnology Co., LTD. The RNA concentration, purity, and integrity number were assessed using the Nanodrop 2000 Bioanalyzer.

cDNA library construction and sequencing
The ribosomal RNA was depleted to isolate mRNA. The mRNA was then fragmented using divalent cations in NEB Fragmentation Buffer following the standard NEB protocol[14]. Utilizing the fragmented mRNA as a template and random oligo-nucleotides as primers, the first cDNA strand was synthesized in an M-MuLV reverse transcriptase system. Subsequently, the RNA strand was degraded by RNaseH, and the second cDNA strand was synthesized in a DNA polymerase I system using dNTPs as substrates. The purified double-stranded cDNA underwent end-repair, A-tailing, and ligation of sequencing adapters.

After the library construction was complete, initial quantification was performed with Qubit2.0 Fluorometer and the library was diluted to 1.5ng/ul. Then, the insert size of the library was tested using Agilent 2100 bioanalyzer. After library check, different libraries were pooled according to the requirements of effective concentration and target data volume for Illumina sequencing (NovaseqX Plus). The raw data have been submitted to the NCBI public database (NCBI.nlm.NIh.gov) with the accession number PRJNA1186106.

Reads processing and assembly
The transcriptome underwent sequencing utilizing the Illumina paired-end RNA-seq methodology. The mean insert size for the paired-end libraries amounted to 300 bp (± 50 bp). Prior to assembly, low-quality reads containing sequencing adaptors, primers, and nucleotides were removed using Cutadapt software (https://cutadapt.readthedocs.io/en/stable/)[15]. Under the Linux operating system, SEQKIT software was used to extract rRNA sequences, and HISAT2 software (https://daehwankimlab.github.io/hisat2/)[16] was used to construct rRNA index and comparison, and unmatched sequences were removed to reduce the proportion of rRNA in the

output. This process reduced deviations from RNA sequencing and assembly. Three vital parameters, specifically Q20 (representing the percentage of bases with a quality score of 20 or above), Q30, and GC content, were evaluated to assess the quality of all reads using the FastQC online tool[17]. The HISAT2 software was utilized to perform a rapid and accurate comparison between Clean Reads and the reference genome, ultimately yielding the positional information of the Reads on the reference genome[18].

Normalization of gene expression levels and differential expression analysis

The differential expression of genes (DEGs) required standardization of gene expression data from different samples to eliminate differences in sequencing depth and gene length between samples. Thus, gene expression was normalized through Fragments Per Kilobase of exon model per Million mapped fragments (FPKM)[19]. The *p*-value were accustomed following the Benjamini and Hochberg's method for regulating the false discovery rate (FDR)[20]. Genes were represented as DEGs when the adjusted *p*-value < 0.05 and |log2 (fold-change)| ≥ 1.

Functional annotation and enrichment of DEGs

In order to delve into the biological roles of DEGs, a comprehensive annotation process was undertaken utilizing several key bioinformatics databases. Specifically, the Clusters of Orthologous Groups (COG), Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases were employed for this purpose. The annotation was carried out using the eggNOG-mapper tool (http://eggnog-mapper.embl.de/) with the default parameters set to ensure a thorough analysis[21].

Following the annotation, a functional enrichment analysis was conducted to pinpoint the GO terms and metabolic pathways that are significantly overrepresented within the set of DEGs. To determine the significance of these enrichments, a statistical threshold was applied, with a *p*-value < 0.05 being considered significant compared to the entire transcriptome background.

Gene set enrichment analysis (GSEA)

To identify genes that may not show significant differences in overall expression level but have significant biological significance, we conducted Gene Set Enrichment Analysis (GSEA). Pathways with a *p*-value < 0.05, false discovery rate (FDR) < 0.25, and normalized enrichment score |NES| > 1 were selected for further analysis.

Expression validation of DEGs from RNA-Seq by qRT-PCR

The quantitative real-time PCR (qRT-PCR) was conducted to confirm the DEGs in the cDNA libraries, which was constructed from midgut of 5th instar larvae. *BmGAPDH* was selected for housekeeping gene. All the primers (Supplementary Table S1) were designed using primer 5.0 software (Premier Biosoft International, Palo Alto, CA) and DNAMAN (Lynnon BioSoft) according to the rigorous principle.

Data processing

Experimental data were statistically analyzed using Student's t-test. Statistical tests and drawing were conducted with Prism 8.0 (GraphPad Prism Inc., San Diego, CA, USA) software. Three PCR replicates of the target gene and housekeeping gene of each sample were performed to obtain the average value, and then the change amount of the target gene in each sample was calculated. The qPCR data using the $2^{-\Delta\Delta Ct}$ method was used to calculate the relative quantity of all the DGEs mRNAs expression in comparison with the reference gene[22], and the data was expressed as the mean ± standard error (SE) ($n = 3$).

Result and discussion

Survey on the weight changes of silkworm

We started rearing from the 5th instar molted silkworm (I5D1), dividing them into a C-Group and S-Group. The S-Group consisted of four different concentrations (6%, 12%, 18%, 24%), and before I5D4, the average body weight of the silkworm in the S-Group was generally higher than that of the C-Group. Moreover, the 18% S-Group had the highest body weight, followed by the 12% S-Group and then the 6% S-Group. The 24% S-Group exhibited the poorest body weight development among all S-Groups (Fig. 1A and Supplementary Table S2). After I5D5, the rate of weight gain in the S-Groups started to decrease. By the end of the 5th instar, all S-Groups had lower weights compared to the C-Group. The weight gain trends of the 12% S-Group and the 18% S-Group exhibited an upward trajectory prior to I5D2, whereas the C-Group, along with the 12% and 18% S-Groups, demonstrated a similar upward trend in weight gain prior to I5D3. Furthermore, all S-Groups experienced a faster rate of weight gain compared to the C-Group before I5D2. In summary, the 12% and 18% S-Groups appear to be the optimal concentrations for promoting developmental growth. It is worth noting that between I5D4-I5D6, the growth rate of 12% S-Group was slightly higher than that of 18% S-Group, and the 18% and 24% treatment groups even showed negative growth in I5D6 (Fig. 1B). And we analyzed the significance of differences between the various groups, and the results are presented in Fig. 1C. Overall, the growth situation of the 12% treatment group was relatively the better among all treatment groups.

In summary, adding carbohydrates during the initial (before I5D3) positively impacted silkworm growth, but may hinder development in later stages. The study of *Phormia regina* (*Meigen*) finds that the feeding duration increases with increasing sucrose concentration and then decreases with further increasing concentration, and the feeding duration is significantly longer in females than in males[23]. The volume uptake rate (nl/s) of sucrose solution by *Thymelicus lineola* is inversely proportional to the square root of viscosity caused by high concentration[24]. Kenneth suggests that initially, eating is initiated by chemical stimulation, but then it stops due to a combination of increased inhibitory feedback from stretch receptors that monitor the crop and intestinal or hemolymph volume[24]. We speculate that this phenomenon may be due to the disruption of carbohydrate metabolism in the gut due to increased glucose levels. Experiments with silkworm fed artificial feed and
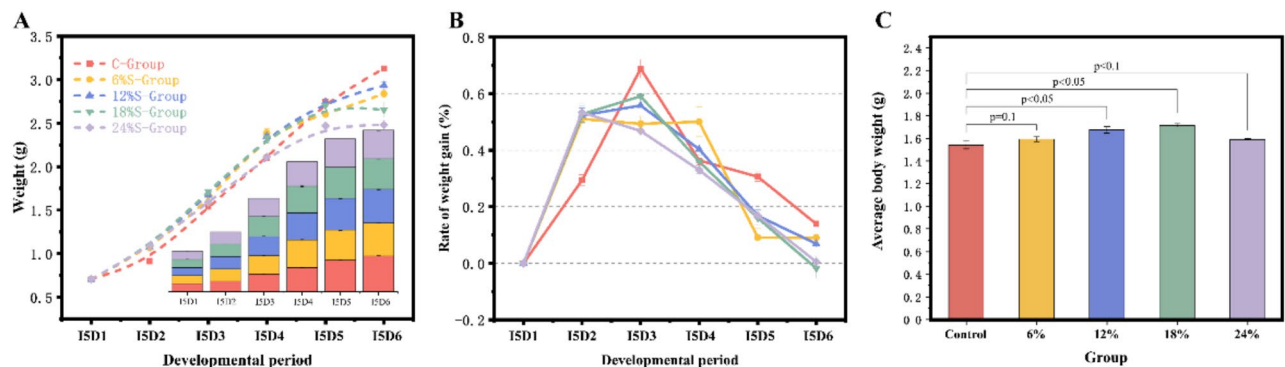
Fig. 1. The effect of sucrose addition on the body weight of silkworms. A: Statistical chart of average body weight of silkworms. B: Statistical chart of body weight growth rate of silkworms. C: Analysis of significant differences in body weight among groups of silkworms. $p \le 0.05$ is considered significant, and $p \le 0.01$ is regarded as extreme significant.

mulberry leaves reveal carbon metabolism as the most significant metabolic pathway for DEGs[25]. Notably, mulberry leaves contain approximately 8% sucrose, while artificial feed primarily consists of protein sources. Simple carbohydrates are rapidly broken down and absorbed in the midgut, unlike proteins which require longer degradation. Silkworm feeds mulberry leaves exhibit predominantly downregulated expression of carbohydrate metabolism genes compared to those fed artificial feed[25]. The hypothesis suggests that early increased sucrose intake leads to carbohydrate accumulation in the silkworm body, which is stored as glycogen and fat. By the mid-to-late stages, the accumulated carbohydrates nearly meet energy requirements, reducing mulberry leaf consumption and adversely affecting silkworm growth and development.

Transcriptome sequencing and assembly
The data processing results were shown in Supplementary Table S3. The quality scores of Q30 exceeded 97% in each sample. Moreover, the mapped ratios in S-Groups and C-Groups were 91.56%, 91.43%, and 91.50% and 89.99%, 90.58%, and 90.73%, respectively, which highly matched with the silkworm genome (Supplementary Table S4). In order to precisely evaluate the gene expression level, the clean reads of genes were normalized to FPKM values. The expression interval of genes was mainly focused on the 0~1 FPKM interval (FI) and the ratio was about 58.73% gene of the C-Group and 59.79% of S-Group (Supplementary Table S5). The Pearson correlation value between the three control samples was 0.958, while that of S-Group was 0.928, respectively; presenting the high repeatability. The results proved that the sequencing data quality was comparatively high.

WGCNA analysis
Weighted correlation network analysis (WGCNA) was performed to identify differential gene expression patterns between different subgroups. Notably, no outlier samples were detected based on average clustering (Fig. 2A). The soft threshold β was set at 9 to determine a scale-free network (Fig. 2B). Genes were assigned to 32 modules, and gray modules included genes that could not be clustered (Fig. 2C). Three gene modules highly correlated with S-Group (sky-blue, white, dark-red) and C-Group (purple, steel-blue, light-yellow) groups were identified. Further, COG, GO and KEGG analyses were performed to identify the potential biological significance of related modules in different subgroups.

Identification and analysis of DEGs
To ascertain the DEGs across various treatment groups, we employed the DESeq approach to compare the gene expression profiles of the C-Group and the S-Group. Utilizing the FPKM values of the genes, we calculated the differential expression levels between the S- and C-Groups, which revealed a substantial number of genes exhibiting altered expression following sucrose exposure. The distribution of these DEGs was visually represented in the form of a volcano plot, as depicted in Fig. 3 and the detailed information of DEGs was listed in Supplementary Table S6. A total of 909 DEGs were obtained between the S-Group and C-Groups. The numbers of up-regulated and down-regulated DEGs were 365 and 544, and the up-regulated genes were higher than down-regulated genes.

Functional annotation of DEGs
We utilized three different databases, namely COG, GO, and KEGG, to investigate the potential physiological functions of the differentially expressed genes (DEGs). Among the DEGs, 22 COG classes were identified (Fig. 4), with over 23% of the DEGs (184 genes) falling into the S category, indicating an unknown function. Additionally, several DEGs were found to be involved in other COG classes, such as O (posttranslational modification, protein turnover, chaperones) with 70 genes, G (carbohydrate transport and metabolism) with 65 genes, J (translation, ribosomal structure and biogenesis) with 64 genes, E (amino acid transport and metabolism) with 42 genes, T (signal transduction mechanisms) with 41 genes, and I (lipid transport and metabolism) with 339 genes. In this study, we selected the 'G' class functions associated with carbohydrate catabolism as our entry point. A total
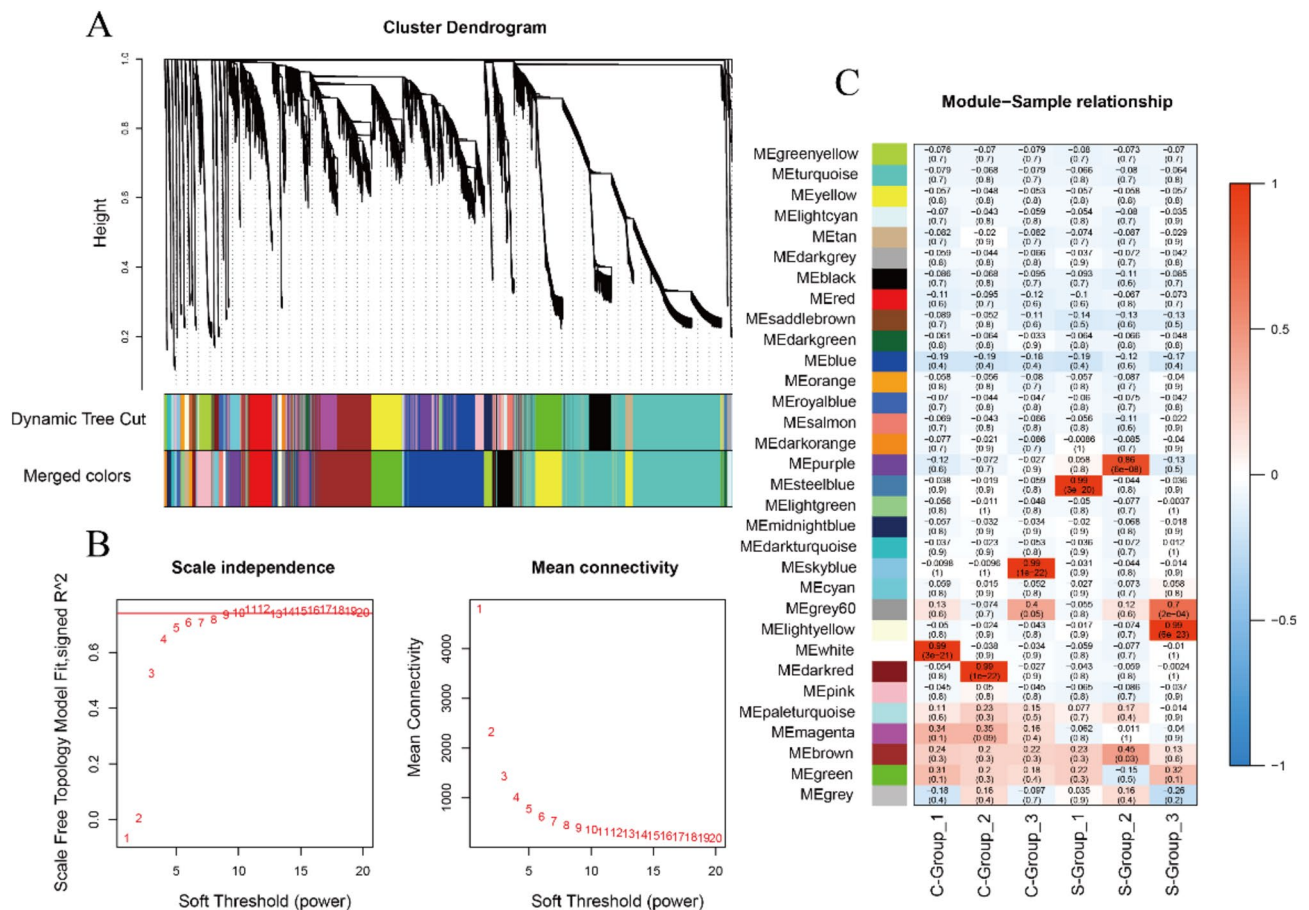
Fig. 2. WGCNA analysis. (A) Clustering dendrograms displayed gene network clustering in sucrose treated samples, with module diagrams obtained via Dynamic Tree Cut. Merged modules had a heterogeneity coefficient < 0.25. (B) Scale-free fit index and average connectivity were analyzed for various soft threshold powers ($\beta = 9$). (C) Eigenvector value correlations were assessed between 32 modules and subtype characteristics.

of 65 genes were primarily characterized as 'Alpha-*Amy*lase domain', 'Glycosyl hydrolase family', and 'UDP-glucuronosyl and UDP-glucosyl transferase'. Among these, genes exhibiting significant differential expression included LOC100379324 (alpha-*Amy*lase precursor, *Amy*), LOC101738628 (*BmSUH*), and LOC101746316 (UDP-glucose 4-epimerase, *GalE*), among others.

GO functional classification and enrichment analysis of DEGs

GO can be divided into three functional classes: molecular function, cellular component and biological process. In S-Group and C-Group, the DEGs were mainly distributed in the 'transmembrane transport' (91 DEGs), 'oxidation-reduction process' (46 DEGs), and 'carbohydrate metabolic process' (29 DEGs) of biological processes. The cellular components were 'cytoplasm' (49 DEGs), 'cytoplasmic parts' (42 DEGs), and 'ribonucleoprotein complex' (22 DEGs). In terms of molecular function, 'transporter activity' (62 DEGs), 'transmembrane transporter activity' (60 DEGs), and 'structural constituent of ribosome' (20 DEGs) were the main categories (Fig. 5A).

From the GO enrichment analysis results, the most significant 30 terms were selected to draw scatter plots for display, as shown in Fig. 5B. The molecular functions were mainly concentrated in transport activities, transmembrane transporter activities, ribosome structure components and so on. The cell components were mainly concentrated in cytoplasm, cytoplasmic parts and ribonucleoprotein complexes. The biological processes were mainly concentrated in the aspects of transmembrane transport, carbohydrate metabolism, oxidation-reduction process and peptide metabolism. It had been observed that numerous genes were concurrently enriched across multiple functional categories. Notably, the biological process of "carbohydrate metabolic process" was one that encompasses the majority of these genes. Furthermore, we had identified that the genes with the highest frequency of occurrence exhibit a striking similarity to the results obtained from COG annotations. Genes such as *Amy* and *BmSUH* had been particularly highlighted in our study. In biological systems, the transmembrane transport subclass constituted a predominant portion of DEGs. The efficacy of transport mechanisms within biological membranes is crucial for maintaining cellular function[26], governed by intricate interactions among membrane lipids, proteins, and carbohydrates. Among these mechanisms, glucose transporter proteins stood
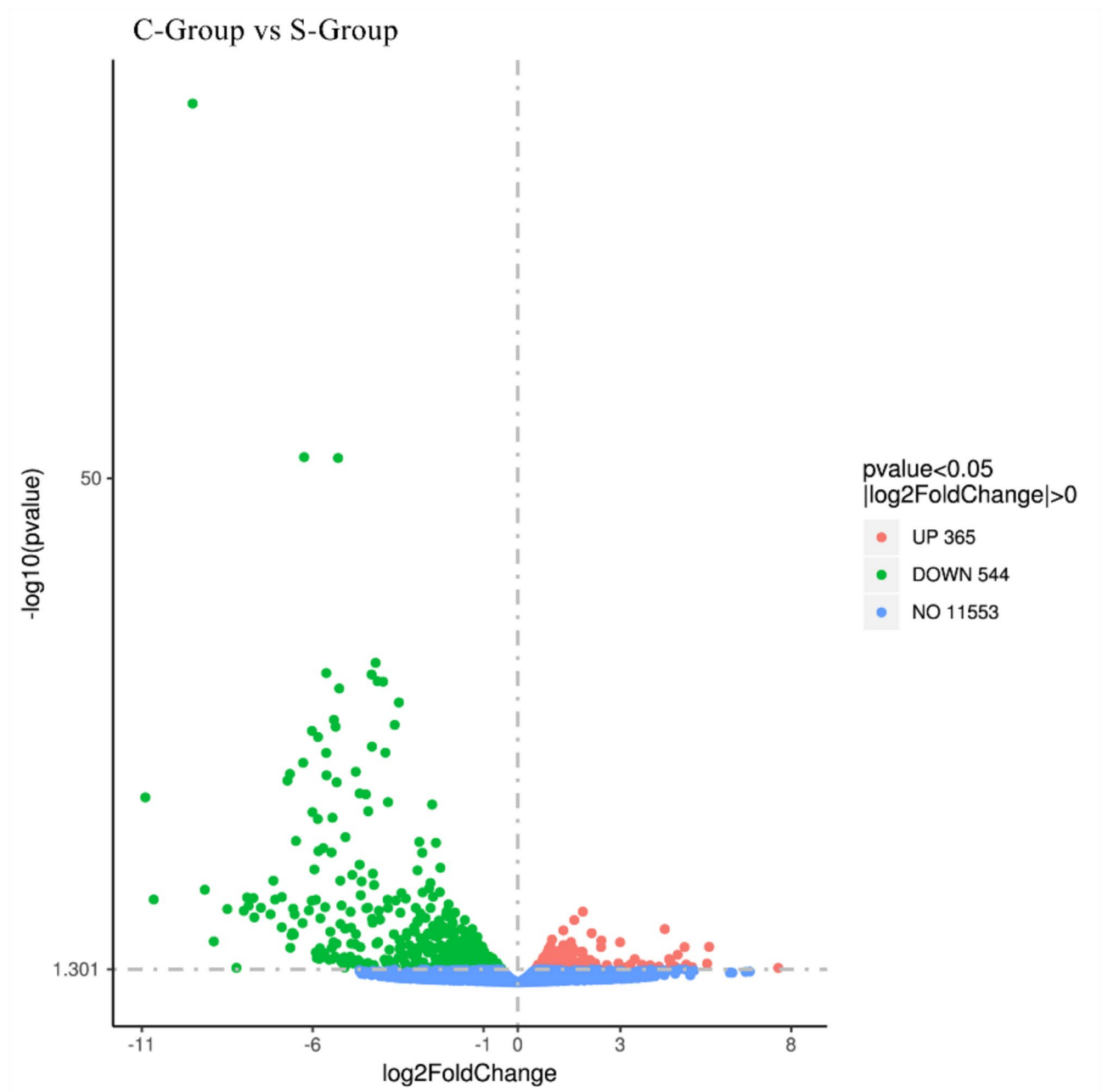
Fig. 3. The volcano plot of DEGs. the horizontal axis corresponded to the fold change in gene expression, while the vertical axis signified the statistical significance of these changes, typically measured by the adjusted *p*-value. The number of DEGs, normalized to a Log_2 fold change, was indicated on the right side of each plot, filtered by a *p*-adj threshold of less than 0.05. Genes that were significantly upregulated were denoted by red points, those downregulated were represented by green points, and genes that did not exhibit significant differential expression are depicted in blue.

out as a pivotal group. Glucose serves as a fundamental energy source for most cells, yet its polar nature and size preclude passive diffusion across cell membranes[27]. Consequently, glucose uptake relies on a family of transport proteins known as glucose transporters[27]. Notably, the categories of transporter protein activity and transmembrane transporter protein activity were enriched with the most prominently expressed genes in molecular function annotations, suggesting that dietary changes significantly affected glucose transport mechanisms, triggering a cascade of effects across various biological domains. We propose that excessive sucrose intake may modify the adipose tissue environment near the intestine during glucose transport to other organs, potentially disrupting cellular glucose transport mechanisms. This disruption may impair glucose translocation, leading to the downregulation of numerous genes associated with transmembrane transport.
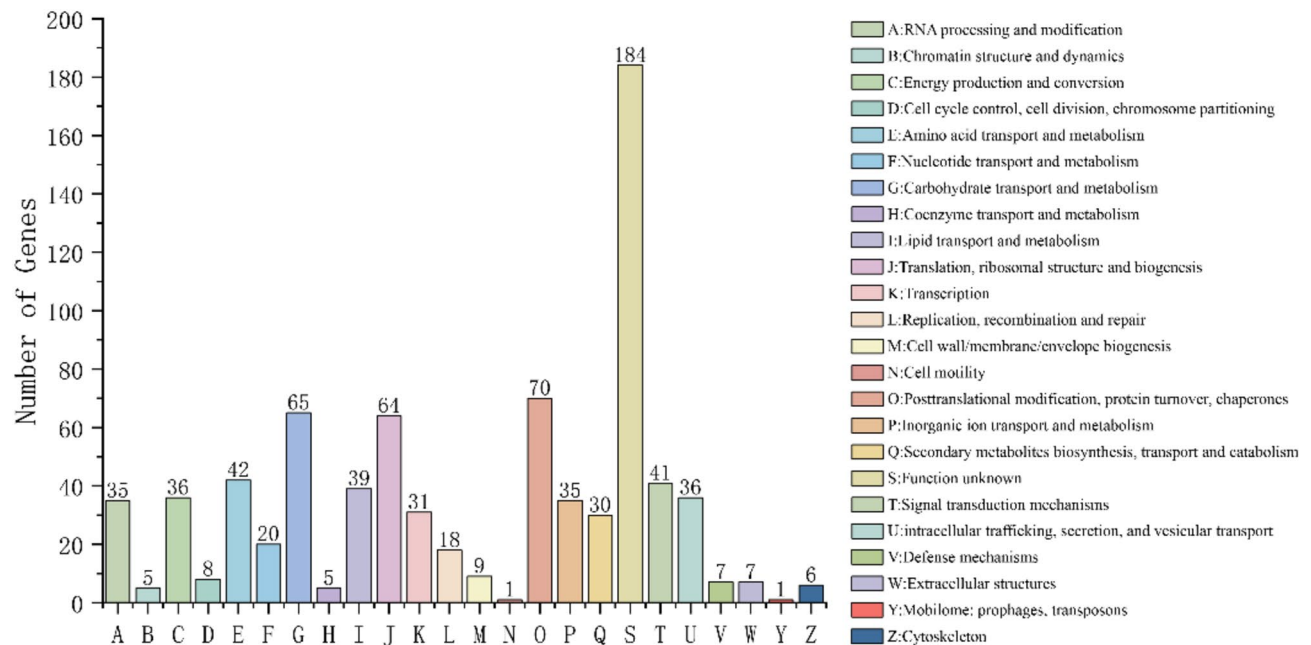
Fig. 4. COG analysis of DEGs.

KEGG functional classification and enrichment analysis of DEGs

Of the 909 DEGs expected to be encoded, 288 genes were mapped to 101 KEGG pathways. The most abundant metabolic category in metabolism was the biosynthesis of cofactors, followed by the mutual conversion of pentose and gluconate, carbon metabolism and drug metabolism (Fig. 6A). The DEGs were classified according to KEGG Pathway annotation. The results showed that DEGs were mainly concentrated in the pathways of ribosome, lysosome, pentose and gluconate interconversion, neuroactive ligand-receptor interaction, and ribosome biogenesis in eukaryotes (Fig. 6B). Through functional analysis, significant enrichment analysis was performed on the metabolic pathways of differentially DEGs, identifying a total of 24 significantly enriched metabolic pathways (Supplementary Table S7). The most prominent pathways include galactose metabolism, pentose and glucuronate interconversions, starch and sucrose metabolism, corresponding to monosaccharides, polysaccharides, and oligosaccharides, respectively. The galactose synthesis pathway describes the process of galactose conversion into sugar intermediates, which is crucial for the maintenance of cell membranes, signal transduction, and cell recognition[28]. Starch and sucrose are the primary dietary carbohydrate sources for herbivorous insects and are successfully degraded in the insect gut through the action of α-Amylase and maltase (α-glucosidase). The degradation of sucrose is typically mediated by α-glucosidase and β-fructosidase[29]. Recently, membrane-bound, sucrose-specific sucrases (*SUHs*) have been detected for the first time in the highly alkaline gut of lepidopteran larvae, which may have differentiated from α-glucosidases to enhance sucrose digestion, providing a major energy source for Lepidoptera[30].

The significant differences in these anabolic pathways could be due to the physiological changes caused by sucrose consumption in silkworm. The midgut of silkworm pH is alkaline (pH > 9), which is a critical regulatory factor for the activity of digestive enzymes. Goblet cells, specialized intestinal epithelial cells, mediate the alkalinization of the midgut lumen through the active transport of K^+ using $K^+/2H^+$ antiporters driven by a V-ATPase pump[31]. The accumulation of sucrose could alter the cellular environment in the lumen of the intestine, thereby affecting the ion balance in the gut, further affecting the activity of various digestive enzymes and inhibiting the absorption of nutrients. Before the I5D4 (the final instar), the weight change in silkworms induced by sucrose is superior to that of normal feeding, with I5D3 being the most significant. This is due to the fact that before the silkworms reach the boundary time, the larvae need to absorb a large amount of carbohydrates (such as sucrose) to complete the construction of their basic body composition. After absorption, the carbohydrates are stored by the fat body as a storage organ. After reaching the boundary time, the nutrients absorbed by the larvae are mainly used for the synthesis of silk protein, with the targeted main absorption substance shifting from a carbon source to a nitrogen source.

Gene set enrichment analysis (GSEA)

GSEA is an algorithm that performs differential expression analysis at the level of gene set[32]. And GSEA allows for the analysis of gene expression patterns within predefined gene sets to uncover relevant biological pathways or functions associated with specific experimental conditions (Fig. 7A). We found that UDP-galactose-4-epimerase (*GalE*) in the glucose anabolic pathway was the conserved gene that made the greatest contribution to the activation of the starch and sucrose anabolic pathways (Fig. 7A).

GalE was the only gene upregulated in the galactose anabolism pathway (Fig. 7B, C). In nearly all mammal, ingested lactose is hydrolysed by lactase, liberating β-D-galactose and D-glucose (Fig. 8)[33]. First, β-D-galactose
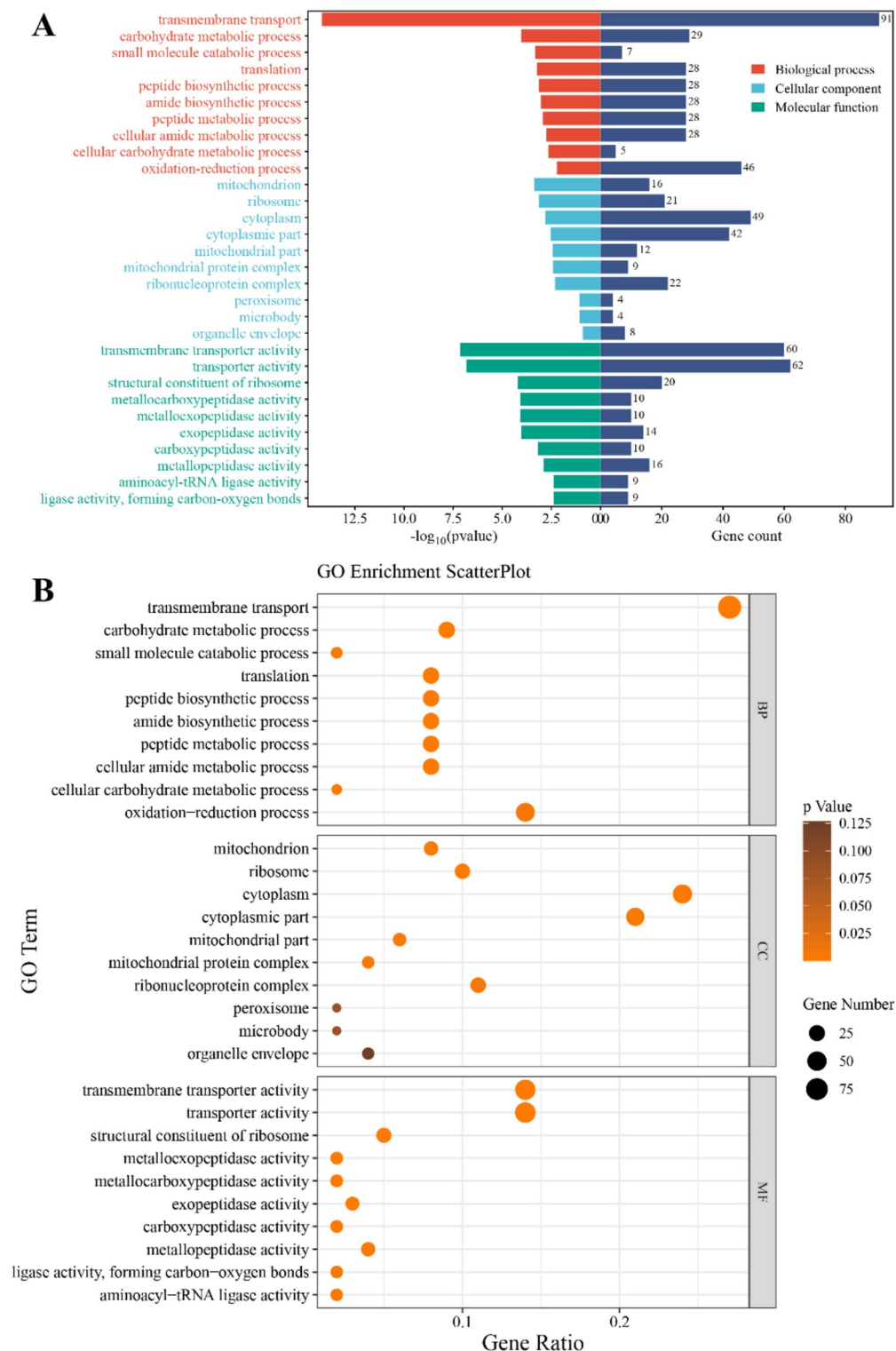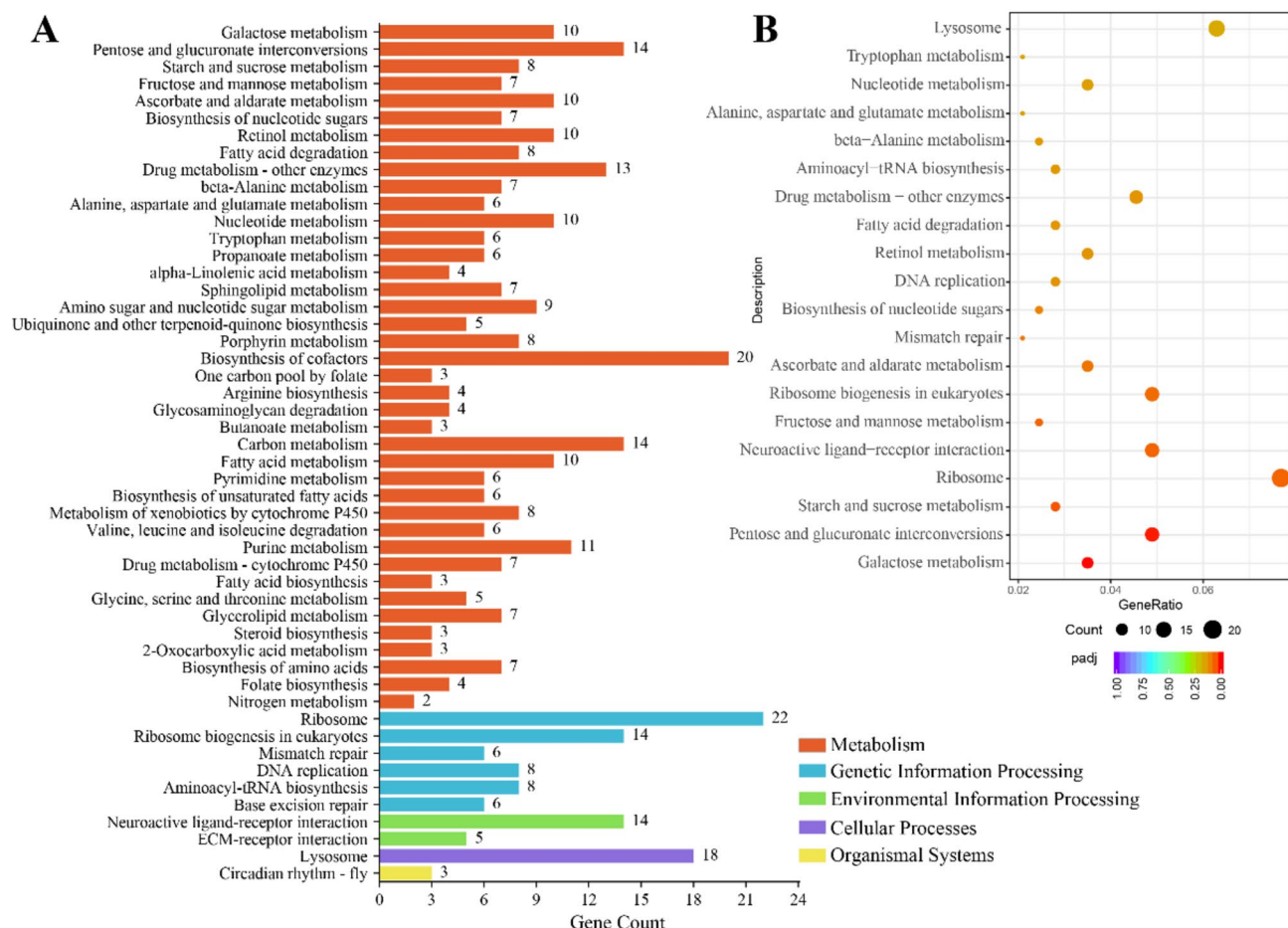
Fig. 5. GO annotation and enrichment analysis of DEGs. (A) GO annotation of DEGs. The horizontal coordinate represented the number of genes annotated to GO entries, the negative axis represented the significance of differences, and the vertical coordinate represented the GO functional classification. (B) GO enrichment analysis of DEGs. The horizontal coordinate was the ratio of the number of differential genes annotated to GO Term to the total number of differential genes; the vertical coordinate was GO Term; the size of the dots represented the number of genes annotated to GO Term; the color from dark to light represented the significance of enrichment.

Fig. 6. KEGG pathway annotation and enrichment analysis of DEGs. (A) KEGG pathway annotation of DEGs. The horizontal coordinate represented the number of genes labeled to the KEGG functional classification, and the vertical coordinate represented the KEGG functional classification. (B) KEGG enrichment analysis of DEGs. The horizontal coordinate was the ratio of the number of DEGs annotated to the KEGG functional classification to the total number of DEGs; the vertical coordinate was KEGG functional classification; the size of the dots represented the number of genes annotated to the KEGG functional classification; the color from red to purple represented the significance of enrichment.

is converted to α-D-galactose by galactose mutarotase (*GALM*). Galactose is then phosphorylated to form galactose-1-phosphate (*Gal-1P*) by galactokinase (*GALK*). Sanders and his colleagues establish a whole-animal model of *GalE* deficiency in Drosophila melanogaster using loss-of-function alleles[34]. Except for the embryonic forebrain ventricles, midgut, Malpighian tubules, and small intestine, the knockdown of *dGalE* in various tissues do not result in a clear phenotype[34]. This suggests that *GalE* plays a crucial role in these tissues and may be involved in glucose anabolism and growth in midgut tissues.

In addition, we identified hexokinase domain-containing 1 (*HKDC1*) as a direct target of TFEB within the glycolytic pathway. *HKDC1* is accountable for phosphorylating glucose in starch and sucrose catabolism processes, in addition to phosphorylating mannose in the metabolism of fructose and mannose[35]. *Maltase A1* is first identified in *D. melanogaster*[36] and probably involved in sugar digestion, subsequently, a homologous gene is also characterized in the silkworm and named *BmSUH*, it is classified as a glycoside hydrolase and, unlike traditional maltase, exhibits specificity towards sucrose[37].

Verification of DEGs with qRT-PCR

To validate the accuracy of DEGs identified from the transcriptome of the high-sucrose treatment group compared to the control group, From the KEGG analysis results, we selected 12 genes associated with four metabolic pathways related to sugar metabolism: starch and sucrose catabolism (*HKDC1*, *Amy and GS [starch]*) (Fig. 9A), fructose and mannose metabolism (*AKR1B1*, *PFKFB2 and GFUS*) (Fig. 9B), galactose catabolism (*BmSUH*, *GALM and GALT*) (Fig. 9C), and amino sugar and nucleotide sugar metabolism (*GlcNAcase2*, *NAL and GalE*) (Fig. 9D). Good consistency between the results of qRT-PCR and the transcriptome confirmed the accuracy and reliability of the sequencing data and revealed the significant difference of these genes during high-sucrose treated.
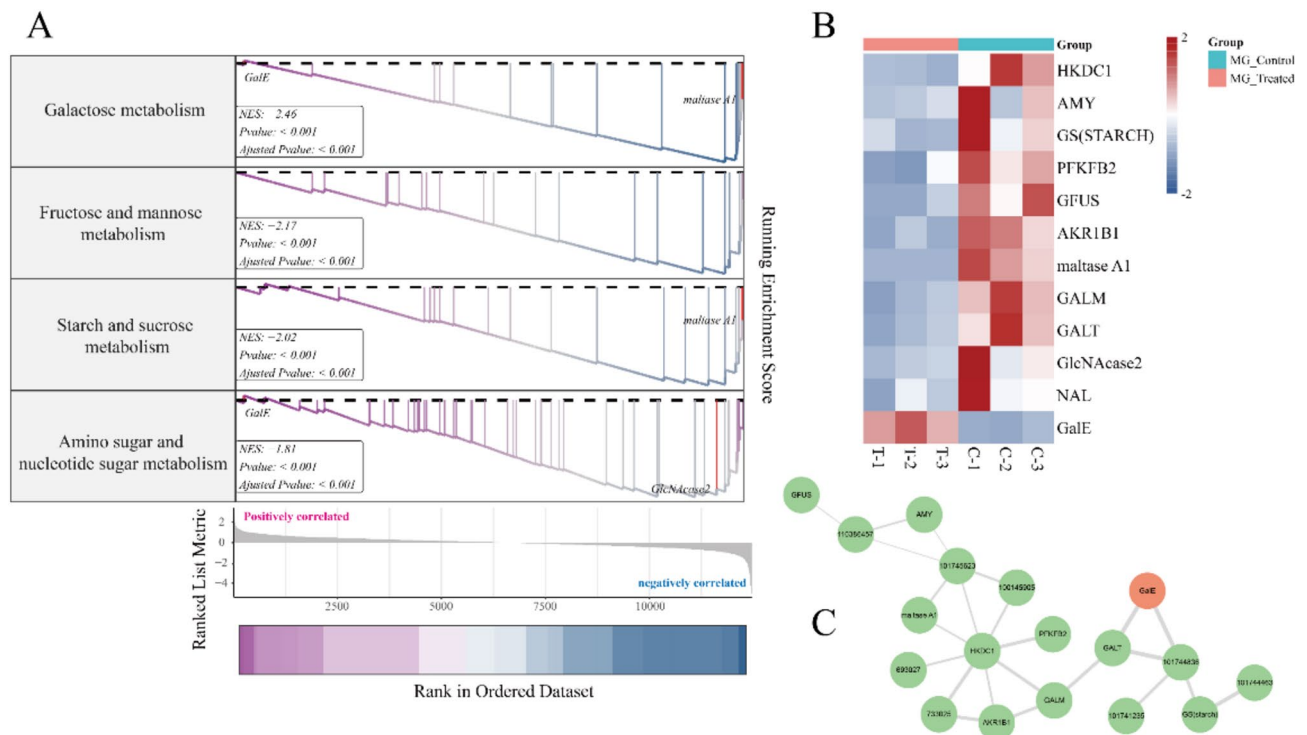
Fig. 7. GSEA, expressional pattern and protein-protein Interaction analysis of glucose catabolism pathways. (A) Utilizing RNA-seq data, conducted GSEA on glucose catabolism-related pathways and genes. (B) The heatmap depicted the expressional pattern of genes associated with glucose catabolism, with a mean score difference not exceeding 0.1. (C)Protein-Protein Interaction (PPI) analysis was executed on the DEGs within the glucose catabolism pathways.

Conclusions

In this study, we conducted a transcriptomic analysis revealing significant changes in carbohydrate metabolism pathways. Metabolic pathways of galactose, starch, and sucrose showed alterations, indicating suppressed synthesis and regulation of digestive enzymes. This suppression leaded to a detrimental cycle of sucrose accumulation, lipid accumulation, inflammation, and reduced lifespan.
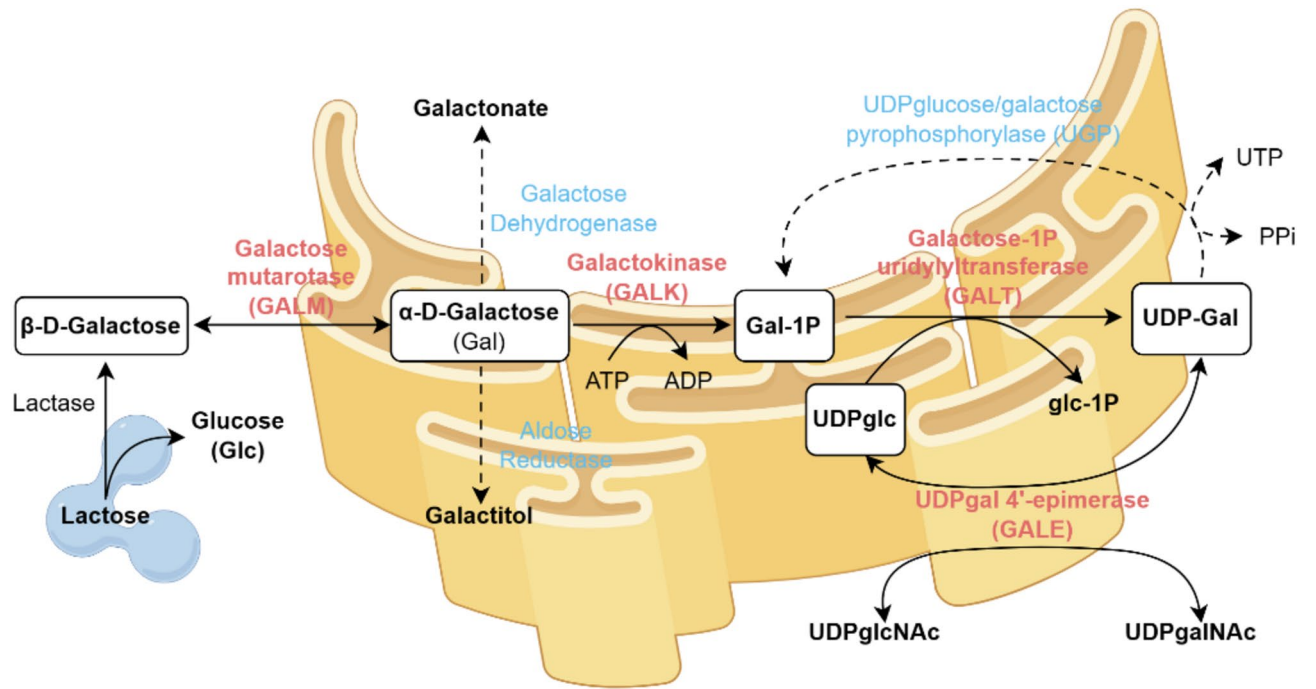
Fig. 8. Major reactions of galactose catabolism. The four enzymes of the Leloir Pathway (*GALM*, *GALK*, *GALT*, and *GALE*) are presented in red font. Enzymes that provide alternative or bypass routes of galactose catabolism are presented in blue font. The four UDP sugars listed are all key substrates for glycosylation. Inherited deficiencies of *GALM*, *GALK*, *GALT*, or *GALE* result in the primary disorders of galactosemia catabolism.
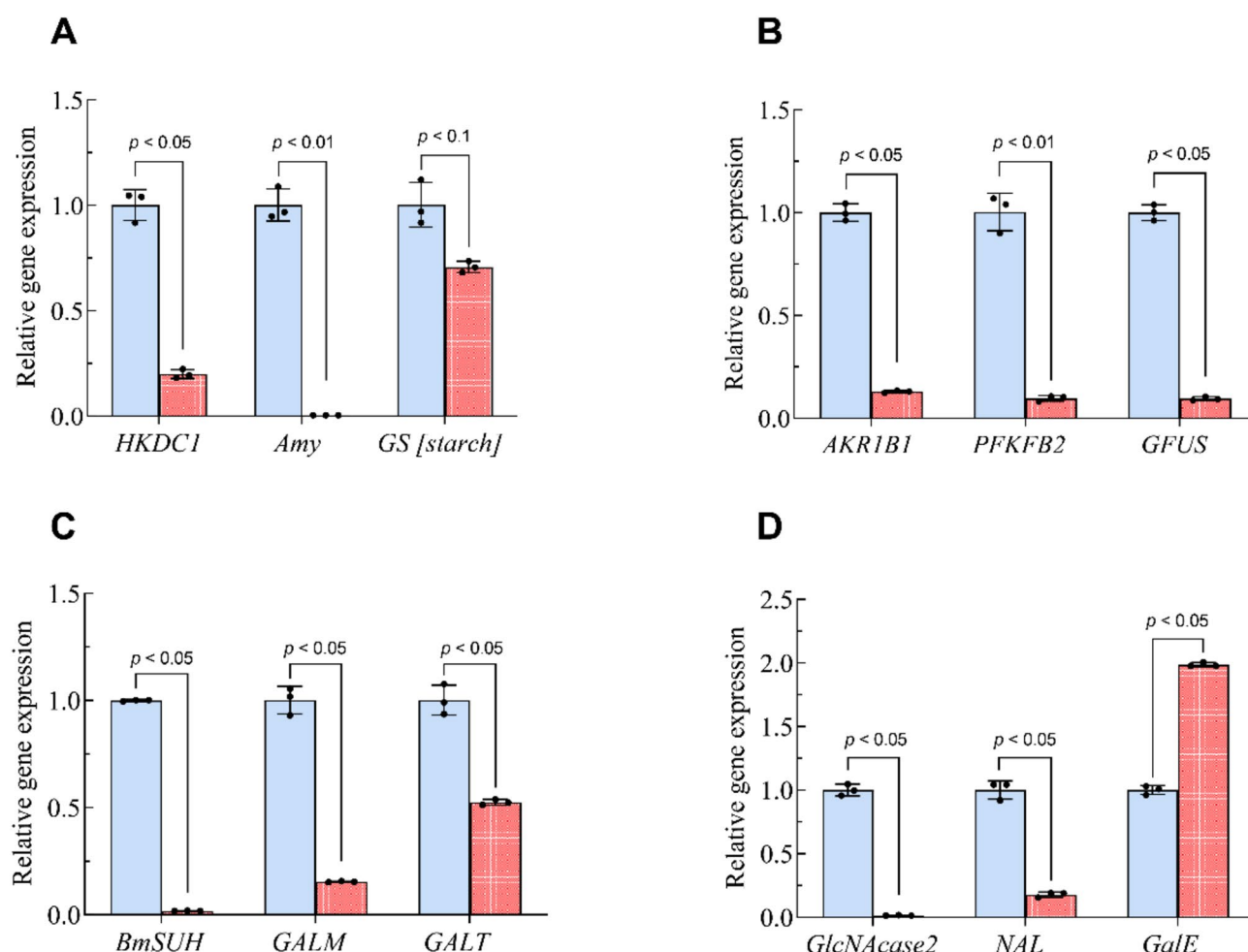
Fig. 9. qRT-PCR analysis of DEGs associated with carbohydrate metabolism. Data are presented as mean values normalized to *BmGAPDH* gene expression. Statistical significance was determined using Student's t-test. $p \leq 0.05$ is considered significant, and $p \leq 0.01$ is regarded as extreme significant.

Data availability

References

1. Xia, Q. et al. Complete resequencing of 40 genomes reveals domestication events and genes in silkworm (*Bombyx*). *Sci. (New York N. Y.)* **326**, 433–436. https://doi.org/10.1126/science.1176620 (2009).
2. Hamamura, Y., Hayashiya, K., Naito, K. I., Matsuura, K. & Nishida, J. U. N. Food selection by silkworm Larvæ. *Nature* **194**, 754–755. https://doi.org/10.1038/194754a0 (1962).
3. Liu, Y. et al. Transcriptome analysis reveals potential antioxidant defense mechanisms in *Antheraea pernyi* in response to zinc stress. *J. Agric. Food Chem.* **66**, 8132–8141. https://doi.org/10.1021/acs.jafc.8b01645 (2018).
4. Zhou, L. et al. Developmental changes for the hemolymph metabolome of silkworm (*Bombyx Mori* L). *J. Proteome Res.* **14**, 2331–2347. https://doi.org/10.1021/acs.jproteome.5b00159 (2015).
5. Xia, Q. et al. Microarray-based gene expression profiles in multiple tissues of the domesticated silkworm, *Bombyx mori*. *Genome Biol.* **8**, R162. https://doi.org/10.1186/gb-2007-8-8-r162 (2007).
6. Thounaojam, B. & Keshan, B. Modulation of gene expression by nutritional state and hormones in *Bombyx* larvae in relation to its growth period. *Gene Expr. Patterns*. **25–26**, 175–183. https://doi.org/10.1016/j.gep.2017.08.003 (2017).
7. Konala, N., Abburi, P., Bovilla, V. R. & Mamillapalli, A. The effect of bovine milk on the growth of *Bombyx mori*. *J. Insect Sci. (Online)*. **13**, 98. https://doi.org/10.1673/031.013.9801 (2013).
8. Hamamura, Y., Kuwata, K. & Masuda, H. Effect of Gallic acid on the growth of the silkworm larvae *Bombyx mori* L. *Nature* **212**, 1386–1387. https://doi.org/10.1038/2121386a0 (1966).
9. Horie, Y. & Nakasone, S. Effects of the levels of fatty acids and carbohydrates in a diet on the biosynthesis of fatty acids in larvae of the silkworm, *Bombyx mori*. *J. Insect. Physiol.* **17**, 1441–1450. https://doi.org/10.1016/0022-1910(71)90153-3 (1971).
10. Eleftherianos, I. et al. Nutrition influences immunity: diet and host-parasite interactions. *Insect Biochem. Mol. Biol.* **175**, 104210. https://doi.org/10.1016/j.ibmb.2024.104210 (2024).

11. Skowronek, P., Wójcik, Ł. & Strachecka, A. Body—multifunctional insect tissue. *Fat.* **12**, 547. https://doi.org/10.3390/insects1206 0547 (2021).
12. Derridj, S., Fiala, V. & Jolivet, E. Increase of European corn borer (*Ostrinia nubilalis*) oviposition induced by a treatment of maize plants with maleic Hydrazide: role of leaf carbohydrate content. *Entomol. Exp. Appl.* **41**, 305–310. https://doi.org/10.1111/j.1570-7 458.1986.tb00542.x (1986).
13. Shukla, E., Thorat, L. J., Nath, B. B. & Gaikwad, S. M. Insect Trehalase: physiological significance and potential applications. *Glycobiology* **25**, 357–367. https://doi.org/10.1093/glycob/cwu125 (2014).
14. Parkhomchuk, D. et al. Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Res.* **37**, e123–e123. https://doi.org/10.1093/nar/gkp596 (2009).
15. Martin, M. J. E. j. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12. https://doi. org/10.14806/ej.17.1.200 (2011).
16. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, stringtie and ballgown. *Nat. Protoc.* **11**, 1650–1667. https://doi.org/10.1038/nprot.2016.095 (2016).
17. Beekman, R. et al. The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. *Nat. Med.* **24**, 868–880. https://doi.org/10.1038/s41591-018-0028-4 (2018).
18. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods.* **5**, 621–628. https://doi.org/10.1038/nmeth.1226 (2008).
19. Meng, X. et al. Transcriptome analysis reveals global gene expression changes of Chilo suppressalis in response to sublethal dose of Chlorantraniliprole. *Chemosphere* **234**, 648–657. https://doi.org/10.1016/j.chemosphere.2019.06.129 (2019).
20. Yoav, B. & Daniel, Y. The control of the false discovery rate in multiple testing under dependency. *Annals Stat.* **29**, 1165–1188. https://doi.org/10.1214/aos/1013699998 (2001).
21. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* **38**, 5825–5829. https://doi.org/10.1093/ molbev/msab293 (2021).
22. Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**, e45. https://doi.or g/10.1093/nar/29.9.e45 (2001).
23. Rachman, N. J. Physiology of feeding preference patterns of female black blowflies (*Phormia regina Meigen*). *J. Comp. Physiol.* **139**, 59–66. https://doi.org/10.1007/BF00666195 (1980).
24. Pivnick, K. A. & McNeil, J. N. Effects of nectar concentration on butterfly feeding: measured feeding rates for *Thymelicus lineola* (Lepidoptera: Hesperiidae) and a general feeding model for adult Lepidoptera. *Oecologia* **66**, 226–237. https://doi.org/10.1007/BF 00379859 (1985).
25. Li, J., Chen, C. & Zha, X. Midgut and head transcriptomic analysis of silkworms reveals the physiological effects of artificial diets. *Insects* **13**, 3. https://doi.org/10.3390/insects13030291 (2022).
26. Kulbacka, J. et al. Cell membrane transport mechanisms: ion channels and electrical properties of cell membranes. *Adv. Anat. Embryol. Cell Biol.* **227**, 39–58. https://doi.org/10.1007/978-3-319-56895-9_3 (2017).
27. Navale, A. M. & Paranjape, A. N. Glucose transporters: physiological and pathological roles. *Biophys. Rev.* **8**, 5–9. https://doi.org/1 0.1007/s12551-015-0186-2 (2016).
28. Kim, S. et al. PubChem 2025 update. *Nucleic Acids Res.* **53**, D1516–D1525. https://doi.org/10.1093/nar/gkae1059 (2025).
29. Terra, W. R. & Ferreira, C. Insect digestive enzymes: properties, compartmentalization and function. *Comp. Biochem. Physiol. Part. B: Comp. Biochem.* **109**, 1–62. https://doi.org/10.1016/0305-0491(94)90141-4 (1994).
30. Li, X. et al. Molecular evolutionary mechanisms driving functional diversification of α-glucosidase in Lepidoptera. *Sci. Rep.* **7**, 45787. https://doi.org/10.1038/srep45787 (2017).
31. Flower, N. E. & Filshie, B. K. Goblet cell membrane differentiations in the midgut of a lepidopteran larva. *J. Cell Sci.* **20**, 357–375. https://doi.org/10.1242/jcs.20.2.357 (1976).
32. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550. https://doi.org/10.1073/pnas.0506580102 (2005).
33. Szilagyi, A. in *Book. Chapter 2 - Digestion, absorption, metabolism, and physiological effects of lactose* (eds. Marcel, P. & Cordula, L.) 49–111 (2019). https://doi.org/10.1016/B978-0-12-811720-0.00002-7.
34. Sanders, R. D., Sefton, J. M. I., Moberg, K. H. & Fridovich-Keil, J. L. UDP-galactose 4′ epimerase (GALE) is essential for development of *Drosophila melanogaster. Dis. Models Mech.* **3**, 628–638. https://doi.org/10.1242/dmm.005058 (2010).
35. Cui, M. *et al.* HKDC1, a target of *TFEB*, is essential to maintain both mitochondrial and lysosomal homeostasis, preventing cellular senescence. *Proc. Natil. Acad. Sci.* **121**, e2306454120. https://doi.org/10.1073/pnas.2306454120 (2024).
36. Miyazaki, T. & Park, E. Y. Structure–function analysis of silkworm sucrose hydrolase uncovers the mechanism of substrate specificity in GH13 subfamily 17 exo-α-glucosidases. *J. Biol. Chem.* **295**, 8784–8797. https://doi.org/10.1074/jbc.RA120.013595 (2020).
37. Wang, H., Kiuchi, T., Katsuma, S. & Shimada, T. A novel sucrose hydrolase from the bombycoid silkworms *Bombyx mori*, Trilocha varians, and Samia cynthia Ricini with a substrate specificity for sucrose. *Insect Biochem. Mol. Biol.* **61**, 46–52. https://doi.org/10.1 016/j.ibmb.2015.04.005 (2015).

Acknowledgements

Author contributions

P.-Z.X. and M.-R.Z. review & editing the main manuscript text, conceptualization, methodology, resources, supervision, project administration and funding acquisition.Z.-B.H., L.-J.Z., X.-Y.C. and J.L. writing the main manuscript text, validation, investigation, data curation and visualization.All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/1 0.1038/s41598-025-00715-x.

Correspondence and requests for materials should be addressed to P.X.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.