scientific data

Check for updates

OPEN

DATA DESCRIPTOR

A comprehensive cell atlas of fall armyworm (*Spodoptera frugiperda*) larval gut and fat body via snRNA-Seq

Chao Sun[1], Yongqi Shao[2]✉ & Junaid Iqbal[2]✉

The midgut and fat body of insects control key physiological processes, including growth, digestion, metabolism, and stress response. Single-nucleus RNA sequencing (snRNA-seq) is a promising way to reveal organ complexity at the cellular level, yet data for lepidopteran insects are lacking. We utilized snRNA-seq to assess cellular diversity in the midgut and fat body of *Spodoptera frugiperda*. Our study identified 20 distinct clusters in the midgut, including enterocytes, enteroendocrine, stem-like cells, and muscle cells, and 27 clusters in the fat body, including adipocytes, hemocytes, and epithelial cells. This dataset, containing all identified cell types in midgut and fat body, is valuable for characterizing the cellular composition of these organs and uncovering new cell-specific biomarkers. This cellular atlas enhances our understanding of cellular heterogeneity of fat and midgut, serving as a basis for future functional and comparative analyses. As the first snRNA-seq study on the midgut and fat body of *S. frugiperda*, it will also support future research, contribute to lepidopteran studies, and aid in developing targeted pest control strategies.

Background & Summary

The fall armyworm (FAW) *Spodoptera frugiperda*, an invasive lepidopteran pest, has a significant adverse impact on food production and is considered a global threat to food security. It is native to Americas and was first discovered in central and west Africa in 2016. Since then, it has invaded several Asian countries[1]. *S. frugiperda* first introduction in China occurred in 2019, and within a year, it rapidly spread throughout the entire country[2]. It is a highly successful pest due to its polyphagous and voracious feeding habits. It is capable of feeding on over 350 plant species, which poses a serious threat to economically valuable crops[3]. An additional contributing factor to its successful invasion is its resistance to pesticides. Notably, in China, this species has exhibited a resistance level 1000 times higher than normal to chlorpyrifos and two times higher to chlorantraniliprole[4]. One fundamental contributing factor in the evolution of cross-resistance is the production of detoxification enzymes. These enzymes play a crucial role in facilitating the transformation of toxic pesticides into non-toxic substances[5,6].

Insects rely on organs such as the midgut and fat body, to carry out vital functions, such as digestion, nutrient absorption, metabolism, and innate immunity. These two organs are particularly important due to their production of detoxification enzymes, antimicrobial peptides (AMPs), and reactive oxygen species (ROS). Additionally, they play a crucial role in regulating the resident microbiota to ensure homeostasis and prevent dysbiosis[7–10]. The insect midgut is the primary site for digestion and nutrient absorption. In addition, the midgut serves as the initial defense against various chemicals and pathogens that are introduced through feeding[11–13]. Typically, the insect's midgut is composed of four types of cells: intestinal stem cells (ISCs), enteroblasts (EB), enterocytes (EC, also called columnar cells), and enteroendocrine (EE) cells[14,15]. However, in lepidopteran insects, the midgut consists of enterocytes/columnar cells (ECs/CCs), enteroendocrine (EE) cells, intestinal stem cells (ISCs), and goblet cells[16]. Each cell type has a specific function. For example, ECs/CCs are responsible for absorbing nutrients and they also play a role in the immune system by secreting enzymes[17]. On the other hand, ISCs are crucial for maintaining the integrity of the midgut by replacing damaged cells, thereby ensuring a continuous immune

[1]Analysis Center of Agrobiology and Environmental Sciences, Zhejiang University, Hangzhou, 310058, China. [2]Institute of Sericulture and Apiculture, College of Animal Sciences, Zhejiang University, Hangzhou, 310058, China. ✉e-mail: yshao@zju.edu.cn; 0623055@zju.edu.cn

	Gut	Fat body
Number of reads	348,567,330	335,872,296
Valid UMIs	99.9%	99.9%
Q30 bases in UMI reads	95.8%	95.3%
Valid barcodes	98%	97.4%
Q30 bases in barcodes	96.6%	96.4%
Q30 bases in RNA reads	94.9%	94.1%
Reads mapped to genome	85.6%	83.5%
Reads mapped confidently to exonic regions	45.5%	42.8%
Reads mapped confidently to genome	67.2%	69.2%

Table 1. Detailed QC of FASTQ files. "Reads mapped confidently" refers to reads uniquely and accurately aligned, either to the genome or to exonic regions.

	Gut	Fat body
Estimated numbers of cells	12,112	7,290
Mean reads per cell	28,779	46,073
Fraction of reads in cells	55.0%	68.9%
Median genes per cell	863	1,206
Total Genes Detected	15,184	15,231
Median UMI counts per cell	1,666	2,874
Sequencing saturation	71.8	73.4

Table 2. Sequencing statistics based on cells.

defense. EE cells act as chemosensory cells, regulating responses to food, nutrients, and metabolites through the synthesis of neuropeptides and peptide hormones[18]. Goblet cells, working in coordination with EC, ensure the high pH of the gut and also produce mucins, which aid in protecting the intestinal lining[19]. On the other hand, the insect fat body functions similarly to the liver of mammals and is the main site for nutrient storage, detoxification, innate immunity, and metabolism[20,21]. It is a relatively large organ located between the integument and the visceral layer that surrounds the gut and reproductive organs. However, the exact cellular composition of the organ remains largely unknown[22]. Mostly, the cell types of the fat body have been identified through morphophysiological analysis and are classified into five types: trophocytes, urocytes, chromatocytes, mycetocytes, and hemoglobin cells[23–25]. These cell types vary across different insect species, but the most abundant and basic cells are known as adipocytes or trophocytes[23,24]. In this advanced era of modern technology, morphophysiological criteria alone are clearly not sufficient for in-depth studies of the function of such important tissues. Therefore, the use of advanced technologies such as single cell RNA sequencing (scRNA-seq) will help to understand the specific role of different cell types in an organ or tissue.

scRNA-seq is a powerful technique that provides insights into the heterogeneity and complexity of transcripts at the cellular level. It enables the identification and characterization of distinct cell types within specific organs or tissues, thereby shedding light on their composition and functions. Notably, there has been a remarkable rise in various scRNA-seq technologies[26–28]. Despite significant advancements, scientists continue to encounter challenges, particularly when studying soft tissues such as fat bodies. These tissues, abundant in fats, are delicate and pose difficulties in isolation due to their fragility, often resulting in damage[29,30]. It is noteworthy that snRNA-seq presents a promising approach for studying these tissues, offering enhanced accuracy in research outcomes[31]. These technologies have been successfully applied to multiple insect tissues, addressing numerous complex questions regarding cell type composition, function, and transcriptional heterogeneity[31]. Among the studied insects, the majority of the research has primarily concentrated on the model insects, while other insects have received relatively less attention[32]. Based on the current knowledge, scRNA-seq and snRNA-seq has been applied to investigate the midgut of *Drosophila*[33,34], silkworms[16], and mosquitoes[35,36]. However, only two studies have utilized these technologies to elucidate the cellular composition of the fat body in *Drosophila*[29] and silkworm[37]. Although *S. frugiperda* is a significant pest, snRNA-seq has not yet been employed to investigate the cellular composition of important organs or to understand how these cells respond to aging or various stressors.

In the current study, we utilized 10x genomics snRNA-seq to reveal the cellular composition of gut and fat body tissues in *S. frugiperda*. We generated snRNA-seq data containing 12,112 and 7,290 cells for gut and fat body respectively (Table 2). These were grouped into 20 and 27 clusters in the gut and fat body, respectively. Further midgut and fat body of *S. frugiperda* cell types were identified using published marker genes. Our primary concern was to identify and characterize the distinct cell types of both tissues. This data can be crucial for future studies aiming to understand the cellular mechanisms underlying insect development, physiology, and responses to environmental stressors. Our findings provide a detailed atlas of the specific cell types in the gut and fat body, paving the way for future research to explore how these cells develop, function, and respond to various environmental challenges. Gaining such knowledge will ultimately contribute to the effective control of *S. frugiperda*.

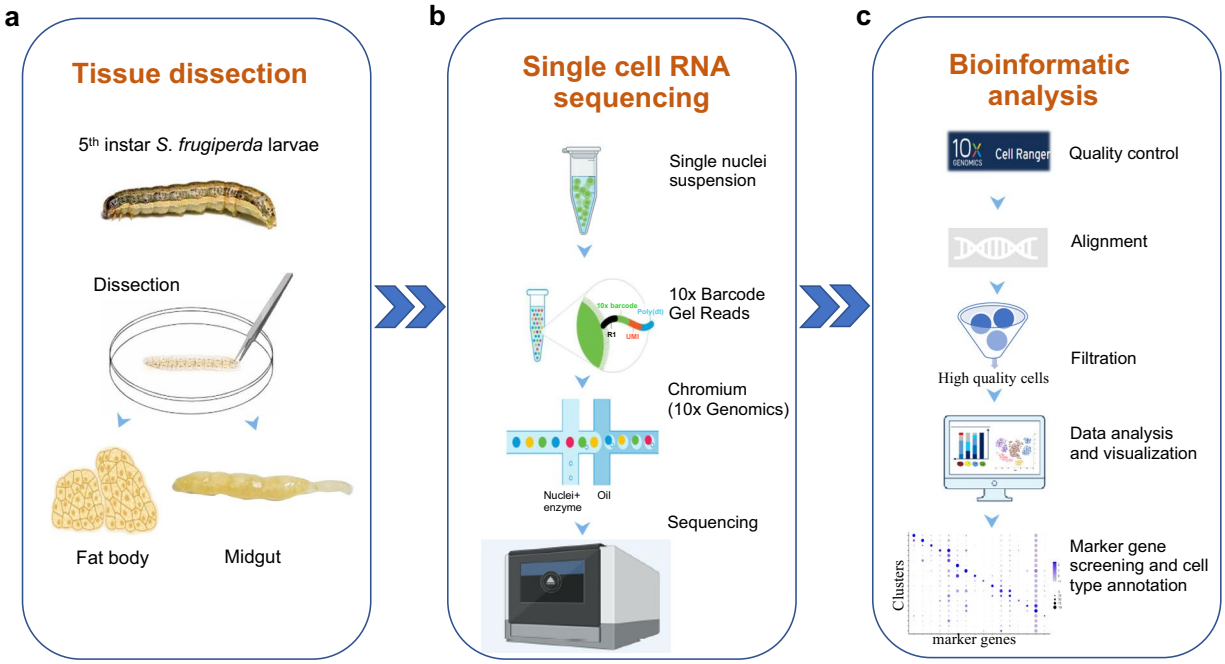

Fig. 1 Schematic illustration of experimental work flow. Starting from dissection of midgut and fat body of 5th instar larvae dissociating target cells from the tissue, cell isolation, RNA extraction, cDNA synthesis by reverse transcriptase, single-nucleus sequencing, expression profile, and cell-type identification.

Methods

Insects. The *S. frugiperda* larvae were fed an artificial diet and kept at a temperature of 25 ± 2 °C with 70% relative humidity in the laboratory. The photoperiod was set to 14 hours of light and 10 hours of darkness (14 L: 10D). Once the larvae reached the 5th instar, the midgut and fat body tissues were carefully dissected for subsequent single-cell sequencing.

Tissue dissociation and preparation of single-cell suspensions. Insect gut and fat body tissues were obtained from healthy 5th instar larvae. At this age, the midgut and fat body are fully mature. Tissue preparation and dissociation were conducted as a single experiment to maintain cell nuclei quality and avoid potential degradation caused by multiple preparations. Tissue was added to 1 mL of pre-chilled lysate, sheared into small pieces of approximately 1 mm, 2 mL of pre-chilled lysate was added, transferred to the Nuclei EZ Prep (NUC101; Sigma-Aldrich), and ground slowly 5 times using Kimble Dounce homogenizer (885300-007). The tissue was incubated on ice for 5 min, followed by the addition of 4 mL of pre-cooled 2% BSA (Bovine serum albumin) dilution, and the reaction was terminated by pipetting with a barrel pipette. Samples were passed through a 30 µm cell sieve, and the filtrate was collected and centrifuged at 300 g for 5 min at 4 °C. The supernatant was discarded. The cells were resuspended with 5 mL of 2% BSA, centrifuged at 4 °C for 5 min at 300 g. The supernatant was discarded and the cells were resuspended with 500 µL of resuspension solution (1x Phosphate-Buffered Saline [PBS], 2% BSA, and 0.1% RNase inhibitor). Stain with Propidium Iodide (PI) and sort to remove debris impurities using a BD FACSAria II flow cytometer. The overall cell viability confirmed to be above 85% through trypan blue exclusion. Cell concentration of single cell suspensions were determined with a CountStar Automated Cell Counter 600–1000 cells/µL.

Chromium 10x genomics library and sequencing. Single-cell suspensions were loaded onto the 10x Chromium system to capture individual cells, following the manufacturer's instructions of the 10x Genomics Chromium Single-Cell 3′ kit (V3). Subsequently, cDNA amplification and library construction steps were carried out according to the standard protocol. The libraries were then sequenced on an Illumina NovaSeq 6000 sequencing system (paired-end multiplexing run, 150 bp) by LC-Bio Technology Co. Ltd. in Hangzhou, China, ensuring a minimum sequencing depth of 20,000 reads per cell.

Barcode assignment, Unique Molecular Identifier (UMI) counting, and alignment. For initial data processing, we utilized the standard 10x Genomics Cell Ranger with default parameters for preliminary analyses. This involved conducting quality control of the FASTQ files, aligning the FASTQ files to the customized medaka reference using STAR26 (https://github.com/alexdobin/STAR), demultiplexing the cellular barcodes, and performing UMI counting. Tables 1, 2 provides a summary of the Cell Ranger quality control and output. ScRNA-seq data were aligned to NCBI genome GCF023101765 AGI-APGP_CSIRO_Sfru_2.0.

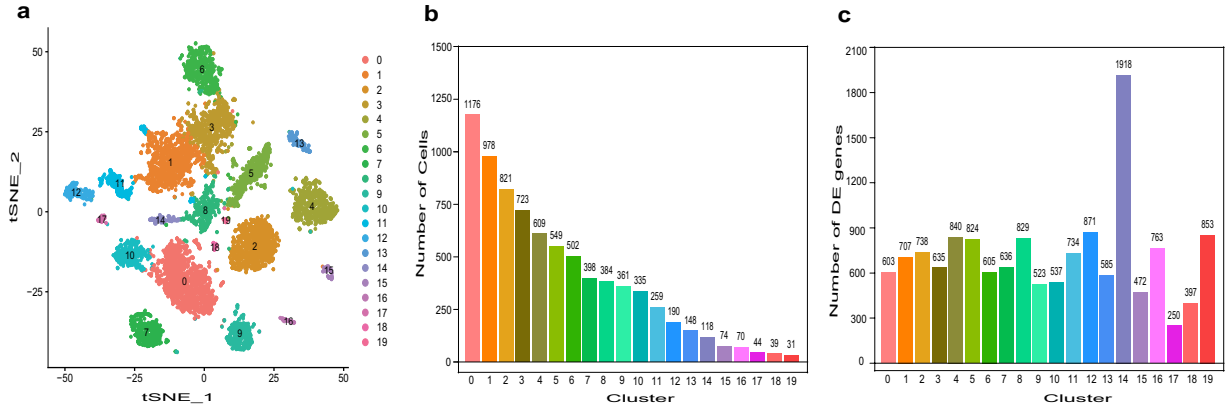

Fig. 2 Single-cell profiling of cell subsets in the midgut of *S. frugiperda* larvae. (**a**) t-Distributed Stochastic Neighbor Embedding (t-SNE) projection representing 20 distinct cell clusters, with different colors representing different clusters. (**b**) Bar graph depicting the number of cells per cluster, where the X-axis represents clusters and the Y-axis represents the number of cells in each cluster. (**c**) Bar graph showing the number of upregulated differentially expressed genes (DEGs) in each cluster, with the X-axis representing clusters and the Y-axis representing the number of DEGs in each cluster.

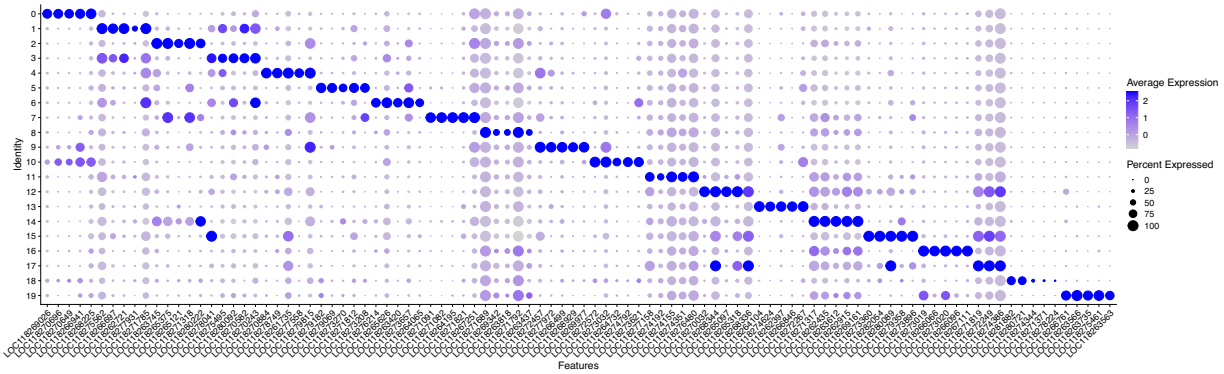

Fig. 3 Top 5 identified differentially expressed genes (DEGs) in each cluster of midgut. The X-axis represents the marker gene accession number, and the Y-axis shows the cluster numbers. The bubble size indicates the proportion of cells within a cluster expressing the marker gene, with larger bubbles representing higher proportions. The bubble color represents the mean expression level of the marker gene in the cluster, with darker bubbles indicating higher average expression levels.

Cell clustering. The Cell Ranger output was loaded into Seurat (version 4.1.0) for dimensional reduction, clustering, and analysis of scRNA-seq data. After expression quality control (QC), we employed a global-scaling normalization method called "LogNormalize" to normalize the gene expression measurements for each cell based on the total expression. To address the extensive technical noise in any single gene for scRNA-seq data, we randomly permuted a subset of the data (1% by default) and re-ran principal component analysis (PCA), constructing a 'null distribution' of gene scores. We repeated this procedure multiple times. We identified 'significant' PCs as those that showed a strong enrichment of low p-value genes. Subsequently, we used Seurat to cluster cells based on their PCA scores, with each PC essentially representing a 'metagene' that combines information across a correlated gene set.

To cluster the cells, we applied modularity optimization techniques, specifically SLM[38], to iteratively group cells together, with the goal of optimizing the standard modularity function. Seurat continues to use t-distributed Stochastic Neighbor Embedding (t-SNE) as a powerful tool to visualize and explore these datasets.

Differentially expressed genes analysis and marker genes analysis. We utilized the likelihood-ratio test (p-value ≤ 0.01, log2FC ≥ 0.26) to identify differential expression in a specific cluster, as compared to all other cells. The differentially expressed genes (DEGs) that were identified were then subjected to GO and KEGG pathway enrichment analysis.

Marker genes for each cluster were identified with the Wilcoxon rank-sum test with default parameters via the FindAllMarkers function in Seurat. This selects marker genes that are expressed in more than 10% of the cells in clusters and average log (Fold Change) of greater than 0.25 (default parameters: 0.26).

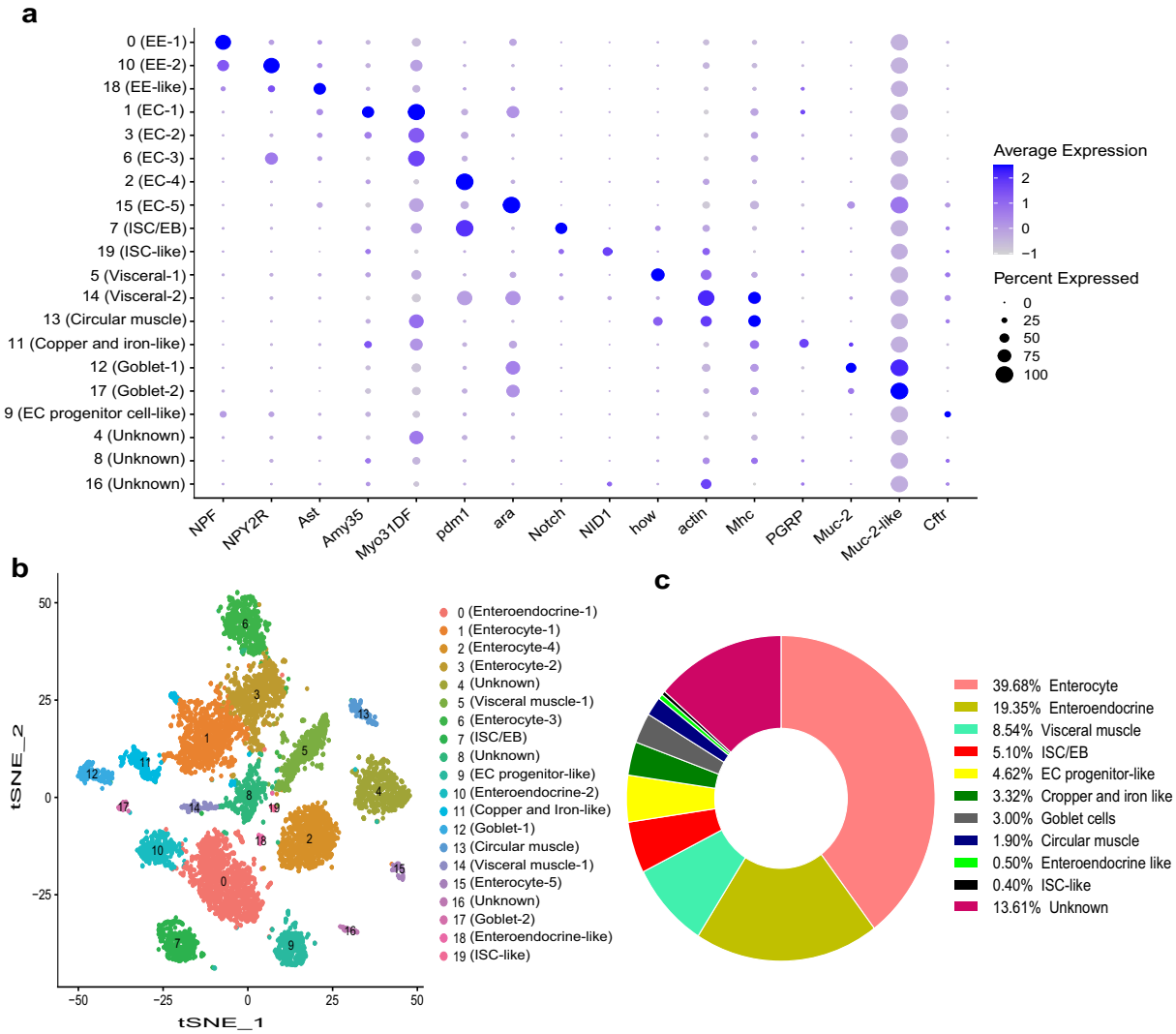

Fig. 4 Identification of specific cell types in midgut tissue using marker genes. (**a**) Dotplot showing the expression levels and the percentage of cells expressing markers. (**b**) t-SNE projection displaying the enteroendocrine cells, enterocytes, intestinal stem cells, muscle cell, goblet cells, visceral muscle and unclassified cells in the *S. frugiperda* midgut. (**c**) Cell number and proportion of each cell type.

Data Records

The snRNA-seq data used in this study can be accessed on NCBI through the BioProject accession numbers PRJNA1129320 and PRJNA1129321 for the midgut and fat body, respectively. The sample descriptions and raw sequencing data can be found under the BioSample accession numbers SAMN42150038 and SAMN42150039, and the Sequence Read Archive (SRA) accession numbers are SRP516946[39] and SRP518472[40] for midgut and fat body, respectively. Additionally, the processed data for the midgut and fat body, including barcodes.tsv, features.tsv, matrix.mtx, marker gene list, and a summary file (web_summary.html), have been uploaded to Figshare[41].

Technical Validation

Sequencing and quality assessment of the data. In this study, we utilized the 10x Genomics Chromium platform to construct snRNA-seq libraries for gut and fat body tissues (Fig. 1). Subsequently, these libraries were further sequenced using the Illumina NovaSeq 6000 sequencing system. Approximately 348.5 and 335.8 million reads were obtained for gut and fat body tissues in the snRNA-seq data respectively. The percentage of valid barcodes detected was 98% for gut tissue and 97.4% for fat body tissues. The distribution of barcodes and the relationship between reads per cell and median genes per cell of both gut and fat body (Fig. S1) indicate robust data quality. Moreover, reads mapped confidently to the genome were 85.6% in the gut and 83.5% in fat body tissue. The detailed sequencing quality of both tissues is listed in Table 1. The estimated numbers of cells acquired for the gut and fat body were 12,112 and 7,290 while after filtration the number of cells reduced to 8,420 and 5,892 in both tissues respectively (Table 2 & Table S1). The Median unique molecular identifier (UMI) counts per cell were 1,666 and 2,874 in the gut and fat body with median genes per cell detected were 863 and 1,206 respectively.

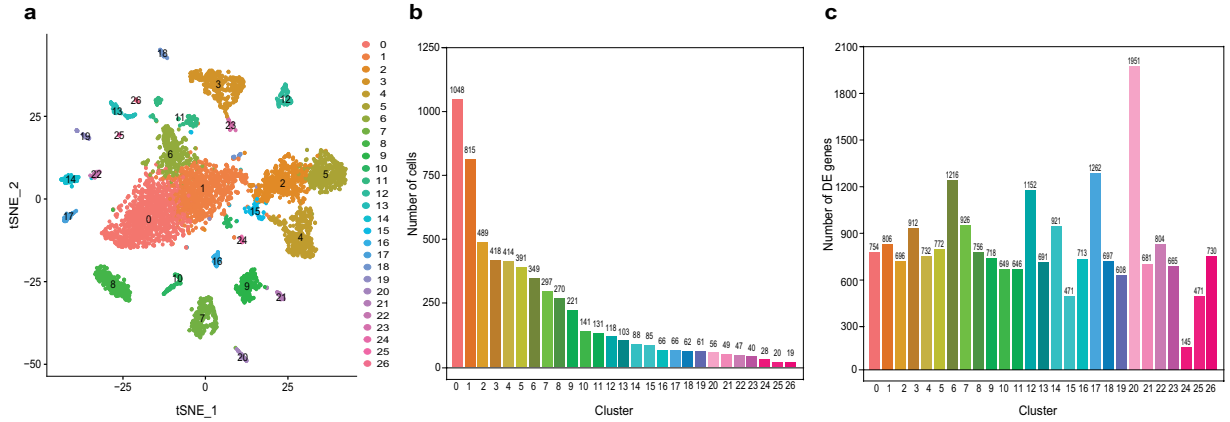

Fig. 5 Single-cell profiling of cell population in the fat body of *S. frugiperda* larvae. (**a**) t-Distributed Stochastic Neighbor Embedding (t-SNE) projection representing 27 distinct cell clusters, with different colors representing different clusters. (**b**) Bar graph depicting the number of cells per cluster, where the X-axis represents clusters and the Y-axis represents the number of cells in each cluster. (**c**) Bar graph showing the number of upregulated differentially expressed genes (DEGs) in each cluster, with the X-axis representing clusters and the Y-axis representing the number of DEGs in each cluster.

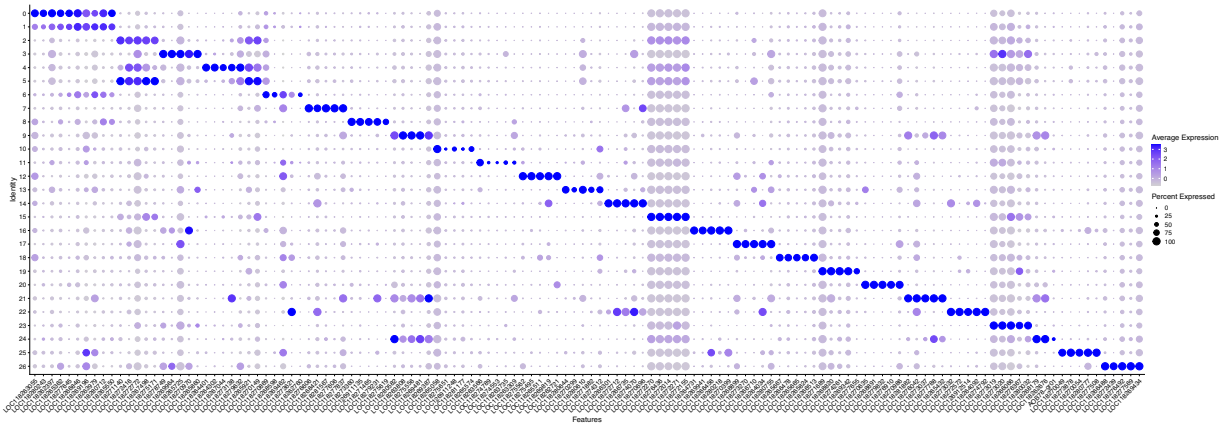

Fig. 6 Top 5 identified differentially expressed genes (DEGs) in each cluster of fat body. The X-axis represents the marker gene accession number, and the Y-axis shows the cluster numbers. The bubble size indicates the proportion of cells within a cluster expressing the marker gene, with larger bubbles representing higher proportions. The bubble color represents the mean expression level of the marker gene in the cluster, with darker bubbles indicating higher average expression levels.

Furthermore, the average number of reads per cell was 28,779 for the gut and 46,073 for the fat body tissues (Table 2). Overall, these results indicate that high-quality data were generated during sequencing.

Midgut cell type identification. Cells of midgut tissue were profiled together into 20 distinct cell clusters and were visualized using t-Distributed Stochastic Neighbor Embedding (*t*-SNE) (Fig. 2a). The total number of cells in each cluster is shown in the Fig. 2b, and the number of cells gradually decreases from cluster 0 to cluster 19. Cluster 0 to 3 combined account for approximately 50% of the total cells in the midgut. The number of DEGs per cluster is shown in Fig. 2c, cluster 14 recorded the highest number of DEGs (1,918) while the lowest recorded in cluster 17 (250). Furthermore, the expression of the top 5 DEGs in each cluster of midgut tissues is shown in the dotplot (Fig. 3). These DEGs need to be confirmed in future research and are proposed to be useful as new marker genes for the identification of specific cell types in *S. frugiperda* midgut.

As there are no recognized marker genes for *S. frugiperda* gut and fat body tissues, we utilized homologous genes to distinguish clusters based on the known marker genes of silkworm and *Drosophila*[16,29,34,37] as well as markers from CellMarker[42]. If the expression of a particular marker gene was $\leq$50% in a subset of cells, then those cells were annotated as a marker gene-like. Based on the lepidopteran midgut morphology, it composed of three main cell types: columnar cells (CCs or ECs), goblet cells (GCs), and stem cells[12]. However, EE have not been extensively studied. In the present study, Cluster 0 and 10 were identified as entero-endocrine cells (EEs) based on the expression of Neuropeptide F (*NPF*) (LOC118270696) and Neuropeptide Y receptor (*NPY2R*) (LOC118272588), while cluster 18 was assigned as EE-like since the percentage expression of Allatostatins (*Ast*)

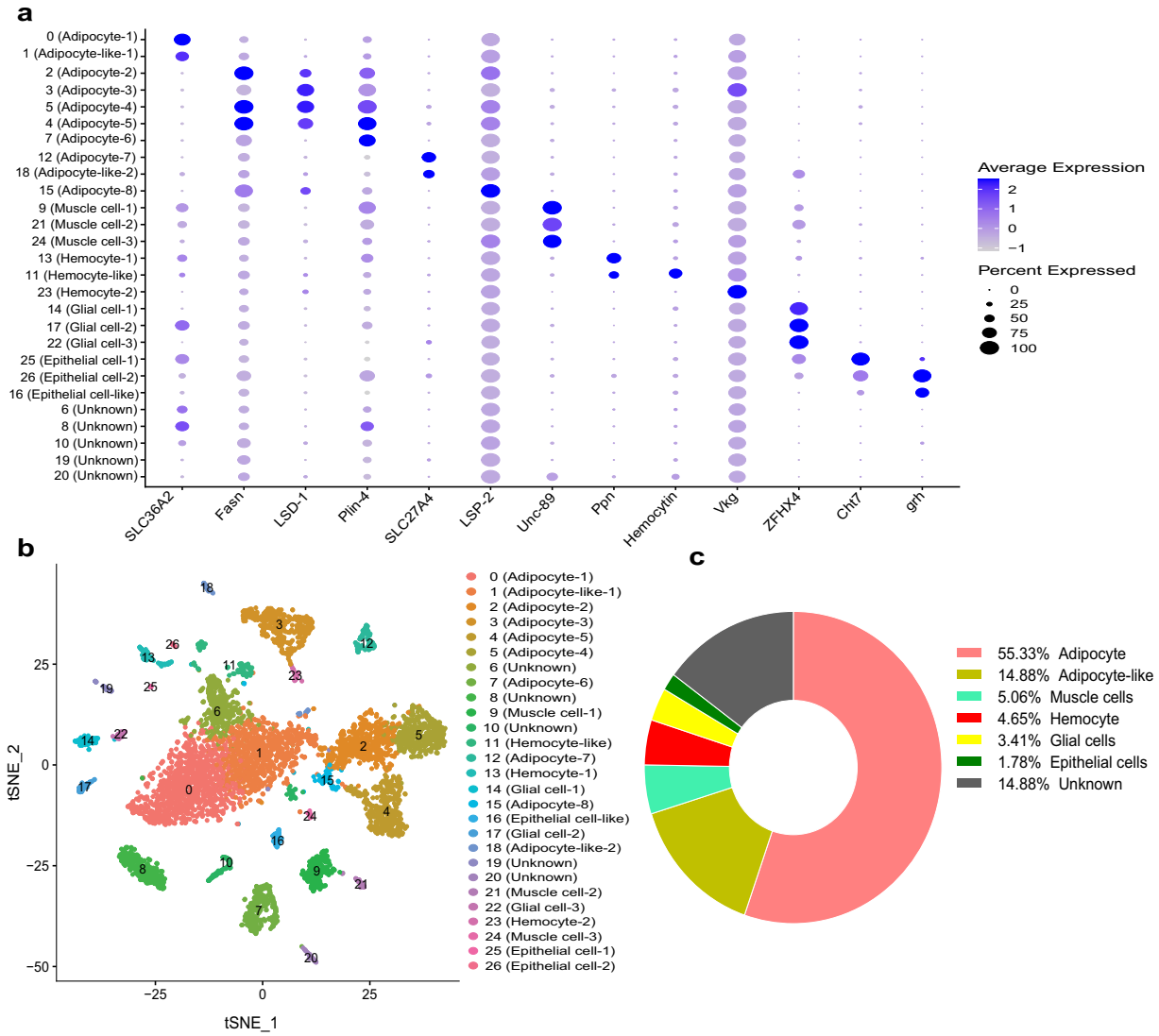

Fig. 7 Identification of specific cell types in fat body using marker genes. (**a**) Dotplot showing the expression levels and the percentage of cells expressing markers. (**b**) t-SNE projection displaying the adipocytes, muscle cells, hemocytes, glial cells, and epithelial cells and unclassified cells in the *S. frugiperda* fat body. (**c**) Cell number and proportion of each cell type.

(LOC118264343) was less than 50% (Fig. 4a). In our data, clusters 1, 2, 3, 6, and 15 are identified as enterocytes based on the expression of *Amy* (LOC118266145), *pdm1* (LOC118265375; also called nubbin), Myosin 31DF (*Myo31DF*) (LOC118269340) and *ara* (LOC118263101). Furthermore, cluster 7 was identified as intestinal stem cells/enteroblast (ISCs/EB) due to the expression of *notch* (LOC118262320), while cluster 19 was classified as ISCs-like based on the presence of DEG *Nidogen 1* (NID1) (LOC118265117). Clusters 5 and 14 were identified as visceral muscle due to high expression of the genes *how* (LOC118269879) and *actin* (LOC118264039) (Fig. 4a). We further classified cluster 13 as circular muscle based on the expression of *Mhc* (LOC118264857), and cluster 11 as copper and iron-like (LOC118274475). Cluster 12 and 17 were classified as goblet cells due to the expression of Muc2 (LOC118263608, LOC118274381). Cluster 9 was identified as EC progenitor cell-like (LOC118265427, Multidrug resistance). However, clusters 4, 8, and 16 cannot be identified using the potential markers listed in the Fig. 4a,b and Table S2. From the landscape of all midgut cell subsets, "enterocytes" represent the main cell type (39.68%), followed by enteroendocrine cells (19.35%) Fig. 3c. These findings are consistent with previous research on fruit fly and silkworm[16,34].

Fat body cell type identification. Research on the insect fat body has shown that it contains not only fat body cells but also neurons, epithelial cells, hemocytes, and other cell types[29]. However, due to technical limitations, specific cell types within fat body have not been further classified[37]. Similar to the gut, identification of cell types in the fat body was done using homologous genes from *Drosophila* and Silkworm, as discussed in the previous section on gut cell identification. The fat body cells of *S. frugiperda* were categorized into 27 distinct clusters (Fig. 5a). The total number of cells in each cluster is shown in Fig. 5b, and the number of cells gradually

decreases from cluster 0 to cluster 26 (Fig. 5b). The number of DEGs recorded per cluster revealed that the highest number of DEGs recorded in cluster 20 (1,951) (Fig. 5c). Furthermore, the expression of the top 5 DEGs in each cluster of fat body tissue is shown in the dotplot (Fig. 6). These DEGs need to be confirmed in future research and are proposed to be useful as new marker genes for the identification of specific cell types in *S. frugiperda* fat body.

In this study, we identified five major cell types in fat body *S. frugiperda* larvae, namely adipocytes, muscle cells, hemocytes, glial cells, and epithelial cells, using published marker genes. Based on the expression of marker genes SLC36A2 (LOC118268686, proton-coupled amino acid transporter-like protein), *Fasn* (LOC118272899), *LSD-1* (LOC118269804), *Plin-4* (LOC118268825), *SLC27A4* (LOC118277430, long-chain fatty acid transport protein (1), and *Lsp-2* (LOC118272270, basic juvenile hormone-suppressible protein (2), cluster 0, 2, 3, 4, 5, 7, 12, 15 were identified as adipocyte. Due to the expression of *SLC36A2* in cluster 1 and *SLA27A4* in cluster 18 being about 50%, these are classified as adipocyte-like (Fig. 7a). As mentioned previously, the fat body comprises not only adipocytes but also various other cell types. Based on the significant expression of *Unc-89* (LOC118272581, Obscurin) in clusters 9, 21, and 24, these clusters were designated as muscle cells. Based on the expression of *Ppn* (LOC118275461), *hemocytin* (LOC118266761), and *Vkg* (LOC118273019, Collagen type IV alpha 1), clusters 11, 13, and 23 were identified as hemocytes (Fig. 7a). Similarly, clusters 14, 17, and 22 were characterized as glial cells due to the high expression of *ZFHX4* (LOC118272755, zinc finger homeobox protein (3). Furthermore, clusters 16, 25, and 26 were identified as epithelial cells based on the expression of *Cht7* (LOC118264314) and *grh* (LOC118269607) (Fig. 7a). Unfortunately, there is not enough evidence to categorize clusters 6, 8, 10, 19, and 20 as identifiable cell types (Fig. 7a,b & Table S2). Consistent with previous studies[26,34], the most abundant cell type in the fat body is adipocytes, which comprise 55.33% of our dataset (Fig. 7c). Our analysis of the snRNA-seq data reveals the diverse cell populations within the midgut and fat body of *S. frugiperda*. This data serves as a foundational platform for the identification of cell types in these important tissues for future investigations, ultimately aiding in the effective control of this notorious pest.

Code availability

No special code was used for the analysis of the current dataset. All of the analyses and image generation were done with the following open-access programs and website tools: Cell ranger version 8.0.0 (https://github.com/10XGenomics/cellranger). Seurat (https://satjalab.org/seurat). SingleCell Data Analysis was performed using the OmicStudio tools created by LC-BIO Co., Ltd (Hangzhou, China) (https://www.omicstudio.cn/cell).

Received: 11 November 2024; Accepted: 21 January 2025;
Published online: 12 February 2025

References

1. Goergen, G., Kumar, P. L., Sankung, S. B., Togola, A. & Tamò, M. First Report of Outbreaks of the Fall Armyworm *Spodoptera frugiperda* (J E Smith) (Lepidoptera, Noctuidae), a New Alien Invasive Pest in West and Central Africa. *PLoS One* **11**, e0165632, https://doi.org/10.1371/journal.pone.0165632 (2016).
2. Yan, Z., Wu, Q.-L., Zhang, H.-W. & Wu, K.-M. Spread of invasive migratory pest *Spodoptera frugiperda* and management practices throughout China. *J Integr. Agric.* **20**, 637–645, https://doi.org/10.1016/s2095-3119(21)63621-3 (2021).
3. Abro, Z. *et al*. Socioeconomic and health impacts of fall armyworm in Ethiopia. *PLoS One* **16**, e0257736, https://doi.org/10.1371/journal.pone.0257736 (2021).
4. Zhang, D. D. *et al*. Insecticide resistance monitoring for the invasive populations of fall armyworm, *Spodoptera frugiperda* in China. *J. Integr. Agric.* **20**, 783–791, https://doi.org/10.1016/S2095-3119(20)63392-5 (2021).
5. Li, X., Schuler, M. A. & Berenbaum, M. R. Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. *Annu. Rev. Entomol.* **52**, 231–253, https://doi.org/10.1146/annurev.ento.51.110104.151104 (2007).
6. David, J. P., Ismail, H. M., Chandor-Proust, A. & Paine, M. J. Role of cytochrome P450s in insecticide resistance: impact on the control of mosquito-borne diseases and use of insecticides on Earth. *Philos Trans R Soc Lond B Biol Sci* **368**, 20120429, https://doi.org/10.1098/rstb.2012.0429 (2013).
7. Ryu, J. H., Ha, E. M. & Lee, W. J. Innate immunity and gut-microbe mutualism in *Drosophila*. *Dev Comp Immunol* **34**, 369–376, https://doi.org/10.1016/j.dci.2009.11.010 (2010).
8. Huang, J. H., Jing, X. & Douglas, A. E. The multi-tasking gut epithelium of insects. *Insect biochemistry and molecular biology* **67**, 15–20, https://doi.org/10.1016/j.ibmb.2015.05.004 (2015).
9. Shao, Y., Mason, C. J. & Felton, G. W. Toward an Integrated Understanding of the Lepidoptera Microbiome. *Annu Rev Entomol* **69**, 117–137, https://doi.org/10.1146/annurev-ento-020723-102548 (2023).
10. Chen, B. *et al*. Comparative shotgun metagenomic data of the silkworm *Bombyx mori* gut microbiome. *Sci Data* **5**, 180285, https://doi.org/10.1038/sdata.2018.285 (2018).
11. Denecke, S., Swevers, L., Douris, V. & Vontas, J. How do oral insecticidal compounds cross the insect midgut epithelium? *Insect Biochem. Mol. Biol* **103**, 22–35, https://doi.org/10.1016/j.ibmb.2018.10.005 (2018).
12. Wu, K. *et al*. Gut immunity in Lepidopteran insects. *Dev Comp Immunol* **64**, 65–74, https://doi.org/10.1016/j.dci.2016.02.010 (2016).
13. Haber, A. L. *et al*. A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339, https://doi.org/10.1038/nature24489 (2017).
14. Nászai, M., Carroll, L. R. & Cordero, J. B. Intestinal stem cell proliferation and epithelial homeostasis in the adult *Drosophila* midgut. *Insect Biochem. Mol. Biol* **67**, 9–14, https://doi.org/10.1016/j.ibmb.2015.05.016 (2015).
15. Hung, R.-J., Li, J. S. S., Liu, Y. & Perrimon, N. Defining cell types and lineage in the *Drosophila* midgut using single cell transcriptomics. *Curr. Opin. Insect. Sci* **47**, 12–17, https://doi.org/10.1016/j.cois.2021.02.008 (2021).
16. Xia, J. *et al*. Single-nucleus sequencing of silkworm larval midgut reveals the immune escape strategy of BmNPV in the midgut during the late stage of infection. *Insect Biochem. Mol. Biol* **164**, 104043, https://doi.org/10.1016/j.ibmb.2023.104043 (2024).
17. Caccia, S., Casartelli, M. & Tettamanti, G. The amazing complexity of insect midgut cells: types, peculiarities, and functions. *Cell Tissue Res* **377**, 505–525, https://doi.org/10.1007/s00441-019-03076-w (2019).
18. Tsakmaki, A., Fonseca Pedro, P., Pavlidis, P., Hayee, B. & Bewick, G. A. ISX-9 manipulates endocrine progenitor fate revealing conserved intestinal lineages in mouse and human organoids. *Mol Metab* **34**, 157–173, https://doi.org/10.1016/j.molmet.2020.01.012 (2020).
19. Awais, M. M., Fei, S., Xia, J., Feng, M. & Sun, J. Insights into midgut cell types and their crucial role in antiviral immunity in the lepidopteran model *Bombyx mori*. *Front. Immunol.* **15**, 1349428, https://doi.org/10.3389/fimmu.2024.1349428 (2024).
20. Gu, Z. Y. *et al*. Differentially expressed genes in the fat body of *Bombyx mori* in response to phoxim insecticide. *Pestic Biochem Physiol* **117**, 47–53, https://doi.org/10.1016/j.pestbp.2014.10.007 (2015).

21. Mao, T. *et al*. Effects of chlorantraniliprole exposure on detoxification enzyme activities and detoxification-related gene expression in the fat body of the silkworm, *Bombyx mori*. *Ecotoxicol Environ Saf* **176**, 58–63, https://doi.org/10.1016/j.ecoenv.2019.03.074 (2019).
22. Arrese, E. L. & Soulages, J. L. Insect fat body: energy, metabolism, and regulation. *Annu Rev Entomol* **55**, 207–225, https://doi.org/10.1146/annurev-ento-112408-085356 (2010).
23. Roma, G. C., Bueno, O. C. & Camargo-Mathias, M. I. Morpho-physiological analysis of the insect fat body: a review. *Micron* **41**, 395–401, https://doi.org/10.1016/j.micron.2009.12.007 (2010).
24. Turgay-İzzetoğlu, G. & Gülmez, M. Characterization of fat body cells at different developmental stages of *Culex pipiens*. *Acta Histochem* **121**, 460–471, https://doi.org/10.1016/j.acthis.2019.04.002 (2019).
25. Skowronek, P., Wójcik, Ł. & Strachecka, A. Fat Body-Multifunctional Insect Tissue. *Insects* **12**, https://doi.org/10.3390/insects12060547 (2021).
26. Shapiro, E., Biezuner, T. & Linnarsson, S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat. Rev. Genet.* **14**, 618–630, https://doi.org/10.1038/nrg3542 (2013).
27. Schwartzman, O. & Tanay, A. Single-cell epigenomics: techniques and emerging applications. *Nat. Rev. Genet.* **16**, 716–726, https://doi.org/10.1038/nrg3980 (2015).
28. Labib, M. & Kelley, S. O. Single-cell analysis targeting the proteome. *Nat. Rev. Chem.* **4**, 143–158, https://doi.org/10.1038/s41570-020-0162-7 (2020).
29. Li, H. *et al*. Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly. *Science* **375**, eabk2432, https://doi.org/10.1126/science.abk2432 (2022).
30. Gupta, V. & Lazzaro, B. P. A robust method to isolate *Drosophila* fat body nuclei for transcriptomic analysis. *Fly (Austin)* **16**, 62–67, https://doi.org/10.1080/19336934.2021.1978776 (2022).
31. Li, H. Single-cell RNA sequencing in *Drosophila*: Technologies and applications. *Wiley Interdiscip. Rev. Dev. Biol.* **10**, e396, https://doi.org/10.1002/wdev.396 (2021).
32. Sun, C., Shao, Y. & Iqbal, J. Insect Insights at the Single-Cell Level: Technologies and Applications. *Cells* **13**, https://doi.org/10.3390/cells13010091 (2023).
33. Guo, X. *et al*. The Cellular Diversity and Transcription Factor Code of *Drosophila* Enteroendocrine Cells. *Cell reports* **29**, 4172–4185 e4175, https://doi.org/10.1016/j.celrep.2019.11.048 (2019).
34. Hung, R. J. *et al*. A cell atlas of the adult *Drosophila* midgut. *PNAS* **117**, 1514–1523, https://doi.org/10.1073/pnas.1916820117 (2020).
35. Cui, Y. & Franz, A. W. E. Heterogeneity of midgut cells and their differential responses to blood meal ingestion by the mosquito, *Aedes aegypti*. *Insect Biochem. Mol. Biol.* **127**, 103496, https://doi.org/10.1016/j.ibmb.2020.103496 (2020).
36. Wang, S. *et al*. A cell atlas of the adult female *Aedes aegypti* midgut revealed by single-cell RNA sequencing. *Sci Data* **11**, 587, https://doi.org/10.1038/s41597-024-03432-8 (2024).
37. Feng, M. *et al*. Single-Nucleus Sequencing of Fat Body Reveals Distinct Metabolic and Immune Response Landscapes in Silkworm Larvae after *Bombyx mori* Nucleopolyhedrovirus Infection. *J Immunol* **211**, 140–153, https://doi.org/10.4049/jimmunol.2300007 (2023).
38. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech.: Theory Exp*. P10008, https://doi.org/10.1088/1742-5468/2008/10/P10008 (2008).
39. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRP516946 (2024).
40. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRP518274 (2024).
41. Iqbal, J. S. Yongqi, A. comprehensive cell atlas of fall armyworm (*Spodoptera frugiperda*) larval midgut and fat body using single-nucleus RNA sequencing. *figshare* https://doi.org/10.6084/m9.figshare.28089326.v2 (2024).
42. Hu, C. *et al*. CellMarker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. *Nucleic Acids Res* **51**, D870–d876, https://doi.org/10.1093/nar/gkac947 (2023).

Acknowledgements

This work was supported by grants from China Agriculture Research System of MOF and MARA (CARS-18), Zhejiang Provincial Natural Science Foundation of China (LZ22C170001), and the National Natural Science Foundation of China (32022081 and 31970483).

Author contributions

Work was planned by Y.S. and executed by C.S. and J.I. C.S. was associated with collection of the sample; J.I. worked on raw data analysis and the draft of article. Y.S. made final revisions to 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/10.1038/s41597-025-04520-z.

Correspondence and requests for materials should be addressed to Y.S. or J.I.