Title:

An 11-point time course midgut transcriptome across 72 h after blood feeding provides detailed temporal resolution of transcript expression in the arbovirus vector, *Aedes aegypti*

Running Title: Fine time course Aedes midgut transcriptome

Author:

Hitoshi Tsujimoto and Zach N. Adelman*

Department of Entomology and Texas A&M Agrilife Research, College Station, TX, USA

* Corresponding author: zachadel@tamu.edu

Zach N Adelman

329A Minnie Belle Heep Center

Dept of Entomology

370 Olsen Blvd

Texas A&M University

College Station, TX 77843

979-458-3107

Abstract:

As the major vector for dengue, Zika, yellow fever, and chikungunya viruses, the mosquito *Aedes aegypti* is one of the most important insects in public health. These viruses are transmitted by bloodfeeding, which is also necessary for the reproduction of the mosquito. Thus, the midgut plays an essential role in mosquito physiology as the center for bloodmeal digestion and as an organ that serves as the first line of defense against viruses. Despite its importance, transcriptomic dynamics with fine temporal resolution across the entire digestion cycle have not yet been reported. To fill this gap, we conducted a transcriptomic analysis of *Ae. aegypti* female midgut across a 72-h bloodmeal digestion cycle for 11 time points with a particular focus on the first 24 h. We obtained over one billion paired-end fragments with an average of 26 million fragments per sample. PCA analysis confirmed that 72 h is indeed a complete digestion cycle. Cluster and GO enrichment analysis showed the orchestrated modulation of thousands of genes to accomplish the midgut's role as the center for digestion as well as nutrient transport with a clear progression with sequential emphasis on transcription, translation, energy production, nutrient metabolism, transport, and finally autophagy by 24-36hr. We further determined that serine proteases likely to be critical for bloodmeal digestion are organized into 5 genomic clusters with multiple examples of co-expression. This study provides a powerful resource for the analysis of genomic features that coordinate the rapid and complex transcriptional program induced by mosquito bloodfeeding.

Keywords: Mosquito, midgut, transcriptome, time-course, bloodmeal, Aedes

Introduction:

Arthropod-borne (arbo-) viruses impose a substantial burden on public health (WHO 2009; WHO 2016; WHO 2018; WHO 2019). As the major vector of dengue, Zika, yellow fever and chikungunya viruses, the mosquito *Aedes aegypti* plays a leading role in arbovirus transmission. Transmission of the viruses occur during bloodfeeding, which is required for reproduction of the mosquito. An *Ae. aegypti* female may take ~2.5 mg of blood, which is ~219% of its body weight (Stobbart 1977) and produces 60-150 eggs per batch. Laboratory strains of *Ae. aegypti* perform this feat (from blood ingestion to maturation of ovaries) in ~72 h. In the meantime, blood-acquired pathogens can invade midgut epithelial cells despite the presence of physical and biochemical barriers, replicate, and disseminate throughout the body of the mosquito, including the salivary glands from which the pathogens will be injected into a new vertebrate host at the next bloodfeeding. Thus, the female midgut is one of the most important organs in a mosquito as the center for bloodmeal digestion, which is directly related to reproduction and the first line of defense against blood-acquired pathogens.

Therefore, a detailed understanding of midgut physiology and the underlying genetic basis for major aspects of digestion and reproduction is necessary to identify novel avenues of preventing pathogen transmission or of killing bloodfeeding mosquitoes. Among many methods, transcriptomic approaches provide a global view of expressed genes in a tissue or organism, which has been useful to determine previously unrecognized genes to exploit for novel control strategies. Previous studies of the *Ae. aegypti* midgut transcriptome demonstrated that the midgut undergoes very complex control of gene expression in response not only to a bloodmeal, but also to pathogens, which may be present in the bloodmeal (Bonizzoni et al. 2012b; Bonizzoni et al. 2012a). Recent studies have added spatial resolution to the female midgut transcriptome. Hixson *et al*. revealed that the subsections of midgut represent quite different transcriptome profiles (Hixson et al. 2022); and another study utilizing single-nuclei RNAseq uncovered that the midgut is comprised of at least 20 different cell types (Cui and Franz 2020). Moreover, Raquin *et al*. analyzed individual midgut RNAseq to determine host factors important for dengue virus (DENV) replication (Raquin et al. 2017), which added resolution at individual level.

A female *Ae. aegypti* undergoes dynamic physiological change in response to a bloodmeal. For instance, about 24% of the bloodmeal weight will be excreted as urine within 30 min after ingestion (Stobbart 1977). Formation of the peritrophic matrix (PM), an extracellular layer that forms around the bloodmeal to protect midgut epithelium and consists of chitin and associated proteins also appears to occur strikingly quick: a structure stained by azan (a stain for fibrous connective materials) began to form between the midgut epithelium and bloodmeal at 50 min post bloodmeal (PBM) (Freyvogel and Staeubli 1965), while expression of two major peritrophic matrix proteins (peritrophins) have been induced within 5 h after bloodmeal uptake (Shao et al. 2005; Devenport et al. 2006). Although the midgut is expected to play a major role in such physiological changes, studies to date have covered only a few time points and mostly relevant to virus infection cycle: 5 h in (Bonizzoni et al. 2011); 1, 4 and 14 days in (Bonizzoni et al. 2012a); 1 and 4 days in (Raquin et al. 2017); 6, 24, 48 h in (Hixson et al. 2022). One study analyzed a 72-h time course over 5 time points (0, 3, 12, 24, 72 hPBM) using the hybridization-based microarray technology (Sanders et al. 2003). However, the number of genes it could analyze (pre-determined 1778 genes) were far from modern sequencing-based technologies in quantity and quality, which makes a fair comparison unreasonable. Thus, the precise timing of changes in the midgut transcriptome over the course of bloodmeal digestion, especially within 24 h remains largely unexplored.

To fill the gap and to uncover the breadth of dynamic change in the transcriptome throughout the digestion cycle, we conducted a transcriptomic study for 11 time points across 72 h upon bloodfeeding. Illumina-based sequencing generated over 2 billion reads, establishing that the midgut undergoes rapid and dynamic change in transcriptome. In fact, we detected a strikingly high number genes that were differentially expressed within 24 hPBM. By analyses of serine proteases and iron/heme-related genes, we show that the expression patterns across the time course are extremely useful to group genes for downstream analysis and make inferences for potential roles in the bloodmeal digestion as well as promoter selection for transgene expression.

These data on the midgut transcriptome should serve as a platform for planning experiments on *Ae. aegypti* bloodmeal digestion and reproduction physiology at genetic and genomic viewpoints. Such studies will stimulate further advancement of *Ae. aegypti* research.

Results and discussion:

Morphology of midgut during the time-course:

We captured images of the midgut with associated organs (Malpighian tubules, hindgut and ovaries) (Fig 1, Fig S1). These images depict a midgut digestion cycle (pre-bloodmeal and from uptake of a bloodmeal to complete clearance of the midgut) that takes 48-72 h, supporting the idea that our time course should provide with a fairly complete view of the change in transcript expression throughout the digestion cycle.

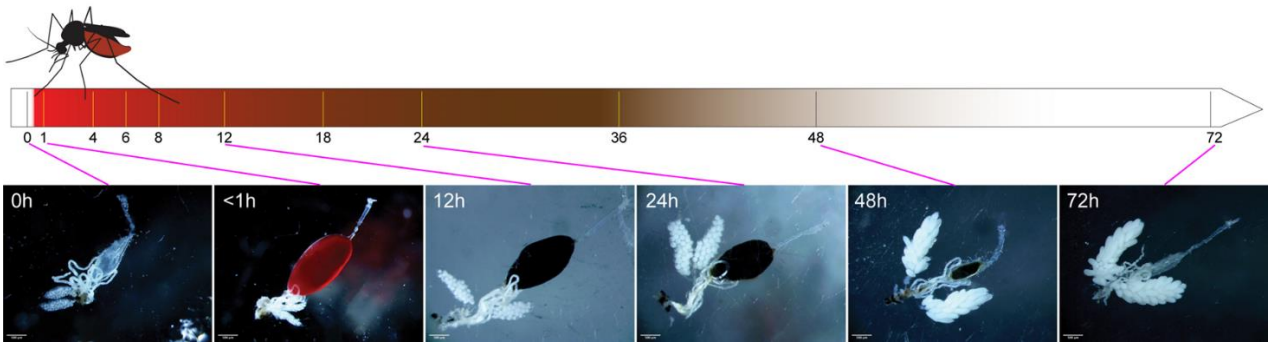

Fig 1: Timeline for the midgut transcriptome sample collection time points with images of the gut system and ovaries for selected time points. Images of all the time points are shown in Fig S1.

RNAseq output and mapping:

From our 4 independent replicates for the 11 time points of pre- and post-bloodmeal, Illumina NovaSeq 6000 paired end sequencing generated 1,160,607,602 fragment pairs (= 2,321,215,204 reads). We obtained a mean of 26,377,445.5 fragments per sample and 105,509,782 fragments per time point. We mapped these fragments using HISAT2 onto the AaegL5 AGWG genome assembly with 78.48% average mapping rate (Fig S3).

Through quantification of expression by StringTie, we detected 16,535 genes were considered expressed in the midgut, of which 13,162 were protein coding genes (89.4% of total in the genome), 3,134 non-coding RNA genes (66.6%), and 260 pseudogenes (68.0%) from the current gene annotation (AaegL5.3).

Principal component analysis (PCA):

As a process of validation for consistency of the replicates and changes in gene expression between time points, we performed principal component analysis (PCA) (Fig 2). PCA showed that the four replicates at all the timepoints were reasonably clustered, suggesting consistency between experiments. The expression changes across the time points showed a circular trace. Surprisingly, 392 genes were differentially expressed even within <1 hPBM compared to the pre-bloodfed control (Table S2). We term it as <1 hPBM for simplicity, but it was in fact around 15-20 min after bloodfeeding (see Materials and Methods), which suggests the response of the female midgut to a bloodmeal is strikingly rapid. The difference between <1 and 4 hPBM appeared to show the largest difference between time points (5115 genes, the highest number between neighboring time points, Table S2), which also suggests a rapid change in transcription in the first 4 h. This suggests that additional time points between <1 and 4 hPBM are warranted, and may fill a useful knowledge gap in understanding midgut digestion physiology. The timepoints that showed the greatest difference from the pre-bloodmeal state (0 hPBM) were 18-24 hPBM, and the expression profile turned to near 0 hPBM at 72 hPBM (Fig 2). The difference between 0 and 72 hPBM (1421 genes differentially expressed) may be due to the state of ovaries. Since we used gravid females, the midgut may require signals from non-gravid ovaries to revert to the pre-bloodmeal state. Thus, the PCA suggests that 72 h is indeed a complete (or close to complete) cycle of blood-digestion in the midgut. Our results differ from the recent results from Hixson et al. (Hixson et al. 2022), which showed 48 h as a cycle. This discrepancy might be because this study used different strain of *Ae. aegypti* ("Liverpool" vs Thai strain), source of blood (citrated sheep blood vs live chicken), temperature (28 vs

29 °C), light cycle (14:10 vs unknown), and number of time points. Moreover, the PCA between the two studies might have used different components as principal factors.

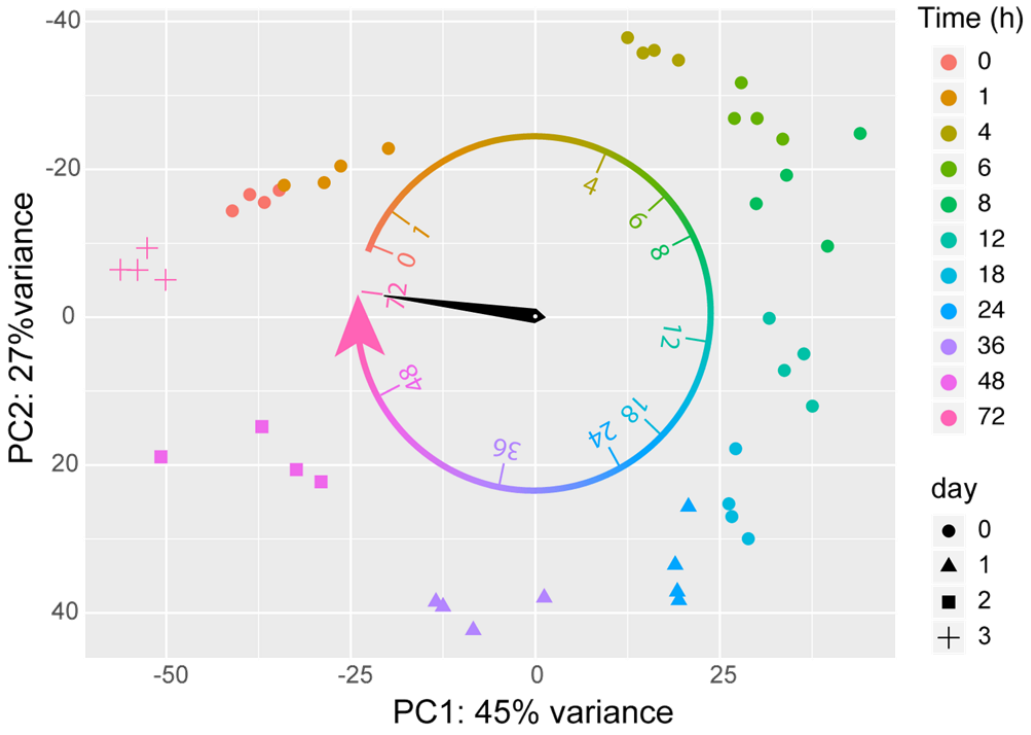

Fig2: Principal component analysis (PCA) of the libraries. Time points (hPBM) are shown by the color scheme and the point shapes. The circular arrow with a clock hand indicates a 72-h cycle.

Time-course analysis and Mfuzz soft clustering:

Through edgeR time-course analysis, 8594 genes showed significant change at FDR < 0.05. These genes were clustered by Mfuzz (Futschik and Carlisle 2005) into 20 clusters, which were manually consolidated into 13 groups by similar expression patterns, which were named by the expression peaks and bottoms: 1) 4h peak, 2) 6h peak, 3) 8h peak, 4) 12h peak, 5) 18-24h peak, 6) 24-36h peak, 7) 36h peak, 8) 48h peak, 9) up (consistently elevated expression relative to pre-bloodmeal state), 10) 6h bottom, 11) 18h bottom, 12) 48h bottom, and 13) down (Fig S3).

Gene ontology (GO) enrichment analyses, with a primary focus on “Biological Process” (BP) terms, indicated a far greater numbers of terms enriched in the first three peak groups (Fig 3 and Table S3).

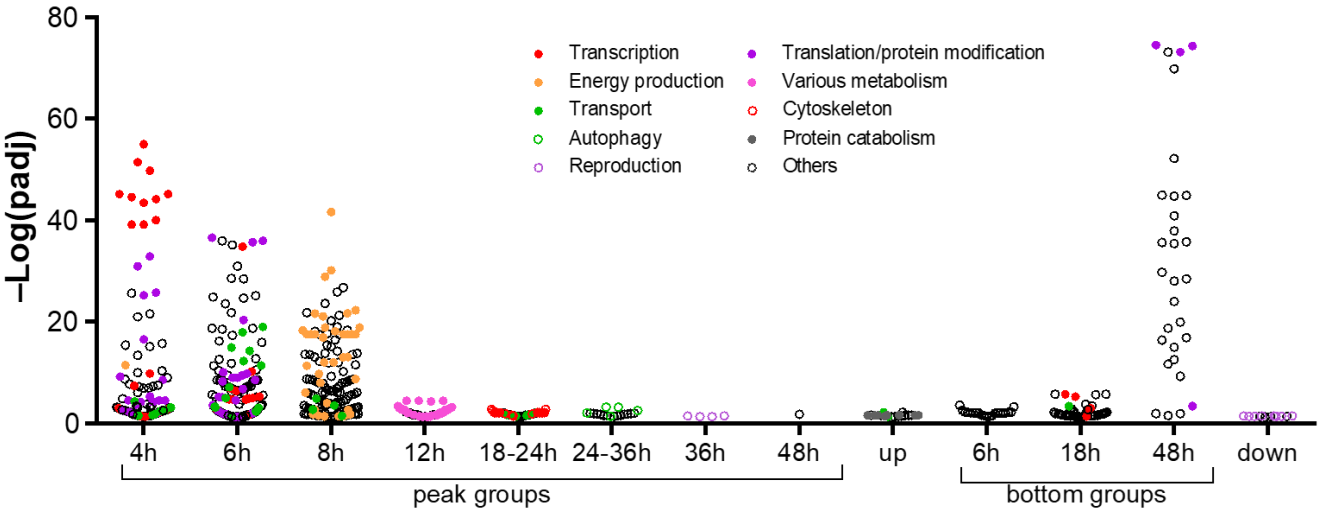

Fig3: –log plot of Benjamini-Hochberg-adjusted p values ($p < 0.05$) for GO enrichment analysis (Biological Process) for each cluster. Colors are applied for the related terms discussed in the text.

The general trend of enriched terms for the peak groups (groups 1-9) showed the transitioning roles of the midgut as the center for bloodmeal digestion (Fig 4, Table S3).

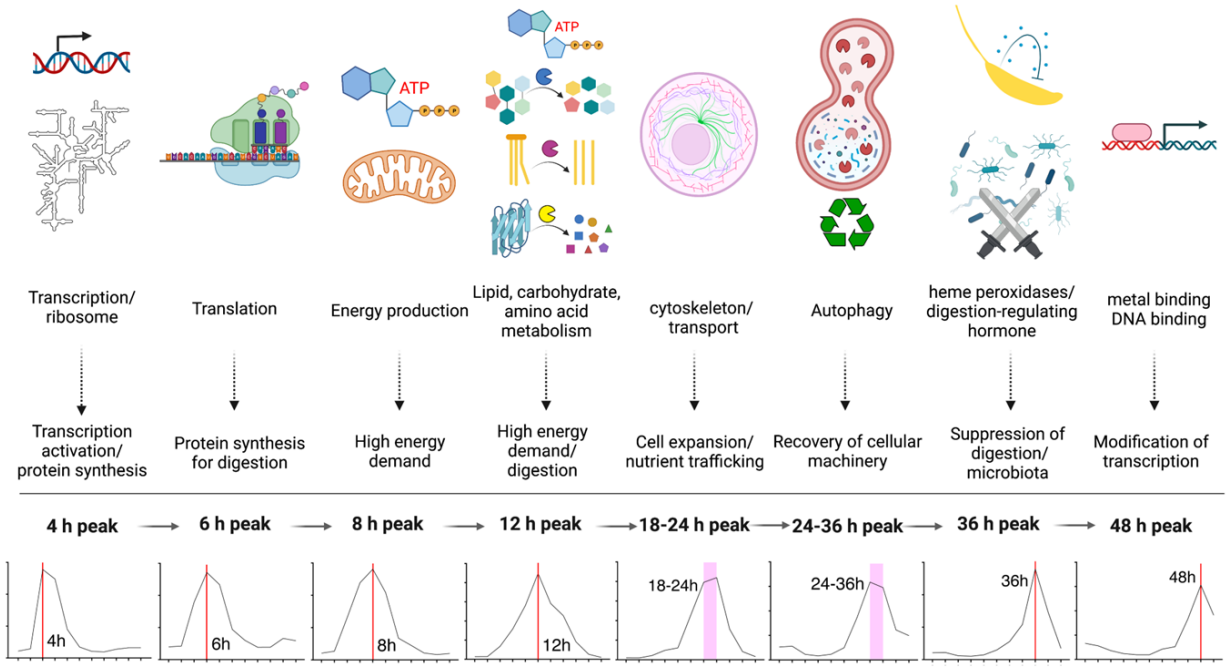

Fig4: GO enrichment for the peak cluster groups (group 1-8). The descriptions below the graphics are directly related to the top GO terms. The description below the arrows are inferred roles. Created with BioRender.com

1) the 4h peak group was enriched with GO terms for RNA metabolism, transcription, and translation, reflecting activation of transcription and protein synthesis. The largest change between time points shown in the PCA (1-4 hPBM) (Fig 2) may be due to this activation of transcription and translation, which suggests the female mosquito undergoes a drastic shift in their physiology to a digestion/reproduction state. 2) The 6h peak group was found to be enriched with GO terms for translation (and protein synthesis), suggesting the system further shifted to protein production to deal with the bloodmeal. 3) The 8h peak group was enriched with GO terms related to energy production, especially oxidative metabolism, suggesting a response to the high demand for energy needed to complete digestion. 4) The 12h peak group was enriched with GO terms for a variety of metabolic processes such as lipid, carbohydrate and amino acid metabolism, oxidoreductase activity, and energy production, which are a potential reflection of the most active time for the bloodmeal digestion. 5) The 18-24h peak group was

enriched with GO terms for the regulation of cytoskeleton (mainly actin) and transport, implying that the midgut is arranging the cellular components to allocate the nutrients from the bloodmeal. 6) The 24-36h peak group was enriched with GO terms associated with autophagy, suggesting recycling and regeneration of cellular components for recovery, may be at the transition of active digestion phase and recovery/regeneration phase. 7) Although the 36h peak group was enriched with GO terms for oogenesis and reproduction, the genes represented by these terms were comprised of heme peroxidases (HPXs) and peptide hormone precursors, one of which is trypsin modulating oostatic factor (TMOF) (Borovsky 2003), which reduces the expression of trypsin in the midgut and in turn reduces oocyte maturation, suggesting a negative regulation of digestion in the midgut at the end of digestion, likely to shut down digestive processes. HPXs may act to suppress the midgut microbiome that had expanded upon bloodmeal, as HPXs mosquitoes appear to have immune functions (Kumar et al. 2010; Barletta et al. 2019). 8) The 48h peak group was enriched with the genes with metal (mostly zinc ion) binding function and nucleic acid binding ("Molecular Function" terms, see Table S3), which suggests modifying transcription by variety of DNA-binding proteins (transcription factors, enhancers, silencers, repressors, chromatin-modifying proteins, etc.). 9) The "up" group was enriched with GO terms related to transport activity and ubiquitin-related functions, suggesting transport and protein turnover are important throughout the cycle as a hub for nutrients from the bloodmeal, implying the midgut as the center not only for digestion, but also for distribution. In addition, terms related to transport are present in other peak groups (Fig 3 and Table S3), which also suggests that transport of nutrients are important in the midgut. 10) to 13) down-regulated gene groups (i.e. "bottom" groups) were enriched with GO terms similar to some peak groups, such as RNA metabolism (18h bottom, as in 4h peak), translation (48h bottom, as in 4h peak and 6h peak), reproduction (down, including peroxidases as in 36h peak). The 6h bottom group contained several terms with "regulation of" (Table S3), which implies turning off the negative regulators of many biological processes. This may be due to role switch for the genes that have similar functions (genes for resting state vs active digestion state).

The clustering analysis overall shows the midgut operates as the center for digestion (and transport) by regulating gene expression in an orchestrated manner to set the stage for most important task for a female mosquito, reproduction. This analysis also suggested that the midgut functions as a distribution center for nutrients obtained from the bloodmeal.

Serine proteases:

To highlight the utility of the temporal resolution provided by this dataset, we focused initially on a group of genes that have a key role in blood digestion and for which exists substantial prior work—serine proteases (SPs). With the midgut's central role for bloodmeal digestion, SPs like trypsins and chymotrypsins have been described to play important roles in bloodmeal digestion in the midgut (Isoe et al. 2009). Indeed, throughout the time course, 8 out of the top 10 most abundantly expressed genes were SPs (Table 1) underscoring their importance in digestion.

Table 1: Top 10 most abundantly expressed genes in the midgut time-course transcriptome.

GeneID	**Product description**	**ID***
AAEL003060	female-specific chymotrypsin	AaCHYMO[a]
AAEL007818	Trypsin 3A1 Precursor (EC 3.4.21.4)	AaET[b]
AAEL013284	late trypsin 1, serine-type endopeptidase	AaLT[b]
AAEL001703	juvenile hormone-regulated chymotrypsin-like serine protease	AaJA15[a]
AAEL001690	serine-type endopeptidase	
AAEL013885	unspecified product	
AAEL007432	serine collagenase 1 precursor, putative	AaSP I[b]
AAEL010196	trypsin	AaSP VI[b]
AAEL010202	trypsin	AaSP VII[b]
AAEL004798	Mucin-like peritrophin	AeIMUC1[c]

GeneIDs with bold font are serine proteases. Product description are from VectorBase; * IDs are from previously described papers: [a] (Brackney et al. 2010), [b] (Isoe et al. 2009), [c] (Rayms-Keller et al. 2000).

The top highly expressed gene, AAEL003060 (adult, female-specific chymotrypsin, AaCHYMO) showed very high transcript levels with a slight reduction trend over time (Fig 5), consistent with the observation by Jiang *et al*. that the expression of this protein is under post-transcriptional control (Jiang et al. 1997). The expression patterns of the early trypsin (AaET, AAEL007818) and the late trypsin (AaLT, AAEL013284) mirror each other, providing a more detailed view of the kinetics of the temporal compartmentalization of these major enzymes (Fig 5). While AaET transcripts were highest at 0 hPBM (pre-bloodfed) and <1 hPBM, Brackney *et al*. described that the protein was detected only at 3-6 hPBM (Isoe et al. 2009), suggesting post-transcriptional regulation occurs with this gene as well. AaLT transcript peaked at 24 hPBM, which corresponds the known protein expression pattern (Isoe et al. 2009). The top 10 list also includes a previously uncharacterized SP gene, AAEL001690. Expression of this gene is highly induced in response to a bloodmeal with a peak at 12-18 hPBM (Fig 5). Functional studies on this gene in regard to bloodmeal digestion and reproduction may provide with further information as a potential target for population reduction strategy.

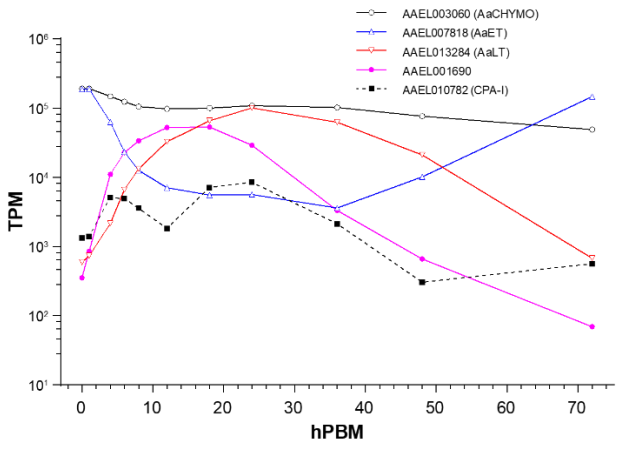

Fig5: Expression of chymotrypsin (AaCHYMO, AAEL003060), early trypsin (AaET, AAEL007818), late trypsin (AaLT, AAEL013284), AAEL001690 and carboxypeptidase-I (AAEL010782, CPA-I, * not a serine protease).

To get a comprehensive view of SPs expressed in the midgut, we obtained a list of all genes with the ontology term GO:0008236 (serine-type peptidase activity) in the *Ae. aegypti* genome at VectorBase. Out of 372 genes, we detected the expression of 346 genes in the midgut time-course transcriptome with non-zero TPM values (Table S4), representing 93% of the serine proteases. However, since many of these genes showed very low expression, we limited the list of genes to those with $\log_{10}(TPM+1) > 1.5$ for at least one time point, which reduced the number to 56. A previous study by Brackney *et al*. described 12 midgut SPs (Brackney et al. 2010). Ten of the 12 genes were found in the current genome annotation (AaegL5.3) by nucleotide sequence similarity. The remaining two genes are in the "AaSP group I" which consists of AaSP II-V due to their high sequence similarity, implying the two genes (AaSP II and V) are allelic forms of other genes in the group (AAEL013623—AaSP IV and AAEL013628—AaSP III) or that the genes have yet to be found. Analysis of the expression patterns shows the presence of the genes that have similar expression patterns to these previously described genes (Fig 6). For example, AaLT (AAEL013284) exhibits a peak around 18-36 hPBM and the lowest at the 0-1 hPBM (as well as 72 hPBM). There are 11 other genes that show similar expression patterns including some previously studied genes (AaSP VI, VII and Aa5G1) and 8 additional genes (also AAEL001690 discussed above) (Fig 5, orange rectangles), which may cooperatively participate in bloodmeal digestion. AaSP I (AAEL007432) shows elevated expression at earlier time points and remained high [$\log_{10}(TPM+1) > 4$] between 4 and 18 hPBM. There were 5 additional genes that exhibit similar expression patterns (Fig 6 purple rectangle), which may be important in initial digestion phase. Finally, there were also SP genes with an expression peak at later time points (36-48 hPBM, blue rectangle in Fig 6), suggesting potentially different roles in digestion. Thus, this study expands the existing knowledge about SPs in the midgut in context of the bloodmeal digestion physiology.

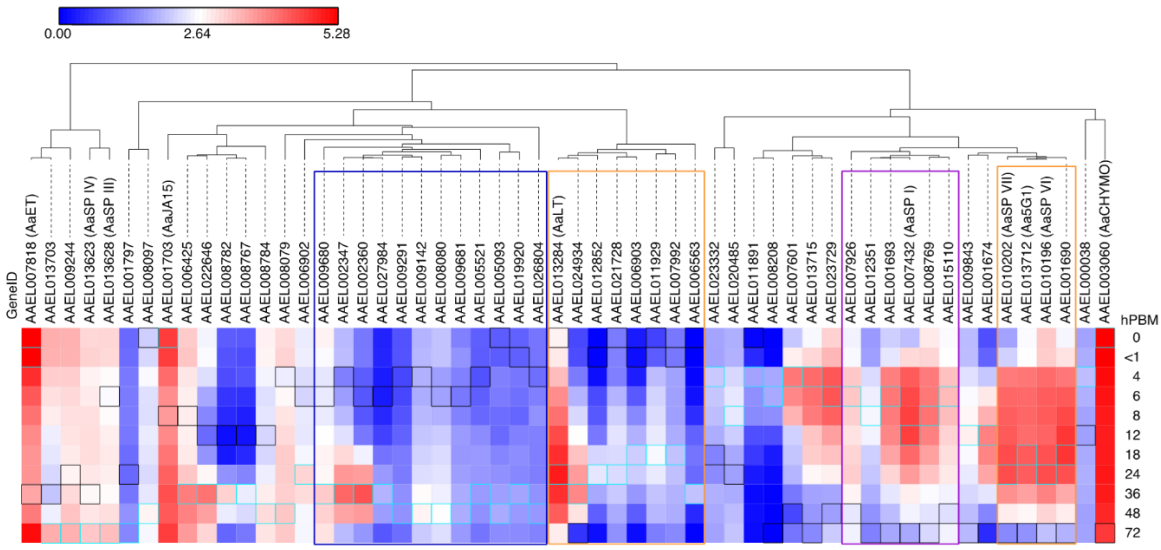

Fig6: Expression patterns of serine proteases with 1.5 < log10 (TPM+1) at least one time point. Genes were hierarchically clustered by expression patterns. Maximum and minimum values across the time course are indicated by cyan and black squares, respectively. See text for the groups indicated by colored rectangles.

Bonizzoni *et al*., suggested that differences in susceptibility for dengue-virus (DENV) between different strains of *Ae. aegypti* may be related to a trypsin, particularly AAEL007818 (AaET). They detected a difference in reduction of AaET at 5 hPBM between DENV-susceptible and DENV-resistant strains (Bonizzoni et al. 2012b), based on the earlier observation that the addition of a trypsin inhibitor in a DENV-containing bloodmeal reduced DENV titer in the mosquito (Molina-Cruz et al. 2005). However, our data such as above suggest that SPs in the midgut microecosystem seem to be more complex and correlation of SPs to DENV infection and dissemination requires further investigation.

To investigate whether transcript expression patterns are associated with either genomic location or sequence similarity (indicative of recent gene duplication) as well as to determine genomic hotspots of midgut SPs, we analyzed the nucleotide sequence similarity (protein-coding region) and locations in the

genome for the top 10% (35 genes) most abundantly expressed SPs (average across the time course). We found 23 of the genes are closely located in the genome and thus form clusters of serine proteases.

On chromosome 1, 5 serine protease genes are clustered in a 191 kb region, with two of these (AAEL013623: AaSP IV and AAEL013628: AaSP III) showing near identical expression patterns. These two genes also share 96% sequence identity, suggesting a recent gene duplication event (Fig 7).

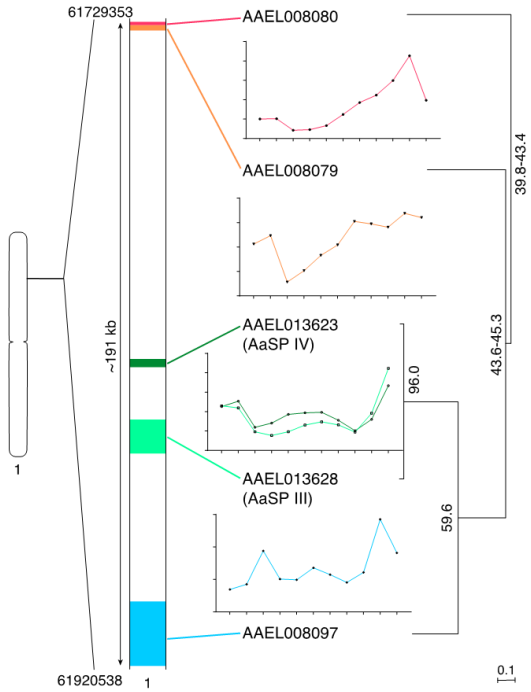

Fig7: Serine protease cluster on chromosome 1, with expression patterns across the 72-h time course and phylogenetic tree to represent sequence similarity. Similarity values (%) are shown on the nodes. Scale bar represents base substitution.

On chromosome 2, three clusters of serine proteases are located within an 88 Mb region (Fig 8A). The most upstream 2 genes (AAEL001690 and AAEL001674) show similar expression patterns, but their sequence similarity is below 50%. Of the three genes downstream, two (AAEL008769 and AAEL001693) share 60.6% sequence identity and expression patterns, but they are intervened by another serine protease gene, AAEL001703 (AaJA15), which shows different expression pattern from the two,

although it is >50% sequence identity to both. The second cluster is located ~69.5 Mb downstream and consists of 4 genes. Although the most upstream gene (AAEL022646) was located >830 kb upstream of other 3, they exhibit similar expression patterns, except for AAEL008784. Notably, AAEL008784 shows a similar expression pattern to AAEL001703, which is the intervening gene in the first cluster, although the expression levels differ ~1.5 orders of magnitude. The final cluster with 3 genes on chromosome 2 was located 16.8 Mb downstream of the second cluster and each of the genes shows similar expression patterns and >60% identity, with overall expression similar to that of the second cluster, but lacks the reduced expression at 12 hPBM (Fig 8B).

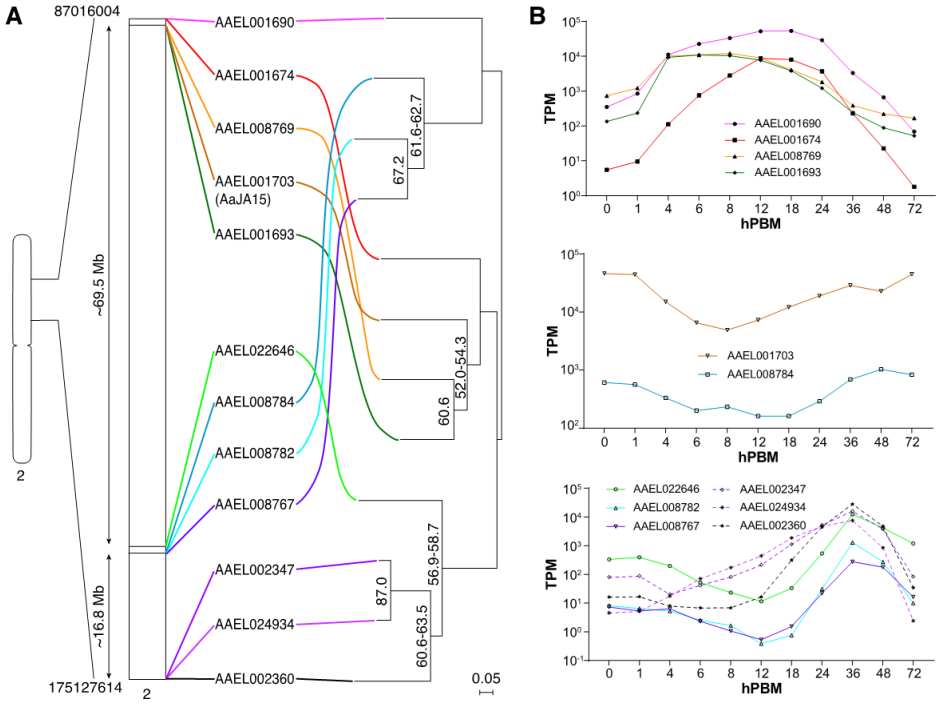

Fig8: (A) Serine protease (SP) clusters on chromosome 2 with phylogenetic tree to represent sequence similarity. Similarity values (%) are shown on the nodes (>50%). Scale bar represents base substitution. (B) Expression patterns across the 72-h time course for the SPs shown in A.

On chromosome 3, 6 genes are clustered within a 119 kb region (Fig 9). Of these, two (AAEL023729 and AAEL013715) and three (AAEL010196: AaSP VI, AAEL013712: Aa5G1, and AAEL010202: AaSP VII) show matched expression patterns. The two subgroups also show high

sequence identity within the subgroups. Interestingly, all six serine protease genes are located within an intron of a gene AAEL013713 (putative β arrestin), and this intron also contains 2 trypsin-like pseudogene sequences (AAEL023056 and AAEL024408). This may be the most (highly-expressed) SP-dense region in the genome.

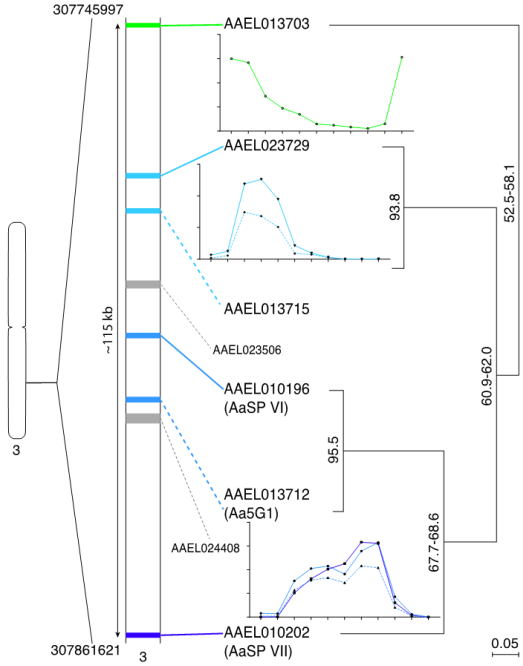

Fig9: Serine protease cluster on chromosome 3, with expression patterns across the 72-h time course (coded by line colors and styles) and phylogenetic tree to represent sequence similarity. Similarity values (%) are shown on the nodes. Intervening pseudogenes are also shown as grey bands. Scale bar represents base substitution.

Moreover, investigation of cis-regulatory elements (promoters) for these genes may lead to future usage for midgut-specific transgene expression, although organ/stage-specific expression needs to be verified. For instance, the promoters empirically verified to date for the midgut-specific expression in *Ae. aegypti* are only from 2 genes: carboxypeptidase A-I (CPA-I) (AAEL010782) (Moreira et al. 2000) and AaET (Zhao et al. 2016). These genes exhibit different expression patterns; the CPA-I transcript

expression has two peaks at 4-6 hPBM and 24 hPBM (Fig 5). Our observation parallels the study on transgene expressing luciferase under CPA-I promoter, particularly luciferase transcript expression (see Fig 4 of the reference), which exhibited strong signal at 8 hPBM and 20-24 hPBM with weaker signal at 12 hPBM (Moreira et al. 2000). Studies with particular interest in transgene expression in the midgut at 12 hPBM may consider avoiding the use of this promoter. AaET expression is robust pre-bloodmeal, but it reduced during the major digestion phase (Fig 4), suggesting AaET promoter is not suitable for bloodmeal-inducible expression of transgenes. Promoters from genes like AaLT and AAEL001690 as well as the genes that showed similar expression patterns (Fig 6) may serve for a more robust blood-meal inducible transgene expression in the midgut.

Iron/heme related genes:

Previously, we conducted screens for putative iron transporters (Tsujimoto et al. 2018; Tsujimoto et al. 2021), peritrophic matrix proteins (Whiten et al. 2018), and putative heme transporters (Eggleston and Adelman 2020) to better understand iron/heme handling from a bloodmeal in *Aedes aegypti*. Using a list of those genes along with ferritins, transferrins and recently determined heme exporters (Wang et al. 2022), we analyzed expression across our midgut transcriptome time course (Table S5).

Putative iron transporters (Fig 10 blue rectangle):

Within ZnT and ZIP family putative iron transporters, AaeZIP13 showed elevated expression peaked at 24hPBM (Fig 10), consistent with our previous observation (Tsujimoto et al. 2018). AaeZnT7 and AaeZIP11 showed relatively lower expression without drastic changes to transcript abundance across the time points, with just a moderate increase around 6 hPBM, 48-72 hPBM for AaeAZnT7, and 12 and 48 hPBM for AaeZIP11. Among genes identified in a screen for putative iron transporters, three (AAEL008406, AAEL011520, and AAEL007758) exhibited TPM higher than 100 (Fig 10, white to shades of red and Table S5). Of these, AAEL008406 (annotated as a cationic amino acid transporter)

displayed a peak of expression around 6-8 hPBM whereas AAEL011520 (SLC45/sugar transporter) and AAEL007758 (unknown function) showed peak expression around 24-36 hPBM (Fig 10 and Table S5).

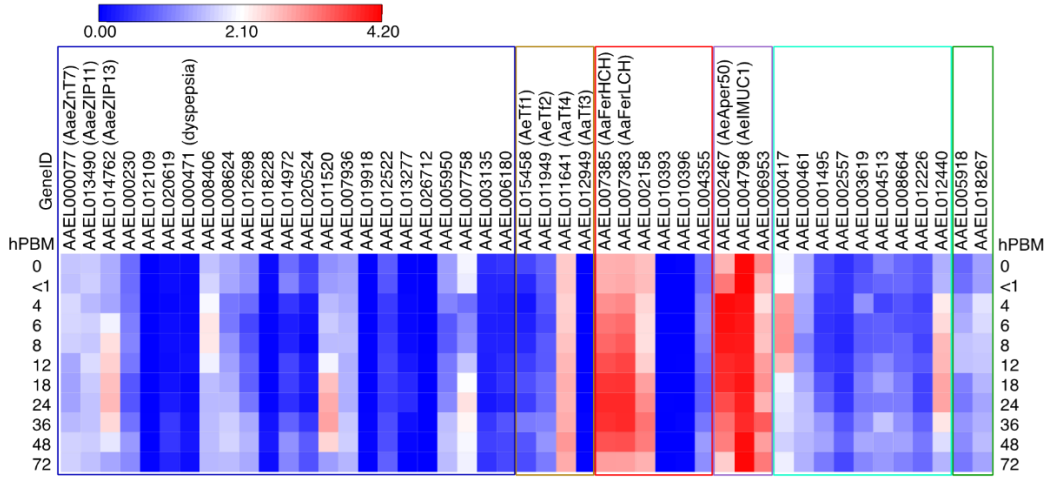

Fig10: Heat map showing expression patterns of putative iron transporters (blue rectangle), transferrins (brown rectangle), ferritins (red rectangle), peritrophins (lavender rectangle), putative heme transporters (cyan rectangle), and verified heme exporters (green rectangle). The genes that showed no expression across all the time points were excluded. See also Table S4.

Transferrins (Fig 10 brown rectangle):

Transferrins (Tfs) are known to be multi-functional iron-binding proteins (Zhou et al. 2009). Four Tf genes are present in the *Ae. aegypti* genome, among which two (AaTf1 and AaTf2) have been functionally studied. AaTf1 transcripts were found in fat bodies and adult female thorax, which were elevated by bloodmeal and bacterial challenge in the ovaries (Harizanova et al. 2005; Zhou et al. 2009). AaTf1 protein was also found in the hemolymph (Harizanova et al. 2005). AaTf2 transcripts are found in fat bodies and ovaries with suggested functions in development and wound response (Zhou et al. 2009). Here we show that unlike AaTf1-3, AaTf4 (AAEL011461) was highly expressed in the midgut ((Harizanova et al. 2005; Zhou et al. 2009) and Fig 10). Functional study on AaTf4 may uncover the role of AaTf4 in iron transport and control of gut microbiota upon bloodmeal.

Ferritins (Fig 10 red rectangle):

Ferritins in general are known to be multimeric proteins that make a large sphere to cage thousands of Fe^{3+} (Pham and Winzerling 2010). Unlike mammalian ferritins, which serve as iron storage proteins, insect ferritins are considered to function as iron shuttle proteins and protectants from oxidative damages caused by Fe^{2+} (Pham and Winzerling 2010). The *Ae. aegypti* genome contains 7 ferritin-like genes, of which no transcript for AAEL000359 was detected in the midgut transcriptome (Table S5). In *Ae. aegypti* ferritin light chain (FerLCH, AAEL007383) and heavy chain (FerHCH, AAEL007385) are well documented. Both proteins have been shown to be induced by a bloodmeal in the midgut (Dunkov et al. 2002). Transcripts of the two genes were induced by ferrous ammonium sulfate (Fe^{2+}), hemin (heme-related compound), and H_2O_2 in whole bodies (Geiser et al. 2003). As expected, the most highly expressed ferritin genes in the midgut were AAEL007383 and AAEL007385 (Fig 10). These genes are closely located by head-to-head orientation, which was determined to contain multiple transcription start sites and documented FerLCH and FerHCH promoters (Geiser et al. 2003; Pham et al. 2003; Pham and Chavez 2005). The expression patterns and levels of these genes are nearly identical, which support that these genes are operated by a common promoter. Their increased expression after bloodmeal may be a response to elevated iron levels to mitigate the oxidative damage by Fe^{2+}. The promoter for the AAEL007383 (AaFerLCH) was demonstrated to be specifically responsive to the intracellular iron levels in a cultured cell line, Aag2 (Tsujimoto et al. 2018). Another ferritin gene, AAEL002158, also exhibits high expression with induction at later stage (after 12 hPBM) than the other two, which implies a distinct role in bloodmeal digestion.

Peritrophins (Fig 10 lavender rectangle):

All three previously described midgut peritrophins were found to be highly expressed across time course (Fig 10). AAEL004798 (AeIMUC1) exhibited very high expression throughout the time course; it is indeed one of the top 10 most expressed genes (Table 1), although its transcript expression was not

induced by bloodfeeding. Since it has been demonstrated to bind a large amount of heme (Devenport et al. 2006), AeIMUC1 may be the major heme-sequestering peritrophin. AAEL006953 also showed varied expression levels across the time course. On the other hand, AAEL002467 (AeAper50) expression was highly induced by a bloodmeal, which agrees with previous report, although our study lacks the 2 hPBM time point, the potential true expression peak (Shao et al. 2005). The response was quite rapid; at <1 hPBM, TPM increased more than 3 times from 0 hPBM, reaching a peak (TPM: 13,025; 25× of 0 hPBM) at 4 hPBM and maintained at least half of the peak level until 18 hPBM (Table S5). Intriguingly, protein expression of AeIMUC1 is post-transcriptionally regulated and exhibited peak at 4 hPBM (Devenport et al. 2006), the high transcript level prior to this time point may be prerequisite for the rapid production of high amount of protein product to protect the midgut. This also implies that at least AeIMUC1 and AeAper50 orchestrate their expression for a rapid formation of PM, which precedes the major digestive phase.

Putative heme transporters (Fig 10 cyan rectangle):

Heme is a component of the major blood protein, hemoglobin. Like iron, heme can cause cellular damage despite also serving as a vital nutrient and a signaling molecule in trace amounts (Eggleston and Adelman 2020). In a previous study based on cell culture transcriptomes, several transporters whose expression levels were responsive to heme were identified (Eggleston and Adelman 2020). Among these, two genes exhibited high (TPM greater than 100) expression (AAEL000417 and AAEL012440) (Fig 10 and Table S5) in our midgut transcriptome time course. Gene AAEL000417 (a member of the monocarboxylate transporter family) showed a rapid increase of transcript with a peak at 6-8 hPBM, while AAEL012440 (annotated as sodium-bile acid cotransporter) showed gradual increase with a peak at 18-24 hPBM. The fact that these genes respond transcriptionally both to heme exposure and to a bloodmeal in the midgut suggests potential roles in blood digestion/detoxification that remain to be determined.

Recently, a heme exporter has been characterized from *Drosophila melanogaster* (Wang et al. 2022). The same study verified that homologs in *Ae. aegypti* also had similar activity. The transcript expression of these genes (AAEL005918 and AAEL018267) in the midgut were not markedly high and were only mildly induced by bloodfeeding, suggesting either that very little heme is exported to the hemolymph or that there are additional heme exporters that remain to be determined that compliment this function.

Conclusion:

This study utilized Illumina sequencing technology to present the midgut transcriptome from pre-bloodfeeding state to the completion of bloodmeal digestion at 72 h in 11 time points in the most important mosquito vector of arboviruses worldwide, the first study to cover a complete digestion cycle in *Ae. aegypti* with such fine temporal resolution. We obtained a total of over 1 billion paired-end fragments, which provided with a depth of more than 26 million fragments per sample. The results supported 72 h as a complete bloodmeal digestion cycle for *Ae. aegypti* (Fig 2). This study uncovered dynamic expression changes across the 72-h time course, which supports the idea that the female midgut performs temporally controlled gene expression as the center for bloodmeal digestion and nutrient distribution (Fig S3, 3-4, Table S4). The data we obtained may be used to analyze expression patterns of genes in the same functional category, such as the serine proteases (SPs) discussed in this manuscript, or other categories of interest such as secreted proteins, membrane-bound transporters, or immune pathways. Our analysis of SPs has revealed a number of previously undescribed SPs that may function synergistically or have different roles in bloodmeal digestion (Fig 6). The view of detailed temporal expression patterns helps categorize (or subcategorize) groups of genes, especially complex sets of genes. Even for a gene of interest in the context of the time course expression may provide us with a new insight, such as expression peaks that may be a component to consider when developing a control strategy (timing of application).

A recent study on transcriptome of all three tagmata (head, thorax, and abdomen) with subsectioned gut system and ovaries uncovered previously understudied spatial resolution of gene expression in the

female *Ae. aegypti* (Hixson et al. 2022). Our study complements theirs by providing finer scale temporal resolution. As noted above, the difference in the length of the digestive cycle between Hixson *et al*. and this study may be reflection of multiple parameters (strain, temperature, etc…). The importance of this difference when analyzing and interpreting RNAseq data during the process of bloodmeal digestion cannot be overstated. Simply put, treatments or conditions that have the effect of slightly slowing or accelerating digestion may result in the appearance of many hundreds or thousands of differentially expressed genes, with no biological significance beyond the alteration of the rate of digestion. Our study provides a road map that can be used to calibrate expression values to identify such shifts, potentially allowing such RNAseq studies to focus on genes more unique to the treatment conditions. Finally, similar studies of based on the transcriptics of other organs with fine time component like this should greatly enhance our understanding of physiology and genetics of bloodfeeding mosquitoes and lead to new research avenues.

Materials and Methods:

Mosquitos:

Ae. aegypti Liverpool strain was maintained in an insectary at Texas A&M Entomology in environmental chambers kept at 27 °C, 70% RH and 14:10 (L:D) cycles. Immature stages were reared with ground TetraMin (Tetra) and adults were fed on defibrinated sheep blood (Colorado Serum Company).

Blood feeding and midgut dissection:

For all the experiments, female mosquitoes 3-4 days post eclosion were used (the females were eclosed in the presence of males). They were maintained on 10% sucrose (granulated sugar from local grocery stores) in water and starved for ~16 h prior to blood feeding. Mosquitoes were fed on citrated sheep blood (Hemostat Laboratories) via an artificial membrane feeder for ~15 min, anesthetized on ice and only engorged ones were selected for midgut collection. Midguts were dissected from 0 (pre-bloodfed/starved, set aside when bloodfeeding was performed), <1 (no recovery from anesthesia), 4, 6, 8, 12, 18, 24, 36, 48

and 72 h post bloodmeal (PBM). The mosquitoes were discouraged from laying eggs on the sugar source (common with 10% sucrose) by increasing the concentration to 30% sucrose. To be consistent with the mosquito's circadian rhythm, blood feeding was performed approximately 1 h after beginning of the light cycle. Blood meal (bolus) was removed from the midguts upon collection. A pool of 30 midguts were collected in a 1.5-mL microcentrifuge tube with TRIzol (Thermo Fisher) and stored at –80 °C for downstream processing. A total of 4 independent replicates were made for each time point.

RNA extraction, library prep and sequencing:

Whole RNA was extracted with TRIzol (Thermo Fisher) standard protocol with half volume format (0.5 mL TRIzol), and initial homogenization in a small volume of TRIzol (20 µL) using an RNase-free pestle (Contes). RNA pellets were resuspended with 16 µL of nuclease-free H_2O and treated by DNase (2 µL of DNase buffer and 2 µL of TURBO DNase; Thermo Fisher) at 37 °C for 1 h. RNA was quantified by a Qubit RNA broad range kit (Thermo Fisher). Integrity of RNA was analyzed by a Fragment Analyzer (Agilent Technology).

Illumina libraries for RNAseq were prepared using NEBNext Ultra II RNA Library Prep kit for Illumina (E7770, NEB) with NEBNext Poly(A) mRNA Magnetic Isolation Module (E7490, NEB) and NEBNext Multiplex Oligos for Illumina (96 Unique index primer pairs) (E6440, NEB). We used 13 min fragmentation time and 9 cycles of PCR enrichment on 1000 ng of total RNA input determined by preliminary library preparation. All libraries were quantified by a Qubit with dsDNA BR kit (Thermo Fisher) and quality-checked by a Fragment Analyzer (Agilent Technology).

Pre-pooling quantification by qPCR, pooling and sequencing on a NovaSeq 6000 2 × 100 bp (paired-end) format were performed by the Genomics and Bioinformatics Service at Texas A&M Agrilife Research. The raw reads (fastq.gz) were deposited in National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) with the BioProject ID: PRJNA639291.

Bioinformatics analysis:

The Illumina reads were mapped, and expression was quantified using HISAT2 and StringTie pipeline (Pertea et al. 2016) using the AaegL5 genome and AaegL5.2 geneset obtained from VectorBase (Giraldo-Calderón et al. 2015). StringTie was used with "-e" and "-A" options. Above processes were performed using the Texas A&M High Performance Research Computing system. A Count table ("gene_count_matrix.csv") for downstream analysis was generated by a python script "prepDE.py" provided in the StringTie manual online (https://ccb.jhu.edu/software/stringtie/index.shtml?t=manual). A TPM (transcripts per kilobase million) table for each library was generated from count tables generated by the -A option of StringTie and the tables for all 44 libraries were combined into one table using a Python script. Pair-wise differential expression analyses and principal component analysis (PCA) were performed using DESeq2 package (Love et al. 2014) on R. The combined TPM table was trimmed using a python script to remove the genes with 0 values across all the time points. The final TPM table is provided as Table S1. and the gene IDs from the list were searched on VectorBase to obtain the information on the gene types (protein-coding, non-coding and pseudogenes).

Time-course analysis was performed following an example shown in edgeR package (Chen et al. 2016) user's guide (June 2020 version), which uses smoothing spline with df=3. Through this, we obtained a list of genes showed significant change across the time course with false-discovery rate (FDR) < 0.05. Then using the FDR<0.05 gene list, we performed soft clustering analysis using an R package "Mfuzz" (Futschik and Carlisle 2005) to obtain gene lists with similar expression patterns. With this analysis, we generated 20 clusters by expression patterns, which were consolidated into 13 groups. Gene Ontology (GO) enrichment analyses was performed for the Biological Process and Molecular Function categories using enrichment analysis option on VectorBase (https://vectorbase.org/vectorbase/app).

Gut system and ovary imaging:

Age-matched (to the mosquitoes used for sequencing) *Ae. aegypti* "Liverpool" females were bloodfed using a membrane feeder and defibrinated sheep blood. Midgut, Malpighian tubules, hindgut and ovaries

were dissected at the same time points post bloodmeal. Unlike RNAseq samples, mosquitoes were bloodfed at different times of the day so that dissection could be performed in the working hours. The dissected organs were photographed using an AmScope MU300 digital camera (AmScope)-equipped stereo-microscope Leica M165FC. Scale bar was added by Fiji (ImageJ) using an image of a hemocytometer grid with the same magnification as a reference.

Serine protease gene family analysis:

To obtain a list of serine proteases in the *Ae. aegypti* genome, we used the Gene Ontology term GO:0008236 (serine-type peptidase activity) on VectorBase. The expression data were retrieved from the master expression table (TPM) using a Python script. To filter SP with reasonable expression levels, the list was filtered to have $\log_{10}(TPM+1) > 1.5$ for at least one time point. To visualize the expression patterns of these genes, Morpheus (https://software.broadinstitute.org/morpheus) was used using hierarchical clustering with default parameters (one minus Pearson correlation and average linkage). Modification of labels and dendrogram as well as high and low value highlighting were done using Illustrator (version 26, Adobe). For analysis of genomic location of the top expressed serine proteases, the top 10% most highly expressed across the time course (means of the time course: 35 genes) were searched on VectorBase to determine the genomic loci. Expression graphs were made using GraphPad Prism (ver. 7, GraphPad). The nucleotide sequences (protein-coding regions) of the highly expressed genes closely located on each chromosome were aligned using Clustal Omega (Sievers et al. 2011) to obtain their sequence similarities and potential phylogenetic relationships by neighbor-joining tree using SeaView (Gouy et al. 2010) (10,000 bootstrap). The chromosomal loci were visualized using PhenoGram (Wolfe et al. 2013). Then, Illustrator (Adobe) was used to combine the PhenoGram output and the gene tree to visualize the relationships of genomic location and sequence similarity.

Putative iron and heme transporters, and heme-binding proteins:

The genes found to have relation to iron transport were obtained from (Tsujimoto et al. 2018) and from Table 1 in (Tsujimoto et al. 2021). The heme transporter candidates were from Table 6 in (Eggleston and Adelman 2020); heme exporters were from (Wang et al. 2022). The peritrophic matrix proteins with chitin binding domains were from (Whiten et al. 2018) (AAEL002495 is AAEL004798 in the current annotation AaegL5.3). Transferrins (Tf) and ferritins were retrieved by text search by "transferrin" and "ferritin" on VectorBase. The list of the genes are in Table S5. The expression values (TPM) for the genes were extracted from the master TPM table (Table S1) using a python script and $\log_{10}$(TPM+1)-transformed on Microsoft Excel. Morpheus (https://software.broadinstitute.org/morpheus) was used to visualize expression patterns after removal of any genes that had zeros for all the time points (AAEL019992 and AAEL000359). Additional labels and markings were made using Illustrator (Adobe).

Acknowledgement:

Fragment Analyzer quality analysis of RNA and the Illumina libraries and NovaSeq sequencing were performed by the Texas A&M AgriLife Genomics and Bioinformatics Service. Portions of this research were conducted with the advanced computing resources provided by Texas A&M High Performance Research Computing.

Funding

This work was funded by the USDA National Institute of Food and Agriculture, Hatch project 1018401 and by Texas A&M AgriLife Research under the Insect Vectored Disease Grant Program to ZNA.

Data Availability

The raw reads (fastq.gz) were deposited in National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) with the BioProject ID: PRJNA639291. All other data is available within the manuscript or as supplemental information.

References

Barletta ABF, Trisnadi N, Ramirez JL, Barillas-Mury C. 2019. Mosquito Midgut Prostaglandin Release Establishes Systemic Immune Priming. *iScience* **19**: 54-62.

Bonizzoni M, Dunn WA, Campbell CL, Olson KE, Dimon MT, Marinotti O, James AA. 2011. RNA-seq analyses of blood-induced changes in gene expression in the mosquito vector species, *Aedes aegypti*. *BMC Genomics* **12**: 82.

Bonizzoni M, Dunn WA, Campbell CL, Olson KE, Marinotti O, James AA. 2012a. Complex modulation of the *Aedes aegypti* transcriptome in response to dengue virus infection. *PloS one* **7**: e50512.

Bonizzoni M, Dunn WA, Campbell CL, Olson KE, Marinotti O, James AA. 2012b. Strain Variation in the Transcriptome of the Dengue Fever Vector, *Aedes aegypti*. *G3 (Bethesda)* **2**: 103-114.

Borovsky D. 2003. Trypsin-modulating oostatic factor: a potential new larvicide for mosquito control. *J Exp Biol* **206**: 3869-3875.

Brackney DE, Isoe J, th WCB, Zamora J, Foy BD, Miesfeld RL, Olson KE. 2010. Expression profiling and comparative analyses of seven midgut serine proteases from the yellow fever mosquito, *Aedes aegypti*. *J Insect Physiol* **56**: 736-744.

Chen Y, Lun AT, Smyth GK. 2016. From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Res* **5**: 1438.

Cui Y, Franz AWE. 2020. Heterogeneity of midgut cells and their differential responses to blood meal ingestion by the mosquito, *Aedes aegypti*. *Insect Biochem Mol Biol* **127**: 103496.

Devenport M, Alvarenga PH, Shao L, Fujioka H, Bianconi ML, Oliveira PL, Jacobs-Lorena M. 2006. Identification of the *Aedes aegypti* peritrophic matrix protein AeIMUCI as a heme-binding protein. *Biochemistry* **45**: 9540-9549.

Dunkov BC, Georgieva T, Yoshiga T, Hall M, Law JH. 2002. *Aedes aegypti* ferritin heavy chain homologue: feeding of iron or blood influences message levels, lengths and subunit abundance. *J Insect Sci* **2**: 7.

Eggleston H, Adelman ZN. 2020. Transcriptomic analyses of *Aedes aegypti* cultured cells and ex vivo midguts in response to an excess or deficiency of heme: a quest for transcriptionally-regulated heme transporters. *BMC Genomics* **21**: 604.

Freyvogel TA, Staeubli W. 1965. The Formation of the Peritrophic Membrane in Culicidae. *Acta Trop* **22**: 118-147.

Futschik ME, Carlisle B. 2005. Noise-robust soft clustering of gene expression time-course data. *J Bioinform Comput Biol* **3**: 965-988.

Geiser DL, Chavez CA, Flores-Munguia R, Winzerling JJ, Pham DQ. 2003. *Aedes aegypti* ferritin. *Eur J Biochem* **270**: 3667-3674.

Giraldo-Calderón GI, Emrich SJ, MacCallum RM, Maslen G, Dialynas E, Topalis P, Ho N, Gesing S, Madey G, Collins FH et al. 2015. VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. *Nucleic acids research* **43**: D707-713.

Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular biology and evolution* **27**: 221-224.

Harizanova N, Georgieva T, Dunkov BC, Yoshiga T, Law JH. 2005. *Aedes aegypti* transferrin. Gene structure, expression pattern, and regulation. *Insect Molecular Biology* **14**: 79-88.

Hixson B, Bing XL, Yang X, Bonfini A, Nagy P, Buchon N. 2022. A transcriptomic atlas of *Aedes aegypti* reveals detailed functional organization of major body parts and gut regional specializations in sugar-fed and blood-fed adult females. *Elife* **11**.

Isoe J, Rascon AA, Jr., Kunz S, Miesfeld RL. 2009. Molecular genetic analysis of midgut serine proteases in *Aedes aegypti* mosquitoes. *Insect Biochem Mol Biol* **39**: 903-912.

Jiang Q, Hall M, Noriega FG, Wells M. 1997. cDNA cloning and pattern of expression of an adult, female-specific chymotrypsin from *Aedes aegypti* midgut. *Insect Biochem Mol Biol* **27**: 283-289.

Kumar S, Molina-Cruz A, Gupta L, Rodrigues J, Barillas-Mury C. 2010. A peroxidase/dual oxidase system modulates midgut epithelial immunity in *Anopheles gambiae*. *Science* **327**: 1644-1648.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Molina-Cruz A, Gupta L, Richardson J, Bennett K, Black W, IV, Barillas-Mury C. 2005. Effect of mosquito midgut trypsin activity on dengue-2 virus infection and dissemination in *Aedes aegypti*. *Am J Trop Med Hyg* **72**: 631-637.

Moreira LA, Edwards MJ, Adhami F, Jasinskiene N, James AA, Jacobs-Lorena M. 2000. Robust gut-specific gene expression in transgenic *Aedes aegypti* mosquitoes. *Proc Natl Acad Sci U S A* **97**: 10895-10898.

Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature protocols* **11**: 1650-1667.

Pham DQ, Shaffer JJ, Chavez CA, Douglass PL. 2003. Identification and mapping of the promoter for the gene encoding the ferritin heavy-chain homologue of the yellow fever mosquito *Aedes aegypti*. *Insect Biochem Mol Biol* **33**: 51-62.

Pham DQ-D, Chavez CA. 2005. The ferritin light-chain homologue promoter in *Aedes aegypti*. *Insect Molecular Biology* **14**: 263-270.

Pham DQD, Winzerling JJ. 2010. Insect ferritins: Typical or atypical? *Biochimica et Biophysica Acta (BBA) - General Subjects* **1800**: 824-833.

Raquin V, Merkling SH, Gausson V, Moltini-Conclois I, Frangeul L, Varet H, Dillies MA, Saleh MC, Lambrechts L. 2017. Individual co-variation between viral RNA load and gene expression reveals novel host factors during early dengue virus infection of the *Aedes aegypti* midgut. *PLoS Negl Trop Dis* **11**: e0006152.

Rayms-Keller A, McGaw M, Oray C, Carlson JO, Beaty BJ. 2000. Molecular cloning and characterization of a metal responsive *Aedes aegypti* intestinal mucin cDNA. *Insect Mol Biol* **9**: 419-426.

Sanders HR, Evans AM, Ross LS, Gill SS. 2003. Blood meal induces global changes in midgut gene expression in the disease vector, *Aedes aegypti*. *Insect Biochem Mol Biol* **33**: 1105-1122.

Shao L, Devenport M, Fujioka H, Ghosh A, Jacobs-Lorena M. 2005. Identification and characterization of a novel peritrophic matrix protein, Ae-Aper50, and the microvillar membrane protein, AEG12, from the mosquito, *Aedes aegypti*. *Insect Biochem Mol Biol* **35**: 947-959.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539.

Stobbart RH. 1977. The control of the diuresis following a blood meal in females of the yellow fever mosquito *Aedes aegypti* (L). *The Journal of experimental biology* **69**: 53-85.

Tsujimoto H, Anderson MAE, Eggleston H, Myles KM, Adelman ZN. 2021. *Aedes aegypti* dyspepsia encodes a novel member of the SLC16 family of transporters and is critical for reproductive fitness. *PLoS Negl Trop Dis* **15**: e0009334.

Tsujimoto H, Anderson MAE, Myles KM, Adelman ZN. 2018. Identification of Candidate Iron Transporters From the ZIP/ZnT Gene Families in the Mosquito *Aedes aegypti*. *Frontiers in Physiology* **9**: 380.

Wang Z, Zeng P, Zhou B. 2022. Identification and characterization of a heme exporter from the MRP family in Drosophila melanogaster. *BMC Biol* **20**: 126.

Whiten SR, Ray WK, Helm RF, Adelman ZN. 2018. Characterization of the adult *Aedes aegypti* early midgut peritrophic matrix proteome using LC-MS. *PLoS One* **13**: e0194734.

WHO. 2009. Dengue Guidelines for diagnosis, treatment, prevention and control. p. 160. WHO.

WHO. 2016. WHO | Zika virus infection – United States of America - United States Virgin Islands.

WHO. 2018. A global strategy to Eliminate Yellow fever Epidemics 2017–2026.

WHO. 2019. Zika epidemiology update. pp. 1-13.

Wolfe D, Dudek S, Ritchie MD, Pendergrass SA. 2013. Visualizing genomic information across chromosomes with PhenoGram. *BioData Min* **6**: 18.

Zhao B, Hou Y, Wang J, Kokoza VA, Saha TT, Wang XL, Lin L, Zou Z, Raikhel AS. 2016. Determination of juvenile hormone titers by means of LC-MS/MS/MS and a juvenile hormone-responsive Gal4/UAS system in *Aedes aegypti* mosquitoes. *Insect Biochem Mol Biol* **77**: 69-77.

Zhou G, Velasquez LS, Geiser DL, Mayo JJ, Winzerling JJ. 2009. Differential regulation of transferrin 1 and 2 in *Aedes aegypti*. *Insect Biochemistry and Molecular Biology* **39**: 234-244.