Dataset Paper

Anopheles gambiae: Metabolomic Profiles in Sugar-Fed, Blood-Fed, and *Plasmodium falciparum*-Infected Midgut

Cody J. Champion,¹ Phanidhar Kukutla,¹ Elizabeth K. K. Glennon,² Bo Wang,² Shirley Luckhart,² and Jiannong Xu¹

¹Department of Biology, New Mexico State University, Las Cruces, NM, USA

²Department of Medical Microbiology and Immunology, University of California, Davis, Davis, CA, USA

Correspondence should be addressed to Shirley Luckhart; sluckhart@ucdavis.edu and Jiannong Xu; jxu@nmsu.edu

Received 10 July 2016; Accepted 21 November 2016; Published 24 July 2017

Academic Editor: Yunping Qiu

Copyright © 2017 Cody J. Champion et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The mosquito midgut is a physiological organ essential for nutrient acquisition as well as an interface that encounters various mosquito-borne pathogens. Metabolomic characterization would reveal biochemical fingerprints that are generated by various cellular processes. The metabolite profiles of the mosquito midgut will provide an overview of the biochemical events in both physiological states and the dynamic responses to pathogen infections. In this study, the midgut metabolic profiles of *Anopheles gambiae* mosquitoes following feeding with sugar, human blood, mouse blood, and *Plasmodium falciparum*-infected human blood were examined. A mass spectrometry system coupled to liquid and gas chromatography produced a time series of metabolites in the midgut at discrete conditions (sugar feeding, 24 h and 48 h post-normal blood and *P. falciparum*-infected blood feeding). Triplicates were included to ensure system validity. A total of 512 individual compounds were identified; 511 were assigned to 8 superpathways and 75 subpathways. The dataset can be used for further inquiry into the metabolic dynamics of sugar and blood digestion and of malaria parasite infection. The dataset is accessible at the repository Dryad.

1. Introduction

Malaria is caused by infection with mosquito-borne parasites of the genus *Plasmodium*. According to the WHO, there were 214 million new cases and an estimated 438,000 deaths worldwide in 2015, with most deaths associated with *P. falciparum* transmitted by *An. gambiae* species complex mosquitoes in Africa [1]. Understanding mosquito biology and vector competence for parasite transmission will facilitate the development of novel intervention strategies for malaria control.

The act of feeding on blood is essential for mosquito reproduction and for malaria parasite transmission. Blood feeding, digestion, and associated physiological responses also affect the microbial community in the gut [2]. As an additional genetic repertoire, the microbiome plays essential roles in various mosquito phenotypes, including reproduction and immunity. “Omics” technologies have been robustly

employed in studying metagenomic interactions in various biological systems [3]. In particular, metabolomics approaches have been used to characterize amino acids, lipids, sugars, cofactors, and other small molecules that are precursors, intermediates, and by-products of reactions and pathways directed by the genetic blueprints in the metagenomic repertoire. In mosquito studies, transcriptomic responses to blood feeding and malaria parasite infection have been well characterized [4–9]. However, comprehensive metabolomic data are limited for medically important arthropods, except for a few reports [10–14]. Metabolomics allows for the identification and quantification of a range of metabolites in a system [15]. In this study, we used a nontargeted metabolomics approach [16, 17] to examine the midgut metabolites of sugar-fed, blood-fed, and *P. falciparum*-infected mosquitoes. This dataset can connect transcriptomic responses to biochemical fluxes related to the blood feeding and digestion and to *P. falciparum* infection.

FIGURE 1: Study design overview. Midgut samples were collected at time points from mosquitoes that were sugar-fed, blood-fed, and *P. falciparum*-infected blood-fed. As a control, uninfected human blood and *P. falciparum*-infected blood were collected as well. Metabolites were determined by Metabolon, Inc.

TABLE 1: Midgut sample collections.

Description	Tissue type	Source lab	Blood source	Sample identifier (n)	
Uninfected blood	Human blood	Luckhart	Human	NB (3)	
<i>P. falciparum</i> -infected blood	Human blood	Luckhart	Human	IB (3)	
Sugar meal	Midgut	Luckhart		SM-L (3)	
Sugar meal	Midgut	Xu		SM-X (3)	
Uninfected blood-fed	Midgut	Luckhart	Human	NBM-L, 24 h (3)	NBM-L, 48 h (3)
Uninfected blood-fed	Midgut	Xu	Mouse	NBM-X, 24 h (3)	NBM-X, 48 h (3)
<i>P. falciparum</i> -infected blood-fed	Midgut	Luckhart	Human	IBM-L, 24 h (3)	IBM-L, 48 h (3)

2. Methodology

The study design is outlined in Figure 1. The mosquito midgut samples were prepared in two laboratories. The midgut samples from the Xu lab (NMSU) were prepared from mosquitoes fed on sugar meals and meals of uninfected mouse blood; and the midgut samples from the Luckhart lab (UCD) were prepared from mosquitoes fed on sugar, on uninfected human blood, and on *P. falciparum*-infected human blood. Samples of uninfected human blood and *P. falciparum*-infected human blood were included as well (Table 1).

An. gambiae G3 mosquitoes were used for the study. All mosquito rearing and feeding protocols were approved and were in accordance with regulatory guidelines and standards set by the Institutional Animal Care and Use Committee of New Mexico State University and the University of California, Davis.

In the Xu lab, mosquitoes were reared under standard insectary conditions of 27–28°C and 80% humidity and a 12 h:12 h light-dark cycle. Larvae were provided with a 1:1 mix of Brewer's yeast and rodent chow (Purina Laboratory Rodent Diet 5001, LabDiet, St. Louis, MO). Adult mosquitoes were provided with 10% sucrose ad libitum. NIH Swiss outbred mice were used as a blood source for egg production.

In the Luckhart lab, *An. gambiae* G3 were reared and maintained at 27–28°C and 80% humidity and a 12 h:12 h light-dark cycle. First instar larvae were provided with a 1:2 mix of Brewer's yeast and Sera® Micron planktonic rearing food, while second to fourth larval instars were provided with Purina® Game Fish Chow pellets. Adult mosquitoes were provided with 10% sucrose ad libitum and allowed to feed on CD1 outbred mice for egg production.

In the Luckhart lab, the procedures for culture of *P. falciparum* strain NF54 MCB and mosquito infection were

described previously [18]. The infection rate for each replicate mosquito cohort was determined by counting *P. falciparum* oocysts in the midgut of a sample of mosquitoes at day 10 after infection. The protocols involving the culture and handling of *P. falciparum* for mosquito infection were approved and were in accordance with regulatory guidelines and standards set by the Biological Safety Administrative Advisory Committee of the University of California, Davis.

In the Xu lab, midgut samples were collected from adult mosquitoes fed on 10% sucrose and were collected at 24 h and 48 h after feeding on uninfected mouse blood. In the Luckhart lab, midgut samples were collected from mosquitoes fed on 10% sucrose, were collected at 24 h and 48 h after feeding on a mixture of uninfected human red blood cells (RBCs) and heat-inactivated human serum (1:1), and were collected at 24 h and 48 h after feeding on *P. falciparum*-infected blood (1:1 human RBCs: serum). Each sample contained 30 midguts at each collecting point, triplicates were collected from three mosquito cohorts, and a total of 30 samples were collected for metabolic profiling at Metabolon, Inc. (Durham, NC). Table 1 summarizes the sample collections.

The MicroLab STAR system (Hamilton Company) was used for sample preparation. For quality control purposes, recovery standards were added before the extraction process. Protein from the samples was removed using both organic and aqueous extraction methods. The extract from each sample was split into two fractions, one of which was used for gas chromatography (GC) and the other was used for liquid chromatography (LC).

For LC/MS analysis, a Waters ACQUITY UPLC and a Thermo-Finnigan LTQ-FT mass spectrometer were used. The sample extract was divided into two identical aliquots. These aliquots were dried and were then reconstituted in acidic or basic LC-compatible solvents. Each of these solvent types contained 11 or more injection standards that had fixed concentrations. The first aliquot was analysed using acidic positive ion optimized conditions. The second aliquot was analysed using basic negative ion optimized conditions. These aliquots were independently injected using separate dedicated columns. If the extract was reconstituted in acidic conditions, a gradient elution using water and methanol containing 0.1% formic acid in each liquid was used. If the extract was reconstituted in basic conditions, a gradient elution containing 6.5 mM ammonium bicarbonate and water/methanol was used. During data collection, the MS analysis alternated between MS and data-dependent MS2 scans using dynamic exclusion.

A Thermo-Finnigan Trace DSQ fast-scanning single-quadrupole mass spectrometer using electron impact ionization was used for GS/MS. Samples were redried under vacuum for at least 24 hours and then were derivatized under dried nitrogen using bis-trimethyl-silyl-trifluoroacetamide (BSTFA). The column used for GC/MS analysis was 5% phenyl and used a temperature ramp from 40° to 300°C in a 16-minute period.

Raw MS files were stored in a database. Data was then examined and quality control limits were applied. Peaks identified were stored in a separate database. Compounds were found by comparison to analytical standards or recurrent

FIGURE 2: Principal component analysis of the dataset. The first two principal components are shown. Data are labelled and coloured by treatment group. Shaded areas indicate 95% confidence intervals. Samples (quality controls) of uninfected and infected human blood cluster outside of the range where the midgut samples (main) group. Additionally, replicates cluster closely indicating that replicate-to-replicate variability was minimal.

unknown entities. Compounds registered into the Laboratory Information Management System (LIMS) were used as a dataset for identification. Also, entities that were classified as unknown could still be identified by their recurrent nature (using both mass spectral and chromatographic characteristics). If a compound was found in both the extraction blank and the sample, it was excluded, unless the signal intensity was at least three times the intensity found in the blank.

A total of 512 individual compounds were identified; 511 were assigned to 8 superpathways and 75 subpathways (Table 2). The dataset is presented as biochemical abundance in each sample (Dataset Item 1). MetaboAnalyst tool suite [19] can be employed for further analysis. Some compounds were detected in certain sample(s) but were not detected in other samples, which yielded no values of the compounds in those samples. A Bayesian PCA (BPCA) method is designed to handle datasets with missing values [20]. Here, we used it to present the patterns of metabolites in the midgut samples in different conditions. Then, the data were log-transformed and followed by data scaling via Pareto scaling [21]. Figure 2 presents a PCA plot, which explains the variance within the dataset and provides an overview of the data with regard to questions related to the conditions under which samples were collected. In our study, blood samples, uninfected human blood (NB), and *P. falciparum*-infected blood (IB) cluster

TABLE 2: Compounds found in at least one sample mapped to a superpathway and subpathway. Abridged to show only subpathways that contain more than 10% of the superpathway compounds.

Superpathway	Compound number	Subpathway	Compound number
Amino acid	105	Cysteine, methionine, SAM, and taurine metabolism	11
		Glutamate metabolism	11
		Glycine, serine, and threonine metabolism	12
		Valine, leucine, and isoleucine metabolism	14
		Other	57
Carbohydrate	41	Amino sugars metabolism	5
		Fructose, mannose, galactose, starch, and sucrose metabolism	7
		Glycolysis, gluconeogenesis, and pyruvate metabolism	13
		Nucleotide sugars and pentose metabolism	16
Cofactors and vitamins	22	Hemoglobin and porphyrin metabolism	2
		Nicotinate and nicotinamide metabolism	7
		Pantothenate and CoA metabolism	3
		Other	10
Energy	8	Krebs cycle	6
		Oxidative phosphorylation	2
Lipid	148	Essential fatty acid (amino, branched, dicarboxylate, ester, monohydroxy, and BCAA metabolism)	29
		Long chain fatty acid	21
		Lysolipid	33
		Sterol/steroid	18
		Other	47
Nucleotide	31	Purine and pyrimidine metabolism	19
		Pyrimidine metabolism	12
Peptide	118	Dipeptide	111
		Other	7
Xenobiotics	38	Benzoate metabolism	11
		Chemical/drug	12
		Food component/plant	7
		Xanthine metabolism	5
		Other	3

distinctively from mosquito midgut samples. In addition, samples of 24 h post-uninfected (NBM 24 h) or post-infected (IBM 24 h) blood feeding are closer to each other, while samples of 48 h post-blood feeding (IBM 48 h, IBM 48 h) are clustered with the samples of sugar-fed (SM-L, SM-X) midguts. The latter is likely due to the completion of blood digestion and the midgut environment returning to sugar-fed conditions. Midgut samples fed on uninfected mouse blood are separated from human blood-fed samples, which is likely attributed to the distinction between human and mouse blood. Overall, the PCA output indicates that variation is largely attributed to the biological conditions, with limited variation induced by nonbiological factors present in the dataset. Figure 3 is a graphical representation of biochemical abundance (rescaled data).

3. Dataset Description

The dataset associated with this Dataset Paper consists of 3 items which are described as follows.

Dataset Item 1 (Table). The original scale of biochemicals, pathway assignment, and abundance. The column Superpathway shows the superpathway the biochemical is associated with and the column Subpathway shows the subpathway the biochemical is associated with. The column Platform presents the platform used for the identification of the biochemical (GC/MS, LC/MS positive, or LC/MS negative); the column Retention Index, the retention index (RI) associated with the biochemical; the column Mass, the mass of species; the column CAS, the Chemical Abstracts Service identifier

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Continued.

FIGURE 3: Graphical representation of biochemical abundance (rescaled data). Each box plot presents the abundance of one compound in all samples. * indicates compounds that have not been officially “plexed” (based on a standard), but we are confident in their identity; # indicates that samples without a PROmeasure (i.e., the media) are lost by this normalization, as well as any compounds that were only detected in these samples.

of the biochemical; the column PubChem, the PubChem identifier of the biochemical; the column KEGG, the KEGG identifier of the biochemical; and the column Group HMDB, the Human Metabolome Database (HMDB) identifier of the biochemical. Columns 12–41 present the biochemical abundance associated with samples. Values are normalized in terms of raw area counts.

Column 1: Biochemical Name
Column 2: Superpathway
Column 3: Subpathway
:
Column 39: NBM-X, 48H1

Column 40: NBM-X, 48H2t

Column 41: NBM-X, 48H3t

Dataset Item 2 (Table). The rescaled biochemicals, pathway assignment, and abundance. The column Superpathway shows the superpathway the biochemical is associated with and the column Subpathway shows the subpathway the biochemical is associated with. The column Platform presents the platform used for the identification of the biochemical (GC/MS, LC/MS positive, or LC/MS negative); the column Retention Index, the retention index (RI) associated with the biochemical; the column Mass, the mass of species; the column CAS, the Chemical Abstracts Service identifier of the biochemical; the column PubChem, the PubChem identifier of the biochemical; the column KEGG, the KEGG identifier of the biochemical; and the column Group HMDB, the Human Metabolome Database (HMDB) identifier of the biochemical. Columns 12–41 present the biochemical abundance associated with samples as denoted in Table 1. Each biochemical in original scale is rescaled to have median equal to 1. Then, missing values are imputed with the minimum.

Column 1: Biochemical Name

Column 2: Superpathway

Column 3: Subpathway

⋮

Column 39: NBM-X, 48H1

Column 40: NBM-X, 48H2t

Column 41: NBM-X, 48H3t

Dataset Item 3 (Table). The amount of protein found in each sample.

Column 1: Sample Identifier

Column 2: Bradford Protein Concentration

4. Concluding Remarks

In the mosquito midgut ecosystem, the metabolites can be derived from the mosquito host, microbial organisms in the gut microbiota, and ingested pathogens. This dataset presents a survey of the metabolomic landscapes in the *An. gambiae* midgut in response to different diets and to human malaria parasite infection. The dataset provides a valuable metabolic reference that will facilitate establishing a connection to transcriptome and proteome data related to the mosquito physiology and interactions with symbiotic microbes and *P. falciparum* in the midgut environment. The dataset allows users to further compare the metabolic dynamics under these conditions in the contexts with their own interest.

Dataset Availability

The dataset associated with this dataset paper is dedicated to the public domain using the CC0 waiver and is available at

<https://doi.org/10.1155/2017/8091749/dataset>. In addition, this dataset is accessible at the repository Dryad.

Disclosure

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health and the National Science Foundation. The current affiliation of Phanidhar Kukutla is as follows: Department of Anaesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Jiannong Xu and Shirley Luckhart designed the study. Phanidhar Kukutla collected midgut samples in the Xu laboratory, and Elizabeth K. K. Glennon and Bo Wang collected midgut samples in the Luckhart laboratory. Cody J. Champion, Jiannong Xu, and Shirley Luckhart analysed the data and Cody J. Champion prepared data deposit. Cody J. Champion, Jiannong Xu, Shirley Luckhart, and Elizabeth K. K. Glennon prepared the manuscript. All authors read and approved the final version of the manuscript.

Acknowledgments

The research reported here was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Awards nos. R01 AI080799 and R01 AI073745 to Shirley Luckhart and SC2GM092789 and SC1AI112786 to Jiannong Xu. Cody J. Champion was supported by the National Science Foundation Graduate Research Fellowship under Grant no. 1144468.

References

- [1] Programme GM, *World Malaria Report 2015*, Geneva, 2015.
- [2] Y. Wang, T. M. Gilbreath III, P. Kukutla, G. Yan, and J. Xu, "Dynamic gut microbiome across life history of the malaria mosquito *Anopheles gambiae* in Kenya," *PLoS ONE*, vol. 6, no. 9, Article ID e24767, 2011.
- [3] G. A. Preidis and P. J. Hotez, "The newest "omics"—metagenomics and metabolomics—enter the battle against the neglected tropical diseases," *PLoS Neglected Tropical Diseases*, vol. 9, no. 2, Article ID e0003382, 2015.
- [4] R. Aguilar, Y. Dong, E. Warr, and G. Dimopoulos, "Anopheles infection responses; laboratory models versus field malaria transmission systems," *Acta Tropica*, vol. 95, no. 3, pp. 285–291, 2005.
- [5] A. N. Dana, Y. S. Hong, M. K. Kern et al., "Gene expression patterns associated with blood-feeding in the malaria mosquito *Anopheles gambiae*," *BMC Genomics*, vol. 6, article 5, 2005.
- [6] O. Marinotti, Q. K. Nguyen, E. Calvo, A. A. James, and J. M. C. Ribeiro, "Microarray analysis of genes showing variable expression following a blood meal in *Anopheles gambiae*," *Insect Molecular Biology*, vol. 14, no. 4, pp. 365–373, 2005.

- [7] F. Oduol, J. Xu, O. Niaré, R. Natarajan, and K. D. Vernick, "Genes identified by an expression screen of the vector mosquito *Anopheles gambiae* display differential molecular immune response to malaria parasites and bacteria," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 21, pp. 11397–11402, 2000.
- [8] A. Padrón, A. Molina-Cruz, M. Quinones et al., "In depth annotation of the *Anopheles gambiae* mosquito midgut transcriptome," *BMC Genomics*, vol. 15, article no. 636, 2014.
- [9] E. Warr, R. Aguilar, Y. Dong, V. Mahairaki, and G. Dimopoulos, "Spatial and sex-specific dissection of the *Anopheles gambiae* midgut transcriptome," *BMC Genomics*, vol. 8, article 37, 2007.
- [10] J. C. Hoxmeier, B. D. Thompson, C. D. Broeckling et al., "Analysis of the metabolome of *Anopheles gambiae* mosquito after exposure to *Mycobacterium ulcerans*," *Scientific Reports*, vol. 5, article 9242, 2015.
- [11] W. Mamai, K. Mouline, C. Blais et al., "Metabolomic and ecdysteroid variations in *Anopheles gambiae* s.l. mosquitoes exposed to the stressful conditions of the dry season in Burkina Faso, West Africa," *Physiological and Biochemical Zoology*, vol. 87, no. 3, pp. 486–497, 2014.
- [12] J. H. M. Oliveira, R. L. S. Gonçalves, G. A. Oliveira, P. L. Oliveira, M. F. Oliveira, and C. Barillas-Mury, "Energy metabolism affects susceptibility of *Anopheles gambiae* mosquitoes to *Plasmodium* infection," *Insect Biochemistry and Molecular Biology*, vol. 41, no. 6, pp. 349–355, 2011.
- [13] C. Rivera-Perez, M. Nouzova, I. Lamboglia, and F. G. Noriega, "Metabolic analysis reveals changes in the mevalonate and juvenile hormone synthesis pathways linked to the mosquito reproductive physiology," *Insect Biochemistry and Molecular Biology*, vol. 51, no. 1, pp. 1–9, 2014.
- [14] M. Villar, N. Ayllón, P. Alberdi et al., "Integrated metabolomics, transcriptomics and proteomics identifies metabolic pathways affected by *Anaplasma phagocytophilum* infection in tick cells," *Molecular and Cellular Proteomics*, vol. 14, no. 12, pp. 3154–3172, 2015.
- [15] G. J. Patti, O. Yanes, and G. Siuzdak, "Innovation: metabolomics: the apogee of the omics trilogy," *Nature Reviews Molecular Cell Biology*, vol. 13, no. 4, pp. 263–269, 2012.
- [16] C. D. Broeckling, A. L. Heuberger, and J. E. Prenni, "Large scale non-targeted metabolomic profiling of serum by ultra performance liquid chromatography-mass spectrometry (UPLC-MS)," *Journal of Visualized Experiments*, no. 73, Article ID e50242, 2013.
- [17] H. G. Gika, G. A. Theodoridis, R. S. Plumb, and I. D. Wilson, "Current practice of liquid chromatography-mass spectrometry in metabolomics and metabonomics," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 87, pp. 12–25, 2014.
- [18] B. Wang, N. Pakpour, E. Napoli et al., "*Anopheles stephensi* p38 MAPK signaling regulates innate immunity and bioenergetics during *Plasmodium falciparum* infection," *Parasites and Vectors*, vol. 8, article 424, 2015.
- [19] J. G. Xia, I. V. Sinelnikov, B. Han, and D. S. Wishart, "MetaboAnalyst 3.0—making metabolomics more meaningful," *Nucleic Acids Research*, vol. 43, no. 1, pp. W251–W257, 2015.
- [20] W. Stacklies, H. Redestig, M. Scholz, D. Walther, and J. Selbig, "pcaMethods—a bioconductor package providing PCA methods for incomplete data," *Bioinformatics*, vol. 23, no. 9, pp. 1164–1167, 2007.
- [21] R. A. van den Berg, H. C. J. Hoefsloot, J. A. Westerhuis, A. K. Smilde, and M. J. van der Werf, "Centering, scaling, and transformations: improving the biological information content of metabolomics data," *BMC Genomics*, vol. 7, article no. 142, 2006.