Selection on time to parasite transmission shapes the host *Anopheles gambiae* transcriptional response and suggests immune evasion

Luís M. Silva [1,2] *

[1] Institute of Biology, University of Neuchâtel, Rue Emilie-Argand 11, 2000 Neuchâtel, Switzerland

[2] Department of Zoology, University of British Columbia, 6270 University Boulevard, Vancouver, BC V6T 1Z4, Canada

* Corresponding author: luis.silva@ubc.ca

ORCID

LMS: 0000-0002-8203-4006

Abstract

Understanding host-parasite interactions is of the utmost importance for the correct disease prediction, prevention and management. Hence, this study assessed the transcriptional response of the primary malaria vector, *Anopheles gambiae*, to infection with several lines of the prominent vector-control parasite, the microsporidian *Vavraia culicis*. These parasitic lines have been selected for early or late transmission within this host. Previous studies extensively described them phenotypically, differing in their virulence, infection dynamics and host exploitation. Using RNA sequencing, gene expression profiles were analyzed in mosquitoes infected with early-selected, late-selected, unselected (reference) *V. culicis* lines and uninfected controls. The results revealed distinct transcriptional changes associated with each parasite line. Early-selected parasites induced a broader immune response than late-selected ones. Differential expression of immune-related genes, including *Toll-interacting protein* and *Protein ERGIC-53*, suggests enhanced immune evasion in late-selected parasites. Additionally, significant changes were observed in pathways related to Golgi membrane function and oxidative stress response, particularly in response to early-selected parasites. These findings highlight the evolutionary pressures shaping host-parasite coevolution and provide insights into how parasite transmission traits can influence mosquito immune response and regulation. This work offers a foundation for future studies on mosquito-microsporidia dynamics and potential applications for vector control strategies, particularly *Plasmodium*.

Introduction

As with any host, mosquitoes are exposed to and harbour numerous parasites [1–6], including microsporidia [7, 8]. Microsporidia are intracellular parasites that infect various hosts and occupy different ecological niches [7, 9–11]. Still, those infecting mosquitoes are predominantly aquatic, as they typically parasitize the aquatic stages of the mosquito life cycle [2]. These parasites have gained attention in the field of vector-borne diseases due to their potential use as biological agents for controlling *Plasmodium* and its mosquito host, *A. gambiae* [12–15]. Studies in both wild and laboratory settings have demonstrated that microsporidians possess *Plasmodium*-inhibitory potential when co-infecting mosquitoes with this parasite [8, 16].

Vavraia culicis is one of the most extensively studied microsporidians infecting mosquitoes and in the context of malaria research [15–18]. Although *A. gambiae* infections by *V. culicis* tend to result in low costs to adult survival, fecundity and developmental time [18, 19], little is known about the molecular basis of the host response to this microsporidian. In contrast, infections by *V. culicis* in the mosquito *Aedes aegypti*, the vector for dengue, are characterized by a weak immune response but significant investment in lectins and CLIP-domain serine proteases, both effectors involved in innate immunity and its signalling [20, 21]. Understanding how *A. gambiae* responds to microsporidian infection at the molecular level can provide valuable insights into host-parasite dynamics, immunity evolution, and the evolutionary pressures exerted by the parasite. Recent research has highlighted the role of host gene expression in response to pathogens, revealing that parasites can modulate or evade host immune responses to facilitate their transmission and survival [22–25]. In mosquitos, immune defences are mediated mainly by innate immunity, with critical pathways such as the Toll, Imd and JAK-STAT pathways involved in response to parasite recognition and elimination [26, 27] — however, the specific role of these pathways in response to microsporidian infections, like *V. culicis*, remains poorly understood. Therefore, this study sought to start by characterizing the *A. gambiae* host response to *V. culicis* at the molecular level.

Despite their small genome [11, 28, 29], microsporidia evolve rapidly [19]. Recent studies have demonstrated that artificial selection for early or late transmission can shape their virulence and several life history traits [19, 30, 31]. In particular, selection for early or late transmission of *V. culicis* in *A. gambiae* has been shown to influence parasite replication, survival outside the host, and virulence [19, 30]. For example, hosts infected with parasites selected for late transmission experienced the most virulent infections, resulting in shortened life cycles for both the host and parasite, increased host exploitation, and greater parasite growth [19]. Interestingly, late-selected parasites did not survive outside the host as well as the early-selected parasites, which could endure the environment for longer [30]. This trade-off is likely due to a better adaptation to the within-host

environment [32]. Thus, these selected lines of *V. culicis* differ significantly in their performance and virulence, particularly within the host and potentially offer variability in parasite strategies that can be examined further.

Considering this, the current study aimed to expand knowledge on the drivers of virulence and transmission evolution. In addition to the infection by the reference parasite, changes specific to selection for early or late transmission were explored through RNA sequencing (RNA-seq) of infected hosts. Through the inclusion of these selected lines, variability in parasite treatments was increased, which may provide insights into which expression patterns are more readily available or likely to evolve in response to different parasite challenges.

Hence, in this study, the transcriptional response of *A. gambiae* to infection by different lines of *V. culicis* - one selected for early transmission, one selected for late transmission and one unselected reference - was characterized. RNA-seq was used to analyze the differential expression of genes in mosquitoes infected or not (Fig. 1). The findings of this study provide the first insights into the *A. gambiae* transcriptional response to infection by *V. culicis*, while also demonstrating how parasite evolution drives changes in gene expression, impacting the mosquito's ability to respond to infection. This study complements previous work on these parasite lines, providing a multidimensional perspective of this particular host-parasite interaction. Moreover, the data presented here have important implications for parasite fitness and vector competence. Understanding the molecular basis of these interactions is crucial for designing effective vector control strategies that exploit the natural evolutionary pressures acting on mosquito-parasite systems [33].

Materials and Methods

1. Experimental model

The Kisumu strain of the mosquito host *Anopheles gambiae (s.s.)* and evolved lines of the microsporidian parasite *Vavraia culicis floridensis* were used for these experiments. These parasite lines had previously been selected for seven generations for: i) early transmission and shorter time within the host; or for ii) late transmission and longer time within the host; or iii) they come from an unselected reference, our laboratory stock [19]. Further details on the selection experiment and its maintenance are described in Silva and Koella (2024). All experiments and rearing were conducted under standard laboratory conditions (26 ± 1ºC, 70 ± 5% relative humidity, and 12 h light/dark).

2. Samples collection

Freshly hatched *A. gambiae* larvae were reared in groups of 50 individuals per petri dish (120 mm diameter × 17 mm) containing 80 ml of distilled water. Larvae were fed daily with Tetramin Baby

fish food according to their age: 0.04, 0.06, 0.08, 0.16, 0.32, and 0.6 mg/larva for 0, 1, 2, 3, 4, and 5 days or older, respectively. On the second day of larval development, individuals were exposed to 10,000 spores per larva of their respective *V. culicis* treatment, or not exposed in the case of the uninfected treatment. At pupation, individuals were transferred into 50 ml falcon tubes with approximately 10 ml of distilled water and were allowed to emerge as adults. Female mosquitoes were transferred to individual cups with humid filter paper and given a constant 6% sucrose solution until death. On day 10 of adulthood, all live female mosquitoes were collected, snap-frozen in liquid nitrogen, and stored at -80ºC until RNA extraction. Transcriptomic profiles were assessed at day 10 of adulthood, as previous studies on these lines have shown this to be when all hosts produce infective spores [19]. Thus, a stronger host response and more pronounced signatures of parasite selection were anticipated for this time point.

3. RNA extractions

RNA extractions were performed using Qiagen's RNAeasy Universal Plus kit following the provided protocol. Before beginning the protocol, ten female mosquitoes per treatment were pooled in a 1.5 mL microcentrifuge tube with 1000 µl of Qiazol reagent (from the RNAeasy Universal Plus kit by Qiagen) briefly mixed, and two stainless steel beads (Ø 2 mm) were added. Each tube was homogenized using a Qiagen TissueLyser LT at a frequency of 30 Hz for three minutes. The Qiagen kit protocol was then followed, resulting in a 50 µl eluted solution of RNA in nuclease-free water at a concentration of approximately 300 ng/µl for each treatment. RNA samples were standardized to 10 ng/µl and delivered on dry ice to the Genomic Technologies Facility (University of Lausanne, Switzerland) for further processing, as described below.

4. RNA sequencing

RNA quality was assessed using a Fragment analyzer (Agilent Technologies). The poly-A tails and adapters were removed from single-end reads, and quality was trimmed with Cutadapt version 4.8 [34]. Reads matching ribosomal RNA sequences were removed with fastq_screen version 0.11.1 [35]. Reads longer than 40nt were aligned against the concatenated genomes of *Anopheles gambiae* (NCBI: GCA_943734735.2) and *Vavraia culicis floridensis* (NCBI: GCA_000192795.1) using STAR version 2.7.10a [36]. The number of read counts per gene locus was summarized with htseq-count version 0.11.2 [37] using gene annotation. RNA-seq libraries were prepared from 250 ng of total RNA with the Illumina Stranded mRNA Prep reagents (Illumina) using a unique dual indexing strategy and following the official protocol, which was automated on a Sciclone liquid handling robot (PerkinElmer). Libraries were quantified by a fluorometric method (QubIT, Life Technologies), and their quality was assessed on a Fragment Analyzer (Agilent Technologies). Sequencing was performed on an Illumina NovaSeq 6000 for 150 cycles single read. Sequencing data were demultiplexed using the bcl2fastq2 Conversion Software (version 2.20, Illumina).

5. Statistical analysis

All the analyses were performed with R version 4.3.0. in RStudio version 3034.06.2+561. The following packages were used for the statistical analysis and diagnosis, and for visualising and plotting the data: "Glimma" [38], "cowplot" [39], "ggrepel" [40], "RColorBrewer" [41], "maptools" [42], "sp" [43], "edgeR" [44], "MASS" [45], "limma" [46], "knitr" [47], "ggplot2" [48] and "dplyr" [49]. Figure panels were created and edited with BioRender.com and Microsoft PowerPoint 16.16.27.

The high quality of the RNA-seq data alignment was verified using RSeQC version 5.0.1 [50]. Alignment rates to both genomes exceed 93.5% for all samples (SI Fig. S1a). A very few number of reads aligned to *V. culicis*, ranging from 127 to 5,393 (SI Fig. S1b). *A. gambiae* samples exhibit good sequencing depth (SI Fig. S1c) and diverse gene representation. The gene body coverage is biased towards the 3' end, suggesting partial RNA degradation, but this pattern is consistent across all samples.

Genes with low counts were filtered out according to the rule of 3 count(s) per million (cpm) in at least 4 sample(s). Library sizes were scaled using TMM normalization and log-transformed into counts per million or CPM (EdgeR package version 3.42.4) [44]. Statistical quality controls were performed through density gene expression distribution, clustering, and sample PCA (SI Fig. S2). Differential expression was computed with limma-trend approach [46] by fitting all samples into one linear model. Next, the following group of comparisons were performed:

- Infected with unselected reference *vs.* uninfected (1 comparison).
- Infected with selected parasite *vs.* infected with unselected reference (10 comparisons).

A moderated t-test was used for each contrast. The Benjamini-Hochberg method computed an adjusted p-value for each comparison, controlling for false discovery rate (FDR, Adj. p). Then, a host gene ontology (GO, biological processes) enrichment analysis was carried out on differentially expressed (DE) genes through Panther 19.0 and "vectorbase.org" using the A. gambiae genome as a reference (Fig. 3 and 4).

Afterwards, a candidate gene analysis was performed (Fig. 4) taking into account the differently expressed biological processes in selected parasites, comparatively to the unselected reference (Fig. 3). For each DE transcript of the GO:0000139 (Golgi membrane), GO:0000302 (Response to reactive oxygen species) and GO:0002376 (Immune system process) classes, a linear model was run with fold-change as a dependent variable and parasite treatment as an explanatory variable. These comparisons

allow us to test for differential fold change (compared to the unselected reference) for early- and late-selected parasites.

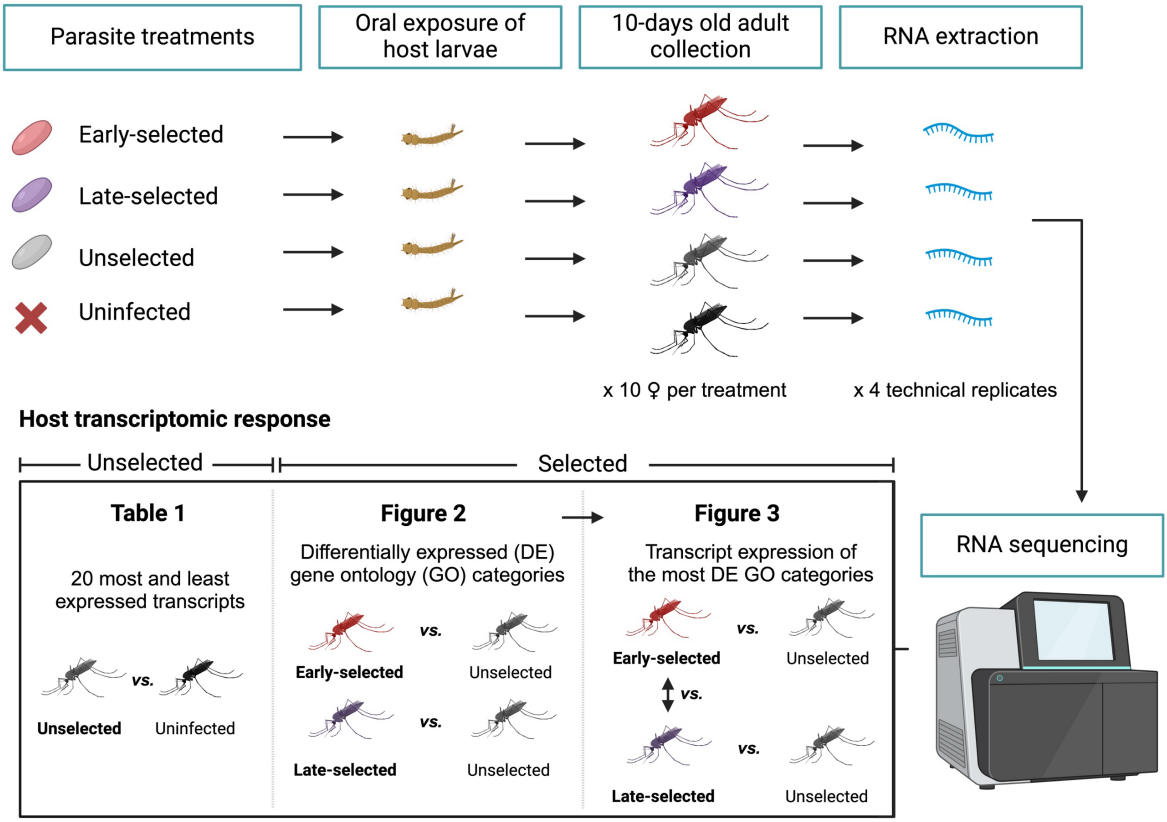

Figure 1. Illustration of the experimental workflow and results. Naive *Anopheles gambiae* larvae were infected, or not, with one of three lines of the parasite *Vavraia culicis,* totalizing in the following treatments: 1) infection with spores selected for early transmission; 2) infection with spores selected for late transmission; 3) infection with unselected reference spores; 4) uninfected. Adult mosquitoes were collected at day 10 in pools of 10 individuals per selection replicate for RNA extraction (10 females x 5 selection replicates). We then assessed differences in host response to the presence/absence of the parasite. First, we compared the 20 most and least expressed transcripts in reference-infected mosquitoes, comparatively to the uninfected ones (Table 1). Second, we categorized all DE transcripts according to their respective gene ontology (GO) biological process. We then presented which GO categories are upregulated/downregulated in infection by the Early- and Late-selected parasite, comparatively to the reference parasite (Fig. 2). Third, we selected candidate GO categories from Figure 2 that differ the most between early- and late-selected and we compared the expression of the transcripts in for each of the categories (Fig. 3) and if they differ between selected regimes.

Results

This study aimed to characterize the transcriptional response of *A. gambiae* to different lines of the microsporidian parasite *V. culicis*. Infections were performed using an unselected reference parasite and two treatment groups, each with five selection replicates, which had been artificially selected for early or late transmission within this mosquito host [19]. Our findings suggest infection by each parasite treatment resulted in distinct transcriptomic profiles, which differed in their host immune response to each parasite. Moreover, we also observed the involvement of several biological processes, namely Golgi apparatus-associated signalling and response to reactive oxygen species (ROS). The following sections describe and discuss these and other findings in detail.

Transcriptional profiles were measured on day 10 of adulthood. Across the twelve samples, 20 to 38 million reads were obtained, covering over 9,900 genes. The reads were successfully aligned to the *A. gambiae* genome (SI Fig. S1), and quality control confirmed their validity (SI Fig. S2).

First, the host transcriptional profile of *A. gambiae* infected with the reference parasite was compared to uninfected controls (Table 1). Out of the 9,900 annotated genes, 863 (8.7%) were significantly down-regulated, 1104 (11.1%) were significantly up-regulated, and 7953 (80.2%) showed no differential expression (DE). A complete list of the transcripts for this comparison can be found in SI Files S1-S11. Compared to uninfected controls, the 20 most up- and down-regulated transcripts were then focused on (Table 1).

For this comparison, although a higher number of up-regulated genes was found, the down-regulated genes exhibited higher fold changes. For instance, the most down-regulated gene was the *Putative uncharacterized protein DDB_G0271606*, with a negative fold change of 1680.95. In contrast, the most up-regulated gene was an uncharacterized gene with a positive fold change of 13.30.

Next, the DE profiles of *A. gambiae* infected by early—and late-selected parasite lines were analyzed and compared to the reference parasite (Fig. 2). After comparing transcript expression to the reference control, transcripts that were uniquely DE in each selected treatment and those shared between treatments were identified (Fig. 2ab).

For down-regulated genes, 81 genes were suppressed by early-selected parasites. In comparison, 71 genes were down-regulated by late-selected parasites, with 28 genes commonly down-regulated in both treatments (Fig. 2a). For up-regulated genes, 113 genes were up-regulated in early-selected parasites, compared to 52 in late-selected parasites, with 77 genes commonly up-regulated in both groups (Fig. 2b). These findings suggest that a higher number of DE genes are driven by early-

selected parasites compared to late-selected ones, at least during the 10th day of host adulthood. However, a notable overlap exists between both treatments (Fig. 2ab).

After categorizing the DE genes into their respective biological processes using Gene Ontology (GO) categories, a heatmap was plotted to show the mean fold-change for each GO category across parasite treatments (Fig. 2c). Most categories exhibited similar fold changes between early- and late-selected parasites. However, three categories stood out as potential candidates for explaining differences between the two selected treatments due to their distinct expression patterns and biological relevance: response to reactive oxygen species (ROS), Golgi membrane, and immune system processes.

Genes with DE and their respective fold-change relative to the reference control were plotted for each candidate GO category (Fig. 3). The Golgi membrane category had the highest number of DE genes and showed the greatest variation between early- and late-selected parasites, followed by the immune process and response to ROS categories. Fold-change differences between the selected treatments were also compared. Only two genes showed significant differences: *Protein ERGIC-53* ($df = 1$, $F = 7.69$, $p = 0.02$; Golgi membrane) and *Toll-interacting protein* ($df = 1$, $F = 9.70$, $p = 0.01$; Immune process). Both genes had a higher fold change in late-selected parasites than in early-selected ones. However, while *Protein ERGIC-53* was up-regulated, *Toll-interacting protein* was down-regulated (Fig. 3).

Table 1. Most up-and down-regulated genes in mosquitoes are infected with the reference parasite compared to those uninfected. The host transcriptomic profile concerning an unselected parasite during infection was compared to the uninfected host profile. The 20 transcripts with the highest and lowest expression, compared to the control, are presented. Fold change and *p*-value were calculated using the transcriptional expression of four technical replicates per treatment. See materials and methods for more details. The *p*-value was adjusted using the Benjamini-Hochberg correction method for each comparison, controlling for a false discovery rate ($p < 0.05$).

#	Gene ID	Gene name	Fold change	Adj. *p*
1	LOC3291958	Putative uncharacterized protein DDB_G0271606	-1680.95	< 0.001
2	LOC133393242	Uncharacterized	-306.13	< 0.001
3	LOC1281554	Protein CIMAP1D	-250.60	< 0.001
4	LOC1271479	Uncharacterized	-246.11	< 0.001
5	LOC133393316	Uncharacterized	-195.71	< 0.001
6	LOC1281083	Arginine kinase 1	-191.24	< 0.001
7	LOC1273786	Serine/arginine-rich splicing factor 2	-190.72	< 0.001
8	LOC133394019	Nucleosome assembly protein 1-like 1	-164.62	< 0.001
9	LOC133391046	Putative cyclin-dependent serine/threonine-protein kinase DDB_G0272797/DDB_G0274007	-161.63	< 0.001
10	LOC1271272	Uncharacterized	-149.47	< 0.001
11	LOC5667228	Uncharacterized	-133.19	< 0.001
12	LOC1269550	Uncharacterized	-126.11	< 0.001
13	LOC1270351	Uncharacterized	-123.92	< 0.001
14	LOC133391930	Uncharacterized	-121.19	< 0.001
15	LOC1277465	Uncharacterized	-119.77	< 0.001
16	LOC133391419	Uncharacterized	-107.89	< 0.001
17	LOC11175695	Uncharacterized	-107.77	< 0.001
18	LOC1278052	Cytosol aminopeptidase	-103.75	< 0.001
19	LOC1267823	Uncharacterized	-95.91	< 0.001
20	LOC4576455	Dynein regulatory complex subunit 4	-89.14	< 0.001
1	LOC1281665	Uncharacterized	13.30	< 0.001
2	LOC1269260	Stress-activated protein kinase JNK	4.89	< 0.001
3	LOC1280522	Neuropeptide FF receptor 2	4.40	0.039
4	LOC4576208	Protein obstructor-E	4.15	< 0.001
5	LOC1271120	Protein regulator of cytokinesis 1	4.10	< 0.001
6	LOC1279262	CLIP domain-containing serine protease B15	4.10	< 0.001
7	LOC1269964	Chymotrypsin-2-like	3.99	0.007
8	LOC1270108	Uncharacterized	3.97	0.020
9	LOC1281383	Ninjurin-2	3.96	0.002
10	LOC1279267	CLIP domain-containing serine protease B15-like	3.89	0.002
11	LOC1277042	Uncharacterized	3.76	< 0.001
12	LOC133392142	Uncharacterized	3.62	0.030
13	LOC1271472	Uncharacterized	3.62	< 0.001
14	LOC4577422	Uncharacterized	3.43	< 0.001
15	LOC133393335	Uncharacterized	3.39	0.005
16	LOC133393411	Uncharacterized	3.33	0.001
17	LOC1280050	Annulin	3.27	< 0.001
18	LOC1269628	Monocarboxylate transporter 12	3.27	0.011
19	LOC1271676	Uncharacterized	3.18	< 0.001
20	LOC1277959	UDP-glucosyltransferase 2	3.12	< 0.001

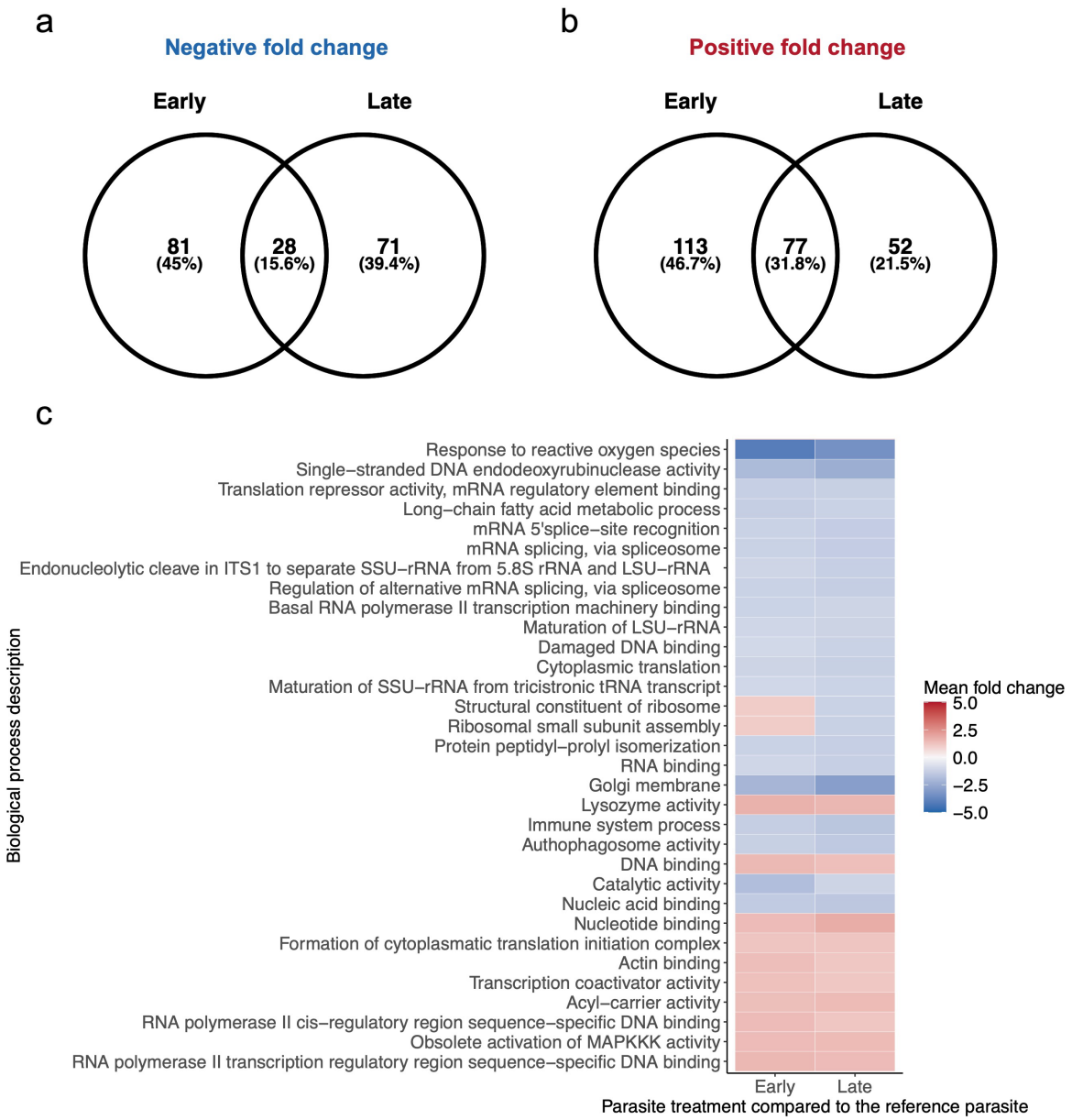

Figure 2. Host transcriptional signatures of infection by selected parasites compared to unselected parasites. Venn diagrams showing the number of differentially expressed (DE) genes (adjusted *p*-value < 0.05) exclusively and commonly found between hosts infected with Early- or Late-selected parasites, in comparison to the unselected reference parasite. Negative (a) and positive (b) differential expression are presented. (c) Heatmap illustrating mean fold change in expression per biological processes families of significantly DE transcripts across all replicates of the same selection regime. Gene ontology annotation was performed using PANTHER 19.0 and VectorBase.

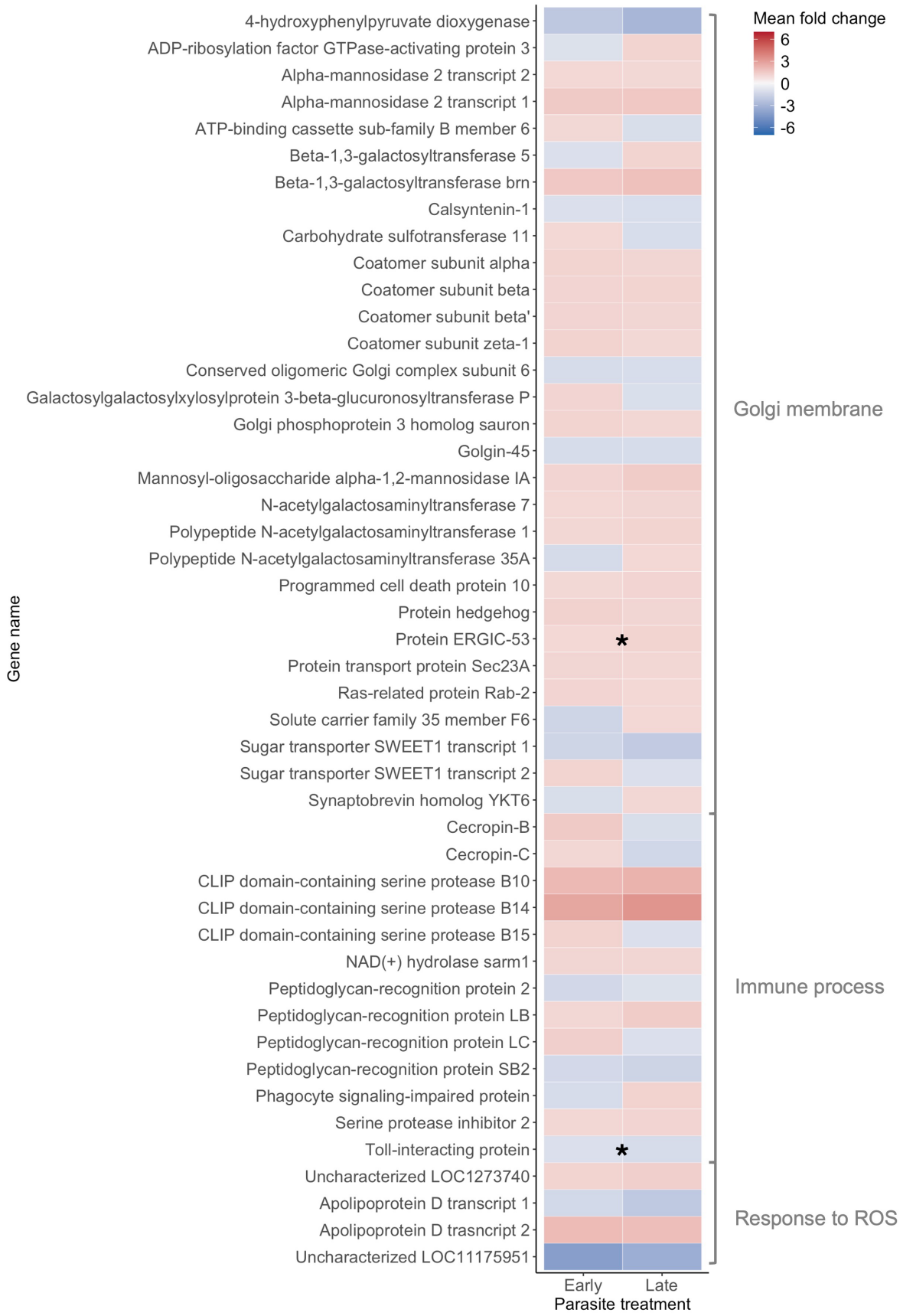
Mean fold change
6
3
0
-3
-6
4-hydroxyphenylpyruvate dioxygenase
ADP-ribosylation factor GTPase-activating protein 3
Alpha-mannosidase 2 transcript 2
Alpha-mannosidase 2 transcript 1
ATP-binding cassette sub-family B member 6
Beta-1,3-galactosyltransferase 5
Beta-1,3-galactosyltransferase brn
Calsyntenin-1
Carbohydrate sulfotransferase 11
Coatomer subunit alpha
Coatomer subunit beta
Coatomer subunit beta'
Coatomer subunit zeta-1
Conserved oligomeric Golgi complex subunit 6
Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase P
Golgi phosphoprotein 3 homolog sauron
Golgin-45
Mannosyl-oligosaccharide alpha-1,2-mannosidase IA
N-acetylgalactosaminyltransferase 7
Polypeptide N-acetylgalactosaminyltransferase 1
Polypeptide N-acetylgalactosaminyltransferase 35A
Programmed cell death protein 10
Protein hedgehog
Protein ERGIC-53
Protein transport protein Sec23A
Ras-related protein Rab-2
Solute carrier family 35 member F6
Sugar transporter SWEET1 transcript 1
Sugar transporter SWEET1 transcript 2
Synaptobrevin homolog YKT6
Cecropin-B
Cecropin-C
CLIP domain-containing serine protease B10
CLIP domain-containing serine protease B14
CLIP domain-containing serine protease B15
NAD(+) hydrolase sarm1
Peptidoglycan-recognition protein 2
Peptidoglycan-recognition protein LB
Peptidoglycan-recognition protein LC
Peptidoglycan-recognition protein SB2
Phagocyte signaling-impaired protein
Serine protease inhibitor 2
Toll-interacting protein
Uncharacterized LOC1273740
Apolipoprotein D transcript 1
Apolipoprotein D trasncript 2
Uncharacterized LOC11175951
Gene name
Golgi membrane
Immune process
Response to ROS
Early
Late
Parasite treatment

Figure 3. Differentially host expressed candidate genes during infection with selected parasites. Fold difference in host expression of Golgi membrane (GO:0000139), response to reactive oxygen species (GO:0000302) and immune system process (GO:0002376) markers when compared to infection with the unselected control. Early- and late-selected fold change parasites was then compared to each other using a linear model (see materials and methods for details). Although the expression of all genes presented differed from the unselected parasite, only the Toll-interacting and ERGIC-53 transcripts' fold change significantly differed between early and late-selected parasites. Late-selected had a lower expression of *Toll-interacting protein* ($df = 1$, $F = 9.70$, $p = 0.01$) and *protein ERGIC-53* ($df = 1$, $F = 7.69$, $p = 0.02$) than early parasites, represented by an asterisk.

Discussion

Our study highlights distinct transcriptional changes when infected with different lines of this parasite (Fig. 1). The findings discussed below provide key insights into this host-parasite interaction and might have consequences for this vector's competence.

1. Mosquito response to reference V. culicis infection

When the host's transcriptomic response to infection was compared (Table 1), a large number of differentially expressed (DE) genes were observed (see File S2 for the complete list of transcripts and respective expression). Such a broad response has been reported in the mosquito *Aedes aegypti* when infected with *V. culicis* [20]. From this response, only a few DE immune markers were identified, with two being notably over-expressed: *Cecropin-B*, an antimicrobial peptide known for its fungicidal properties *in vitro* [51, 52], and *Ninjurin-2*, a non-apoptotic cell death protein induced by the Toll pathway and which plays a key role in the dipteran response to microbial infections [53, 54].

Significant over-expression of regeneration markers was also detected. For instance, *Annulin* was found to be responsible for peritrophic matrix formation in the gut and early resilience to infection [55], and *Obstructor-E*, a member of a multigene family, was shown to regulate insect cuticle development [56]. Additionally, *Rab11*, essential for circulatory iron internalization, was over-expressed [57], further validating previous findings on the role of iron during infection with *V. culicis*, where iron sequestration by the parasite has been hypothesized [31].

An overall down-regulation of sensory genes was observed, including *Prestin*, which increases gut pH [58] and *CIMAP1D*, involved in sound-sensory function [59]. *Arginine kinase 1*, known to participate in ATP-to-ADP conversion and consequent energy release in certain bacteria [60], was found to be severely down-regulated in infected mosquitoes (Table 1). Interestingly, most Mucin proteins were up-regulated, except for *Mucin-6*, which was down-regulated. Mucin proteins are recognized for their

role in epithelial lubrification, immune modulation and inflammation [61, 62]. Notably, *Mucin-6* is exclusive to the *A. gambiae* gut, which has a protective function against parasites such as *Plasmodium spp* [63].

2. Mosquito response to selected lines of V. culicis

Next, the focus is turned to infections caused by parasite lines that have undergone selection for early or late transmission within *A. gambiae* [19] (Fig. 2 and 3). As expected, fewer genes were responsible for differences between selected and unselected parasites, i.e., 299 for early-selected and 228 for late-selected (Fig. 2ab), than unselected and uninfected, i.e., 1967 genes. Notably, a broader transcriptional response was induced by early-selected parasites than by late-selected ones (when they are compared to the reference control), as indicated by the high number of DE genes in both selection treatments (Fig. 2ab). Despite this, a core set of genes was found to be commonly up- or down-regulated across both parasite treatments (Fig. 2ab).

To better understand the differences in transcriptomic profiles from infections with selected parasites, DE genes were categorized into biological processes using GO analysis (Fig. 2c). Most of the differences between early- and late-selected parasites were grouped into three GO biological processes: Golgi membrane, immune processes, and response to reactive oxygen species (ROS). These findings are unsurprising, given the biological relevance of these categories for infection [64–66], particularly from the host's perspective. Equally interestingly, the distinct host transcriptomic profile observed in response to infection between the early and late selected parasites indicates that selection did affect not only phenotypic traits, as shown in previous studies [19, 30, 31], but also drove significant parasite evolution, causing a substantial shift in the host's response to the selected parasites (Fig. 2 and 3).

The three GO categories were treated as candidate categories, along with their associated genes, which are believed to play crucial roles in modulating host-parasite dynamics and therefore warrant further study (Fig. 3). First, the Golgi membrane was considered, as it has recently gained attention for its critical role in innate immune signalling, preceding other organelles such as mitochondria and the endoplasmic reticulum (ER) during infection [64]. As a result, the apparent difference observed in the host response to early- or late-selected parasite treatments was unsurprising (Fig. 3). The genes associated with the Golgi membrane include coatomers, galactosyltransferases, and mannosidases, which are involved in vesicle transport and carbohydrate metabolism with the ER. The richness and differential expression of Golgi membrane-associated genes suggest that some upregulation of effectors involved in glycosylation and trafficking processes in the Golgi apparatus occurs across both parasite treatments (Fig. 3). One gene, *ERGIC-53*, stood out due to its significantly different fold-change between host responses to infections by early-selected and late-selected parasites. The protein

encoded by *ERGIC-53* is involved in glycoprotein trafficking. Early transmission-selected parasites seem to induce higher production of ERGIC-53 and, consequently, increased investment in glycoprotein trafficking compared to late-selected parasites (Fig. 3). Although the precise role of ERGIC-53 in the immune response of *A. gambiae* remains unclear, it can be hypothesized, based on existing literature, that glycoproteins (e.g., peptidoglycan recognition proteins or PGRPs) serve as essential information markers for the host, vital for accurate recognition and elimination of parasites [67, 68].

Second, variation in the expression of immune-related genes was analyzed, including *peptidoglycan-recognition proteins* (PGRPs), *CLIP domain serine proteases* and *cecropins*, among others (Fig. 3). A notable difference was observed between the host response to early- and late-selected parasite treatments, with many of the DE immune related-genes being downregulated upon infection with late-selected parasites (Fig. 3). Interestingly, both *cecropin-B* and *cecropin-C* were overexpressed in response to early-selected parasites compared to the unselected reference, suggesting a more robust antimicrobial response. However, this pattern was not observed in response to late-selected parasites, but rather the opposite.

Although it was hypothesized that the immune response to late-selected parasites would be stronger due to their higher virulence, time and replication within the host, the data suggests otherwise. Two hypotheses are proposed to explain this finding. The first hypothesis involves the recognition of *V. culicis* as a virulent parasite. As discussed in the manuscript, early-selected parasites induce a larger Golgi membrane response, potentially leading to better recognition of PGRPs than infection by late-selected parasites. This could explain the rapid up-regulation of immune markers in response to early-selected parasites. The second hypothesis concerns the balance between the costs and benefits of the immune response and its association with the presence/absence of damage-associated molecular patterns (DAMPs) [69]. As with any other life history trait, investment in immunity is expected to trade off with other physiological traits [70]. As suggested by Rolff and Lazzaro (2012), the type of intensity of the immune response may result from the recognition of a combination of molecules, alerting the host to the parasite (PGRPs) and/or to the damage inflicted by the infection (DAMPs). If so, infection by the more virulent late-selected parasite could lead to a more costly and damaging infection, potentially overwhelming the host. In this scenario, the host may prioritize reproductive fitness over survival, which aligns with previous observations of increased reproductive investment during infection by late-selected parasites [19]. This phenomenon has been documented in various infections as predicted by life history theory [70]. In extreme cases, the investment shift to fecundity can be lifelong due to a shortening of the host's survival. In this specific case, we refer to this phenomenon as terminal investment, which may occur when a host invests its resources into fecundity due to an inability to survive the infection [71–73]. It is unclear if this shift in fecundity and immunity

is temporary or observed throughout the host life, given that both fecundity and RNA-seq were performed for the same infection time point and did not represent the whole infection process. It is noteworthy to say the individuals in this study were not blood-fed, necessary for their reproduction. Although we cannot confirm there is a trade-off with fecundity or other life history traits, we are aware *A. gambiae* invests in its first clutch of eggs before having a blood meal, which would represent at least a partial investment in fecundity. Moreover, from this category, the *Toll-interacting protein* was found to have significantly different fold changes between the infection by each of the two selection parasite treatments (Fig. 3). This gene, which produces a protein important for parasite recognition [74–77], was slightly more up-regulated in response to early-selected parasites than to late-selected ones (Fig. 3). This supports the idea that early-selected parasites may be more recognizable by the host's immune system. However, the extent to which parasite selection has affected the presence or absence of specific PGRPs that enable host recognition remains unclear. Nonetheless, it is reasonable to assume that late-selected parasites might have evolved higher levels of immune evasion within *A. gambiae*.

Third, only a few genes were found to be DE in response to ROS, showing mixed expression patterns (Fig. 3). Overall, the early-selected parasites elicited a stronger host response to ROS than late-selected ones, suggesting higher oxidative stress within the host. This is likely due to a heightened immune response to this infection treatment.

Conclusions

In conclusion, the transcriptional response of *A. gambiae* to infection by *V. culicis* has revealed curious and significant insights into this host-parasite interaction and its evolutionary history. Moreover, the findings in this study further explain many of the results measured in previous studies concerning this particular *V. culicis* selection lines for early and late transmission [19, 30, 31]. Differential expression of genes related to immune response, Golgi membrane function, and oxidative stress were observed, with distinct patterns emerging between infections with reference to unselected, early-selected and late-selected parasite lines. Overall, early-selected parasites elicited a broader and more intense immune response, characterized by the upregulation of immune-related genes such as cecropin and PGRPs and enhanced glycoprotein trafficking through the Golgi apparatus. In contrast, late-selected parasites appeared to trigger a weaker immune response and, therefore, potentially evolved immune evasion mechanisms.

Altogether, the results of this study demonstrate that selection pressures on *V. culicis* not only influence its phenotypic traits but also drive significant changes in the host gene expression, revealing complex host-parasite co-evolutionary dynamics. Further investigation is needed to elucidate, confirm

and quantify the specific molecular mechanisms involved, particularly in relation to immune evasion strategies and the balance between host survival and reproductive fitness. Additionally, it is important to position this study within the context of the microsporidian's ability to interfere with *Plasmodium* development, as previously mentioned. Vector-borne parasites, such as *Plasmodium spp.*, transmitted by this host, have long transmission cycles within their vectors. It is reasonable to expect that co-infection of Plasmodium spp. and microsporidia may inadvertently select for late-transmitted microsporidian parasites, which could have significant consequences for the spread of both parasites within a population. Although the mechanism by which microsporidians suppress *Plasmodium* development remains unclear, the data from this study suggest that it may not result from collateral damage caused by a host immune response directed against the microsporidian. Instead, resource competition between the parasites might be the source of this inhibitory relationship. Moreover, the lack of immunity towards late-transmitted *V. culicis* might highlight the potential of this parasite as a biological vector-control as it suggests selection to reduce its prevalence in a population is low. Further research is required to understand this interaction fully, but our study stands as a foundation for them. These findings provide valuable insights that could inform future vector control strategies, particularly in the context of the evolutionary pressures exerted by parasites on mosquito populations and the diseases they vector.

Author contributions

LMS designed and conducted the experiments, analyzed the data and wrote the manuscript.

Acknowledgements

I thank Tiago G. Zeferino, Paula Escuer Pifarré, and Amjad Khalaf for their advice and experimental support. I also thank the Lausanne Genomic Technologies Facility (University of Lausanne, Switzerland) team, Johann Weber, Julien Marquis, Sandra Calderon Copete, and Viviane Praz for the library preparation, sequencing, and data analysis. I especially thank Jacob C. Koella for constant feedback, support, and critical comments on the first draft of the manuscript. The project was supported by the Basler Stiftung für biologische Forschung and LMS by SNF grant 310030_192786.

References

[1] Silva LM, Acerbi G, Amann M, et al. Exposure to Pseudomonas spp. increases Anopheles gambiae insecticide resistance in a host-dependent manner. *bioRxiv* 2023; 2011–2023.

[2] Ranasinghe HAK, Amarasinghe LD. Naturally occurring microbiota associated with mosquito breeding habitats and their effects on mosquito larvae. *Biomed Res Int*; 2020.

[3] Dong Y, Manfredini F, Dimopoulos G. Implication of the mosquito midgut microbiota in the

defense against malaria parasites. *PLoS Pathog* 2009; 5: e1000423.

[4] Caragata EP, Tikhe C V, Dimopoulos G. Curious entanglements: interactions between mosquitoes, their microbiota, and arboviruses. *Curr Opin Virol* 2019; 37: 26–36.

[5] Duncan AB, Agnew P, Noel V, et al. The consequences of co-infections for parasite transmission in the mosquito A edes aegypti. *J Anim Ecol* 2015; 84: 498–508.

[6] Bozic J, Capone A, Pediconi D, et al. Mosquitoes can harbour yeasts of clinical significance and contribute to their environmental dissemination. *Environ Microbiol Rep* 2017; 9: 642–648.

[7] Becnel JJ, Weiss LM. *Microsporidia: pathogens of opportunity*. John Wiley & Sons Incorporated, 2014.

[8] Herren JK, Mbaisi L, Mararo E, et al. A microsporidian impairs Plasmodium falciparum transmission in Anopheles arabiensis mosquitoes. *Nat Commun* 2020; 11: 1–10.

[9] Keeling PJ, Fast NM, Law JS, et al. Comparative genomics of microsporidia. *Folia Parasitol (Praha)* 2013; 52: 8–14.

[10] Keesey IW, Koerte S, Khallaf MA, et al. Pathogenic bacteria enhance dispersal through alteration of Drosophila social communication. *Nat Commun* 2017; 8: 265.

[11] Keeling PJ, Fast NM. Microsporidia: biology and evolution of highly reduced intracellular parasites. *Annu Rev Microbiol* 2002; 56: 93–116.

[12] Koella JC, Lorenz L, Bargielowski I. Microsporidians as evolution-proof agents of malaria control? *Adv Parasitol* 2009; 68: 315–327.

[13] Bukhari T, Pevsner R, Herren JK. Microsporidia: a promising vector control tool for residual malaria transmission. *Front Trop Dis* 2022; 3: 957109.

[14] Koella JC, Saddler A, Karacs TPS. Blocking the evolution of insecticide-resistant malaria vectors with a microsporidian. *Evol Appl* 2012; 5: 283–292.

[15] Lorenz LM, Koella JC. The microsporidian parasite Vavraia culicis as a potential late life–acting control agent of malaria. *Evol Appl* 2011; 4: 783–790.

[16] Bargielowski I, Koella JC. A possible mechanism for the suppression of Plasmodium berghei development in the mosquito Anopheles gambiae by the microsporidian Vavraia culicis. *PLoS One* 2009; 4: e4676.

[17] Vyas-Patel N. The Suppression of Plasmodium berghei in Anopheles coluzzii infected later with Vavraia culicis. *bioRxiv* 2023; 2002–2023.

[18] Zeferino TG, Koella JC. Host-specific effects of a generalist parasite of mosquitoes. *Sci Rep* 2024; 14: 18365.

[19] Silva LM, Koella JC. Complex interactions in the life cycle of a simple parasite shape the evolution of virulence. *bioRxiv* 2024; 2024.01.28.577571.

[20] Desjardins CA, Sanscrainte ND, Goldberg JM, et al. Contrasting host–pathogen interactions and genome evolution in two generalist and specialist microsporidian pathogens of mosquitoes. *Nat Commun* 2015; 6: 7121.

[21] Kanost MR, Jiang H. Clip-domain serine proteases as immune factors in insect hemolymph. *Curr Opin insect Sci* 2015; 11: 47–55.

[22] Clayton AM, Dong Y, Dimopoulos G. The Anopheles innate immune system in the defense against malaria infection. *J Innate Immun* 2014; 6: 169–181.

[23] Dos-Santos ALA, Carvalho-Kelly LF, Dick CF, et al. Innate immunomodulation to trypanosomatid parasite infections. *Exp Parasitol* 2016; 167: 67–75.

[24] McNeilly TN, Nisbet AJ. Immune modulation by helminth parasites of ruminants: implications for vaccine development and host immune competence. *Parasite*; 21.

[25] Arias-Rojas A, Frahm D, Hurwitz R, et al. Resistance to host antimicrobial peptides mediates resilience of gut commensals during infection and aging in Drosophila. *Proc Natl Acad Sci* 2023; 120: e2305649120.

[26] Yassine H, Osta MA. Anopheles gambiae innate immunity. *Cell Microbiol* 2010; 12: 1–9.

[27] Osta MA, Christophides GK, Vlachou D, et al. Innate immunity in the malaria vector Anopheles gambiae: comparative and functional genomics. *J Exp Biol* 2004; 207: 2551–2563.

[28] Nakjang S, Williams TA, Heinz E, et al. Reduction and expansion in microsporidian genome evolution: new insights from comparative genomics. *Genome Biol Evol* 2013; 5: 2285–2303.

[29] Khalaf A, Francis O, Blaxter ML. Genome evolution in intracellular parasites: Microsporidia and Apicomplexa. *J Eukaryot Microbiol* 2024; e13033.

[30] Silva LM, Koella JC. Virulence evolution: thinking outside of the host. *bioRxiv* 2024; 2005–2024.

[31] Silva LM, Vallat A, Koella JC. Mechanisms of host exploitation by a microsporidian parasite. *in prep*.

[32] Silva LM, King KC, Koella JC. Dissecting transmission to understand parasite evolution.

[33] Waterhouse RM, Kriventseva E V, Meister S, et al. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science (80-)* 2007; 316: 1738–1743.

[34] Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 2011; 17: 10–12.

[35] Wingett SW, Andrews S. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research*; 7.

[36] Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013; 29: 15–21.

[37] Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *bioinformatics* 2015; 31: 166–169.

[38] Su S, Law CW, Ah-Cann C, et al. Glimma: interactive graphics for gene expression analysis. *Bioinformatics* 2017; 33: 2050–2052.

[39] Wilke CO, Wickham H, Wilke MCO. Package 'cowplot'. *Streamlined plot theme plot Annot 'ggplot2*; 1.

[40] Slowikowski K, Schep A, Hughes S, et al. Package ggrepel. *Autom position non-overlapping text labels with 'ggplot2*.

[41] Neuwirth E, Neuwirth ME. *Package 'RColorBrewer'*. CRAN 2011-06-17 08: 34: 00. Apache License 2.0, 2011.

[42] Bivand R, Lewin-Koh N, Pebesma E, et al. Package 'maptools'. *Packag 'maptools'*.

[43] Pebesma E, Bivand R, Pebesma ME, et al. Package 'sp'. *Compr R Arch Netw*; 9.

[44] Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *bioinformatics* 2010; 26: 139–140.

[45] Ripley B, Venables B, Bates DM, et al. Package 'mass'. *Cran r* 2013; 538: 113–120.

[46] Ritchie ME, Phipson B, Wu DI, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; 43: e47–e47.

[47] Xie Y. knitr: a comprehensive tool for reproducible research in R. In: *Implementing reproducible research*. Chapman and Hall/CRC, 2018, pp. 3–31.

[48] Wickham H. *ggplot2: elegant graphics for data analysis*. springer, 2016.

[49] Wickham H, François R, Henry L, et al. Package 'dplyr'. *A Gramm Data Manip R Packag version*; 8.

[50] Wang L, Wang S, Li W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics* 2012; 28: 2184–2185.

[51] De Lucca AJ, Bland JM, Jacks TJ, et al. Fungicidal and binding properties of the natural peptides cecropin B and dermaseptin. *Med Mycol* 1998; 36: 291–298.

[52] De Lucca AJ, Bland JM, Vigo CB, et al. D-cecropin B: proteolytic resistance, lethality for pathogenic fungi and binding properties. *Med Mycol* 2000; 38: 301–308.

[53] Broderick S, Wang X, Simms N, et al. Drosophila Ninjurin A induces nonapoptotic cell death.

[54] Thomas T, De T Das, Sharma P, et al. Structural and functional prediction analysis of mosquito Ninjurin protein: Implication in the innate immune responses in Anopheles stephensi. *Int J Mosq Res* 2014; 1: 60–65.

[55] Wang Z, Wilhelmsson C, Hyrsl P, et al. Pathogen entrapment by transglutaminase—a conserved early innate immune mechanism. *PLoS Pathog* 2010; 6: e1000763.

[56] Behr M, Hoch M. Identification of the novel evolutionary conserved obstructor multigene family in invertebrates. *FEBS Lett* 2005; 579: 6827–6833.

[57] Wilcke M, Johannes L, Galli T, et al. Rab11 regulates the compartmentalization of early endosomes required for efficient transport from early endosomes to the trans-golgi network. *J Cell Biol* 2000; 151: 1207–1220.

[58] Hirata T, Czapar A, Brin L, et al. Ion and solute transport by Prestin in Drosophila and Anopheles. *J Insect Physiol* 2012; 58: 563–569.

[59] Senthilan PR, Piepenbrock D, Ovezmyradov G, et al. Drosophila auditory organ genes and genetic hearing defects. *Cell* 2012; 150: 1042–1054.

[60] Cozzone AJ. ATP-dependent protein kinases in bacteria. *J Cell Biochem* 1993; 51: 7–13.

[61] Dam TK, Brewer CF. Multivalent lectin—carbohydrate interactions: energetics and mechanisms of binding. *Adv Carbohydr Chem Biochem* 2010; 63: 139–164.

[62] Kufe DW. Mucins in cancer: function, prognosis and therapy. *Nat Rev Cancer* 2009; 9: 874–885.

[63] Shen Z, Dimopoulos G, Kafatos FC, et al. A cell surface mucin specifically expressed in the midgut of the malaria mosquito Anopheles gambiae. *Proc Natl Acad Sci* 1999; 96: 5610–5615.

[64] Tao Y, Yang Y, Zhou R, et al. Golgi apparatus: an emerging platform for innate immunity. *Trends Cell Biol* 2020; 30: 467–477.

[65] Bahia AC, Oliveira JHM, Kubota MS, et al. The role of reactive oxygen species in Anopheles aquasalis response to Plasmodium vivax infection. *PLoS One* 2013; 8: e57014.

[66] Andreyev AY, Kushnareva YE, Starkova NN, et al. Metabolic ROS signaling: to immunity and beyond. *Biochem* 2020; 85: 1650–1667.

[67] Davicino RC, Eliçabe RJ, Di Genaro MS, et al. Coupling pathogen recognition to innate immunity through glycan-dependent mechanisms. *Int Immunopharmacol* 2011; 11: 1457–1463.

[68] Sperandio M, Gleissner CA, Ley K. Glycosylation in immune cell trafficking. *Immunol Rev* 2009; 230: 97–113.

[69] Lazzaro BP, Rolff J. Danger, microbes, and homeostasis. *Science (80-)* 2011; 332: 43–44.

[70] Stearns SC. Trade-offs in life-history evolution. *Funct Ecol* 1989; 3: 259–268.

[71] Hudson AL, Moatt JP, Vale PF. Terminal investment strategies following infection are dependent on diet. *J Evol Biol* 2020; 33: 309–317.

[72] Clutton-Brock TH. Reproductive effort and terminal investment in iteroparous animals. *Am Nat* 1984; 123: 212–229.

[73] Duffield KR, Bowers EK, Sakaluk SK, et al. A dynamic threshold model for terminal investment. *Behav Ecol Sociobiol* 2017; 71: 1–17.

[74] Burns K, Clatworthy J, Martin L, et al. Tollip, a new component of the IL-1RI pathway, links IRAK to the IL-1 receptor. *Nat Cell Biol* 2000; 2: 346–351.

[75] Zhang G, Ghosh S. Negative regulation of toll-like receptor-mediated signaling by Tollip. *J Biol Chem* 2002; 277: 7059–7065.

[76] Didierlaurent A, Brissoni B, Velin D, et al. Tollip regulates proinflammatory responses to interleukin-1 and lipopolysaccharide. *Mol Cell Biol*.

[77] Valanne S, Wang J-H, Rämet M. The Drosophila toll signaling pathway. *J Immunol* 2011; 186: 649–656.

Supplementary information

Figure S1. Host RNA-sequencing and parasite gene counts.

Figure S2. Quality control clustering and PCAs.

File S1. Transcript expression and statistical contrasts for reference infection *vs.* no infection.

File S2. Transcript expression and statistical contrasts for early-selected (selection replicate 1) infection *vs.* reference infection.

File S3. Transcript expression and statistical contrasts for early-selected (selection replicate 2) infection *vs.* reference infection.

File S4. Transcript expression and statistical contrasts for early-selected (selection replicate 3) infection *vs.* reference infection.

File S5. Transcript expression and statistical contrasts for early-selected (selection replicate 4) infection *vs.* reference infection.

File S6. Transcript expression and statistical contrasts for early-selected (selection replicate 5) infection *vs.* reference infection.

File S7. Transcript expression and statistical contrasts for late-selected (selection replicate 1) infection *vs.* reference infection.

File S8. Transcript expression and statistical contrasts for late-selected (selection replicate 2) infection *vs.* reference infection.

File S9. Transcript expression and statistical contrasts for late-selected (selection replicate 3) infection *vs.* reference infection.

File S10. Transcript expression and statistical contrasts for late-selected (selection replicate 4) infection *vs.* reference infection.

File S11. Transcript expression and statistical contrasts for late-selected (selection replicate 5) infection *vs.* reference infection.