2022-04-18

The Role of Bacterial Genotype in the Persistence of the Microbiota of Drosophila melanogaster

Sarah J. Gottfredson
Brigham Young University

The Role of Bacterial Genotype in the Persistence of the Microbiota of *Drosophila melanogaster*

Sarah J. Gottfredson

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

John M. Chaston, Chair
Paul B. Frandsen
Jonathon T. Hill

Department of Plant and Wildlife Sciences

Brigham Young University

ABSTRACT

The Role of Bacterial Genotype in the Persistence of the Microbiota of *Drosophila melanogaster*

Sarah J. Gottfredson
Department of Plant and Wildlife Sciences, BYU
Master of Science

In this work we use the fruit fly *Drosophila melanogaster* as a model to identify bacterial genes that help bacteria to persist in their animal hosts. Early work on this model system established that dietary replenishment drives the composition of the *D. melanogaster* gut microbiota, and subsequent research has shown that some bacterial strains can colonize the fly for much longer than the flow of bulk diet through the gut. In this work we reveal that bacterial genes influence bacterial persistence by studying the correlation between bacterial genotype and persistence in the *D. melanogaster* gut microbiota. We performed an initial assay with 7 bacterial strains to establish that different bacterial strains persist differently independent of ingestion in the fly. We then repeated the assay with 41 different strains of bacteria in order to perform a metagenome wide association (MGWA) to find distinct bacterial genes that are significantly correlated with persistence. Based on the MGWA, we tested if 44 mutants from 6 gene categories affect bacterial persistence in the flies. We identified that transposon insertions in four flagellar genes (*fliF*, *flgH*, *fliI*, and *flgE*), one urea carboxylase gene, one phosphatidyl inositol gene, one bacterial secretion gene, and one antimicrobial peptide (AMP) resistance gene each significantly lowered colonization forming units (CFUs) that resulted from plating the gut content in *Drosophila melanogaster.* Follow-up experiments with the flagellar gene mutants revealed that each significant flagellar mutant was non-motile compared with the wild type. Taken together, these results reveal that there are bacterial genes that are involved in mechanisms, like bacterial motility, that help bacteria to persist in the fly gut.

Keywords: microbiome, microbiota, flagella, persistence, colonization, drosophila

ACKNOWLEDGEMENTS

I would first like to acknowledge my mentor, advisor, principal investigator, and teacher, John Chaston, who has introduced me to and guided me through my five years at BYU and without whom I may have never accomplished what I have. I would like to also acknowledge my committee, Paul Frandsen and Jonathan Hill, and our lovely graduate program advisor, Mariellen Tuckett, for supporting me through this unconventional journey that has been getting my Master of Science. I would like to thank my friends and family for encouraging me and helping me every step of the way. I would especially like to acknowledge and thank my loving husband who has stayed up many long nights with me as I worked on experiments and writing.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

The Role of Bacterial Genotype in the Persistence of the Microbiota of *Drosophila melanogaster*

Sarah J. Gottfredson, John M. Chaston

Department of Plant and Wildlife Sciences, Brigham Young University
Master of Science

ABSTRACT

In this work we use the fruit fly *Drosophila melanogaster* as a model to identify bacterial genes that help bacteria to persist in their animal hosts. Early work on this model system established that dietary replenishment drives the composition of the *D. melanogaster* gut microbiota, and subsequent research has shown that some bacterial strains can colonize the fly for much longer than the flow of bulk diet through the gut. In this work we reveal that bacterial genes influence bacterial persistence by studying the correlation between bacterial genotype and persistence in the *D. melanogaster* gut microbiota. We performed an initial assay with 7 bacterial strains to establish that different bacterial strains persist differently independent of ingestion in the fly. We then repeated the assay with 41 different strains of bacteria in order to perform a metagenome wide association (MGWA) to find distinct bacterial genes that are significantly correlated with persistence. Based on the MGWA, we tested if 44 mutants from 6 gene categories affect bacterial persistence in the flies. We identified that transposon insertions in four flagellar genes (*fliF*, *flgH*, *fliI*, and *flgE*), one urea carboxylase gene, one phosphatidyl inositol gene, one bacterial secretion gene, and one antimicrobial peptide (AMP) resistance gene each significantly lowered colonization forming units (CFUs) that resulted from plating the gut content in *Drosophila melanogaster*. Follow-up experiments with the flagellar gene mutants revealed that each significant flagellar mutant was non-motile compared with the wild type.

Taken together, these results reveal that there are bacterial genes that are involved in mechanisms, like bacterial motility, that help bacteria to persist in the fly gut.

INTRODUCTION

Drosophila melanogaster is a model for microbiome research

Drosophila melanogaster is one of the best studied genetic models in existence. Genetic research on *D. melanogaster* has been done for over a century and as a result the data and information on *D. melanogaster* genetics are very thorough and expansive (Jennings, 2011). One recent area of study that has gained attention in relation to *D. melanogaster* is that of microbiome studies. Microbiome studies in *D. melanogaster* have shown that the microbiota plays an important role in the health of *D. melanogaster* and is known to have many diverse effects on *D. melanogaster* phenotype and behavior (Pais et al., 2018; Sharon et al., 2010; Storelli et al., 2011). Some of the major phenotype effects of the microbiota on the host that have been studied include life history traits, like fecundity (Matthews et al., 2021), lifespan (Matthews et al., 2020), and starvation resistance (Judd et al., 2018).

The microbiota of Drosophila melanogaster is well characterized

D. melanogaster is well-suited to study host-microbe interactions for several different reasons. One reason is that the *D. melanogaster* gut microbiome, residing primarily in the foregut and crop of the fly, is well characterized and dominated by a few genera of bacteria (Dodge et al., 2021; Pais et al., 2018; Wong et al., 2011). The *D. melanogaster* microbiome generally contains 1-30 taxa of bacteria, mainly consisting of acetic acid bacteria and lactic acid bacteria (Broderick and Lemaitre, 2012). When compared to the over 500 taxa and diverse

genera that dominate vertebrates (Ley et al., 2008; Muegge et al., 2011), *D. melanogaster* is a much simpler microbiome model. Another reason is that the *D. melanogaster* microbiota is readily manipulated in lab conditions. *D. melanogaster* can be made axenic, mono-associated with specific bacterial strains, or poly-associated with several specific bacterial strains with high success (Koyle et al., 2016). The previous, expansive research done in *D. melanogaster* is another reason why *D. melanogaster* is a well-suited model for studying the microbiome (Douglas, 2018).

Early studies established a paradigm that the Drosophila melanogaster microbiome is established through ingestion of bacteria in the diet

Despite all the existing microbiome studies in *D. melanogaster*, how and why the microbiota establishes in the host remains poorly defined. Bacteria in the environment is introduced into the gut through eating via horizontal transfer. Although the diet introduces bacteria to the gut, it is highly likely that other host and microbial factors play a role in the establishment of the microbiota. An important early study suggested that the microbiome is established as flies ingest microbes in their diet, and that the microbial community is thereafter maintained by continuous consumption in the diet (Blum et al., 2013). Food travels through the entirety of the *D. melanogaster* gut in less than an hour (Lemaitre and Miguel-Aliaga, 2013), and the general suggestion was that the microbiota could only be present in the gut while in the bulk flow of food during this short transit time. At the same time, a separate study showed that the identity and abundance of the fly microbiota is inconstant within and across generations (Wong et al., 2013). Taken together, the primary initial conclusion was that the microbiota of *D. melanogaster* lives transiently in and does not colonize the fly gut.

Later studies on the Drosophila melanogaster microbiome showed bacteria can stably colonize the gut

Later work studying the colonization of the *D. melanogaster* gut has refined this early view. It is now understood that some bacterial strains colonize their hosts and others do not, and that bacteria from wild flies generally colonize their hosts better than congeneric laboratory strains of bacteria. Some of these works show that bacterial isolates can proliferate in the fly gut, allowing for stable association with the host independent of continuous uptake through diet (Ma and Leulier, 2018; Obadia et al., 2018; Pais et al., 2018). One major flaw with solely attributing bacterial establishment in the host to diet is that the bacteria content and abundance of diet does not match the bacterial content and abundance of the host microbiome. Another study showed that uric acid degradation genes and flagellar genes are primarily present in bacteria isolated from wild, but not laboratory*, D. melanogaster* lines, and suggested that these genes might be important in processes that are important for bacteria in wild-caught flies—such as colonization of the fly (Winans et al., 2017). Finally, the foregut is the region of the gut that is most abundantly colonized by bacteria (Dodge et al., 2021; Pais et al., 2018). The goal of my study is to extend the previous work by defining the bacterial genetic factors that influence persistence—the amount of time a bacterial strain resides in the fly gut, independent of continuous inoculation—of the microbiota of *D. melanogaster*.

Metagenome-wide association studies can function as a surrogate genetic screen

In order to identify bacterial genetic factors that may influence bacterial persistence, a metagenome-wide association (MGWA) can be used as a surrogate genetic screen (Chaston et

al., 2014). An MGWA compares a set of phenotypes and genomes to identify candidate genes that cause the change in phenotype. The MGWA can be used instead of a traditional forward genetic screen via random mutagenesis because the usage of different bacterial species creates diversity in phenotype and genotype that normally comes from mutagenesis (St Johnston, 2002). Since an MGWA looks at multiple species, it can provide phylogenetic information and identify important genes from multiple species, whereas a traditional genetic screen cannot provide either. Furthermore, the number of experimental measures needed to perform an MGWA is orders of magnitude smaller than the depth required in a traditional mutagenesis screen, as few as 30 or 40 different bacterial treatments can provide sufficient genetic resolution in an MGWA (Chaston et al., 2014), whereas a traditional reverse genetic approach would normally screen thousands of bacterial mutants. Two major drawbacks of using MGWA is that model used in the MGWA can greatly bias the resulting predictions, and that the MGWA results are predictions and must be validated by follow-up mutant analysis. Despite these limitations, numerous studies have demonstrated that MGWA can be a suitable way to identify candidate bacterial genes that affect a specific phenotype while avoiding the complexity and time-intensity of a traditional genetic screen (Chaston et al., 2014; Judd et al., 2018; Matthews et al., 2020; White et al., 2018).

Major goals of this study

In this study, I aim to define some bacterial genes that play a role in bacterial persistence of the microbiome in *D. melanogaster* and the effect those genes have on bacterial persistence. The MGWA can identify candidate genes for effect on persistence, mutants of these genes can be tested, and then further investigation of these genes can give insight into how they affect persistence. If distinct genes are identified, then it supports the idea that diet is not the only factor

that influences bacterial persistence in *D. melanogaster* and will provide key genetic insights to help define how microbes establish and persist within a model animal host.

MATERIALS AND METHODS

Bacterial and fly cultures

The fly stock was originally obtained from Mariana Wolfner at Cornell University and is a Wolbachia-free stock of Canton-S *Drosophila Melanogaster* flies. The stock flies were raised in an incubator on a 12-h light-dark cycle at 25°C. They were raised on a yeast-glucose (YG) diet that contains 10% brewer's yeast, 10% glucose, 1% agar, 0.084% propionic acid, and 0.08% phosphoric acid.

Stocks of bacterial strains were stored at -80°C. The bacterial strains were streaked for isolation onto clade-specific media plates and incubated at 30°C for 2-3 days. The different media types were: mMRS (Criterion C5932), LB (Apex 11-119), and potato dextrose (Sigma-Aldrich 70139-500G). Aerobic strains were placed in the incubator while anaerobic strains were put in carbon dioxide-flooded containers that were sealed and put in the incubator. One colony was then removed from the plates and placed in a tube of 5mL of clade-specific media broth and incubated at 30°C for 1-2 days. If the strains were aerobic the tubes of liquid broth were grown under oxic conditions by shaking. Aerotolerant strains were raised under microoxic conditions by remaining static. The bacteria were then diluted in a 1:8 dilution four times and normalized to OD_{600} of .01.

Axenic and mono-associated flies

All flies used in the persistence assay were derived as bacteria-free embryos before they were inoculated with bacteria. Fly eggs were made axenic by removing the chorion layer of eggs. To do this, stock flies were allowed to lay eggs for 18-20 hours on a plate made of 10% brewer's yeast, 10% glucose, 1% agar, and grape juice. The eggs were then collected and washed with a 0.6% hypochlorite solution twice for 2.5 minutes each. They were then washed three times with double distilled, autoclaved water. Then, 40-60 eggs were transferred into 50 mL vials containing 7.5 mL of autoclaved YG diet.

To mono-associate the flies, 50 µL of normalized bacteria were inoculated to axenic eggs in the sterile diet. The fly vials were then placed in a tray and put in an incubator at 25 ° C with a 12-hour light-dark cycle.

Persistence assay

Bacterial persistence with the flies was measured using an assay that frequently transferred adult flies to a sterile diet. Four days post bulk eclosion of the flies, 4 female flies from each vial were transferred under carbon dioxide anesthesia into separate wells of a 96 well plate with 150 µL of sterile diet at the bottom. The flies were then transferred to new 96-well plates containing sterile food 3 times a day (8AM, 1PM, 6PM) for 2 days. After the last transfer, the flies were placed in 1.7 mL microcentrifuge tubes with 150 µL of PBS and 150 µL of ceramic beads and homogenized in a GenoGrinder for 2 minutes at 1750rpm. The contents of the microcentrifuge tubes were then dilution plated and cultured in an incubator at 30°C until colonies were large enough to count (around 2-3 days), each colony was counted as one colony forming unity (CFU) and used as a measure of persistence. If the colonies were too dense to

count, then the 160 was used as the CFU number. The first analysis was performed with 7 strains to collect preliminary data (Table 1), and a Kruskal-Wallis test and pairwise Wilcoxon tests between sexes for each strain were performed to assess if the CFU per fly were significantly different. The second analysis was performed with 41 different strains for use in an MGWA (Table 2), and a Kruskal-Wallis test with post-hoc all-against-all-pairwise Wilcoxon tests were performed to determine significance groups of CFU per fly between all strains. The third analysis was performed with 44 mutants identified through the MGWA (Table 3), and a Kruskal-Wallis test was performed to test if each mutant was significantly different from the wild-type control. In each experiment, each treatment had triplicate vials in each of three separate experiments. Vials were discarded from the analysis if they were contaminated or the vial density was less than 30 flies. A vial was determined to be contaminated if undiluted aliquots bore more than 5 CFU of an unexpected colony morphology.

Metagenome-wide association

A metagenome-wide association (MGWA) was performed to predict bacterial genes that influence persistence. In order to perform the MGWA, amino acid sequences were obtained from GenBank for the exact strains we phenotyped. The amino acid sequences of 55 bacterial genomes (Table 4) were clustered into orthologous groups (OGs) using OrthoMCL (Li et al., 2003) with an inflation factor of 1.5. The MGWA was then performed using the R package, MAGNAMWAR (Sexton et al., 2018). The inputs for MAGNAMWAR were the clusters of orthologous groups assignments and the CFU per fly at the end of the persistence assay. The MGWA associated OG presence-absence patterns with bacterial persistence levels using a

Wilcoxon test. Resultant p-values were Bonferroni corrected and we set an arbitrary significance threshold of $p < 0.01$.

A KEGG enrichment analysis was then done to find functional categories enriched among the significant OGs. BlastKOALA (Kanehisa et al., 2016) used to assign KEGG functions to a representative sequence from each OG. Pathway significance was then determined by an FDR-corrected chi-square test.

Motility assay

A motility assay was performed by placing 1µL of OD_{600} normalized bacteria in PBS on mMRS plate with 2 g of agar per liter, replacing the normal plates. The plates were then left at room temperature for 48-72 hours and the diameter of the halo that forms on the plates was measured using a ruler. The tests were performed on flagellar mutants, urea carboxylase mutants, and the wild-type strain *Acetobacter fabarum* (Table 5). A linear mixed-effect model and ANOVA was run on the data in R. A Dunnett test was then run in R to test for significant differences between the mean diameter of each mutant versus the wild-type.

RESULTS

Serial transfers show differential bacterial colonization abundances

A high throughput assay was required in order to get the replication needed to find biological significance in an MGWA. Our goal was to identify a regime of serial transfers that would allow us to measure bacterial persistence across a range from highly proficient to completely deficient. We detected a wide range of CFU abundances in flies that were colonized from birth with seven different bacterial strains and then, when 3 day-old-adults, serially

transferred to sterile diets six times total over two days (Fig. 1). The abundances of bacteria in male and female flies reared with the seven bacterial strains from three different families varied in their abundance in the flies from 0 to > 88,000 CFU fly^{-1} (Kruskal Wallis (KW) $\chi^2_{13,380} = 191.09$, $p < 1 \times 10^{-15}$). There were no significant difference between the CFU abundances in male and female flies (Fig. 1), a somewhat unexpected finding because female flies usually bear higher bacterial loads than male flies. The finding that male and female flies bore comparable bacterial loads following frequent transfer to sterile diets suggests that, for these tested strains, the live and persistent microorganisms may occupy similar spaces and niches between the two sexes. We focused on just one sex in our subsequent assays and elected to study female flies, in part to make our results comparable with other analyses that measured life history traits of mono-associated females flies (Judd et al., 2018; Matthews et al., 2021, 2020). Overall, these results confirm that our assay allowed us to detect a range of persistence phenotypes by different bacterial strains that did not vary significantly with the sex of the flies.

Identification of bacterial genes significantly associated with persistence

To predict bacterial genes that contribute to bacterial persistence in the flies, we measured CFU loads of 41 different bacterial strains in the flies after six serial transfers and statistically associated the bacterial loads with bacterial gene presence-absence patterns. The 41 different strains showed a wide range of CFU abundances in the flies (Kruskal Wallis (KW) $\chi^2_{6,380} = 176.04$, $p < 1 \times 10^{-15}$), providing excellent strain-level phenotypic variation (Fig 2). Then, we performed an MGWA to identify bacterial genes whose presence was associated with this variation in CFU counts. We measured the association between bacterial persistence and 12,105 orthologous groups (OGs) that were collectively spread across 3,760 phylogenetic

distribution groups (PDGs, a unique set of taxa in which an OG is present). We determined that the presence-absence patterns of 385 OGs were statistically associated with bacterial persistence in the flies (Bonferroni-corrected $p < 0.01$).

From the OGs that were significantly associated with changes in bacterial persistence, we selected a subset to focus on in a mutant analysis. All genes tested were from a mutant library we have access to (White et al., 2018). We chose to test genes from 4 enriched pathways from the MGWA. We also chose to test flagellar assembly genes that appeared in our MGWA but were not significant because we previously found flagellar genes as possibly significant in persistence (Winans et al., 2017). We assigned the 385 significant KEGG IDs from the MGWA to KEGG pathways and performed a KEGG enrichment analysis which resulted in 10 enriched pathways (Table 6). This led us to choosing genes from 4 pathways: phosphatidyl inositol signaling system, bacterial secretion system, nicotinate and nicotinamide metabolism, and cationic antimicrobial peptide (CAMP) resistance. Taken together, this provided us with a total of 44 mutants to use in follow-up experiments and analysis (Table 3).

Bacterial genes involved in persistence

We measured the persistence phenotype of bacterial mutants for genes identified in the MGWA as a step towards validating MGWA predictions and identifying bacterial gene candidates that influence bacterial persistence with the flies. We identified 10 mutants (Kruskal Wallis (KW) $\chi^2_{56,1239} = 189.32$, $p < 10^{-15}$) in 8 different genes across 5 different categories of genes that significantly influenced bacterial persistence with the flies (Fig. 3). Nine of the mutants conferred a lower persistence phenotype, and one mutant, myo-inositol-monophosphatase, conferred a higher persistence phenotype. Six of the significant mutants were

flagellar assembly mutants, representing four flagellar genes. These efforts confirm that multiple different bacterial pathways influence bacterial persistence with the flies, and especially implicate bacterial flagellar genes as possible effectors of this phenotype.

Motility plays a role in persistence

In order to determine if the motility plays a role in bacterial persistence, motility tests were performed on all flagellar mutants, the wild-type, and a few non-flagellar mutants. Non-flagellar mutants were included to ensure that the mutagenesis performed to make the mutants did not render all mutants non-motile. All flagellar mutants that were tested were non-motile, while the wild-type *A. fabarum* and the non-flagellar mutants were all motile (ANOVA, linear mixed-model, $p < 5.176 \times 10^{-15}$, f-value = 13.662) (Fig. 4, Table 7). The non-motile flagellar mutants all create proteins that assemble into different parts of the flagella (Fig. 5). There was not a clear pattern to which flagellar assembly genes were significant in bacterial persistence, which may suggest redundancy for some functions, that some of our transposon insertion mutants incompletely inactivated the corresponding protein product, or that there simply is no pattern in which bacterial flagellar genes affect the persistence phenotype. The flagellar mutants being non-motile indicates that bacteria being able to move plays a role in bacterial persistence in flies.

DISCUSSION

Persistence of the microbiota

This work adds to the growing knowledge of how the microbiome is established and persists in *Drosophila melanogaster*. Host genes and behavior play a major role in introducing bacteria to the gut. Identifying bacterial genes involved in persistence shows that the host alone

is not responsible for the establishment and persistence of the gut microbiota. Interestingly, the poorest colonizing *Acetobacter* was isolated from a wild fly, contradicting the thought that wild fly isolates persist better than isolates from other sources. All this taken together is a beginning from which to further understand microbial persistence in hosts.

Flagellar genes and motility

The exact role that flagellar genes play in bacterial persistence with the fly is still unknown. However, it is likely that flagellar motility increases persistence because flagellar mutants that did not persist well with the flies were also non-motile. These results suggest that being motile may enhance bacterial persistence in the flies, perhaps by enabling bacteria to find favorable niches through bacterial taxis within the *Drosophila* gut. A favorable niche could be a region of the fly that has better quantity or quality of nutrients, less immune stress, or decreased competition from other microorganisms. The abundance of B12 (cobalamin) and B1 (thiamine) influence gut colonization (Costliow and Degnan, 2017; Frye et al., 2021; Goodman et al., 2009; Romine et al., 2017). Being motile could potentially allow bacteria to outcompete other non-motile bacteria for these important vitamins by moving to where these vitamins are located. Other studies have shown that bacteria with intact, motile flagella better colonize their hosts (Barbosa et al., 2017; Gorski et al., 2009; Nachamkin et al., 1993). It is also possible that these flagellar genes are involved in bacterial secretion. Bacterial secretion could create a favorable environment or be beneficial for the host. Bacterial surface components, including flagella, are also known to interact directly with the gut epithelium, which could also promote persistence (Liu et al., 2020). One thing to consider is that the four significant flagellar genes were *fliI, fliF, flgE,* and *flgH,* while the other flagellar genes tested were not significant.

Other significant genes

In addition to flagellar motility genes, we identified other genes that were significantly associated with variation in bacterial persistence with the flies. One AMP resistance gene, *mprF*, is involved in resistance to multiple antimicrobial peptides from the host and other competing microorganisms. AMP resistance has already been identified as a mechanism for stable association of gut commensals in the human gut during periods of host inflammation (Cullen et al., 2015). Pathogens are also able to stably colonize the gut due to AMP resistance (Goto et al., 2017), so it is possible that beneficial microbes also stably colonize the gut through AMP resistance. Some AMP resistance genes, such as *degP* (although not significant in our analysis) have been shown to help bacteria adjust to survive at high temperatures by decreasing temperature-sensitive growth (Strauch et al., 1989). Overall, AMP resistance can help microbes to colonize the gut by protecting against host antimicrobial peptides and helping the bacteria to adjust to a changing environment.

Of 3 bacterial secretion genes we tested, only *secB,* which is involved in the quorum sensing, protein export, and bacterial secretion system pathways, was significantly associated with variation in bacterial persistence with the flies. *SecB* specifically exports proteins, but is also involved in stress-responsive type II toxin-antitoxin (TA) systems (Guillet et al., 2019) which has been shown to be important in niche-specific colonization of *Escherichia coli* in humans (Norton and Mulvey, 2012). Another gene, *yajC,* is also involved in these pathways but not shown to be significant. This leads to the idea that only particular parts of the pathways are important in persistence. Interestingly, the sec bacterial secretion system is used in pathogenic factor to secrete virulence factors (Green and Mecsas, 2016), hinting at a possible interaction

with the host. Bacterial secretion can also aid the bacteria in establishing a hospitable niche by secreting products such as AMP resistance products which are known substrates of SecB (Sala et al., 2014).

The other two genes found to be significant were urea carboxylase and myo-inositol-1(-or 4)-monophosphatase. Urea carboxylase is involved in arginine biosynthesis, atrazine degradation, and metabolic pathways. Myo-inositol-1(-or 4)-monophosphatase is involved in streptomycin biosynthesis, inositol phosphate metabolism, metabolic pathways, biosynthesis of secondary metabolites, and phosphatidylinositol signaling system. Since these genes are not closely related to other significant genes and are involved in many different pathways, further research would be required to understand the role they play in bacterial persistence.

Uses and tools

Understanding the bacterial genes that influence bacterial persistence in flies can be useful in order to engineer microbes with increased or decreased persistence to use in experiments. For example, if microbes with differential persistence are used to study their effect on a specific host phenotype, using engineered bacteria with more equal persistence levels in the fly can remove the confounding variable of persistence differences on host phenotype. If a bacteria with a low persistence confers a smaller effect on host phenotype, then it cannot be determined if the smaller effect was due to that bacteria's interaction with the host or if it was due to being present in the fly for a shorter amount of time. Adding or removing genes that affect bacterial persistence in microbes can allow for control of bacterial persistence level to control for the effect of bacterial persistence on host phenotype.

Applicability to other organisms

Understanding how gut commensals persist in *Drosophila* is the first step to understanding bacterial persistence in other organisms. Mice models have also been used to study the microbiome and factors that affect persistence and colonization of the gut (Hooper et al., 1999; Lee et al., 2013). Combining the knowledge of bacterial persistence and colonization in *Drosophila* and mice can be useful to studying higher mammalian organisms as well, such as chimpanzees (Degnan et al., 2012) and even humans (Cullen et al., 2015). By understanding how bacteria persist in different organisms can help us to understand how to manipulate the microbiota. In particular, manipulating the microbiota in humans can lead to novel medical treatments and more effective probiotics.

Future directions

Although bacterial genes were identified that affect bacterial persistence in *Drosophila,* further studies could test more genes, more bacteria, and different hosts. Furthermore, the prospective pathways highlighted here should be studied further to understand the exact role they play in bacterial persistence. It should also be examined whether the pathways as a whole or parts of the pathways are important in persistence. Additionally, hypothetical proteins could also be studied to determine if undiscovered proteins play a role in persistence, since hypothetical proteins have already been identified that affect intake of B12, a known factor in bacterial persistence and colonization (Wexler et al., 2018). Overall, this study is only a small part in beginning to understand the role bacterial genes play in persistence.

CONCLUSION

This work adds to the growing studies surrounding the *Drosophila* microbiota and its' establishment and persistence. Showing that there are specific bacterial genes that affect the bacterial persistence in the fly gut adds to the understanding of how the fly gut microbiota establishes and persists. Specifically, this work provides possible genes, pathways, and mechanisms that can be further researched to understand the specific role they may play in bacterial persistence, particularly flagellar motility.

LITERATURE CITED

Barbosa, F. de O., Freitas Neto, O.C. de, Batista, D.F.A., Almeida, A.M. de, Rubio, M. da S., Alves, L.B.R., Vasconcelos, R. de O., Barrow, P.A., Berchieri Junior, A., 2017. Contribution of flagella and motility to gut colonisation and pathogenicity of Salmonella Enteritidis in the chicken. Braz. J. Microbiol. 48, 754–759. https://doi.org/10.1016/j.bjm.2017.01.012

Blum, J.E., Fischer, C.N., Miles, J., Handelsman, J., 2013. Frequent Replenishment Sustains the Beneficial Microbiome of Drosophila melanogaster. mBio 4. https://doi.org/10.1128/mBio.00860-13

Broderick, N., Lemaitre, B., 2012. Gut-associated microbes of Drosophila melanogaster. Gut Microbes 3, 307–321. https://doi.org/10.4161/gmic.19896

Chaston, J.M., Newell, P.D., Douglas, A.E., 2014. Metagenome-Wide Association of Microbial Determinants of Host Phenotype in Drosophila melanogaster. mBio 5, e01631-14. https://doi.org/10.1128/mBio.01631-14

Costliow, Z.A., Degnan, P.H., 2017. Thiamine Acquisition Strategies Impact Metabolism and Competition in the Gut Microbe Bacteroides thetaiotaomicron. mSystems 2, e00116-17. https://doi.org/10.1128/mSystems.00116-17

Cullen, T.W., Schofield, W.B., Barry, N.A., Putnam, E.E., Rundell, E.A., Trent, M.S., Degnan, P.H., Booth, C.J., Yu, H., Goodman, A.L., 2015. Antimicrobial peptide resistance mediates resilience of prominent gut commensals during inflammation. Science 347, 170–175. https://doi.org/10.1126/science.1260580

Degnan, P.H., Pusey, A.E., Lonsdorf, E.V., Goodall, J., Wroblewski, E.E., Wilson, M.L., Rudicell, R.S., Hahn, B.H., Ochman, H., 2012. Factors associated with the diversification

of the gut microbial communities within chimpanzees from Gombe National Park. Proc. Natl. Acad. Sci. 109, 13034–13039. https://doi.org/10.1073/pnas.1110994109

Dodge, R., Jones, E.W., Zhu, H., Obadia, B., Martinez, D.J., Wang, C., Aranda-Díaz, A., Aumiller, K., Liu, Z., Voltolini, M., Brodie, E.L., Huang, K.C., Carlson, J.M., Sivak, D.A., Spradling, A.C., Ludington, W.B., 2021. A gut commensal niche regulates stable association of a multispecies microbiota. https://doi.org/10.1101/2021.09.30.462663

Douglas, A.E., 2018. The Drosophila model for microbiome research. Lab Anim. 47, 157–164. https://doi.org/10.1038/s41684-018-0065-0

Frye, K.A., Piamthai, V., Hsiao, A., Degnan, P.H., 2021. Mobilization of vitamin B12 transporters alters competitive dynamics in a human gut microbe. Cell Rep. 37, 110164. https://doi.org/10.1016/j.celrep.2021.110164

Goodman, A.L., McNulty, N.P., Zhao, Y., Leip, D., Mitra, R.D., Lozupone, C.A., Knight, R., Gordon, J.I., 2009. Identifying genetic determinants needed to establish a human gut symbiont in its habitat. Cell Host Microbe 6, 279–289. https://doi.org/10.1016/j.chom.2009.08.003

Gorski, L., Duhé, J.M., Flaherty, D., 2009. The use of flagella and motility for plant colonization and fitness by different strains of the foodborne pathogen Listeria monocytogenes. PloS One 4, e5142. https://doi.org/10.1371/journal.pone.0005142

Goto, R., Miki, T., Nakamura, N., Fujimoto, M., Okada, N., 2017. Salmonella Typhimurium PagP- and UgtL-dependent resistance to antimicrobial peptides contributes to the gut colonization. PloS One 12, e0190095. https://doi.org/10.1371/journal.pone.0190095

Green, E.R., Mecsas, J., 2016. Bacterial Secretion Systems – An overview. Microbiol. Spectr. 4, 10.1128/microbiolspec.VMBF-0012–2015. https://doi.org/10.1128/microbiolspec.VMBF-0012-2015

Guillet, V., Bordes, P., Bon, C., Marcoux, J., Gervais, V., Sala, A.J., Dos Reis, S., Slama, N., Mares-Mejía, I., Cirinesi, A.-M., Maveyraud, L., Genevaux, P., Mourey, L., 2019. Structural insights into chaperone addiction of toxin-antitoxin systems. Nat. Commun. 10, 782. https://doi.org/10.1038/s41467-019-08747-4

Hooper, L.V., Xu, J., Falk, P.G., Midtvedt, T., Gordon, J.I., 1999. A molecular sensor that allows a gut commensal to control its nutrient foundation in a competitive ecosystem. Proc. Natl. Acad. Sci. 96, 9833–9838. https://doi.org/10.1073/pnas.96.17.9833

Jennings, B.H., 2011. Drosophila – a versatile model in biology & medicine. Mater. Today 14, 190–195. https://doi.org/10.1016/S1369-7021(11)70113-4

Judd, A.M., Matthews, M.K., Hughes, R., Veloz, M., Sexton, C.E., Chaston, J.M., 2018. Bacterial Methionine Metabolism Genes Influence Drosophila melanogaster Starvation Resistance. Appl. Environ. Microbiol. 84, e00662-18. https://doi.org/10.1128/AEM.00662-18

Kanehisa, M., Sato, Y., Morishima, K., 2016. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. J. Mol. Biol. 428, 726–731. https://doi.org/10.1016/j.jmb.2015.11.006

Koyle, M.L., Veloz, M., Judd, A.M., Wong, A.C.-N., Newell, P.D., Douglas, A.E., Chaston, J.M., 2016. Rearing the Fruit Fly Drosophila melanogaster Under Axenic and Gnotobiotic Conditions. JoVE J. Vis. Exp. e54219. https://doi.org/10.3791/54219

Lee, S.M., Donaldson, G.P., Mikulski, Z., Boyajian, S., Ley, K., Mazmanian, S.K., 2013. Bacterial colonization factors control specificity and stability of the gut microbiota. Nature 501, 426–429. https://doi.org/10.1038/nature12447

Lemaitre, B., Miguel-Aliaga, I., 2013. The Digestive Tract of Drosophila melanogaster | Annual Review of Genetics. Annu. Rev. Genet. 47, 377–404. https://doi.org/10.1146/annurev-genet-111212-133343

Ley, R.E., Lozupone, C.A., Hamady, M., Knight, R., Gordon, J.I., 2008. Worlds within worlds: evolution of the vertebrate gut microbiota. Nat. Rev. Microbiol. 6, 776–788. https://doi.org/10.1038/nrmicro1978

Li, L., Stoeckert, C.J., Roos, D.S., 2003. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. Genome Res. 13, 2178–2189. https://doi.org/10.1101/gr.1224503

Liu, Q., Yu, Z., Tian, F., Zhao, J., Zhang, H., Zhai, Q., Chen, W., 2020. Surface components and metabolites of probiotics for regulation of intestinal epithelial barrier. Microb. Cell Factories 19, 23. https://doi.org/10.1186/s12934-020-1289-4

Ma, D., Leulier, F., 2018. The importance of being persistent: The first true resident gut symbiont in Drosophila. PLOS Biol. 16, e2006945. https://doi.org/10.1371/journal.pbio.2006945

Matthews, M.K., Malcolm, J., Chaston, J.M., 2021. Microbiota Influences Fitness and Timing of Reproduction in the Fruit Fly Drosophila melanogaster. Microbiol. Spectr. 9, e00034-21. https://doi.org/10.1128/Spectrum.00034-21

Matthews, M.K., Wilcox, H., Hughes, R., Veloz, M., Hammer, A., Banks, B., Walters, A., Schneider, K.J., Sexton, C.E., Chaston, J.M., 2020. Genetic Influences of the Microbiota

on the Life Span of Drosophila melanogaster. Appl. Environ. Microbiol. 86, e00305-20. https://doi.org/10.1128/AEM.00305-20

Muegge, B.D., Kuczynski, J., Knights, D., Clemente, J.C., González, A., Fontana, L., Henrissat, B., Knight, R., Gordon, J.I., 2011. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science 332, 970–974. https://doi.org/10.1126/science.1198719

Nachamkin, I., Yang, X.H., Stern, N.J., 1993. Role of Campylobacter jejuni flagella as colonization factors for three-day-old chicks: analysis with flagellar mutants. Appl. Environ. Microbiol. 59, 1269–1273. https://doi.org/10.1128/aem.59.5.1269-1273.1993

Norton, J.P., Mulvey, M.A., 2012. Toxin-Antitoxin Systems Are Important for Niche-Specific Colonization and Stress Resistance of Uropathogenic Escherichia coli. PLOS Pathog. 8, e1002954. https://doi.org/10.1371/journal.ppat.1002954

Obadia, B., Keebaugh, E.S., Yamada, R., Ludington, W.B., Ja, W.W., 2018. Diet influences host–microbiota associations in Drosophila. Proc. Natl. Acad. Sci. 115, E4547–E4548. https://doi.org/10.1073/pnas.1804948115

Pais, I.S., Valente, R.S., Sporniak, M., Teixeira, L., 2018. Drosophila melanogaster establishes a species-specific mutualistic interaction with stable gut-colonizing bacteria. PLOS Biol. 16, e2005710. https://doi.org/10.1371/journal.pbio.2005710

Romine, M.F., Rodionov, D.A., Maezato, Y., Anderson, L.N., Nandhikonda, P., Rodionova, I.A., Carre, A., Li, X., Xu, C., Clauss, T.R.W., Kim, Y.-M., Metz, T.O., Wright, A.T., 2017. Elucidation of roles for vitamin B12 in regulation of folate, ubiquinone, and methionine metabolism. Proc. Natl. Acad. Sci. 114, E1205–E1214. https://doi.org/10.1073/pnas.1612360114

Sala, A., Bordes, P., Genevaux, P., 2014. Multitasking SecB chaperones in bacteria. Front. Microbiol. 5.

Sexton, C.E., Smith, H.Z., Newell, P.D., Douglas, A.E., Chaston, J.M., 2018. MAGNAMWAR: an R package for genome-wide association studies of bacterial orthologs. Bioinforma. Oxf. Engl. 34, 1951–1952. https://doi.org/10.1093/bioinformatics/bty001

Sharon, G., Segal, D., Ringo, J.M., Hefetz, A., Zilber-Rosenberg, I., Rosenberg, E., 2010. Commensal bacteria play a role in mating preference of Drosophila melanogaster. Proc. Natl. Acad. Sci. U. S. A. 107, 20051–20056. https://doi.org/10.1073/pnas.1009906107

St Johnston, D., 2002. The art and design of genetic screens: Drosophila melanogaster. Nat. Rev. Genet. 3, 176–188. https://doi.org/10.1038/nrg751

Storelli, G., Defaye, A., Erkosar, B., Hols, P., Royet, J., Leulier, F., 2011. Lactobacillus plantarum promotes Drosophila systemic growth by modulating hormonal signals through TOR-dependent nutrient sensing. Cell Metab. 14, 403–414. https://doi.org/10.1016/j.cmet.2011.07.012

Strauch, K.L., Johnson, K., Beckwith, J., 1989. Characterization of degP, a gene required for proteolysis in the cell envelope and essential for growth of Escherichia coli at high temperature. J. Bacteriol. 171, 2689–2696. https://doi.org/10.1128/jb.171.5.2689-2696.1989

Wexler, A.G., Schofield, W.B., Degnan, P.H., Folta-Stogniew, E., Barry, N.A., Goodman, A.L., 2018. Human gut Bacteroides capture vitamin B12 via cell surface-exposed lipoproteins. eLife 7, e37138. https://doi.org/10.7554/eLife.37138

White, K.M., Matthews, M.K., Hughes, R.C., Sommer, A.J., Griffitts, J.S., Newell, P.D., Chaston, J.M., 2018. A Metagenome-Wide Association Study and Arrayed Mutant

Library Confirm Acetobacter Lipopolysaccharide Genes Are Necessary for Association with Drosophila melanogaster. G3 Bethesda Md 8, 1119–1127. https://doi.org/10.1534/g3.117.300530

Winans, N.J., Walter, A., Chouaia, B., Chaston, J.M., Douglas, A.E., Newell, P.D., 2017. A genomic investigation of ecological differentiation between free-living and Drosophila-associated bacteria. Mol. Ecol. 26, 4536–4550. https://doi.org/10.1111/mec.14232

Wong, A.C.-N., Chaston, J.M., Douglas, A.E., 2013. The inconstant gut microbiota of Drosophila species revealed by 16S rRNA gene analysis. ISME J. 7, 1922–1932. https://doi.org/10.1038/ismej.2013.86

Wong, C.N.A., Ng, P., Douglas, A.E., 2011. Low-diversity bacterial community in the gut of the fruitfly Drosophila melanogaster. Environ. Microbiol. 13, 1889–1900. https://doi.org/10.1111/j.1462-2920.2011.02511.x

FIGURES

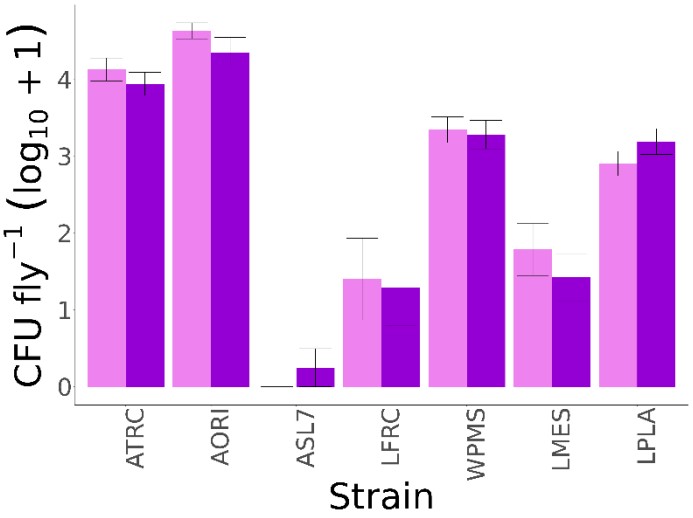

Figure 1. Seven strains show difference in bacterial persistence in flies. This bar plot shows the difference in CFU abundances by bacterial strain and fly sex. Significant differences between the sexes were determined by a Wilcoxon test. Table 1 reports the strain names of the 4-character codes.

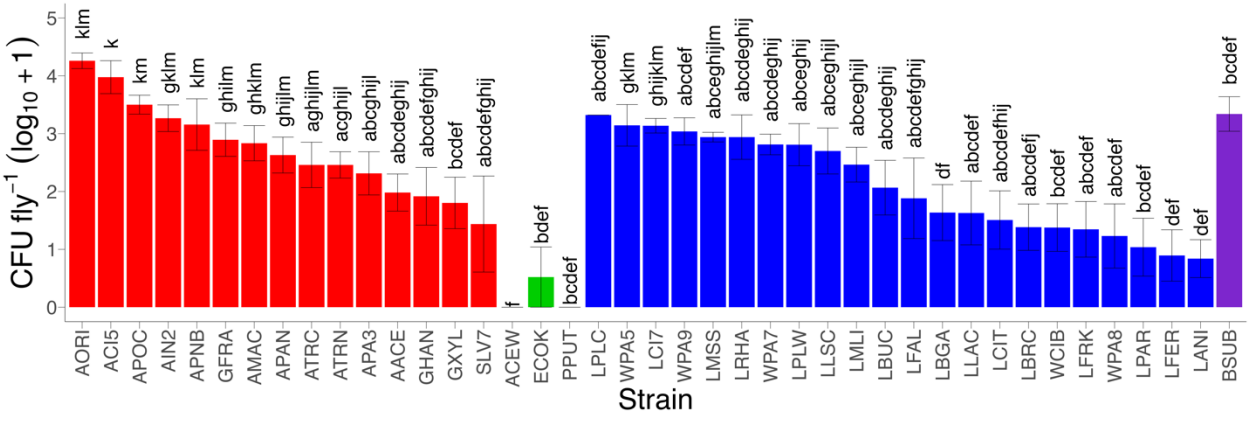

Figure 2. Differences in bacterial persistence in the fly of 41 different bacterial strains. This bar plot shows the $\log_{10}$+1 transformed CFU abundances per female fly of 41 different bacterial strains. Shading matches bacterial groups: acetic acid bacteria (red), gammaproteobacteria (green), lactic acid bacteria (blue), non-lactic acid firmicutes (purple). Different letters above each bar show significant differences in bacterial persistence between strains by a Kruskal-Wallis test followed by a post-hoc Dunn test. Table 2 reports the strain names of the 4-character codes.

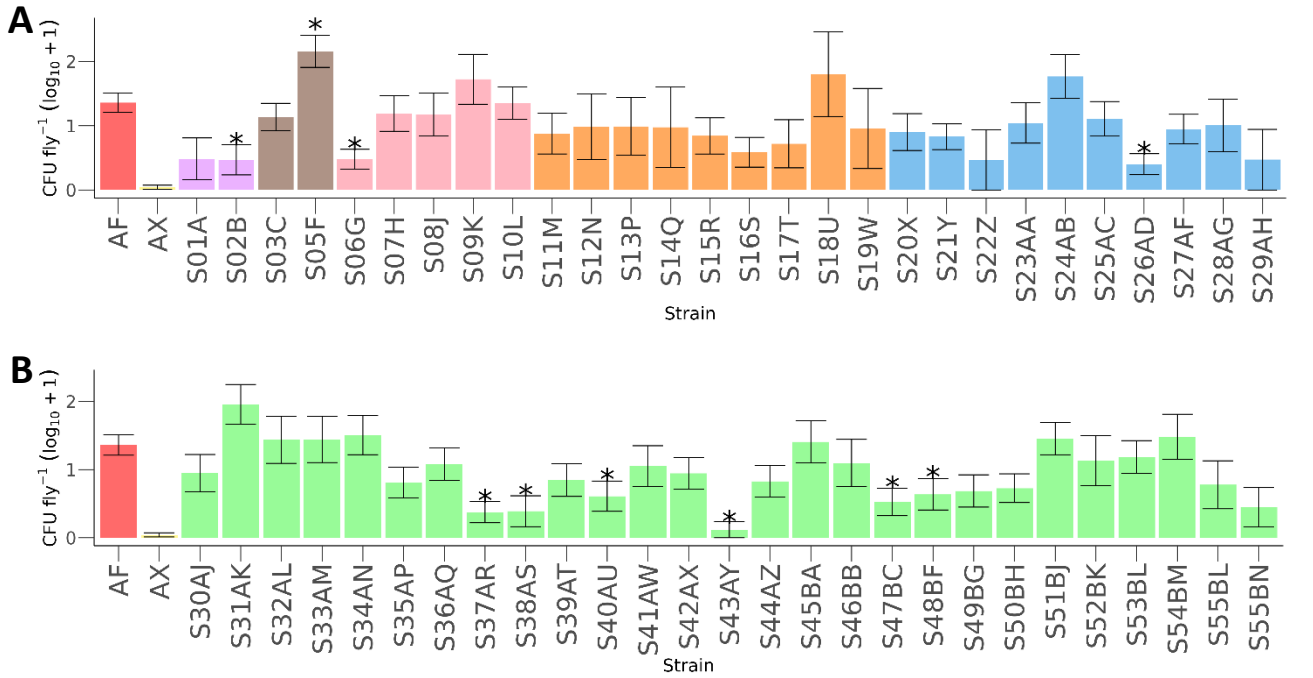

Figure 3. Differences in bacterial persistence in the fly of bacterial mutants against the wild-type show 10 significant mutants in bacterial persistence. This bar plot shows the $\log_{10}$+1 transformed CFU abundances per fly of the mutants tested in the persistence assay. Asterisks indicate that the bacterial mutants had a significantly different persistence phenotype from the wild-type bacterial strain (Acetobacter fabarum, AF), determined by a Dunnett test. (A) All non-flagellar mutants testing in the assay. Shading matches type of mutants tested: wild-type Acetobacter fabarum (red), bacteria-free flies (yellow), urea carboxylase (purple), phosphatidylinositol (brown), bacterial secretion (pink), nicotinate metabolism (orange), AMP resistance (blue). (B) All flagellar mutants tested in the assay, shown in green.

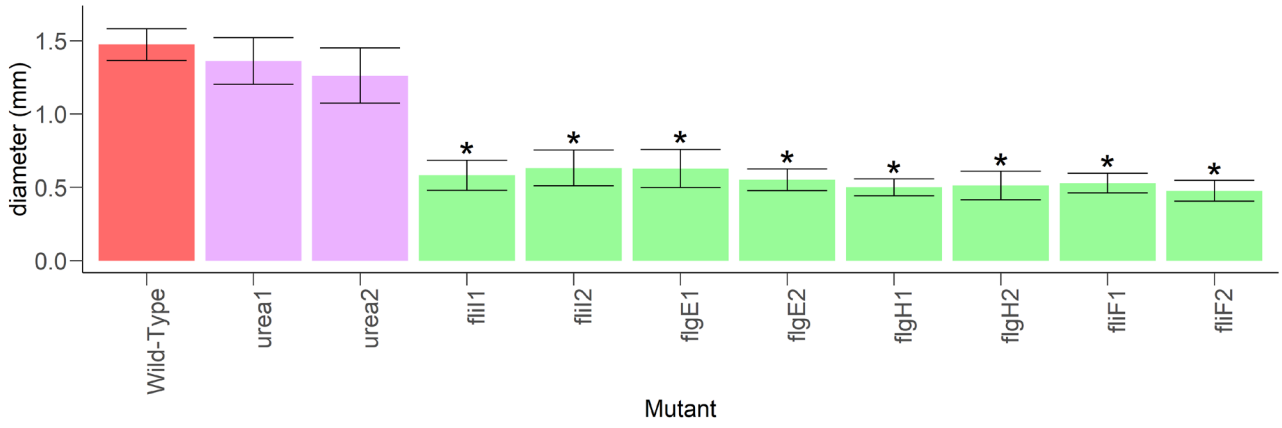

Figure 4. Motility tests reveal flagellar assembly mutants are non-motile. This bar plot shows the mean diameter measured on the halos of the motility tests for each mutant. Shading matches type of mutants tested: wild-type Acetobacter fabarum (red), bacteria-free flies (yellow), urea carboxylase (purple), flagellar assembly (green). Asterisks show an average diameter significantly different from the wild-type, determined by a Dunnett test based on a linear mixed-model.

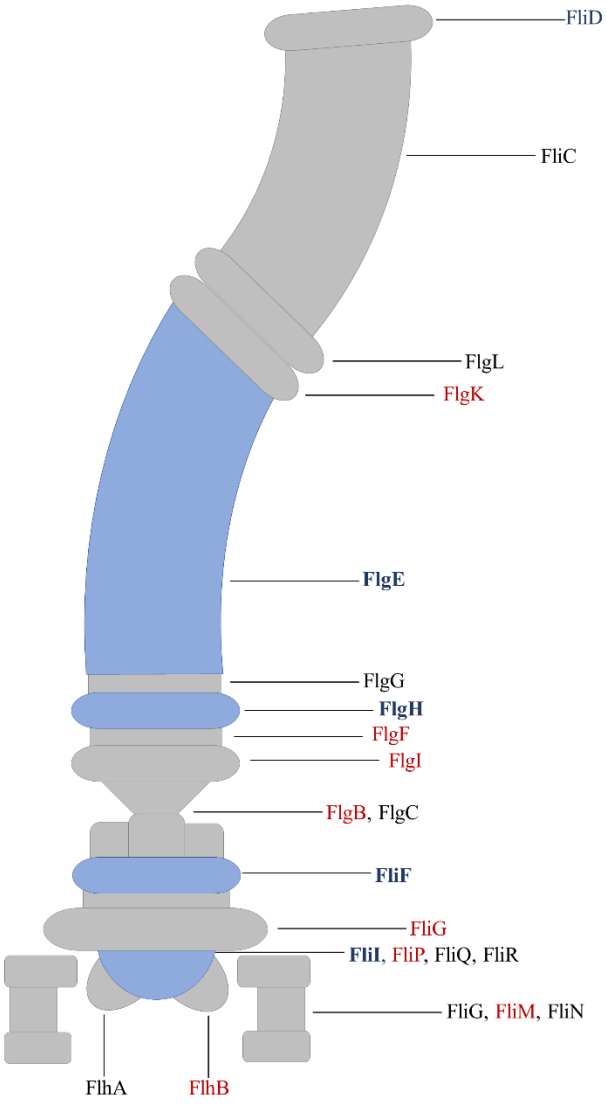

Figure 5. Significant flagellar assembly genes shown on a flagella diagram. This diagram shows the 4 genes found to be significant in persistence and where the proteins they code for assemble in the flagella. Color of text corresponds to if the gene was tested but not found significant (red), tested and found significant (blue), or not tested (black).

TABLES

Table 1. Bacterial strains used in an initial 7-strain persistence assay

Strain Name	**Abbreviation**	**Medium**
Acetobacter tropicalis DMCS_006	ATRC	MRS
Lactobacillus plantarum DMCS_001	LPLA	MRS
Acetobacter orientalis strain DmW_045	AORI	MRS
Acetobacter sp. SLV-7 DmW_125131	ASL7	MRS
Weissella paramesenteroides DmW_107100	WPMS	MRS
Lactobacillus fructivorans DMCS_002	LFRC	MRS
Leuconostoc mesenteroides	LMES	MRS

Table 2. Bacterial strains used in a 41-strain persistence assay

Strain Name	**Abbreviation**	**Medium**
Acetobacter aceti NBRC 14818	AACE	MRS
Acetobacter indonesiensis DmW_046	AIN2	MRS
Acetobacter malorum DmCS_005	AMAC	MRS
Acetobacter orientalis DmW_045	AORI	MRS
Acetobacter orientalis DmW_048	ACEW	MRS
Acetobacter pasteurianus 3p3	APA3	MRS
Acetobacter pasteurianus NBRC 101655	APAN	MRS
Acetobacter pasteurianus NBRC1 06471	APNB	MRS
Acetobacter pomorum DmCS_004	APOC	MRS
Acetobacter sp. DmW_043	ACI5	MRS
Acetobacter sp. SLV-7 DmW_125	SLV7	MRS
Acetobacter tropicalis DmCS_006	ATRC	MRS
Acetobacter tropicalis NBRC 101654	ATRN	MRS
Bacillus subtilis str. 168	BSUB	LB
Escherichia coli str. K—12 substr. MG1655	ECOK	LB
Fructilactobacillus fructivorans KCTC 3543	LFRK	MRS
Gluconobacter frateurii NBRC 181659	GFRA	POT
Komagataeibacter hansenii ATCC 23769	GHAN	POT
Komagataeibacter medellinensis NBRC 3288	GXYL	POT
Lacticaseibacillus paracasei str. IBB3423	LPAR	MRS
Lacticaseibacillus rhamnosus GG	LRHA	MRS
Lactiplantibacillus plantarum DmCS_001	LPLC	MRS
Lactiplantibacillus plantarum WCFS1	LPLW	MRS
Lactobacillus brevis gravesensis ATCC 27305	LBGA	MRS
Lactobacsillus bucheri NRRL B-30929	LBUC	MRS
Lactococcus lactis	LLAC	MRS
Lactococcus lactis subsp. Cremoris SK11	LLSC	MRS
Leuconostoc citreum DmW_111	LCIT	MRS

Leuconostoc citreum str. CBA3627	LCI7	MRS
Leuconostoc fallax KCTC 3537	LFAL	MRS
Leuconostoc mesenteroides subsp. Suionicum str. LT-38	LMSS	MRS
Levilactobacillus brevis DmCS_003	LBRC	MRS
Ligilactobacillus animalis KCTC 3501 DSM 20602	LANI	MRS
Limosilactobacillus fermentum ATCC 14931	LFER	MRS
Liquorilactobacillus mali KCTC 3596, DSM 20444	LMLI	MRS
Pseudomona putida F1	PPUT	LB
Weissella cibaria DmW_103	WCIB	MRS
Weissella paramesenteroides DmW_107	WPA7	MRS
Weissella paramesenteroides DmW_109	WPA9	MRS
Weissella paramesenteroides DmW_115	WPA5	MRS
Weissella paramesenteroides DmW_118	WPA8	MRS

Table 3. Mutants tested in the persistence assay. Asterisks show significant difference in CFU per fly from the wild-type.

Code	Gene	Kegg	P-Value
Urea Carboxylase			
S01A	E6.3.4.6; urea carboxylase [EC:6.3.4.6]	K01941	6.00E-06
S02B*	E6.3.4.6; urea carboxylase [EC:6.3.4.6]	K01941	6.00E-06
Phosphatidylinositol Signaling			
S03C	E3.1.3.25; myo-inositol-1(or 4)-monophosphatase [EC:3.1.3.25]	K01092	0.002246476
S05F*	E3.1.3.25; myo-inositol-1(or 4)-monophosphatase [EC:3.1.3.25]	K01092	0.002246476
Bacterial Secretion			
S06G*	6666666.222946.peg.857 *secB;* preprotein translocase subunit SecB	K03071	0.04025775
S07H	*yajC*; preprotein translocase subunit YajC	K03210	0.020034512
S08J	*yajC*; preprotein translocase subunit YajC	K03210	0.020034512
S09K	*tolC*; outer membrane protein	K12340	0.023560031
S10L	*tolC*; outer membrane protein	K12340	0.023560031
Nicotinic Acid			
S15R	*pncC*; nicotinamide-nucleotide amidase [EC:3.5.1.42]	K03743	1
S16S	*pncC*; nicotinamide-nucleotide amidase [EC:3.5.1.42]	K03743	1
S17T	*nadD*; nicotinate-nucleotide adenylyltransferase [EC:2.7.7.18]	K00969	1
AMP Resistance			
S20X	*degP*; serine protease Do [EC:3.4.21.107]	K04771	0.023560031
S21Y	*degP*; serine protease Do [EC:3.4.21.107]	K04771	0.023560031
S24AB	*degP*; serine protease Do [EC:3.4.21.107]	K04771	0.023560031
S25AC	*degP*; serine protease Do [EC:3.4.21.107]	K04771	0.023560031
S26AD*	*mprF*; phosphatidylglycerol lysyltransferase [EC:2.3.2.3]	K14205	0.047856172

S27AF	*mprF*; phosphatidylglycerol lysyltransferase [EC:2.3.2.3]	K14205	0.047856172
Flagellar Assembly			
S30AJ	*fliP*; flagellar biosynthetic protein FliP	K02419	0.423453997
S31AK	*fliM*; flagellar motor switch protein FliM	K02416	1
S32AL	*fliM*; flagellar motor switch protein FliM	K02416	1
S33AM	*fliH*; flagellar assembly protein FliH	K02411	1
S34AN	*fliH*; flagellar assembly protein FliH	K02411	1
S35AP	*fliG*; flagellar motor switch protein FliG	K02410	1
S36AQ	*fliG*; flagellar motor switch protein FliG	K02410	1
S37AR*	*fliF*; flagellar M-ring protein FliF	K02409	1
S38AS*	*fliF*; flagellar M-ring protein FliF	K02409	1
S39AT	*fliI*; flagellum-specific ATP synthase [EC:3.6.3.14]	K02412	1
S40AU*	*fliI*; flagellum-specific ATP synthase [EC:3.6.3.14]	K02412	1
S41AW	*flgK*; flagellar hook-associated protein 1 FlgK	K02396	1
S42AX	*flgK*; flagellar hook-associated protein 1 FlgK	K02396	1
S43AY*	*flgE*; flagellar hook protein FlgE	K02390	1
S44AZ	*flgE*; flagellar hook protein FlgE	K02390	1
S45BA	*flgF*; flagellar basal-body rod protein FlgF	K02391	1
S46BB	*flgF*; flagellar basal-body rod protein FlgF	K02391	1
S47BC*	*flgH*; flagellar L-ring protein precursor FlgH	K02393	1
S48BF*	*flgH*; flagellar L-ring protein precursor FlgH	K02393	1
S49BG	*flgI*; flagellar P-ring protein precursor FlgI	K02394	1
S50BH	*flgI*; flagellar P-ring protein precursor FlgI	K02394	1
S51BJ	*flgB*; flagellar basal-body rod protein FlgB	K02387	1
S52BK	*flhB*; flagellar biosynthetic protein FlhB	K02401	1
S53BL	*flhB*; flagellar biosynthetic protein FlhB	K02401	1
S54BM	*flgA*; flagella basal body P-ring formation protein FlgA	K02386	1
S55BN	f*lgA*; flagella basal body P-ring formation protein FlgA	K02386	1

Table 4. Bacterial strains used to cluster OGs

Code	**Accession**	**Strain name**
afa2	JOPD01	*Acetobacter* sp. DsW_54
lplc	JOJT01	*Lactiplantibacillus plantarum* DmCS_001
efav	NC_004668.1	*Enterococcus faecalis* V583
lrha	NC_013198.1	*Lactobacillus rhamnosus* GG
efog	NC_017316.1	*Enterococcus faecalis* OG1RF
lbga	ACGG01	*Lactobacillus brevis* subsp. gravesensis ATCC 27305
lfer	ACGI01	*Lactobacillus fermentum* ATCC 14931
khan	ADTV01	*Komagataeibacter hansenii* ATCC 23769
ehor	AFHR01	*Enterobacter hormaechei* ATCC 49162
pbur	AKKL01	*Providencia burhodogranaria* DSM 19968

lani	AYYW01	*Lactobacillus animalis* KCTC 3501
atr2	JOMM01	*Acetobacter tropicalis* DmW_042
ain3	JOMP01	*Acetobacter indonesiensis* DmW_046
acew	JOOY01	*Acetobacter orientalis* DmW_048
ama3	JOPG01	*Acetobacter malorum* DsW_057
llac	JRFX01	*Lactococcus lactis* BPL1
lcit	NDXG01	*Leuconostoc citreum* DmW_111
1ab8	VHPI01	*Weissella paramesenteroides* DmW_109
kobo	CADT01	*Komagataeibacter oboediens* 174Bp2
ke5p	CADS01	*Komagataeibacter europaeus 5P3*
wcib	NDXJ01	*Weissella cibaria* DmW_103
1cb7	VHPE01	*Weissella paramesenteroides* DmW_115113
aace	SLZP01	*Acetobacter aceti NBRC 14818*
aci5	JOMN01	*Acetobacter sp. DmW_043*
amac	JOJU01	*Acetobacter malorum DmCS_005*
aori	JOMO01	*Acetobacter orientalis DmW_045*
apa3	CADQ01	*Acetobacter pasteurianus 3P3*
apan	AP014881.1	*Acetobacter pasteurianus NBRC 101655*
apnb	BDER01	*Acetobacter pasteurianus NBRC 106471 or LMG1262t*
apoc	JOKL01	*Acetobacter pomorum DmCS_004*
asl5	VHOZ01	Acetobacter sp. SLV-7 DmW_125
atrc	JOKM01	*Acetobacter tropicalis DmCS_006*
atrn	BABS01	*Acetobacter tropicalis NBRC 101654*
bsub	NC_000964.3	*Bacillus subtilis subsp. subtilis str.168*
cint	JOPB01	*Acetobacter fabarum* DmL_052
ecok	NC_000913.3	*Escherichia coli str. K-12 substr. MG1655*
galb	JOPF01	*Gluconobacter sp DsW_056*
gfra	BADZ02	*Gluconobacter frateurii NBRC 101659*
kmed	NC_016037.1	*Komagataeibacter medellinensis NBRC 3288*
lbrc	JOKA01	*Lactobacillus brevis DmCS_003*
lbuc	NC_015428.1	*Lactobacillus buchneri NRRLB-30929*
lc37	JADAXK01	Leuconostoc citreum DmW_137
lfal	AEIZ01	*Leuconostoc fallax KCTC 3537*
lfrc	JOJZ01	*Lactobacillus fructivorans DmCS_002*
lfrk	QSLO01	*Lactobacillus fructivorans KCTC 3543*
llcs	NEQN01	Lactococcus lactis subsp. cremoris SK11 strain SK11
lmli	AYYH01	*Lactobacillus mali KCTC 3596 = DSM 20444*
lpar	NDXH01	*Lactobacillus paracasei* DmW181
lplw	NC_004567.2	*Lactiplantibacillus plantarum WCFS1*

lsui	JADAXL01	*Leuconostoc suionicum* strain DmW_098
pput	NC_009512.1	*Pseudomonas putida F1*
wp07	VHPP01	*Weissella paramesenteroides* DmW_107
wp15	VHPE01	*Weissella paramesenteroides* DmW_115
wp18	VHPB01	*Weissella paramesenteroides* DmW_118

Table 5. Mutants used in a motility test

Code	**Mutant**
fliI1	*fliI*; flagellum-specific ATP synthase [EC:3.6.3.14]
fliI2	*fliI*; flagellum-specific ATP synthase [EC:3.6.3.14]
flgE1	*flgE*; flagellar hook protein FlgE
flgE2	*flgE*; flagellar hook protein FlgE
flgH1	*flgH*; flagellar L-ring protein precursor FlgH
flgH2	*flgH*; flagellar L-ring protein precursor FlgH
fliF1	*fliF*; flagellar M-ring protein FliF
fliF2	*fliF*; flagellar M-ring protein FliF
urea1	E6.3.4.6; urea carboxylase [EC:6.3.4.6]
urea2	E6.3.4.6; urea carboxylase [EC:6.3.4.6]

Table 6. KEGG enrichment analysis

Pathway	**Map ID**	**P-Value**
Staphylococcus aureus infection	map05150	2.22E-06
Biofilm formation - *Pseudomonas aeruginosa*	map02025	4.11E-06
D-Alanine metabolism	map00473	4.92E-05
Nicotinate and nicotinamide metabolism	map00760	0.000538481
Bacterial secretion system	map03070	0.001866839
Phosphatidylinositol signaling system	map04070	0.026694856
Atrazine degradation	map00791	0.026694856
Cationic antimicrobial peptide	map01503	0.043958764
Vancomycin resistance	map01502	0.045734966
Arginine biosynthesis	map00220	0.0499426
ABC transporters	map02010	0.054002894
Carbon metabolism	map01200	0.092982686
Cyanoamino acid metabolism	map00460	0.165663116
Microbial metabolism in diverse environments	map01120	0.210107772
Pyruvate metabolism	map00620	0.224024463
Starch and sucrose metabolism	map00500	0.236241006
Ubiquinone and other terpenoid-quinone biosynthesis	map00130	0.283802339

Fructose and mannose metabolism	map00051	0.29254434
Sesquiterpenoid and triterpenoid biosynthesis	map00909	0.294291935
Purine metabolism	map00230	0.294936952
Glycerophospholipid metabolism	map00564	0.316860478
Metabolic pathways	map01100	0.31759792
Phosphotransferase system	map02060	0.332549435
Glycolysis / Gluconeogenesis	map00010	0.334759596
Alanine, aspartate and glutamate metabolism	map00250	0.335579782
Biosynthesis of amino acids	map01230	0.354854388
Inositol phosphate metabolism	map00562	0.405464272
Phenylpropanoid biosynthesis	map00940	0.422567115
Adipocytokine signaling pathway	map04920	0.422567115
Nonribosomal peptide structures	map01054	0.422567115
Primary immunodeficiency	map05340	0.422567115
Quorum sensing	map02024	0.424748916
Aminoacyl-tRNA biosynthesis	map00970	0.487462362
Biosynthesis of secondary metabolites	map01110	0.490427682
Amoebiasis	map05146	0.533772341
Ferroptosis	map04216	0.533772341
RNA transport	map03013	0.533772341
D-Glutamine and D-glutamate metabolism	map00471	0.533772341
Methane metabolism	map00680	0.548573261
Cysteine and methionine metabolism	map00270	0.56188657
Two-component system	map02020	0.612085352
PPAR signaling pathway	map03320	0.630322761
Ether lipid metabolism	map00565	0.630322761
Thyroid hormone signaling pathway	map04919	0.630322761
Galactose metabolism	map00052	0.648871238
Carbon fixation in photosynthetic organisms	map00710	0.719297119
Isoquinoline alkaloid biosynthesis	map00950	0.789847436
Histidine metabolism	map00340	0.824202537
Novobiocin biosynthesis	map00401	0.856820964
Citrate cycle	map00020	0.8864225
Glycerolipid metabolism	map00561	0.889710631
Lipopolysaccharide biosynthesis	map00540	0.912115434
Tropane, piperidine and pyridine alkaloid biosynthesis	map00960	0.917204789
Glycine, serine and threonine metabolism	map00260	0.937809317
Pantothenate and CoA biosynthesis	map00770	0.951708627
Biofilm formation - *Vibrio cholerae*	map05111	0.965817248

Arginine and proline metabolism	map00330	0.983585619
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	0.99386282
2-Oxocarboxylic acid metabolism	map01210	0.99386282
Fatty acid metabolism	map01212	0.99386282
Phenylalanine metabolism	map00360	1
Nucleotide excision repair	map03420	1
Tyrosine metabolism	map00350	1
Protein export	map03060	1
Porphyrin and chlorophyll metabolism	map00860	1
Fatty acid biosynthesis	map00061	1
Pentose phosphate pathway	map00030	1
Streptomycin biosynthesis	map00521	1
Glyoxylate and dicarboxylate metabolism	map00630	1
Butanoate metabolism	map00650	1
Glutathione metabolism	map00480	1
Peroxisome	map04146	1
Thermogenesis	map04714	1
Taurine and hypotaurine metabolism	map00430	1
Base excision repair	map03410	1
Terpenoid backbone biosynthesis	map00900	1
Riboflavin metabolism	map00740	1
Fatty acid degradation	map00071	1
Selenocompound metabolism	map00450	1

Table 7. Results of the Dunnett test on motility measures

Mutant compared to wild type	**Estimate**	**Std. Error**	**Z Value**	**P-Value**
fliI1 – AF	-0.89111	0.07158	-12.450	<1E-04
flgE1 – AF	-0.84556	0.07158	-11.813	<1E-04
flgH1 – AF	-0.97333	0.07158	-13.598	<1E-04
fliF1 – AF	-0.9444	0.07158	-13.195	<1E-04
flgE2 – AF	-0.9222	0.07158	-12.884	<1E-04
fliI2 – AF	-0.84111	0.07158	-11.751	<1E-04
flgH2 – AF	-0.96111	0.07158	-13.428	<1E-04
fliF2 – AF	-0.90993	0.06148	-14.801	<1E-04
urea1 – AF	-0.11222	0.07158	-1.568	0.6253
urea2 - AF	-0.21111	0.07158	-2.949	0.0287