COMMUNITY DYNAMICS AND NUTRITIONAL BENEFITS OF THE *DROSOPHILA* GUT MICROBIOTA

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Chun Nin (Adam) WONG

August 2013

COMMUNITY DYNAMICS AND NUTRITIONAL BENEFITS OF THE *DROSOPHILA* GUT MICROBIOTA

Chun Nin (Adam) WONG, Ph. *D.*

Cornell University 2013

Recent advances in high throughput sequencing have provided important insights on the diversity and functional capabilities of gut microbiota in various animals. Despite tremendous sampling efforts in mammalian systems, the community dynamics and assembly patterns of gut microbiota are poorly understood, and experimental demonstration of their nutritional benefits remains largely absent.

To address these issues, this study develops *Drosophila* as a model system to study: 1) the diversity of the gut microbiota, by characterizing the gut microbiota composition of laboratory *Drosophila melanogaster* and other *Drosophila* species across phylogeny using high-throughput sequencing of the 16S rRNA gene, 2) the nutritional benefits of gut microbiota under different dietary regimes, by comparing the performance and nutritional responses between conventional and axenic (i.e. microbe-free) flies onto diets of systematically-varied nutrient (yeast-glucose) content. Results from this project demonstrate that *Drosophila* has a low-diversity gut bacterial community that is amenable for studying gut microbiota functions. The taxonomic composition appears to be inconstant, with no evidence for core taxa or co-evolution between the host and its microbiota. However, elimination of the gut microbiota results in prolonged host development and nutritional response to diet. The gut bacteria promote host health under conditions of nutritional stress resulting from

unbalanced diet by increasing micronutrient (vitamins B) availability and/or reducing excessive dietary sugar. Future investigations will include examining the nutritional functions of individual gut bacteria via re-associations with axenic flies, and testing congruence between taxonomic and functional (microbiome) profiles of the gut microbiota in response to changing diet.

BIOGRAPHICAL SKETCH

Adam joined Cornell in 2009 to pursue his PhD degree of Entomology and Microbiology. Prior to Cornell studies, Adam has always been interested in understanding the impact of microbes on animal health and wellbeing. During his undergraduate training in Molecular Cell Biology at the University of York, UK, Adam earned valuable insights on microbial physiology through undertaking an one-year internship at GlaxoSmithKline R&D with Dr Richard Hall and Dr Jo Jones, where he manipulated bacteria (*Escherichia coli* and *Bacillus subtilis*) to generate recombinant proteins for drug testing. After finishing his undergraduate studies, Adam conducted his Masters research in the Medical School of University of Hong Kong with Dr John Nicholls, investigating influenza pathogenesis in humans. He was inspired by the observation that infection outcome of influenza was dependent on both the virus and host types. He also contributed to the development of an anti-influenza drug (DAS181) that works by interfering host-microbe interaction (in this case, elimination of host cell receptors required for viral entry) currently under clinical trial. To further expand his interest and expertise on host-microbe interaction, Adam joined the lab of Dr Angela Douglas at Cornell University to pioneer research on the *Drosophila* gut microbiota. Throughout the past 4/4.5 years, Adam's work focused on the ecology and nutritional functions of the fly gut microbiota.

Dedicated to my family and fiancée for their unconditional love and support.

ACKNOWLEDGMENTS

I would like to thank my advisor Dr Angela Douglas for giving me the opportunity to come to Cornell and pioneer this exciting gut microbiota project. Angela has been extremely patient and supportive all the time. I have learnt a lot from her scientific expertise (knowledge, critical thinking, writing etc).

A special thanks to Dr Brian Lazzaro and Dr Esther Angert, who have truly inspired my thinking during my PhD journey. I particularly enjoyed discussing various scientific issues and ways to address these issues with them. Without their advice, my second publication in ISME J would not have happened. They are the best committee members I could have hoped for.

I also want to thank all members in the Douglas lab, who have helped my research in various ways. Especially, thanks to Dr John Chaston, Dr Adam Dobson and Dr XiangFeng Jing for their contributions on three important publications, and technical support from Sara Hermann, Stephanie Westmiller and Jean Yoon.

My collaborators and friends at Cornell deserve a lot of credits, thank you. They include Patrick Ng, everyone from the Hong Kong Christian Fellowship (HKCF), and my basketball buddies.

Additionally, my research at Cornell was made possible by the Department of Entomology; I appreciate their generous financial support and providing a great research environment for my work.

Finally, and most importantly, I want to give most credits to my family and fiancée Ruby Lee. Without them, I will not reaching this stage of my scientific journey today.

TABLE OF CONTENTS

CHAPTER 1

BACKGROUND AND RESEARCH OBJECTIVES

Animals are constantly associated with micro-organisms. The vast majority of animal-associated microbes are commensals or mutualists, and many play an important role in shaping host physiology (Kostic et al., 2013; McFall-Ngai et al., 2013; Sommer and Backhed, 2013) and behavior (Diaz Heijtz et al., 2011; Ezenwa et al., 2012).

Over tha past decades, research on animal-microbe symbiosis has focused on several non-mammalian and mammalian systems. Among these systems, the best-studied include the pea aphid (*Acyrthosiphon pisum*) and its endosymbiotic bacterium *Buchnera aphidicola*, the tsetse fly (*Glossina morsitans*) and its bacterial symbiont *Wigglesworthia glossinidia* the bobtailed squid (*Euprymna scolopes*) and its bioluminescent symbiont *Vibrio fischeri*, ruminant animals and protozoa (Hungate, 1943; Hobso and Stewart, 1997). Studies using these models have demonstrated key nutritional and immune functions of symbionts to host animals, and yielded valuable insights on their transmission/selection mechanisms (Douglas, 2011; McFall-Ngai et al., 2012; Rio et al., 2012).

Nevertheless, there has been a general consensus that a different form of animal-microbe symbiosis prevails in nature – i.e. animal tissues, especially the intestinal tract, colonized by communities of bacteria, known as the “microbiota”. Host associations with the gut microbiota differ from the single associations introduced in the previous paragraph in a number of ways; for instances, the microbial communities in the gut may involve complex synergistic and/or competitive microbe-microbe interactions (Denou et al., 2009; Flint et al., 2007), generally lack stringent transmission routes (Koenig et al., 2011), and the gut (unlike specialized symbiotic

organs) is frequently accessed by allochthonous (non-resident) microorganisms, including pathogens (Littman and Pamer, 2011).

Pioneers of research on the gut microbiota in animals have included Cleveland, L.R (1943) and Hungate R (1966), who studied the microorganisms in wood roaches and ruminant animals, respectively. The opportunity for gut microbiota research has been transformed in recent years by the development and widespread application of next-generation sequencing (Chaston et al., 2011; Zaneveld et al., 2011). The composition of the gut microbiota in humans, as well as other mammalian systems, has been studied extensively by high-throughput genetic surveys (including the Human Microbiome [http:// commonfund.nih.gov/hmp] and MetaHIT [(http://www.metahit.eu] projects). It is now accepted that a healthy human gut harbors ~10^{14} prokaryotic cells representing >500 species (Eckburg et al., 2005; Kuczynski et al., 2010). Collectively, these gut bacteria express a diverse array of genes ("microbiome", ~150 times of a human genome, Qin et al., 2012) that exert important nutritional and immune functions, such as energy extraction from indigestible dietary polysaccharides (Flint et al., 2008; Robert and Bernalier-Donadille, 2003; Xu et al., 2003), synthesis of key nutrients [short-chain fatty acids (SCFA), vitamins and essential amino acids (LeBlanc et al., 2013; Topping and Clifton, 2001; Wong et al., 2006)], production of antimicrobial compounds and modulators of immune system (Bouskra et al., 2008; Chung et al., 2012; Ivanova et al., 2009; Rea et al., 2007).

Currently, gut microbiota research in humans and other mammalian systems faces several challenges: a large fraction of the microbiota remains uncultivated, thus most diversity and functional studies has relied heavily on culture-independent genetic

approaches. The sampling depth of these approaches is often compromised by the sheer diversity of the mammalian microbiota. Furthermore, each individual host appears to have its unique gut microbiota (Costello et al., 2009). Altogether, it has not been feasible to couple microbiota functions to specific taxa. The inadequate sampling depth and substantial variation among individuals have obstructed our understanding on the ecological principles that govern the assembly of the microbiota, and how pertubation of a "healthy" microbiota promotes obesity and certain metabolic diseases (Costello et al., 2012). This also resulted in some debatable claims, such as the existence of a phylogenetic core microbiota (Lozupone et al., 2012).

The central theme of this project is to use *Drosophila* as a model to study host-gut microbiota symbiosis. *Drosophila* has been a powerful genetics model in revealing developmental and immune mechanisms. Many developmental/immune/metabolic genes and pathways are conserved between the fly and human (Pandey and Nichols, 2011; Rajan and Perrimon, 2013; Reiter et al., 2001). Previous reports have suggested that the *Drosophila* gut is colonized by a low diversity bacterial community, using culture-dependent and shallow sampling methods (Corby-Harris et al., 2007; Cox and Gilmore, 2007; Ren et al., 2007; Ryu et al., 2008). The community structure and functions of the fly gut microbiota are largely unknown prior to this project.

The aims of this project are two-fold: 1) to determine the composition and diversity of the gut microbiota in drosophilid flies, using high-resolution sequencing methods, 2) to use laboratory *Drosophila melanogaster* as a model to study the impact of gut microbiota on host nutritional physiology, by comparison of the response of flies containing and experimentally deprived of their microbiota to diets of different composition. This involves establishing an optimal procedure for experimental

elimination of the fly gut microbiota. Ultimately, results from this work will contribute fundamental knowledge on the community dynamics of the gut microbiota and the mechanisms underlying their impact to host physiology, and this can be applied to enhance our understanding of higher organisms, including humans.

References

Bouskra, *D.*, Brezillon, C., Berard, M., Werts, C., Varona, R., Boneca, I.G., Eberl, G., 2008. Lymphoid tissue genesis induced by commensals through NOD1 regulates intestinal homeostasis. Nature 456, 507-510.

Chaston, J.M., Suen, G., Tucker, S.*L.*, Andersen, *A.*W., Bhasin, *A.*, Bode, E., Bode, H.B., Brachmann, *A.*O., Cowles, C.E., Cowles, K.N., Darby, C., de Leon, *L.*, Drace, K., Du, Z., Givaudan, *A.*, Herbert Tran, E.E., Jewell, K.*A.*, Knack, J.J., Krasomil-Osterfeld, K.C., Kukor, R., Lanois, *A.*, Latreille, P., Leimgruber, N.K., Lipke, C.M., Liu, R., Lu, X., Martens, E.C., Marri, P.R., Medigue, C., Menard, M.*L.*, Miller, N.M., Morales-Soto, N., Norton, S., Ogier, J.C., Orchard, S.S., Park, *D.*, Park, Y., Qurollo, B.*A.*, Sugar, *D.*R., Richards, G.R., Rouy, Z., Slominski, B., Slominski, K., Snyder, H., Tjaden, B.C., van der Hoeven, R., Welch, R.*D.*, Wheeler, C., Xiang, B., Barbazuk, B., Gaudriault, S., Goodner, B., Slater, S.C., Forst, S., Goldman, B.S., Goodrich-Blair, H., 2011. The entomopathogenic bacterial endosymbionts Xenorhabdus and Photorhabdus: convergent lifestyles from divergent genomes. Plos One 6, e27909.

Chung, H., Pamp, S.J., Hill, J.*A.*, Surana, N.K., Edelman, S.M., Troy, E.B., Reading, N.C., Villablanca, E.J., Wang, S., Mora, J.R., Umesaki, Y., Mathis, *D.*, Benoist, C., Relman, *D.A.*, Kasper, *D.L.*, 2012. Gut immune maturation depends on colonization with a host-specific microbiota. Cell 149, 1578-1593.

Cleveland, L. R. 1934. The wood-feeding roach Cryptocercus, its protozoa, and the symbiosis between protozoa and roach. Mem. Amer. Acad. Arts Sci. 17:185-342.

Corby-Harris, V., Pontaroli, *A.*C., Shimkets, *L.*J., Bennetzen, J.*L.*, Habel, K.E., Promislow, *D.*E., 2007. Geographical distribution and diversity of bacteria associated with natural populations of Drosophila melanogaster. Appl Environ Microbiol 73, 3470-3479.

Costello, E.K., Lauber, C.*L.*, Hamady, M., Fierer, N., Gordon, J.I., Knight, R., 2009. Bacterial community variation in human body habitats across space and time. Science 326, 1694-1697.

Costello, E.K., Stagaman, K., Dethlefsen, *L.*, Bohannan, B.J., Relman, *D.A.*, 2012. The application of ecological theory toward an understanding of the human microbiome. Science 336, 1255-1262.

Cox, C.R., Gilmore, M.S., 2007. Native microbial colonization of Drosophila melanogaster and its use as a model of Enterococcus faecalis pathogenesis. Infect Immun 75, 1565-1576.

Denou, E., Rezzonico, E., Panoff, J.M., Arigoni, F., Brussow, H., 2009. A Mesocosm of Lactobacillus johnsonii, Bifidobacterium longum, and Escherichia coli in the mouse gut. DNA Cell Biol 28, 413-422.

Diaz Heijtz, R., Wang, S., Anuar, F., Qian, Y., Bjorkholm, B., Samuelsson, *A.*,

Hibberd, M.*L.*, Forssberg, H., Pettersson, S., 2011. Normal gut microbiota modulates brain development and behavior. Proc Natl Acad Sci U S A 108, 3047-3052.

Douglas, *A.*E., 2011. Lessons from studying insect symbioses. Cell Host Microbe 10, 359-367.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, *L.*, Sargent, M., Gill, S.R., Nelson, K.E., Relman, *D.A.*, 2005. Diversity of the human intestinal microbial flora. Science 308, 1635-1638.

Ezenwa, V.O., Gerardo, N.M., Inouye, *D.*W., Medina, M., Xavier, J.B., 2012. Microbiology. Animal behavior and the microbiome. Science 338, 198-199.

Flint, H.J., Bayer, E.*A.*, Rincon, M.T., Lamed, R., White, B.*A.*, 2008. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. Nat Rev Microbiol 6, 121-131.

Flint, H.J., Duncan, S.H., Scott, K.P., Louis, P., 2007. Interactions and competition within the microbial community of the human colon: links between diet and health. Environ Microbiol 9, 1101-1111.

Hobson, P.N. and C.S. Stewart., (1997). The Rumen Microbial Ecosystem, 2nd edition. New York: Springer. ISBN 0-7514-0366-0.

Hungate, R.E. 1943. Quantitative analyses on the cellulose fermentation by termite protozoa. Ann Entomol Soc Am 36:730-739.

Hungate, R. E. 1966. The Rumen and its Microbes. London: Academic Press.

Ivanova, E.V., Perunova, N.B., Valyshev, *A.*V., Valysheva, I.V., Bukharin, O.V., 2009. [Species characteristic and factors of persistence of gut bifidoflora during healthy state and dysbiosis]. Zh Mikrobiol Epidemiol Immunobiol, 89-93.

Koenig, J.E., Spor, *A.*, Scalfone, N., Fricker, *A.D.*, Stombaugh, J., Knight, R., Angenent, *L.*T., Ley, R.E., 2011. Succession of microbial consortia in the developing infant gut microbiome. Proc Natl Acad Sci U S A 108 Suppl 1, 4578-4585.

Kostic, *A.D.*, Howitt, M.R., Garrett, W.S., 2013. Exploring host-microbiota interactions in animal models and humans. Genes Dev 27, 701-718.

Kuczynski, J., Costello, E.K., Nemergut, *D.*R., Zaneveld, J., Lauber, C.*L.*, Knights, *D.*, Koren, O., Fierer, N., Kelley, S.T., Ley, R.E., Gordon, J.I., Knight, R., 2010. Direct sequencing of the human microbiome readily reveals community differences. Genome Biol 11, 210.

LeBlanc, J.G., Milani, C., de Giori, G.S., Sesma, F., van Sinderen, *D.*, Ventura, M., 2013. Bacteria as vitamin suppliers to their host: a gut microbiota perspective. Curr Opin Biotechnol 24, 160-168.

Littman, *D.*R., Pamer, E.G., 2011. Role of the commensal microbiota in normal and pathogenic host immune responses. Cell Host Microbe 10, 311-323.

Lozupone, C.*A.*, Stombaugh, J.I., Gordon, J.I., Jansson, J.K., Knight, R., 2012. Diversity, stability and resilience of the human gut microbiota. Nature 489, 220-230.

McFall-Ngai, M., Hadfield, M.G., Bosch, T.C., Carey, H.V., Domazet-Loso, T., Douglas, *A.*E., Dubilier, N., Eberl, G., Fukami, T., Gilbert, S.F., Hentschel, U., King, N., Kjelleberg, S., Knoll, *A.*H., Kremer, N., Mazmanian, S.K., Metcalf, J.*L.*, Nealson, K., Pierce, N.E., Rawls, J.F., Reid, *A.*, Ruby, E.G., Rumpho, M., Sanders, J.G., Tautz, *D.*, Wernegreen, J.J., 2013. Animals in a bacterial world, a new imperative for the life sciences. Proc Natl Acad Sci U S A 110, 3229-3236.

McFall-Ngai, M., Heath-Heckman, E.*A.*, Gillette, *A.A.*, Peyer, S.M., Harvie, E.*A.*, 2012. The secret languages of coevolved symbioses: insights from the Euprymna scolopes-Vibrio fischeri symbiosis. Semin Immunol 24, 3-8.

Pandey, U.B., Nichols, C.*D.*, 2011. Human disease models in Drosophila melanogaster and the role of the fly in therapeutic drug discovery. Pharmacol Rev 63, 411-436.

Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., Liang, S., Zhang, W., Guan, Y., Shen, *D.*, Peng, Y., Zhang, *D.*, Jie, Z., Wu, W., Qin, Y., Xue, W., Li, J., Han, *L.*, Lu, *D.*, Wu, P., Dai, Y., Sun, X., Li, Z., Tang, *A.*, Zhong, S., Li, X., Chen, W., Xu, R., Wang, M., Feng, Q., Gong, M., Yu, J., Zhang, Y., Zhang, M., Hansen, T., Sanchez, G., Raes, J., Falony, G., Okuda, S., Almeida, M., LeChatelier, E., Renault, P., Pons, N., Batto, J.M., Zhang, Z., Chen, H., Yang, R., Zheng, W., Yang, H., Wang, J., Ehrlich, S.*D.*, Nielsen, R., Pedersen, O., Kristiansen, K., 2012. A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature 490, 55-60.

Rajan, *A.*, Perrimon, N., 2013. Of flies and men: insights on organismal metabolism from fruit flies. BMC Biol 11, 38.

Rea, M.C., Clayton, E., O'Connor, P.M., Shanahan, F., Kiely, B., Ross, R.P., Hill, C., 2007. Antimicrobial activity of lacticin 3,147 against clinical Clostridium difficile strains. J Med Microbiol 56, 940-946.

Reiter, *L.*T., Potocki, *L.*, Chien, S., Gribskov, M., Bier, E., 2001. A systematic analysis of human disease-associated gene sequences in Drosophila melanogaster. Genome Res 11, 1114-1125.

Ren, C., Webster, P., Finkel, S.E., Tower, J., 2007. Increased internal and external bacterial load during Drosophila aging without life-span trade-off. Cell Metab 6, 144-152.

Rio, R.V., Symula, R.E., Wang, J., Lohs, C., Wu, Y.N., Snyder, *A.*K., Bjornson, R.*D.*, Oshima, K., Biehl, B.S., Perna, N.T., Hattori, M., Aksoy, S., 2012. Insight into the transmission biology and species-specific functional capabilities of tsetse (Diptera:

glossinidae) obligate symbiont Wigglesworthia. MBio 3.

Robert, C., Bernalier-Donadille, *A.*, 2003. The cellulolytic microflora of the human colon: evidence of microcrystalline cellulose-degrading bacteria in methane-excreting subjects. FEMS Microbiol Ecol 46, 81-89.

Ryu, J.H., Kim, S.H., Lee, H.Y., Bai, J.Y., Nam, Y.*D.*, Bae, J.W., Lee, *D.*G., Shin, S.C., Ha, E.M., Lee, W.J., 2008. Innate immune homeostasis by the homeobox gene caudal and commensal-gut mutualism in Drosophila. Science 319, 777-782.

Sommer, F., Backhed, F., 2013. The gut microbiota--masters of host development and physiology. Nat Rev Microbiol 11, 227-238.

Topping, *D.L.*, Clifton, P.M., 2001. Short-chain fatty acids and human colonic function: roles of resistant starch and nonstarch polysaccharides. Physiol Rev 81, 1031-1064.

Wong, J.M., de Souza, R., Kendall, C.W., Emam, *A.*, Jenkins, *D.*J., 2006. Colonic health: fermentation and short chain fatty acids. J Clin Gastroenterol 40, 235-243.

Xu, J., Bjursell, M.K., Himrod, J., Deng, S., Carmichael, *L.*K., Chiang, H.C., Hooper, *L.*V., Gordon, J.I., 2003. A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. Science 299, 2074-2076.

Zaneveld, J.R., Parfrey, *L.*W., Van Treuren, W., Lozupone, C., Clemente, J.C., Knights, *D.*, Stombaugh, J., Kuczynski, J., Knight, R., 2011. Combined phylogenetic and genomic approaches for the high-throughput study of microbial habitat adaptation. Trends Microbiol 19, 472-482.

CHAPTER 2

LOW-DIVERSITY BACTERIAL COMMUNITY IN THE GUT OF THE FRUITFLY *DROSOPHILA MELANOGASTER*[1]

Abstract

The bacteria in the fruitfly *Drosophila melanogaster* of different life stages was quantified by 454 pyrosequencing of 16S rRNA gene amplicons. The sequence reads were dominated by 5 operational taxonomic units (OTUs) at ≤ 97% sequence identity that could be assigned to *Acetobacter pomorum, A. tropicalis, Lactobacillus breivs, L. fructivorans* and *L. plantarum*. The saturated rarefaction curves and species richness indices indicated that the sampling (85 000–159 000 reads per sample) was comprehensive. Parallel diagnostic PCR assays revealed only minor variation in the complement of the five bacterial species across individual insects and three *D. melanogaster* strains. Other gut-associated bacteria included 6 OTUs with low %ID to previously reported sequences, raising the possibility that they represent novel taxa within the genera *Acetobacter* and *Lactobacillus*. A developmental change in the most abundant species, from *L. fructivorans* in young adults to *A. pomorum* in aged adults was identified; changes in gut oxygen tension or immune system function might account for this effect. Host immune responses and disturbance may also contribute to the low bacterial diversity in the *Drosophila* gut habitat.

[1] Presented with minor modifications from the originally published article:
Wong, ACN., Ng, P. and Douglas, AE. (2011). Low-diversity bacterial community in the gut of the fruitfly *Drosophila melanogaster*. *Environmental Microbiology*, 13: 1889–1900.
All supplementary information can be found at:
http://onlinelibrary.wiley.com/doi/10.1111/j.1462-2920.2011.02511.x/suppinfo
Ng, P contributed to the bioinformatics for evaluating candidate novel taxa.

Introduction

Healthy animals are a habitat for microorganisms, most of which are benign or beneficial (Wilson, 2005; Douglas, 2010). Mammals and other vertebrates appear to support many more microbial species than most invertebrates. For example, the gut microbiota in an individual mammal comprises > 1000 taxa, most of which are unique to each host individual (Dethlefsen et al., 2007; Ley et al., 2008; Costello et al., 2009; Qin et al., 2010). The diversity of the gut microbiota in most invertebrates that have been studied is apparently one to two orders of magnitude lower than in the mammals (Dillon and Dillon, 2004; Dunn and Stabb, 2005; Behar et al., 2008; Lehman et al., 2009; Morales-Jimenez et al., 2009; Grunwald et al., 2010; Robinson et al., 2010). Nevertheless, vertebrates, especially mammals, have been the subject of far greater sampling effort than invertebrates, raising the possibility that this difference between vertebrates and invertebrates may be partly artefactual. Two further issues affect the interpretation of data on the diversity of the microbiota in animal guts. First, the composition of the gut microbiota can vary with diet, and developmental age and physiological condition of the animal host (e.g. Dethlefsen et al., 2007; Lehman et al., 2009; Sharon et al., 2010). Second, microorganisms recovered from the gut comprise two ecologically distinct groups: the autochthonous (resident) taxa and the allochthonous (non-resident) forms that are ingested with, and pass through, the gut with the food. The allochthonous microbes can artificially inflate both the reported microbial diversity in an individual host, and among-host variation in microbial diversity, especially where the animals sampled utilize different sources of food. The purpose of this study was to determine the diversity of the gut bacteria of the fruitfly *Drosophila melanogaster* using 454 pyrosequencing of PCR-generated

amplicons from the 16S rRNA gene. We used *Drosophila* raised on an axenic diet of fixed composition, to minimize the diversity of allochthonous taxa, and sampled the animals across the full life cycle, to establish the total diversity and how it varies with life stage. Our analysis builds on previous research, which has identified various taxa, including *Lactobacillus*, Enterococcus and *Acetobacter* associated with *Drosophila* [Corby-Harris et al., 2007; Cox and Gilmore, 2007; Ren et al., 2007; Ryu et al., 2008; also see the review (Crotti et al., 2010) of *Acetobacter* as insect symbionts]. Our study overcomes three key limitations of previous studies: all may have failed to detect low-abundance taxa through shallow sampling using limited Sanger sequencing of cloned 16S rRNA gene sequences; most were conducted on the whole insect, making it impossible to identify the bacteria specifically associated with the gut; and several studies did not attempt to limit the incidence of allochthonous taxa.

Experimental procedures

The experimental material

Drosophila melanogaster was reared at 25°C with a 12 h:12 h light–dark cycle on autoclaved yeast-glucose medium [Y-G diet, comprising Brewer's yeast and glucose (both at 83 g l^{-1}, from MP Biomedicals), agar (10 g l^{-1}, from Frutarom) and preservatives (0.04% phosphoric acid, 0.42% propionic acid, from Sigma)], and transferred to fresh medium weekly. Outbred populations of strains Canton-S and Oregon-R had been maintained on Y-G diet for at least 18 years. Strain Ithaca-83 is an isofemale line established from a single female collected at Littletree Orchard, Newfield, New York in 2004, and maintained on Y-G diet since collection.

The experimental samples comprised: guts (from proventriculus to rectum, excluding

Malpighian tubules) dissected from third-instar larvae and adults; whole first- to second- ('early') instar larvae (< 48 h after hatching: these insects were too small for gut dissections); pupae (which lack a gut); and eggs (< 20 h after deposition). All samples except the eggs were surface-sterilized in 10% sodium hypochlorite solution, followed by three rinses in sterile distilled water. Gut dissections were conducted in sterile Ringer's solution on clean glass slides with sterilized forceps, using a dissecting microscope at × 7 magnification. This sampling design followed preliminary experiments that confirmed the presence of bacteria in all surface-sterilized samples except eggs (data not shown), consistent with published evidence that bacteria are borne within larvae, pupae and adults, but not internal to the eggshell (Bakula, 1969). All experiments used reagent-only controls comprising a drop of Ringer's solution treated as for dissections (including swirling the dissection instruments in the solution), but without *D. melanogaster* materials.

DNA isolation

For pyrosequencing, total genomic DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen, Valencia, California, USA) following a protocol modified from the manufacturer's instructions for Gram-positive bacteria. Briefly, samples were hand-homogenized in 20 mM Tris-HCl (pH 8.0), 2 mM sodium EDTA, 1.2% Triton® X-100 containing 20 mg lysozyme ml^{-1}. The homogenates were incubated at 37°C for 1.5 h with a 5 min bead-beating in a Disruptor Genie® using 0.1 mm glass beads (Scientific Industries) at 45 min. Pilot experiments confirmed that this treatment disrupted Gram-positive bacteria including Bacillus and *Lactobacillus*, and achieved 10–50% greater yield than lysozyme digestion without bead-beating (data not shown).

All DNA samples were quantified by Nanodrop 1000 (Thermo Scientific) and the PCR products for pyrosequencing were analysed by Agilent 2100 Bioanalyser.

Multiplex 454 pyrosequencing of 16S rRNA gene sequences

Each DNA sample comprised three biological replicates of *D. melanogaster* strain Canton S: 100 eggs, 50 early-instar larvae, guts from 50 third-instar larvae, 30 pupae, and guts from 50 each of male and female adults at 3–7 days and 3–5 weeks post eclosion. The variable region 2 (V2) of the bacterial 16S rRNA gene was amplified with the general 16S rRNA gene primers 27F (5′-AGAGTTTGATCMTGGCTCAG-3′) and 338R (5′-TGCTGCCTCCCGTAGGAGT-3′), with the sample-specific 27F primer bearing a multiplex identifier (MID) sequences and all 27F and 338R primers modified with 5′-Adaptor A and 5′-Adaptor B sequences, respectively, for pyrosequencing (Roche) (Table S3A). PCRs for the biological samples and reagent control were conducted in triplicate with 0.6 U Platinum®Taq DNA Polymerase (Invitrogen) in 1 × PCR buffer, 2 mM MgCl2, 8 pmol each primer, 0.24 mM dNTP and c. 100 ng of DNA sample in 25 µl final volume, at 94°C for 10 min followed by 25 cycles of 94°C for 1 min, 58°C for 1 min and 72°C for 1 min. DNA from an aliquot of each PCR reaction was purified using the Agencourt Ampure® SPRI kit and quantified using the Quant-iT™PicoGreen® Kit. Each reaction product was diluted to 1 × 109 molecules μl^{-1}, except MID-1 (egg DNA) and MID-9 (reagent-only control), which were diluted to 1 × 108 molecules μl^{-1}. Equal volumes of the three reaction products per sample were mixed together and diluted to 1 × 107 (samples 2–8) or 1 × 106 (MID-1 and MID-9) molecules μl^{-1} for emulsion PCR at one copy per bead using only ‘A’ beads for unidirectional sequencing. Beads were subjected to sequencing on

one full plate of the 454 GS-FLX pyrosequencing instrument using standard Titanium chemistry.

Pyrosequencing flowgrams were converted to sequence reads using 454 Life Science software (http://www.454.com). Reads with ambiguous nucleotides (N) and < 270 nucleotides after the forward primer, and mismatches with the 16S rRNA gene primers were excluded in the initial filtering. To ensure accurate determination of microbial diversity, the data were processed with Perl scripts (Kunin and Hugenholtz, 2010) (http://pyrotagger.jgi-psf.org/release) modified to remove reads with 0.2% per-base error probability (≥ 3% of bases with Phred scores < 27). The remaining sequences were trimmed to 270 nucleotides, dereplicated and clustered into OTUs with 93%, 95% and 97% sequence identity (ID) thresholds. The most abundant unique sequence of each OTU cluster was selected as representative, aligned by p-clustalw at BioHPC (http://cbsuapps.tc.cornell.edu.proxy.library.cornell.edu/clustalw.aspx) and subjected to chimera check by the Mallard algorithm (Ashelford et al., 2006). Taxonomy of the non-chimaeric sequences was assigned by NCBI StandAlone blast (megablast program) using the nucleotide (nt) database (13 June 2010) with default settings. Identified reads were counted and distributed to their respective MID samples. Phylotypes with < 10 reads or fewer reads than in the reagent-only control were interpreted as contaminants, and removed. The richness [Chao1, abundance-based coverage estimators (ACE) and Jackknife] and diversity (Simpson's and Shannon) indices for each biological sample were calculated using R. Rarefaction curves were generated using Analytic Rarefaction v1.3. (http://www.uga.edu/~strata/software/index.html). The Bonferroni-corrected Poisson probability of occurrence of 454 reads with %ID ≤ 98% to the blast top hits in each

biological sample were calculated using R. A pyrosequencing error rate of 0.3% was used as it was suggested that pyrosequencing errors can be reduced 0.25% (i.e. up to 3 bp per kb) after discarding reads with ambiguous bases (N) (Huse et al., 2007). All non-chimaeric 454 sequences are deposited in the short read archive at NCBI, Accession No. SRA023605.3.

PCR assays

Taxon-specific 16S rRNA gene primers were designed for *A. tropicalis, A. pomorum, L. breivs*, *L. fructivorans* and *L. plantarum* (Table S3B) using Primer3 software and unique regions identified from alignments of full 16S rRNA gene sequences. Preliminary experiments confirmed that the primers generated no detectable cross-amplification between species (data not shown). PCRs were performed as above with 65°C annealing temperature and 35 cycles. PCR products were separated by gel electrophoresis using 1% agarose gel and visualized with SYBR®Safe (Invitrogen), and their identities were confirmed by Sanger sequencing.

Specific 16S rRNA gene primers were designed for QRT-PCR of the dominant bacteria (Table S3B). The reactions were conducted in triplicate, with a reagent-only negative control, in C1000™ Thermal cycler (Bio-Rad) with 1 × Brilliant III Ultra-Fast QPCR Master Mix (Agilent Technologies), 8 pmol each primer and c. 100 ng in 20 µl volume, under a thermal profile of 95°C for 10 min, then 35 amplification cycles of 95°C for 10 s, 60°C for 30 s and dissociation cycle of 95°C for 10 s, 60°C for 5 s then brought back to 95°C. Fold differences of bacterial genes were calculated by the ΔΔCt method (Livak and Schmittgen, 2001). The dissociation curve confirmed that every reaction yielded a single PCR product with the predicted Tm. QRT-PCR assays

were used to check the repeatability of pyrosequencing data for dominant bacterial species in *D. melanogaster*. Samples comprised DNA samples from the 454 pyro-sequencing experiment (November 2009), and Canton-S flies (June 2010: 10 replicate samples of five pupae and five guts from third-instar larvae and adults at 3–7 days, 2–3 weeks and 4–5 weeks post eclosion). Bacterial relative abundances were compared for *A. pomorum*/*L. fructivorans* in adults, *L. fructivorans*/*L. plantarum* in third-instar larvae and *A. tropicalis*/*L. fructivorans* in pupae.

Pairwise comparisons of %ID of 16S rRNA gene sequences

A non-redundant set of (near-)full 16S rRNA gene sequences for 15 species of *Acetobacter* (79 sequences) and 102 species of *Lactobacillus* (1082 sequences) was collected from Greengenes (http://greengenes.lbl.gov), Ribosomal Database Project (RDP; http://rdp.cme.msu.edu.proxy.library.cornell.edu) and Silva (http://www.arb-silva.de). Species with a single sequence, unidentified species and species without binomial nomenclature were excluded from the analysis. The remaining sequences were trimmed to 1270 bp, and the V2 region was isolated in silico and trimmed to 270 bp from position 48–318. For each species, all possible pairwise alignments were obtained, and %ID between every sequence pair was calculated using algorithm of (Needleman and Wunsch, 1970) for the (near-full) 16S rRNA and V2 sequences. The lowest value of %ID for each species was adopted as a measure of the total sequence variation for that species.

Results

Pyrosequencing data

The 454 pyrosequencing analysis of 16S rRNA gene amplicons from the dissected guts of *D. melanogaster* strain Canton S produced 923 109 reads, with an average length of 361 nucleotides (including the multiplex identifier 'MID' and primer sequences), after quality filtering and removal of chimaeric sequences. The reads could be assigned to 720 operational taxonomic units (OTUs) at 93% sequence identity threshold, 894 and 1135 OTUs at 95% and 97% threshold, and 8935 OTUs at 99% threshold. A substantial number of the OTUs identified were represented by just one to several reads in both the experimental samples and the reagent-only control. These were interpreted as contaminants and they were discarded, leaving 808 483 reads that were distributed among the samples as follows: *D. melanogaster* eggs (0.2%), early-instar larvae (10.6%), pupae (13.9%) and guts from third-instar larvae (13.4%), 3- to 7-day-old males (14.0%) and females (19.7%), 3- to 5-week-old males (10.5%) and females (17.5%). Altogether, the reads yielded 122 OTUs at the 97% identity threshold recommended for accurate diversity estimation (Kunin et al., 2010). For each sample, the rarefaction curves tended towards saturation at similar numbers of clusters at 97%, 95% and 93% pairwise ID thresholds (Figure. 1.1). Subsequent analysis was, therefore, conducted at 97% I*D.* All values of richness indices (Chao1, ACE and Jackknife) equaled the number of OTUs (Table 1.1), confirming the conclusion from rarefaction analysis that sampling of each life stage had reached saturation. The third-instar larvae bore the most species-rich bacterial community, comprising 71 OTUs. The egg surface had the most diverse bacterial community by both Simpson's and Shannon indices (Table 1.1), including 19/28 (68%) unique

clusters. In all other samples, five OTUs (clusters 1, 2, 5, 6 and 7) accounted for > 80% of all reads (Table S1), and OTUs unique to one life stage were rare (early instars, 3- to 5-week-old adults) or absent (3- to 7-day-old adults). Exceptionally, 12/30 (40%) of OTUs in pupae and 48/71 (68%) OTUs in third-instar larvae were unique.

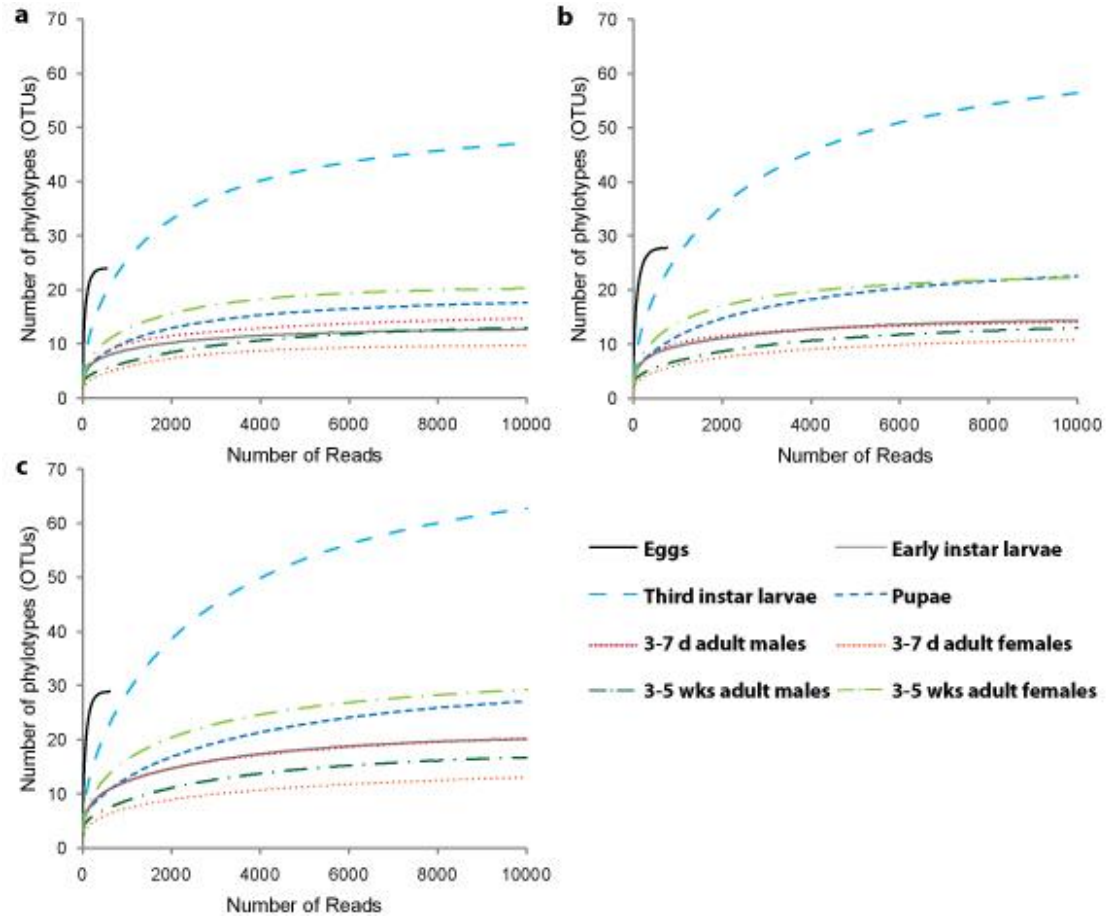


Figure. 1.1 Rarefaction curves of OTUs clustered at different %ID across life stages of *D. melanogaster* Canton-S. (A) 93%, (B) 95%, (C) 97%.

Table 1.1 Richness and diversity estimation of the 16S rRNA gene libraries from the pyrosequencing analysis.

Sample	Reads	Number of OTUs*	Species richness indices			Species diversity indices	
			Chao1	ACE	Jackknife	Shannon	Simpson
Egg surface	1798	28	28	28	28	2.62	0.88
Early instars	86038	21	21	21	21	0.87	0.37
Third instars	112382	71	71	71	71	1.40	0.58
Pupae	108609	30	30	30	30	1.47	0.70
3-7 days old adult males	113614	19	19	19	19	1.26	0.59
3-7 days old adult females	159309	15	15	15	15	0.35	0.15
3-5 weeks old adult males	85095	17	17	17	17	0.72	0.32
3-5 weeks old adult females	141761	31	31	31	31	1.02	0.43

*The operational taxonomic units (OTUs) were defined with pair-wise 97%ID.

Taxonomic composition of bacteria identified by pyrosequencing

At the phylum level, Firmicutes and Proteobacteria accounted for the vast majority of reads (> 97%) in the larvae, pupae and adults, and 66% of the reads for the eggs.

Actinobacteria, Bacteroides and Cyanobacteria were also detected (Table 1.2a).

The five OTUs dominating most samples (see above) corresponded to *Acetobacter* and *Lactobacillus* species: *A. pomorum*, *A. tropicalis*, *L. breivs*, *L. fructivorans* and *L. plantarum* (Table 1.2b, Table S1). The relative abundance of these taxa varied with developmental age (Table 1.2b). *Lactobacillus fructivorans* accounted for > 60% of the reads in early-instar larvae and 3- to 7-day-old adults (both sexes); *L. plantarum* dominated the gut bacteria of third-instar larvae; and *A. tropicalis* and *A. pomorum* were strongly represented in pupae and 3- to 5-week-old adults respectively. These species were detected in eggs at varying abundance (Table S1): *L. fructivorans* (21%), *A. pomorum* (14%), *A. tropicalis* (2%), *L. breivs* (1.5%) and *L. plantarum* (four reads, which was below the cut-off for contaminants). The sequences of the five OTUs were submitted to NCBI GenBank (accession HQ173707–HQ173711).

Pupae bore appreciable numbers of Staphylococcus, accounting for 16% of the reads, of which > 99% were assigned to Staphylococcus sp. K6-17B (Table S1D), while Staphylococcus represented < 0.1% of reads in all other life stages.

Table 1.2 Abundance of 16S rRNA gene amplicons in *D. melanogaster* samples, expressed as % of total in each life stage.

(a) Bacterial phyla

Phylum	% of total sequence reads in each life stage							
	Eggs	Early instar larvae	3rd instar larvae	Pupae	3-7 d old males	3-7 d old females	3-5 week old males	3-5 week old females
Actinobacteria	23.28	0.00	0.10	0.00	0.00	0.00	0.12	0.00
Bacteroidetes	5.91	0.03	2.19	0.00	0.00	0.00	0.00	0.03
Cyanobacteria	3.88	0.00	0.14	0.00	0.00	0.00	0.00	0.00
Firmicutes	35.32	85.97	87.83	43.67	80.45	93.60	15.52	20.29
Proteobacteria	30.34	14.01	9.66	56.31	19.55	6.40	84.37	79.67
Other	1.27	0.00	0.09	0.02	0.00	0.00	0.00	0.00

(b) Bacterial species

Species	% of total sequence reads in each life stage						
	Early instar larvae	3rd instar larvae	Pupae	3-7 d old males	3-7 d old females	3-5 week old males	3-5 week old females
Acetobacter pomorum	8.97	1.75	8.98	5.79	5.81	81.65	74.43
Acetobacter tropicalis	5.03	3.56	47.31	13.74	0.60	2.72	4.18
Lactobacillus brevis	1.94	22.42	3.11	15.13	1.03	7.05	2.42
Lactobacillus fructivorans	80.30	4.30	3.28	61.01	92.50	7.60	10.22
Lactobacillus plantarum	3.73	60.90	21.44	4.31	0.07	0.82	7.65
Staphylococcus sp.	0.00	0.00	15.73	0.00	0.00	0.00	0.00
Other	0.04	7.07	0.15	0.02	0.00	0.16	1.09

Candidate novel bacterial taxa

The %ID between some 454 reads and the blast top hits was less than 97% (Table S1 and Table 1.3). Two approaches were adopted to assess whether these low %IDs were likely a consequence of sequencing error. First, the polymorphisms were confirmed not to be in homopolymeric regions, which are common sites of 454 sequencing error. Second, the Bonferroni-corrected Poisson probabilities were calculated for each biological sample. At %IDs of 96% or less, the probability of the polymorphism arising by sequencing error was ≤ 0.0002 (Table S2). These data suggest that the low %ID of the clusters in Table 1.3 are not the result of sequencing error.

Table 1.3 Phylotype clusters with low % sequence identity (≤ 97%) to the top hit sequences in the NCBI database.

Phylotype cluster	Accession	BLAST top hit	Score	E value	%ID	Candidate novel taxa	
						Minimum V2 %ID	97% full 16S sequence cutoff
Cluster16042	EU096229.1	*Acetobacter pomorum strain EW816*	433	3.00E-118	95.9	+	
Cluster932	EU096229.1	*Acetobacter pomorum strain EW816*	411	1.00E-111	94.1	+	+
Cluster5070	FJ915625.1	*Acetobacter tropicalis strain IMAU30060*	444	1.00E-121	96.3		
Cluster7664	FJ915625.1	*Acetobacter tropicalis strain IMAU30060*	281	1.00E-72	92.2	+	+
Cluster3222	X76330.1	*Lactobacillus fructivorans (DSM 20203 T)*	436	2.00E-119	96.6		
Cluster668	X76330.1	*Lactobacillus fructivorans (DSM 20203 T)*	381	1.00E-102	92.8	+	+
Cluster467	X76330.1	*Lactobacillus fructivorans (DSM 20203 T)*	385	8.00E-104	91.9	+	+
Cluster8879	X76330.1	*Lactobacillus fructivorans (DSM 20203 T)*	438	6.00E-120	95.9	+	
Cluster94	X76330.1	*Lactobacillus fructivorans (DSM 20203 T)*	390	2.00E-105	92.5	+	+
Cluster1982	AB289116.1	*Lactobacillus fructivorans strain JCM 1198*	320	2.00E-84	93.6	+	+
Cluster4458	GU415690.1	*Lactobacillus brevis clone CX018*	455	6.00E-125	97.0		
Cluster115	GU415690.1	*Lactobacillus brevis clone CX018*	375	5.00E-101	92.3		+
Cluster1522	AB025971.1	*Lactobacillus brevis*	331	1.00E-87	94.4		+
Cluster1494	GU430842.1	*Lactobacillus plantarum clone OCR057*	442	5.00E-121	96.3		
Cluster77	FJ532361.1	*Lactobacillus plantarum strain 14W*	405	6.00E-110	93.8		+
Cluster4922	AB362740.1	*Lactobacillus plantarum strain: NRIC 1749*	326	5.00E-86	93.5		+
Cluster6807	AB362740.1	*Lactobacillus plantarum strain: NRIC 1749*	357	2.00E-95	90.5		+
Cluster1601	GU125508.1	*Lactobacillus plantarum strain: IMAU80086*	394	1.00E-106	93.5		+

To investigate the possibility that the sequences might represent novel taxa, the %ID between each cluster and its top blast hit (Table 1.3) was compared with pairwise %ID comparisons among publicly available sequences representing the same bacterial species in the 16S rRNA databases (Text S1). The minimum values of pairwise %ID

of the V2 region among publicly available sequences of *A. pomorum* and *A. tropicalis* are 99.6% and 94.5% respectively; equivalent values for *L. breivs, L. fructivorans* and *L. plantarum* are 80.8%, 98.5% and 61.8% (Figure. 1.2A and B). The variation in minimum %ID could be explained by its significant negative regression on the number of publicly available sequences (Figure. 1.2A and B), which can be attributed to inadequate sampling at high %ID and possible mis-identifications, especially at low %I*D.*

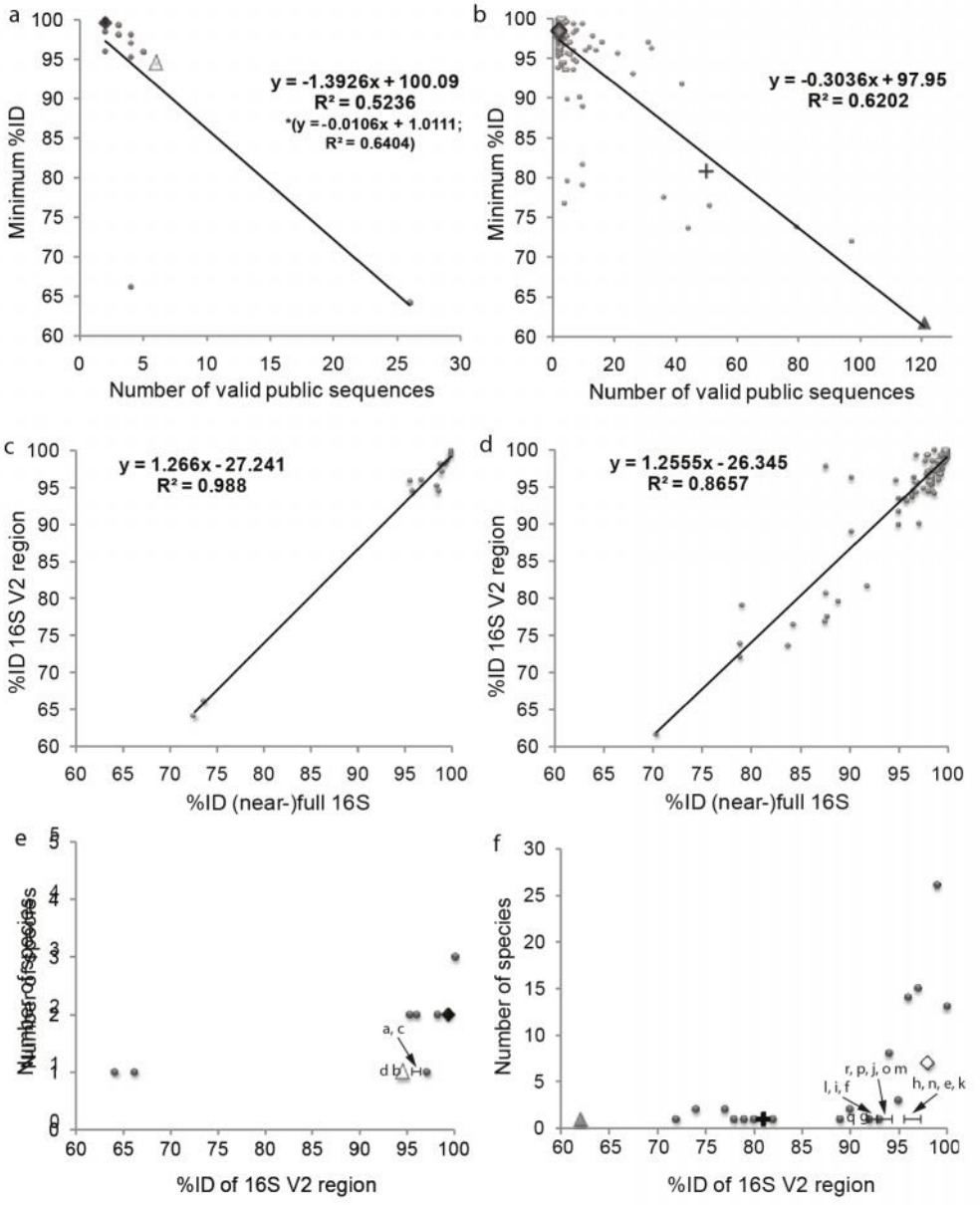


Figure. 1.2 Identification of candidate novel taxa from pairwise comparisons of %ID among 16S rRNA gene sequences.

We adopted two criteria to investigate whether the sequences in Table 1.3 might be candidate novel taxa. The first criterion applied the minimum %ID obtained for publicly available sequences of the target bacterial species as the cut-off value (Figure.

2A and B). When applied to the 18 clusters listed in Table 1.3, this criterion yielded eight clusters, three of *Acetobacter* and five of *Lactobacillus* (Figure. 2C and D, summarized in Table 1.3).

The second criterion was based on the widely used 97% ID of the full 16S rRNA gene as a cut-off threshold to define bacterial species (Drancourt et al., 2004; Drancourt and Raoult, 2005; Janda and Abbott, 2007). A 97% ID of the (near-)full 16S rRNA gene is equivalent to 95.6% ID of the V2 region for *Acetobacter* (96.0% after excluding the outlier species with very low %ID) and 95.4% ID for the V2 region of *Lactobacillus* (whether or not outliers are included) (Figure. 1.2E and F). To be conservative, we rounded down these values to 95% ID cut-off for both species. The 14 clusters in Table 1.3 with < 95% ID to the top hit overlapped with the eight clusters identified by the first criterion [Figure. 1.2C and D, yielding six sequences as representative of candidate novel taxa (Table 3): one each related to *A. pomorum* (cluster932, NCBI GenBank accession HQ168004) and *A. tropicalis* (cluster7664, HQ168006), and four clusters related to *L. fructivorans* (cluster94, HQ168011; cluster467, HQ168009; cluster668, HQ168008; cluster1982, HQ168012)]. The candidate novel species accounted for 0.1–0.8% of the total reads in a sample (calculated from data in Table S1).

QRT-PCR and diagnostic PCR analyses

QRT-PCR conducted on adult flies of different ages in June 2010 confirmed the change in relative number of 16S rRNA gene copies of *L. fructivorans* and *A. pomorum*, from dominance by *L. fructivorans* sequences in young adults to *A. pomorum* sequences in old flies, identified by 454 analysis in November 2009 (Figure.

1.3A). The dominance of 16S rRNA gene copies of *L. plantarum* and *A. tropicalis* in the third-instar larvae and pupae, respectively, were not observed in June 2010, suggesting those life stage-specific effects are not consistent (Figure. 1.3B and C).

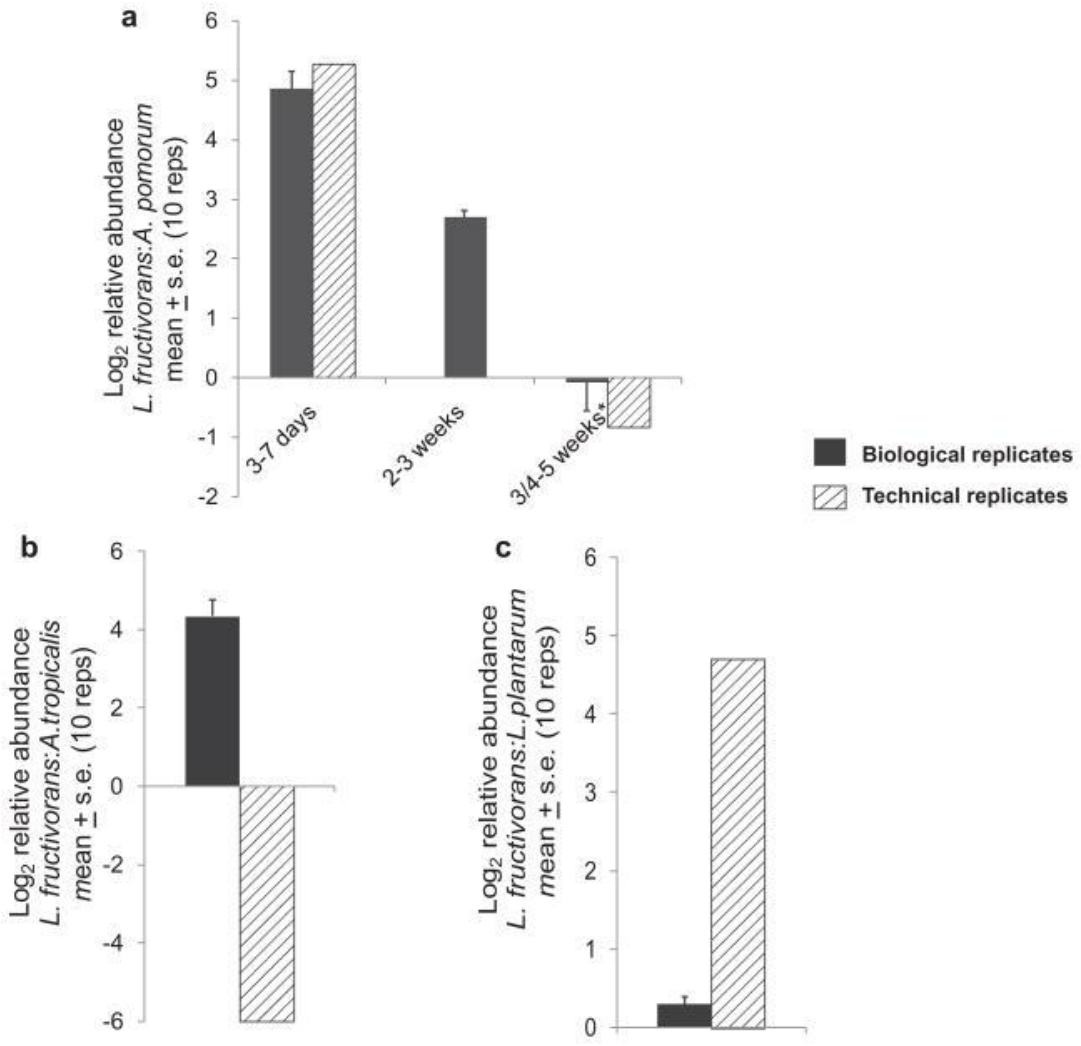


Figure. 1.3 QRT-PCR of relative abundance of 16S rRNA gene sequence in technical replicates of the pyrosequencing experiment (November 2009) and independent biological samples (June 2010).

Each DNA sample used for 454 sequencing comprised many insects of strain Canton-S. To assess the prevalence of 16S rRNA gene copies of the various bacterial taxa in individual insects, the guts from five adult males and females were tested for the five dominant bacteria by diagnostic end-point PCR. All Canton S flies were positive for every bacterium, apart from one female which yielded a negative result for *L. fructivorans* (Figure. S1A). Gut samples from *D. melanogaster* strains Oregon-R and Ithaca-83 also bore *A. pomorum*, *A. tropicalis*, *L. fructivorans* and *L. plantarum*, but were negative for *L. breivs* (Figure. S1B).

Discussion

Previous research on the microbiota of *D. melanogaster* (Ryu et al., 2006; Corby-Harris et al., 2007; Cox and Gilmore, 2007; Ren et al., 2007) employed relatively shallow sampling strategies that would not have detected low-abundance bacteria, and, apart from Ryu et al. (2008), sampled whole insects. These sampling limitations are overcome in this study by the pyrosequencing of dissected guts. The saturation of the rarefaction curves and species richness indices for all samples of larval, pupal and adult flies (Figure. 1.1, Table 1.1) suggests that the entire gut microbiota had been sampled effectively. Nevertheless, it is formally possible that the microbial diversity was underestimated because either the general primers used in this study failed to amplify sequences from certain bacteria, or the amplification of very low-abundance sequences in the template was consistently inadequate for detection. Despite these caveats, which are common to any study founded on PCR, the data indicate that the bacterial community of the *Drosophila* studied here is, indeed, small, with 17–71 OTUs at 97% ID detected, and dominated by just five species in the two genera,

Acetobacter and *Lactobacillus*. The Shannon index of diversity, at 0.35–1.47 (Table 1.1), is lower than values for the microbiota in many habitats, including soils (2.4–3.7) (Fierer and Jackson, 2006), coral-associated assemblages (1.54–3.33) (Garren et al., 2009) and vertebrate gut communities [e.g. 4.29 in ostrich caecum (Matsui et al., 2010)], and it overlaps with values (0.8–1.7) obtained for the gut microbiota in the butterfly, Pieris rapae (Robinson et al., 2010).

The diversity obtained from our inventory of the bacteria in the gut habitat of *Drosophila* is orders of magnitude lower than in the mammalian gut habitat, validating the general pattern in the literature (see Introduction). Low-diversity communities are generated in habitats with extreme disturbance regimes or inhospitable conditions in which few organisms can grow (Grime, 1977; Connell, 1978). The *D. melanogaster* gut is a transient and disturbed environment at multiple spatiotemporal scales, and arguably more so than in mammals. The larval gut persists for about 4 days before dissolution at metamorphosis, followed by the development of the adult gut and its colonization by bacteria; and the lifespan of the adult gut is 4–5 weeks. Additional sources of disturbance include the passage of food, the elimination of the cuticle lining the foregut and hindgut at each larval moult, and sloughing of gut epithelial cells by a process that is accelerated by the presence of microorganisms (Buchon et al., 2009). Features of animal guts that render them inhospitable to many microorganisms include active enzymes (proteases, lysozyme, etc.) and unfavourable oxygen tensions or pH. The oxygen tension in the *D. melanogaster* gut has not been studied directly, but its colonization by *Acetobacter*, which require molecular oxygen, and *Lactobacillus*, which is intolerant of fully oxic conditions (Yamada and Yukphan, 2008; Ljungh and Wadstrom, 2009), suggests that the conditions in the *D. melanogaster* gut are either

microaerobic or spatially variable with respect to oxygen tension. The composition of the gut microbiota may also be influenced by the composition of the food ingested by the insect host (see Introduction). Of particular relevance to the data obtained here, the *Drosophila* used in this analysis had been reared on a nutritionally complex diet of yeast extract, fortified with glucose and supplemented with organic acid preservatives for many generations. This regime is predicted to have exerted a strong and consistent selection pressure, for example against taxa intolerant of the organic acids, and favouring taxa at a competitive advantage in high-glucose environments. Further research is needed to understand the detail of interactions between diet and composition of the gut microbiota for *Drosophila* and other animals.

The low bacterial diversity in the *Drosophila* gut habitat is evident at the within-species level as well as higher taxonomic levels, such that the same OTU at 97% ID is the most abundant representative for each of the five dominant species in every host life stage from early-instar larvae to aged adults. The additional OTUs of each species (Table S1) may represent low-abundance taxa present in many or all individual hosts, or taxa that dominate a few hosts but are absent from most individuals. Low-abundance 'cryptic' taxa have been reported in a various symbiotic systems, including rhizobia in legume root nodules (Denison and Kiers, 2004) and dinoflagellate Symbiodinium in corals (Baker et al., 2004). They may be competitively inferior to the dominant OTU under the prevailing conditions, but become dominant under different circumstances, as reported, for example in coral hosts (Venn et al., 2008). Such shuffling of microbial symbionts can be advantageous to the host, offering insurance against failure of the previous dominant to tolerate or deliver services under different environmental conditions (Douglas, 2010). Alternatively, the minor OTUs

may be deleterious to the host, acting as opportunistic pathogens when controls over their growth and division are relaxed. For example, Gluconobacter morbifer is generally occurs at low abundance in *D. melanogaster* guts, but it proliferates rapidly in immunocompromised flies to become the dominant gut inhabitant with deleterious consequences for the insect (Roh et al., 2008; Ryu et al., 2008). (This species was not detected in our study.)

Central to the design of this study was the variation in the gut bacteria with developmental age and stage of *D. melanogaster*. The pyrosequencing and QRT-PCR analyses concur that the bacterial composition changed with increasing adult age from dominance of 16S rRNA gene sequences of *L. fructivorans* to *A. pomorum* sequences (Table 1.2b and Figure 1.3). *Acetobacter*, unlike *Lactobacillus*, grows rapidly under fully aerobic conditions, raising the possibility that the conditions in the *D. melanogaster* gut become more oxic in ageing insects. Immunological dysfunction associated with ageing can also affect the composition of the gut microbiota, as illustrated by elevated Bacteroides populations in elderly people with persistent activation of the NF-κB transcription factor that plays a central role in innate immunity (Claesson et al., 2011). In this study, two further developmental changes in relative abundance of 16S rRNA gene sequences were identified by pyrosequencing: to high levels of *L. plantarum* sequences in third-instar larvae and *A. tropicalis* sequences in pupae (Table 1.2b). Although confirmed by QRT-PCR of technical replicates, these results were not reproduced in separate biological samples (Figure. 1.3). In the absence of any overt variation in culture conditions, these data point to potentially important sources of environmental variation that remain to be identified.

This study is based exclusively on 16S rRNA gene sequence data. It should be

interpreted with caution in that information on the complement and expression of genes mediating bacterial colonization and proliferation in the gut environment is entirely lacking. This limitation is potentially significant because functionally distinct bacteria with identical or near-identical 16S sequence are known (Scanlan et al., 2009), and differences in gene sequence or expression can have far-reaching phenotypic consequences. For example, gene expression levels are important determinants of the abundance of Leptospirillum bacteria in natural biofilms in acid mine drainage (Denef et al., 2010), and the host range of symbiotic Vibrio is determined by a single regulatory gene (Mandel et al., 2009). These considerations raise the possibility that both the divergent representatives of *Acetobacter* and *Lactobacillus* species in *D. melanogaster* (Table 3) and the bacteria that can confidently be allocated to known species by 16S criteria may be genetically distinct from free-living conspecifics in the content, sequence or regulation of protein-coding genes.

In conclusion, this comprehensive analysis of 16S rRNA gene diversity indicates that the *D. melanogaster* gut bears a low-diversity bacterial community. Further research focusing on the functional traits of the bacteria is critically important to establish the scale of evolutionary change and diversification of protein-coding genes associated with life in an animal gut.

References

Ashelford KE, Chuzhanova NA, Fry JC, Jones AJ, Weightman AJ. New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. Appl Environ Microbiol. 2006;72:5734–5741.

Baker AC, Starger CJ, McClanahan TR, Glynn PW. Coral reefs: corals' adaptive response to climate change. Nature. 2004;430:741.

Bakula M. The persistence of a microbial flora during postembryogenesis of Drosophila melanogaster. J Invertebr Pathol. 1969;14:365–374.

Behar A, Yuval B, Jurkevitch E. Gut bacterial communities in the Mediterranean fruit fly (Ceratitis capitata) and their impact on host longevity. J Insect Physiol. 2008;54:1377–1383.

Buchon N, Broderick NA, Poidevin M, Pradervand S, Lemaitre B. Drosophila intestinal response to bacterial infection: activation of host defense and stem cell proliferation. Cell Host Microbe.2009;5:200–211.

Claesson MJ, Cusack S, O'Sullivan O, Greene-Diniz R, de Weerd H, Flannery E, et al. Microbes and Health Sackler Colloquium: composition, variability, and temporal stability of the intestinal microbiota of the elderly. Proc Natl Acad Sci USA. 2011;108:4586–4591.

Connell JH. Diversity in tropical rain forests and coral reefs. Science. 1978;199:1302–1310.

Corby-Harris V, Pontaroli AC, Shimkets LJ, Bennetzen JL, Habel KE, Promislow DE. Geographical distribution and diversity of bacteria associated with natural populations ofDrosophila melanogaster. Appl Environ Microbiol. 2007;73:3470–3479.

Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Bacterial community variation in human body habitats across space and time. Science. 2009;326:1694–1697.

Cox CR, Gilmore MS. Native microbial colonization of Drosophila melanogaster and its use as a model of Enterococcus faecalis pathogenesis. Infect Immun. 2007;75:1565–1576.

Crotti E, Rizzi A, Chouaia B, Ricci I, Favia G, Alma A, et al. Acetic acid bacteria, newly emerged symbionts of insects. Appl Environ Microbiol. 2010;76:6963–6970.

Denef VJ, Kalnejais LH, Mueller RS, Wilmes P, Baker BJ, Thomas BC, et al. Proteogenomic basis for ecological divergence of closely related bacteria in natural acidophilic microbial communities. Proc Natl Acad Sci USA. 2010;107:2383–2390.

Denison RF, Kiers ET. Lifestyle alternatives for rhizobia: mutualism, parasitism, and forgoing symbiosis. FEMS Microbiol Lett. 2004;237:187–193.

Dethlefsen L, McFall-Ngai M, Relman D*A.* An ecological and evolutionary perspective on human–microbe mutualism and disease. Nature. 2007;449:811–818.

Dillon RJ, Dillon VM. The gut bacteria of insects: nonpathogenic interactions. Annu Rev Entomol. 2004;49:71–92.

Douglas AE. The Symbiotic Habit. Princeton, NJ, USA: Princeton University Press; 2010.

Drancourt M, Raoult *D.* Sequence-based identification of new bacteria: a proposition for creation of an orphan bacterium repository. J Clin Microbiol. 2005;43:4311–4315.

Drancourt M, Berger P, Raoult *D.* Systematic 16S rRNA gene sequencing of atypical clinical isolates identified 27 new bacterial species associated with humans. J Clin Microbiol.2004;42:2197–2202.

Dunn AK, Stabb EV. Culture-independent characterization of the microbiota of the ant lionMyrmeleon mobilis (Neuroptera: Myrmeleontidae) Appl Environ Microbiol. 2005;71:8784–8794.

Fierer N, Jackson RB. The diversity and biogeography of soil bacterial communities. Proc Natl Acad Sci US*A.* 2006;103:626–631.

Garren M, Raymundo L, Guest J, Harvell CD, Azam F. Resilience of coral-associated bacterial communities exposed to fish farm effluent. PLoS ONE. 2009;4:e7319.

Grime JP. Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. Am Nat. 1977;111:1169–1194.

Grunwald S, Pilhofer M, Holl W. Microbial associations in gut systems of wood- and bark-inhabiting longhorned beetles [Coleoptera: Cerambycidae] Syst Appl Microbiol. 2010;33:25–34.

Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM. Accuracy and quality of massively parallel DNA pyrosequencing. Genome Biol. 2007;8:R143.

Janda JM, Abbott S*L.* 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. J Clin Microbiol. 2007;45:2761–2764.

Kunin V, Hugenholtz P. PyroTagger: a fast, accurate pipeline for analysis of rRNA amplicon pyrosequence data. Open J. 2010;1:1–8.

Kunin V, Engelbrektson A, Ochman H, Hugenholtz P. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. Environ Microbiol.2010;12:118–123.

Lehman RM, Lundgren JG, Petzke LM. Bacterial communities associated with the digestive tract of the predatory ground beetle, Poecilus chalcites, and their modification by laboratory rearing and antibiotic treatment. Microb Ecol. 2009;57:349–358.

Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI. Worlds within worlds: evolution of the vertebrate gut microbiota. Nat Rev Microbiol. 2008;6:776–788.

Livak KJ, Schmittgen T*D.* Analysis of relative gene expression data using real-time quantitative PCR and the 2(−Delta Delta C(T)) Method. Methods. 2001;25:402–408.

Ljungh A, Wadstrom T, editors. Lactobacillus Molecular Biology: From Genomics to Probiotics.Norfolk, UK: Caister Academic Press; 2009.

Mandel MJ, Wollenberg MS, Stabb EV, Visick KL, Ruby EG. A single regulatory gene is sufficient to alter bacterial host range. Nature. 2009;458:215–218.

Matsui H, Kato Y, Chikaraishi T, Moritani M, Ban-Tokuda T, Wakita M. Microbial diversity in ostrich ceca as revealed by 16S ribosomal RNA gene clone library and detection of novelFibrobacter species. Anaerobe. 2010;16:83–93.

Morales-Jimenez J, Zuniga G, Villa-Tanaca L, Hernandez-Rodriguez C. Bacterial community and nitrogen fixation in the red turpentine beetle, Dendroctonus valens LeConte (Coleoptera: Curculionidae: Scolytinae) Microb Ecol. 2009;58:879–891.

Needleman SB, Wunsch C*D.* A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol. 1970;48:443–453.

Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, et al. A human gut microbial gene catalogue established by metagenomic sequencing. Nature. 2010;464:59–65.

Ren C, Webster P, Finkel SE, Tower J. Increased internal and external bacterial load duringDrosophila aging without life-span trade-off. Cell Metab. 2007;6:144–152.

Robinson CJ, Schloss P, Ramos Y, Raffa K, Handelsman J. Robustness of the bacterial community in the cabbage white butterfly larval midgut. Microb Ecol. 2010;59:199–211.

Roh SW, Nam YD, Chang HW, Kim KH, Kim MS, Ryu JH, et al. Phylogenetic characterization of two novel commensal bacteria involved with innate immune homeostasis in Drosophila melanogaster. Appl Environ Microbiol. 2008;74:6171–6177.

Ryu JH, Ha EM, Oh CT, Seol JH, Brey PT, Jin I, et al. An essential complementary role of NF-κB pathway to microbicidal oxidants in Drosophila gut immunity. EMBO J. 2006;25:3693–3701.

Ryu JH, Kim SH, Lee HY, Bai JY, Nam YD, Bae JW, et al. Innate immune homeostasis by the homeobox gene caudal and commensal-gut mutualism in Drosophila. Science. 2008;319:777–782.

Scanlan DJ, Ostrowski M, Mazard S, Dufresne A, Garczarek L, Hess WR, et al. Ecological genomics of marine picocyanobacteria. Microbiol Mol Biol Rev. 2009;73:249–299.

Sharon G, Segal D, Fingo JM, Hefetz A, Zilber-Rosenberg I, Rosenberg E. Commensal bacteria play a role in mating preference of Drosophila melanogaster. Proc Natl Acad Sci USA.2010;107:20051–20056.

Venn AA, Loram JE, Trapido-Rosenthal HG, Joyce DA, Douglas AE. Importance of time and place: patterns in abundance of symbiodinium clades a and b in the tropical sea anemone condylactis gigantea. Biol Bull. 2008;215:243–252.

Wilson M. Microbial Inhabitants of Humans. Cambridge, UK: Cambridge University Press; 2005.

Yamada Y, Yukphan P. Genera and species in acetic acid bacteria. Int J Food Microbiol.2008;125:15–24.

CHAPTER 3

THE INCONSTANT GUT MICROBIOTA OF *DROSOPHILA* SPECIES REVEALED BY 16S RRNA GENE ANALYSIS[2]

Abstract

The gut microorganisms in some animals are reported to include a core microbiota of consistently associated bacteria that is ecologically distinctive and may have coevolved with the host. The core microbiota is promoted by positive interactions among bacteria, favoring shared persistence; its retention over evolutionary timescales is evident as congruence between host phylogeny and bacterial community composition. This study applied multiple analyses to investigate variation in the composition of gut microbiota in drosophilid flies. First, the prevalence of five previously described gut bacteria (*Acetobacter* and *Lactobacillus* species) in individual flies of 21 strains (10 *Drosophila* species) were determined. Most bacteria were not present in all individuals of most strains, and bacterial species pairs co-occurred in individual flies less frequently than predicted by chance, contrary to expectations of a core microbiota. A complementary pyrosequencing analysis of 16S rRNA gene amplicons from the gut microbiota of 11 *Drosophila* species identified 209 bacterial operational taxonomic units (OTUs), with near-saturating sampling of

[2] Presented with minor modifications from the originally published article:
Wong, ACN., Chaston, JM. and Douglas, AE. (2013). The inconstant gut microbiota of *Drosophila* species revealed by 16S rRNA gene analysis. *The ISME Journal*, Epub 30 May.
All supplementary information can be found at:
http://www.nature.com/ismej/journal/vaop/ncurrent/suppinfo/ismej201386s1.html?url=/ismej/journal/vaop/ncurrent/full/ismej201386a.html
Chaston, JM contributed to analyses of the 454 data and relationship between *Drosophila* phylogeny and microbiota composition.

sequences, but none of the OTUs was common to all host species. Furthermore, in both of two independent sets of *Drosophila* species, the gut bacterial community composition was not congruent with host phylogeny. The final analysis identified no common OTUs across three wild and four laboratory samples of *D. melanogaster*. Our results yielded no consistent evidence for a core microbiota in *Drosophila*. We conclude that the taxonomic composition of gut microbiota varies widely within and among *Drosophila* populations and species. This is reminiscent of the patterns of bacterial composition in guts of some other animals, including humans.

Introduction

The animal gut is a habitat for microorganisms, which are generally acquired orally with food. Nevertheless, the gut microbiota does not simply reflect the microorganisms in the food, but can be dominated by bacteria that are taxonomically distinct from bacteria in other environments (Ley et al., 2008b; Tamames et al., 2010; Chandler et al., 2011). The distinctiveness of the gut microbiota can be attributed to the ecological conditions in the gut, including regions with extreme pH or redox potential, biologically active compounds (for example, digestive enzymes, immune effectors) and disturbance (for example, bulk flow of food, production of mucus or other extracellular secretions, epithelial cell turnover) (Karasov and Douglas, 2013). Furthermore, the gut is a living habitat, and coevolutionary interactions between the microbiota and the animal have been predicted, potentially resulting in the evolutionary divergence of gut-associated microorganisms from their free-living relatives, and codiversification of the microbiota and animal host (Dethlefsen et al., 2007; Walter et al., 2011). Sustained codiversification results incongruence between

host phylogeny and composition of the gut microbiota.

A subset of the gut microbiota has been reported to be shared among host individuals within various animal species, including *Anopheles* mosquitoes, the honey bee *Apis mellifera*, zebrafish *Danio rerio* and the laboratory mouse (Mohr and Tebbe, 2006; Martinson et al., 2011; Roeselers et al., 2011; Wang et al., 2011; Pedron et al., 2012; Tang et al., 2012). This subset has been described as the core microbiota (Hamady and Knight, 2009; Shade and Handelsman, 2012). Nevertheless, substantial temporal and among-individual variation in composition of the microbiota has been reported in some animals (Robinson et al., 2010; Caporaso et al., 2011; Lozupone et al., 2012; The Human Microbiome Project Consortium, 2012), and it has been suggested that high variability in species composition may be characteristic of some microbial communities in animals and other habitats (Burke et al., 2011).

The presence and abundance of microorganisms in a host can also be influenced by ecological relationships among the gut microorganisms. The interactions may be antagonistic (competition (−/−), amensalism (−/0)) or positive (commensalism (+/0), mutualism (+/+)). Positive interactions would promote the persistence of a core microbiota, while negative interactions would reduce microbial co-occurrence, potentially leading to variation in microbiota composition among host individuals. Specific instances of competition, metabolite cross-feeding and other among-microbe interactions are known, (for example, Coyne et al., 2005; Donohoe et al., 2011; Rosenthal et al., 2011), but the overall contribution of positive and negative interactions to the microbial community has rarely been considered. Exceptionally, Faust et al. (2012) found that most interactions in the human microbiota are negative, suggesting that processes such as competition and niche differentiation may be

important determinants of community structure in this system.

The purpose of this study was to investigate whether drosophilid flies have a core set of gut-associated bacterial taxa. The gut microbiota in these insects has been reported to include Proteobacteria (especially *Acetobacteraceae* and Enterobacteriaceae) and Firmicutes of the order Lactobacillales (notably *Lactobacillus* and *Enterococcus* species). Despite regional variation in conditions (pH, redox potential and so on) in the gut (Shanbhag and Tripathi, 2009), bacteria occur in the crop, midgut and hindgut, with densities up to 106 cells per fly (Corby-Harris et al., 2007; Cox and Gilmore, 2007; Ren et al., 2007; Roh et al., 2008; Sharon et al., 2010; Chandler et al., 2011; Storelli et al., 2011; Wong et al., 2011). Elimination of the gut bacteria can result in delayed larval development, altered lifespan and changes in nutrient allocation attributable to disruption in insect insulin signaling (Brummel et al., 2004; Shin et al., 2011; Storelli et al., 2011; Ridley et al., 2012). An important caveat to our understanding is whether the gut microbiota includes a common phylogenetic subset. Cox and Gilmore (2007) noted three taxa, *Acetobacter aceti*, *A. pasteurianus* and *Enterococcus faecalis*, in two laboratory strains and one wild population, but Corby-Harris et al. (2007) described 74 taxa that were 'unevenly spread' among wild populations of *D. melanogaster*. Chandler et al. (2011) reported that members of Enterobacteriaceae and Lactobacillales are very widely distributed, but apparently not universal, across 20 populations of multiple species. The shallow sampling available to the Sanger sequencing of the 16S rRNA gene used in these studies raises the possibility that some invariant taxa were undetected. This caveat can be addressed by high throughput sequencing as the bacterial communities in *D. melanogaster* are of low diversity, with saturation of rarefaction curves at <20 000 pyrosequencing reads of

16S rRNA gene amplicons (Wong et al., 2011).

In addition to the gut microbiota, some drosophilids possess bacteria, notably *Wolbachia* and *Spiroplasma*, which colonize internal organs, especially the gonads (Mateos et al., 2006). These vertically transmitted bacteria can cause reproductive distortion, and confer protection against natural enemies (Hedges et al., 2008; Werren et al., 2008). They often have intermediate prevalence in populations and species, and do not contribute to the microbiota in the gut lumen (O'Neill et al., 1997; Jaenike et al. 2010).

The specific aims of this study on the gut microbiota of drosophilid flies were twofold. First, we tested for a common subset of the gut microbiota by two complementary methods: taxon-specific PCR assays of bacteria previously shown to account for >90% of the bacteria in *D. melanogaster* (Wong et al., 2011) and pyrosequencing of the total bacterial community. Second, we investigated two ecological patterns likely associated with a core microbiota: positive co-occurrence of different bacteria in individual flies and congruence between host phylogeny and bacterial community composition. Most experiments were conducted on flies in laboratory culture. This enabled us to use aseptically dissected guts (not feasible with field-collected flies), giving assurance that the bacteria scored were members of the gut microbiota. Supplementary whole-body analyses of field-collected *D. melanogaster* compared the microbiota in wild and laboratory flies of one species.

This first comprehensive analysis of the gut microbiota in multiple *Drosophila* species revealed that the composition of the gut microbiota is remarkably inconstant, and does not vary in concordance with host phylogeny. In this respect, we found no evidence of microbial taxa that are shared in all *Drosophila* hosts.

Experimental procedures

Drosophila samples

Samples of adult *Drosophila* were derived from: 11 *Drosophila* species reared at Cornell University on Y-G diet (Brewer's yeast (MP Biomedicals, Santa Ana, CA, USA) and glucose (Sigma, St Louis, MO, USA) (both at 83 g l^{-1}), agar (10 g l^{-1} (Genesee Scientific, San Diego, CA, USA) and preservatives (0.04% phosphoric acid, 0.42% propionic acid (Sigma); seven *Drosophila* species maintained at University of Rochester on Formula 4–24 (Carolina Biological Supply Company); and samples of *D. melanogaster* adults (mixed age and sex) collected from three USA sites and fixed immediately in 70% ethanol (Table 2.1 and Supplementary Table S1).

Table 2.1 *Drosophila* species and number of bacterial OTUs identified from 454 sequence analysis of 16S rRNA gene amplicons.

Species	*Strain*	*Source*	*Number of OTUs*[a]
Set -1: reared on Y-G diet at Cornell University			
D. ananassae	DSSC #14024-0371.13	Hawaii	42
D. erecta	DSSC #14021-0224.01	Not known	17
D. melanogaster	Canton S	Not known	41,48[b]
D. persimilis	DSSC #14011-0111.42	Port Coquitlam, British Columbia, Canada	35
D. pseudoobscura	SD02	Not known	130
D. santomea	DSSC #1402 -0271.01	San Tome and Principe Island	46
D. sechellia	DSSC #14021-0248.03	Cousin Island, Seychelles	62
D. simulans	DSSC # 14021-0251.001	Georgetown, Guyana	36
D. virilis	DSSC #15010--1051.87	Not known	64
D. willistoni	DSSC # 14030-0811.24	Guadaloupe Island, France	48
D. yakuba	DSSC # 14021-0261.01	Liberia	62
Set-2: reared on Formula 4-24 at Rochester University			
D. deflecta	15130-2018.00	Skunk cabbage-feeder, Princeton NJ	318
D. duncanii	NJ-1 2009	Mushroom-feeder, New Jersey	108
D. falleni	Pittsford 2010-1	Mushroom-feeder, Pittsford, NY	223
D. munda	SWRS 2005	Mushroom-feeder, SW Research Station, Portal, AZ	195
D. neotestacea	W+S+ Pittsford 2007	Mushroom-feeder, Pittsford, NY	71
D. quinaria	Pittsford 2010-1	Skunk cabbage-feeder, Pittsford, NY	204,224[b]
D. suboccidentalis	Diamond Lake 2005-19	Mushroom-feeder, Diamond Lake, Oregon	194
Set 3: fresh field-collected			
D. melanogaster	NY1	Decaying vegetation at Pittsford, NY (09/7/10)	178
D. melanogaster	NY2	Apples at Apple Farm, Victor, NY (09/7/10)	61
D. melanogaster	AZ	Bananas at SW Research Station, AZ (09/10/11)	110

Abbreviation: OTUs - Operational taxonomic units, defined with pair-wise 97% sequence identity
[a] Samples comprised dissected guts for set-1 and set-2, and whole insect bodies for set-3
[b] Two technical replicates were analyzed, with both values displayed.

DNA isolation

Total genomic DNA was extracted from isolated adult fly guts or whole-bodies (age and sex varying with experiment, as below) by the method of Cenis et al. (1993). Guts from surface-sterilized flies were dissected in sterile Ringer's solution as previously described (Wong et al., 2011). Samples were homogenized in 180 µl lysis buffer (20 mM Tris-HCl, pH 8.0, 2 mM sodium EDTA, 1.2% Triton-X 100, 20 mg ml^{-1} lysozyme) and incubated at 37 °C for 1.5 h, with brief bead-beating at 45 min in a Disruptor Genie using 0.1 mm glass beads (Scientific Industries, Bohemia, NY, USA). Twenty microlitres 10 × extraction buffer (2 M Tris-HCl, pH 8.5, 2.5 M NaCl, 0.25M EDTA, 5% w/v SDS) and 10 µl proteinase K (20 mg ml^{-1}) were added, samples were

incubated at 55 °C for 1 h and precipitated with 100 µl 3 M sodium acetate (pH 5.2). The supernatant was mixed with equal volume 100% ice-cold isopropanol and incubated at room temperature for 30 min before centrifugation for 30 min at 18 000 g. After discarding the supernatant, each pellet was washed in 500 µl 70% ice-cold ethanol, dried and resuspended in 20 µl sterile water.

End-point PCR assays of bacterial prevalence

L. breivs, L. fructivorans, L. plantarum, Acetobacter pomorum and *A. tropicalis* in the guts of individual flies were scored by end-point PCRs using taxon-specific 16S rRNA gene primers (Supplementary Table S2a). The experimental samples were five; 5–7-days-old and 4–5-weeks-old adults of both sexes, run in parallel with positive controls comprising DNA from pure culture of the corresponding bacteria and sterile water as negative control. The PCR reactions were as in Wong et al. (2011). PCR products were separated by gel electrophoresis using 1% agarose gel and visualized with SYBRSafe (Invitrogen, Carlsbad, CA, USA). Sanger's sequencing confirmed the identity of representative bands.

Multiplex 454 pyrosequencing of 16S rRNA gene sequences

Each sample comprised 50 guts (laboratory-reared flies) or 10 bodies of *D. melanogaster* (laboratory strain ZH26, wild samples), with a drop of Ringer's solution treated as for dissections but without insect material as the negative control. The laboratory fly samples comprised approximately equal numbers of males and females, and were of similar age range within set-1 and ZH26 (5–10-days-old) and set-2 (a broad age distribution for every species); the wild flies were of unknown age. 16S

rRNA amplicons of the V2 16S rRNA region were prepared as previously described (Wong et al., 2011), with primers mentioned in Supplementary Table S2b. Equal amounts (ng) of three reaction products per sample were mixed and purified using the QIAquick PCR purification kit (QIAGEN, Valencia, CA, USA), followed by Pico-Green quantification. Emulsion PCR was conducted at 1.5 copies per bead using only 'A' beads for unidirectional 454 GS-FLX pyrosequencing with standard Titanium chemistry.

Pyrosequencing flowgrams were converted to sequence reads using 454 Life Science software (www.454.com). Reads with ambiguous nucleotides (N) and <270 nucleotides after the forward primer, and mismatches with the 16S rRNA gene primers were excluded in the initial filtering. Multiplexed samples from two half-plate runs were combined before downstream analyses by modifying the barcodes in the fna files and concatenating the two fna and qual files into a single fna and qual file, respectively. The QIIME 1.4.0 virtualbox package was used to split the multiplexed sequences, discard chimeras, denoise the data, bin sequences at 97% sequence identity and make taxonomy calls to genus level (Caporaso et al., 2010). Default parameters were used except that the denoising cutoff was set to retain doubletons, and the RDP classifier was applied using a custom Greengenes database to assign class through genus designations. Species identity of each operational taxonomic units (OTU) was assigned by local BLAST (Stand-alone MEGABLAST program) with the 16S Microbial database (June 2012). OTUs with either single reads or fewer reads than in the negative controls were excluded. For comparison, OTU tables were generated in Pyrotagger (http://pyrotagger.jgi-psf.org/release). Reads assigned to *Wolbachia* were excluded because, first, this bacterium is not a member of the gut microbiota (it has

weak tropism for the gut, and does not inhabit the gut lumen); and, second, the *D. ananassae* genome includes laterally acquired *Wolbachia* sequences (Dunning Hotopp et al., 2007), such that *Wolbachia* reads are a measure of host DNA in the gut samples (*D. ananassae* accounted for 90% of *Wolbachia* reads across all gut samples assayed). For consistency, *Wolbachia* reads were removed from data sets for whole-body samples of wild flies. Reads assigned to *Wolbachia* are shown as 'excluded sequences' in Supplementary Table S3, and the minimal effect of their exclusion on our analysis is indicated by PCA plots in Supplementary Figure S1. The samples included technical replicates for two *Drosophila* species (*D. melanogaster* in set-1, *D. quinaria* in set-2). PCA plots of the bacterial communities were created using pcaMethods (Stacklies et al., 2007) in R (R Development Core Team, 2012), following log-transformation of number of reads per OTU. Correlation matrices derived from the OTU tables were used to create dendrograms of the bacterial communities using pvclust (CRAN.R-project.org/package=pvclust) and ape (Paradis et al., 2004) in R, and compared with *Drosophila* phylogenetic trees built in BioEdit from a clustalX alignment of concatenated DNA sequences obtained from NCBI. Trees were manipulated in FigTree v1.3.1. Graphical taxonomy networks were created using the make_otu_network.py QIIME script and visualized as an unweighted forced-directed layout with Cytoscape v2.8.2 (Smoot et al., 2011) using default QIIME instructions. The analyses shown were conducted with the full microbiota; the patterns were equivalent when OTUs representing <1% or <0.1% of reads were excluded (data not shown).

Bacterial co-occurrence analyses

The likelihood of co-occurrence of bacterial species in individual flies was analysed by C-score test (Stone and Roberts, 1990) using data obtained by PCR with taxon-specific primers for each fly. C-score calculates the mean number of instances where two bacterial species co-occur, across all fly species pairs. The computed C-score is significantly greater than the null distribution if the bacteria co-occur less frequently than predicted by chance, and less than the null distribution for positive co-occurrence. The prevalence data sets were arranged in presence-absence matrices with the five bacterial species as rows and individual flies as columns. The most appropriate null model for these data, in which the presence/absence of each bacterial species in each fly is known, is the 'fixed-fixed' null model (SIM9 of Gotelli, 2000). The observed data matrices were compared with 5000 randomly generated matrices using EcoSim 7.72 (Gotelli and Entsminger, 2012).

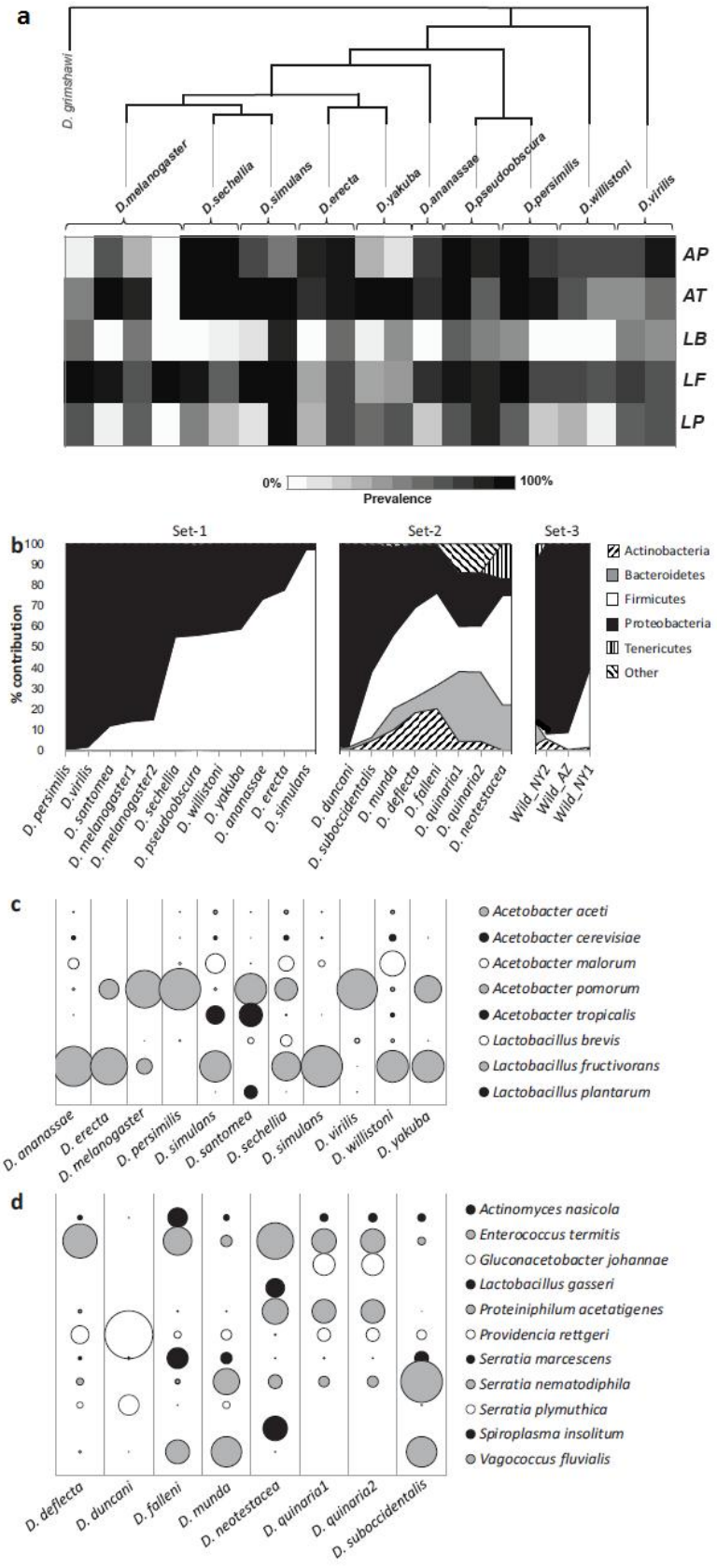
Results

Prevalence of bacteria in laboratory Drosophila populations

Our first approach to investigate the taxonomic composition of the gut microbiota in *Drosophila* was to score for five bacterial taxa in individual flies of 21 strains in 10 *Drosophila* species (Figure 2.1a). The five bacteria have previously been shown to account for >90% of the bacteria in multi-individual samples *D. melanogaster* strain Canton-S in our laboratory (Wong et al., 2011). No bacterial taxon was detected in every individual of every fly strain. One bacterium, *L. fructivorans*, was detected in at least one fly of every *Drosophila* strain; *A. pomorum*, *A. tropicalis* and *L. plantarum* were detected in every strain except *D. melanogaster* ZH26 (strain-4 in Figure 2.1a)

and *L. breivs* was detected in 13 (62%) of the strains. Overall, the frequency of each bacterium did not vary significantly with age (5–7-days-old versus 4–5-weeks-old) or sex ($P>0.05$), but the frequency of *A. pomorum*, *L. breivs* and *L. plantarum* varied significantly among strains ($P<0.001$).

Figure 2.1 Analysis of the composition of bacterial communities in *Drosophila* species. (a) Prevalence of 5 bacterial taxa in 21 *Drosophila* strains (*Drosophila* strain details provided in Supplementary Table S1). (b) Abundance of bacterial phyla in pyrosequence analyses. (c) Abundance of dominant species in *Drosophila* species set-1 based on 97% similarity OTU assignments. (d) Abundance of dominant species in *Drosophila* species set-2 based on 97% similarity OTU assignments.


a
D. grimshawi
D.melanogaster
D.sechellia
D.simulans
D.erecta
D.yakuba
D.ananassae
D.pseudoobscura
D.persimilis
D.willistoni
D.virilis
AP
AT
LB
LF
LP
0%
100%
Prevalence
b
Set-1
Set-2
Set-3
% contribution
Actinobacteria
Bacteroidetes
Firmicutes
Proteobacteria
Tenericutes
Other
D. persimilis
D.virilis
D. santomea
D. melanogaster1
D. melanogaster2
D. sechellia
D. pseudoobscura
D. willistoni
D. yakuba
D. ananassae
D. erecta
D. simulans
D. duncani
D. subocccidentalis
D. munda
D. deflecta
D. falleni
D. quinaria1
D. quinaria2
D. neotestacea
Wild_NY2
Wild_AZ
Wild_NY1
c
Acetobacter aceti
Acetobacter cerevisiae
Acetobacter malorum
Acetobacter pomorum
Acetobacter tropicalis
Lactobacillus brevis
Lactobacillus fructivorans
Lactobacillus plantarum
D. ananassae
D. erecta
D. melanogaster
D. persimilis
D. simulans
D. santomea
D. sechellia
D. simulans
D. virilis
D. willistoni
D. yakuba
d
Actinomyces nasicola
Enterococcus termitis
Gluconacetobacter johannae
Lactobacillus gasseri
Proteiniphilum acetatigenes
Providencia rettgeri
Serratia marcescens
Serratia nematodiphila
Serratia plymuthica
Spiroplasma insolitum
Vagococcus fluvialis
D. deflecta
D. duncani
D. falleni
D. munda
D. neotestacea
D. quinaria1
D. quinaria2
D. subocccidentalis

In a few of the 21 *Drosophila* strains, every individual scored positive for a bacterial taxon: nine (43%) strains for *A. tropicalis*, five (24%) for *L. fructivorans*, four (19%) for *A. pomorum*, one (5%) for *L. plantarum* and none for *L. breivs*. Furthermore, each of the bacteria was at intermediate prevalence (that is, at least one fly scored positive and one fly scored negative) in more than half of the 21 strains (ranging from 52% for *A. tropicalis* to 90% for *L. plantarum*). By the criterion of diagnostic PCR assay, most of the five bacteria are not members of the core microbiota in most of the *Drosophila* strains, and none was core to every strain.

To investigate the pattern of occurrence of the five bacteria across the individual flies, the data set was analysed by C-scores. The C-score for the full data set, 4114.5, was significantly higher than expected by chance ($P<0.001$), indicating that the bacterial species co-occurred less often than in random distribution. Significantly elevated C-scores were also obtained for young males ($P<0.001$) and females ($P<0.002$), and old males ($P<0.01$), indicative of segregation among the bacteria in these samples. The C-score for old females was not significant ($P>0.05$). In general, significant scores were associated with negative relationships between *L. fructivorans* and *Acetobacter* species. The observed segregation among these bacteria would tend to hinder the assembly of a core microbiota.

One *D. melanogaster* strain, ZH26, was unique; in that every fly was colonized with only one of the five tested taxa: *L. fructivorans* (Figure 1a). In a complementary 454 analysis (Supplementary Table S3a), *L. fructivorans* accounted for >99% of the 55 683 reads, confirming the PCR data and indicating that strain ZH26 does not bear a highly divergent bacterial community. This colonization status was not consistent across fly generations: when the five taxon-specific PCR assays were repeated on the same stock

of ZH26 9 months later, all five bacteria were universally present; but *L. breivs* and *L. plantarum* were absent after a further 3 months (Table 2.2, Supplementary Figure S2a).

Table 2.2 Temporal variation in frequency of bacterial taxa in individual adult *D. melanogaster*.

Taxon-specific PCR assay	*Strain ZH26 (10 flies per sample)*				*Strain Canton S (10 flies per sample)*			
	Number of flies positive for bacterium			X^2 *(2 df)*	*Number of flies positive for bacterium*			X^2 *(2 df)*
	Feb 2011 (time-0)	*Nov 2011 (9 months)*	*Apr 2012 (12 months)*		*May 2010 (time-0)*	*Feb 2011 (9 months)*	*Feb 2012 (21 months)*	
A.pomorum	0	10	10	10, p=0.007*	10	10	6	1.23, p=0.54
A.tropicalis	0	10	10	10, p=0.007*	10	4	10	3, p=0.223
L.brevis	0	10	0	20, p <0.001*	10	5	0	10, p=0.007*
L.fructivorans	10	10	10	0, p=1	9	10	9	0.07, p=0.966
L.plantarum	0	10	0	20, p<0.001*	10	9	0	10, p=0.007*

To assess whether variability in the composition of the gut bacteria was unique to ZH26, we determined the prevalence of the five dominant gut taxa in *D. melanogaster* strain Canton-S (in which the five taxa were originally identified (Wong et al., 2011). All five bacteria were detected, but none was universally present, in the three samples of 10 flies analysed over 21 months. The prevalence of every bacterium shifted between the three sampling periods, and *L. breivs* and *L. plantarum* varied between being present in all and none of the 10 flies tested (Table 2.2, Supplementary Figure

S2b). We conclude that variation in bacterial prevalence is not unique to strain ZH26. These results indicate that the five bacteria previously identified as major constituents of the gut microbiota under our laboratory rearing conditions are not universally present in all individual flies, and they vary in prevalence across generations.

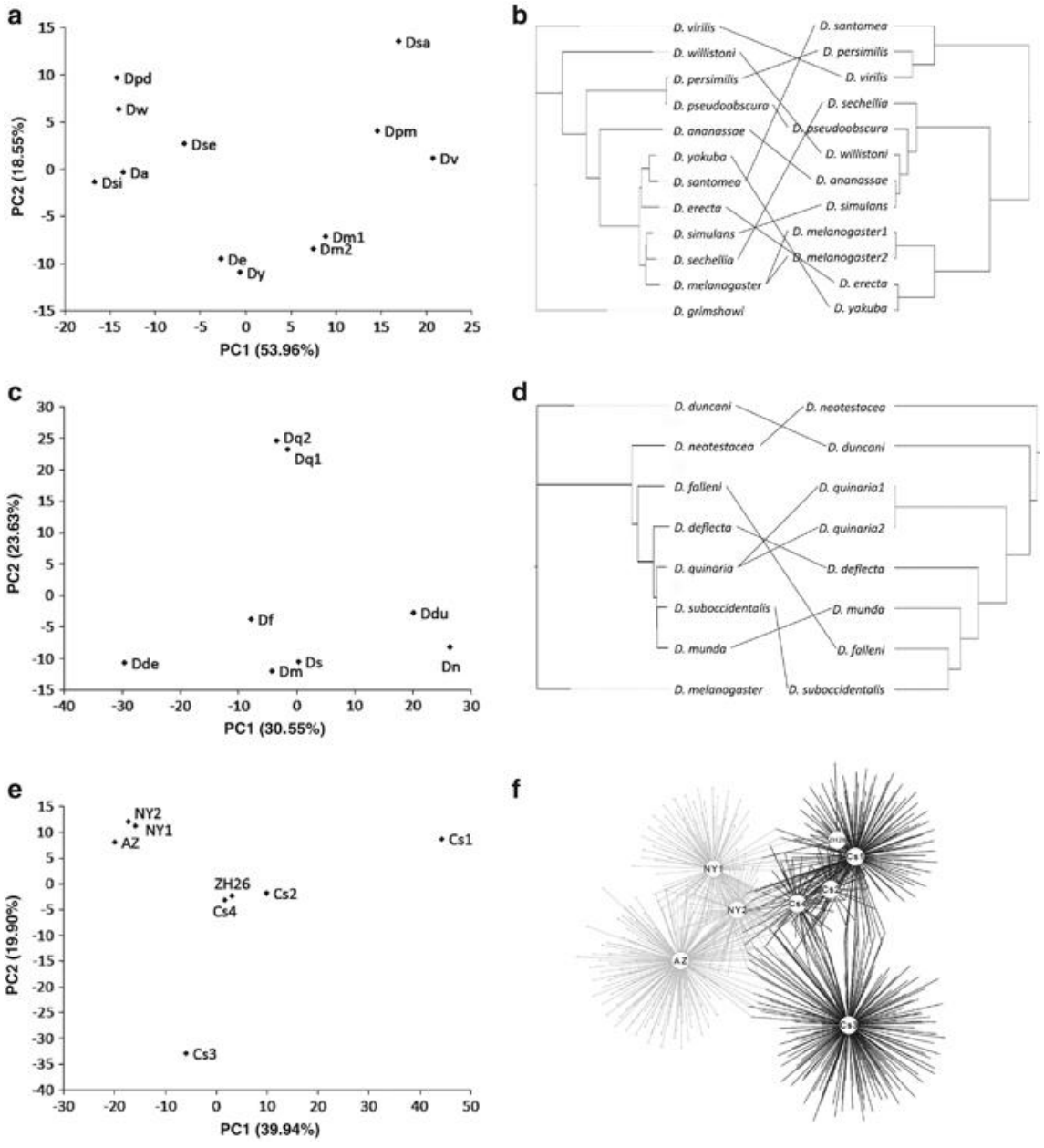
Pyrosequencing of bacterial communities in Drosophila

As an alternative approach to investigate the bacterial communities in *Drosophila* guts, we quantified the total gut microbiota by pyrosequencing 16S rRNA gene amplicons from three independent sets of drosophilid flies (Table 2.1). In total, 26 811–62 138 reads of 16S rRNA gene amplicons per sample were identified in QIIME, after quality filtering and removal of chimeras and single reads (Supplementary Table S3b–d). All the rarefaction curves tended to saturation (Supplementary Figure S3), indicating that the OTUs were representative of the total bacterial community in each sample. Close correspondence in the number and identity of the OTUs between two technical replicate samples (samples of the same genomic DNA) were obtained for both *D. melanogaster* in set-1 (Pearson's correlation coefficient $r=0.998$, $P<0.0001$) and *D. quinaria* in set-2 ($r=0.959$, $P<0.0001$) (Supplementary Figure S4), showing that random sampling effects, which have constrained the reproducibility of pyrosequencing data in certain complex bacterial communities (Zhou et al., 2011), were not significant in this study. The combined data for set-1 and set-2 (Supplementary Table S3e) were also processed by Pyrotagger, an alternative program used in our previous research on the gut microbiota of *D. melanogaster* (Wong et al., 2011). The correlation between the outputs of QIIME and Pyrotagger was highly significant for numbers of reads (Pearson's correlation coefficient, $r=0.988$, $P<0.001$) and OTUs ($r=0.972$, $P<0.001$), although, on average,

18% fewer reads were obtained by Pyrotagger than QIIME (Supplementary Table S4). All 16S rRNA gene amplicon reads in set-1 could be assigned to two phyla: Proteobacteria and Firmicutes (Figure 2.1b). Two genera, *Lactobacillus* (Firmicutes) and *Acetobacter* (α-Proteobacteria) accounted for 94–100% of the reads in every sample (Supplementary Table S3b). The most abundant bacterium in every *Drosophila* species was either *L. fructivorans* OTU179 or *A. pomorum* OTU630, which accounted for up to 63% and 82%, respectively, of all 16S reads per sample (Figure 2.1c). Nonetheless, none of the 209 OTUs or 124 bacterial species were present in every *Drosophila* species (Supplementary Table S3b). We conclude that no bacterial taxon at the level of OTU or species is present at detectable levels in all 11 *Drosophila* species.

The data for set-1 were investigated by PC*A.* Phylogenetically related *Drosophila* species were not clustered by the first two axes, which together accounted for 73% of the variance (Figure 2.2a), or any other axis combination tested. The implication that the bacterial communities were not patterned according to host phylogeny was confirmed by the poor correspondence between the phylogenetic relationship among the 11 *Drosophila* species and the relatedness of host-associated gut bacterial community taxonomic composition (Figure 2.2b). Furthermore, the bacterial communities could not be differentiated between fly samples possessing and lacking *Wolbachia* (Supplementary Figure S1).

Figure 2.2 Relationship between bacterial community composition and *Drosophila* species based on 97% similarity OTU assignments. (a) Principal components analysis (PCA) of the bacterial community and (b) correspondence between dendrograms of bacterial communities and the phylogeny of *Drosophila* in set-1. (c) PCA and (d) dendrogram correspondence in *Drosophila* set-2. (e) PCA and (f) bipartite graph of *D. melanogaster* from wild (AZ, NY1, NY2: see Table 1c) and laboratory (CS1-4 (Canton-S isolates) and ZH26: see legend to Supplementary Table S3f). (Abbreviations in (a) and (c) indicate species name, as provided in (b) and (d), respectively, for example, Dq is *D. quinaria*, Dde is *D.* deflecta).


a
PC2 (18.55%)
PC1 (53.96%)
Dsa
Dpd
Dw
Dpm
Dse
Dv
Da
Dsi
Dm1
Dm2
De
Dy
b
D. virilis
D. willistoni
D. persimilis
D. pseudoobscura
D. ananassae
D. yakuba
D. santomea
D. erecta
D. simulans
D. sechellia
D. melanogaster
D. grimshawi
D. santomea
D. persimilis
D. virilis
D. sechellia
D. pseudoobscura
D. willistoni
D. ananassae
D. simulans
D. melanogaster1
D. melanogaster2
D. erecta
D. yakuba
c
PC2 (23.63%)
PC1 (30.55%)
Dq2
Dq1
Df
Ddu
Dde
Ds
Dm
Dn
d
D. duncani
D. neotestacea
D. fallleni
D. deflecta
D. quinaria
D. suboccidentalis
D. munda
D. melanogaster
D. neotestacea
D. duncani
D. quinaria1
D. quinaria2
D. deflecta
D. munda
D. fallleni
D. suboccidentalis
e
PC2 (19.90%)
PC1 (39.94%)
NY2
NY1
AZ
Cs1
ZH26
Cs2
Cs4
Cs3
f
NY1
NY2
AZ
Cs1
Cs2
Cs4
Cs3

As an independent test for the relationship between bacterial community composition and host phylogeny, we investigated the bacterial community in guts dissected from *Drosophila* species of set-2 (Table 2.1). These bacteria included representatives of seven phyla (Figure 2.1b, Supplementary Table S3c) and were dominated by Enterococcus termitis OTU659 and Vagococcus fluvialis OTU4 in the Firmicutes (Lactobacillales), and Providencia rettgeri OTU937 and Serratia nematodiphila OTU3 in the γ-Proteobacteria (Figure 2.1d). Ten (1%) of the 997 OTUs were detected in all seven *Drosophila* species (Supplementary Table S3c), accounting in total for 1–70% of the reads (median 9%), but the prevalence of these OTUs among the individual flies contributing to each samples (that included both sexes and a broad age range) is unknown. As with set-1, the relationship among bacterial communities did not map onto the phylogeny of their *Drosophila* hosts (Figures 2.2c and d).

Our final analysis tested for bacterial OTUs or species shared across field-collected and laboratory samples of a single *Drosophila* species, *D. melanogaster*. The three field-collected samples included representatives of three bacterial phyla: Firmicutes, Proteobacteria and at <5%, Actinobacteria (Figure 2.1b). The dominant Firmicutes included Leuconostoc mesenteroides OTU5 and Lactococcus lactis OTU121, and the abundant Proteobacteria were *Acetobacteraceae*, specifically Gluconobacter japonicus OTU4 and Gluconobacter albidus OTU6 and the γ-proteobacterium Tatumella ptyseos OTU1 (Supplementary Table S3d).

The bacterial communities in the three wild samples were compared with five data sets for laboratory cultures of *D. melanogaster*. No OTU or species was detected in each sample (Supplementary Table S3f), offering no support for bacterial taxa universally present in the guts of *D. melanogaster*. The wild samples grouped together closely on

the first two axes of the PCA, and were separated from the laboratory samples (Figure 2.2e). The difference between wild and laboratory samples and the greater variability among laboratory samples are confirmed by the bipartite graph, in which the edges connect each host sample node to every bacterial OTU in that sample (Figure 2.2f).

Discussion

Immigration with food and emigration with feces are important processes shaping the microbial community in animal guts, including variation in community composition among host individuals and over time within one host. Despite this continual flux of microorganisms through the gut habitat, a subset of the microorganisms is consistently recovered from certain animal taxa. This subset, sometimes described as the 'core microbiota', is of special interest because it is predicted to be ecologically-distinctive and may have coevolved with the host.

The concept of the core microbiota has been applied in multiple ways. In some studies, specific bacterial taxa has been detected in all samples, each of which comprised multiple hosts, but the prevalence of the bacteria in each individual was not tested (Mohr and Tebbe, 2006; Martinson et al., 2011; Roeselers et al., 2011; Wang et al., 2011). Other investigations have tested individual hosts, often with study-specific criteria for a core, for example, relaxation of the detected prevalence of the bacteria to 80 or 50% of hosts, or use of variable or low (<97%) OTU-call cutoffs (Qin et al., 2010; Boissiere et al., 2012; Moran et al., 2012; Nelson et al., 2012; Salonen et al., 2012). Such relaxation can be justified for technical reasons, including the artifactual inflation of community diversity from contamination, error in sequencing and sequence alignment and incomplete sampling, especially for highly diverse bacterial

communities (Huse et al., 2010; Kunin et al., 2010; Sun et al., 2012; Wylie et al., 2012). Nevertheless, the variation in criteria adopted across different studies and animal systems hinders systematic analysis of the degree of partner fidelity between animals and their gut microbiota.

Despite these general difficulties, consistent patterns in the taxonomic diversity of the gut microbiota in drosophilid flies are emerging. The bacterial communities are very predictable at high-phylogenetic levels, dominated by one, two or all of the order Lactobacillales (phylum Firmicutes) and the families Enterobacteriaceae and *Acetobacteraceae* (phylum Proteobacteria) (References in Introduction); but they vary irregularly at the level of genus, species and OTU. This inconstancy is evident at multiple phylogenetic scales of the host, among-species, within-species and even within single laboratory lines, and no OTU was detected in every sample analysed in this study. Although technical artifacts can inflate among-sample differences (see above), the severity of these limitations is much ameliorated in this study of the *Drosophila* system by the use of whole-gut samples and the near-saturation of sequence reads.

The composition of the gut microbiota can also be affected strongly by rearing conditions. In particular, repeated environmental perturbations (including variation in food consumed) in the field may prevent the realization of the full core microbiota in some individuals, while laboratory-reared animals may not have access to key members of the core microbiota occurring in the natural habitat. In this study, individual microbial taxa were not generally found to be shared universally, either within or among drosophilid species in laboratory conditions. In particular, the data do not substantiate the common bacterial taxa found by Cox and Gilmore (2007) across

laboratory and field conditions for one species, *D. melanogaster*. Our results complement and extend the research of Chandler et al. (2011), in which shallow sampling with Sanger sequencing failed to yield a common subset of bacterial OTUs among field samples of multiple *Drosophila* species.

The incongruence between drosophilid phylogeny and bacterial community composition suggests weak partner fidelity, and that a consistent microbiota does not operate across evolutionary timescales in this system. The implication is that, in terms of taxonomic composition, the gut microbiota in *Drosophila* has neither coevolved with the host over evolutionary time, nor tracked evolutionary changes in gut physiology that may vary according to phylogenetic relatedness between different host taxa. In this respect, *Drosophila* appears to parallel mammals, for which no phylogenetic pattern in the composition of the gut microbiota has been found (Ley et al., 2008a; Muegge et al., 2011). Our results differ from the evidence for congruence between host phylogeny and gut microbiota composition obtained, for example, for bacterial community composition in laboratory cultures of jewel wasps Nasonia (Brucker and Bordenstein, 2012), and wild populations of both great apes/humans (Ochman et al., 2010) and termites (Hongoh et al., 2005); and the genotypes of one bacterial species, *Lactobacillus* reuteri, in studies that included inbred lab mice and rats (Oh et al., 2010; Frese et al., 2011). An important issue for future work is the ecological factors that dictate the variation in the congruence of host-microbiota phylogenies across different animal groups.

The inconstancy in *Drosophila* gut microbiota composition raises two broad issues: the population processes that dictate whether a microbial community includes a consistent subset, and how taxonomic composition influences the functional traits of

the bacterial community. A consistent subset of the microbiota would be promoted by high rates of transmission between conspecific hosts (including parent-to-offspring and among kin), and by extended residence time within individual hosts. As gut microorganisms are routinely shed in feces, residence time of an ingested microorganism and its descendants is shaped by the relationship between the rates of proliferation and emigration (Savage, 1977; Costello et al., 2012). Variation in these key ecological parameters among different animal groups has yet to be compared systematically. The second issue, the relationship between taxonomic and functional diversity of animal-associated bacterial communities, has been studied in mammalian gut associations, with evidence that taxonomically diverse bacterial communities can be functionally equivalent, for example, that a taxonomically-variable microbiota can potentially support a core microbiome (Turnbaugh et al., 2008; Muegge et al., 2011; The Human Microbiome Project Consortium, 2012; Morgan et al., 2013). Research to date on *Drosophila* has focused on the relationship between taxonomic composition of the bacteria and host phenotype. There is persuasive evidence that individual members of the gut microbiota vary in their impact on the phenotype of *D. melanogaster* (Shin et al., 2011; Storelli et al., 2011), but the effects of natural variation among bacteria on the phenotype and fitness of *Drosophila* in laboratory culture and field remain to be studied.

Relevant to these considerations, multiple aspects of insect function can be altered by experimental elimination of the gut microbiota, including intestinal cell proliferation, nutrient content, metabolic rate, insulin signaling, larval developmental rates and lifespan (Brummel et al., 2004; Buchon et al., 2009; Shin et al., 2011; Storelli et al., 2011; Ridley et al., 2012). These data suggest *Drosophila* is adapted to the presence of

microorganisms in the gut, even though the taxonomic composition of the microbiota is variable and partner fidelity is weak at timescales ranging from a few generations in a single laboratory culture to millions of years of *Drosophila* evolution. As noted above, this may be reminiscent of evolution with a core microbiome rather than a core microbiota. Furthermore, many drosophilid flies are additionally associated with vertically-transmitted bacteria (especially *Wolbachia* and Spiroplasma) localized to the reproductive organs and other internal tissues. It is an open question whether these bacteria may influence the composition and function of the spatially distinct gut microbiota. A full understanding of the ecology of the inconstant gut microbiota of drosophilids will require further research on the interaction of host traits with the composition and activities of the bacterial taxa.

References

Boissiere A, Tchioffo MT, Bachar D, Abate L, Marie A, Nsango SE et al (2012). Midgut microbiota of the malaria mosquito vector Anopheles gambiae and interactions with Plasmodium falciparum infection. PLoS Pathog 8: e1002742.

Brucker RM, Bordenstein SR. (2012). The roles of host evolutionary relationships (genus: Nasonia) and development in structuring microbial communities. Evolution 66: 349–362.

Brummel T, Ching A, Seroude L, Simon AF, Benzer S. (2004). Drosophila lifespan enhancement by exogenous bacteria. Proc Natl Acad Sci USA 101: 12974–12979.

Buchon N, Broderick NA, Chakrabarti S, Lemaitre B. (2009). Invasive and indigenous microbiota impact intestinal stem cell activity through multiple pathways in Drosophila. Genes Dev 23: 2333–2344.

Burke C, Thomas T, Lewis M, Steinberg P, Kjelleberg S. (2011). Composition, uniqueness and variability of the epiphytic bacterial community of the green alga Ulva australis. ISME J 5: 590–600.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al (2010). QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7: 335–336.

Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J et al (2011). Moving pictures of the human microbiome. Genome Biol 12: R50.

Cenis JL, Perez P, Fereres *A.* (1993). Identification of aphid (Homoptera, Aphididae) species and clones by random amplified polymorphic DN*A.* Ann Entomol Soc Amer 86: 545–550.

Chandler JA, Lang JM, Bhatnagar S, Eisen JA, Kopp *A.* (2011). Bacterial communities of diverse Drosophila species: ecological context of a host-microbe model system. PLoS Genet 7: e1002272.

Corby-Harris V, Pontaroli AC, Shimkets LJ, Bennetzen JL, Habel KE, Promislow DE. (2007). Geographical distribution and diversity of bacteria associated with natural populations of Drosophila melanogaster. Appl Environ Microbiol 73: 3470–3479.

Costello EK, Stagaman K, Dethlefsen L, Bohannan BJ, Relman D*A.* (2012). The application of ecological theory toward an understanding of the human microbiome. Science 836: 1255–1262.

Cox CR, Gilmore MS. (2007). Native microbial colonization of Drosophila melanogaster and its use as a model of Enterococcusfaecalis pathogenesis. Infect Immun 75: 1565–1576.

Coyne MJ, Reinap B, Lee MM, Comstock LE. (2005). Human symbionts use a host-like pathway for surface fucosylation. Science 307: 1778–1781.

Dethlefsen L, McFall-Ngai M, Relman D*A.* (2007). An ecological and evolutionary perspective on human-microbe mutualism and disease. Nature 449: 811–818.

Donohoe DR, Garge N, Zhang X, Sun W, O'Connell TM, Bunger MK et al (2011). The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. Cell Metab 13: 517–526.

Dunning Hotopp JC, Clark ME, Oliveira DC, Foster JM, Fischer P, Munoz Torres MC et al (2007). Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science 317: 1753–1756.

Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J et al (2012). Microbial co-occurrence relationships in the human microbiome. PLoS Comput Biol 8: e1002606.

Frese SA, Benson AK, Tannock GW, Loach DM, Kim J, Zhang M et al (2011). The evolution of host specialization in the vertebrate gut symbiont. PLoS Genet 7: e1001314.

Gotelli NJ, Entsminger G*L*. (2012), EcoSim 7.72.Acquired Intelligence, Inc. http://www.uvm.edu/~ngotelli/EcoSim/EcoSim.html.

Gotelli NJ. (2000). Null model analysis of species co-occurrence patterns. Ecology 81: 2606–2621.

Hamady M, Knight R. (2009). Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. Genome Res 19: 1141–1152.

Hedges LM, Brownlie JC, O'Neill SL, Johnson KN. (2008). *Wolbachia* and virus protection in insects. Science 322: 702.

Hongoh Y, Deevong P, Inoue T, Moriya S, Trakulnaleamsai S, Ohkuma M et al (2005). Intra- and interspecific comparisons of bacterial diversity and community structure support coevolution of gut microbiota and termite host. Appl Environ Microbiol 71: 6590–6599.

Huse AM, Welch DM, Morrison HG, Sogin M*L*. (2010). Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environ Microbiol 12: 1889–1898.

Jaenike J, Unckless R, Cockburn SN, Boelio LM, Perlman SJ. (2010). Adaptation via symbiosis: recent spread of a Drosophila defensive symbiont. Science 329: 212–215.

Karasov WH, Douglas AE. (2013). Gastrointestinal physiology. In: JW H, Garland T, Wang T (eds.) APS Handbook of Physiology: Comparative and Evolutionary

Physiology. Wiley-Blackwell, Hoboken, NJ.

Kunin V, Engelbrektson A, Ochman H, Hugenholz P. (2010). Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. Environ Microbiol 12: 118–123.

Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS et al (2008a). Evolution of mammals and their gut microbes. Science 320: 1647–1651.

Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI. (2008b). Worlds within worlds: evolution of the vertebrate gut microbiota. Nat Rev Microbiol 6: 776–788.

Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. (2012). Diversity, stability and resilience of the human gut microbiota. Nature 489: 220–230.

Martinson VG, Danforth BN, Minckley RL, Rueppell O, Tingek S, Moran N*A*. (2011). A simple and distinctive microbiota associated with honey bees and bumble bees. Mol Ecol 20: 619–628.

Mateos M, Castrezana SJ, Nankivell BJ, Estes AM, Markow TA, Moran N*A*. (2006). Heritable endosymbionts of Drosophila. Genetics 174: 363–376.

Mohr KI, Tebbe CC. (2006). Diversity and phylotype consistency of bacteria in the guts of three bee species (Apoidea) at an oilseed rape field. Environ Microbiol 8: 258–272.

Moran NA, Hansen AK, Powell JE, Sabree Z*L*. (2012). Distinctive gut microbiota of honey bees assessed using deep sampling from individual worker bees. PLoS One 7: e36393.

Morgan XC, Segata N, Hutttenhower C. (2013). Biodiversity and functional genomics in the human microbiome. Trends Genet 29: 51–58.

Muegge BD, Kuczynski J, Knights D, Clemente JC, Gonzalez A, Fontana L et al (2011). Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science 332: 970–974.

Nelson TM, Rogers TL, Carlini AR, Brown MV. (2012). Diet and phylogeny shape the gut microbiota of Antarctic seals: a comparison of wild and captive animals. Environ Microbiol 15: 1132–1145.

Ochman H, Worobey M, Kuo CH, Ndjango JB, Peeters M, Hahn BH et al (2010). Evolutionary relationships of wild hominids recapitulated by gut microbial communities. PLoS Biol 8: e1000546.

Oh PL, Benson AK, Peterson DA, Patil PB, Moriyama EN, Roos S et al (2010). Diversification of the gut symbiont Lactobacillusreuteri as a result of host-driven evolution. ISME J 4: 377–387.

O'Neill SL, Hoffmann AA, Werren JH (eds).. (1997) Influential Passengers—Inherited Microorganisms and Arthropod Reproduction. Oxford University Press: Oxford, UK.

Paradis E, Claude J, Strimmer K. (2004). APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20: 289–290.

Pedron T, Mulet C, Dauga C, Frangeul L, Chervaux C, Grompone G et al (2012). A crypt-specific core microbiota resides in the mouse colon. MBio 3: e00116–12.

Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C et al (2010). A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464: 59–65.

R Development Core Team (2012) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria, ISBN 3-900051-07-0.

Ren C, Webster P, Finkel SE, Tower J. (2007). Increased internal and external bacterial load during Drosophila aging without life-span trade-off. Cell Metab 6: 144–152.

Ridley EV, Wong AC, Westmiller S, Douglas AE. (2012). Impact of the resident microbiota on the nutritional phenotype of Drosophila melanogaster. PLoS One 7: e36765.

Robinson CJ, Schloss P, Ramos Y, Raffa K, Handelsman J. (2010). Robustness of the bacterial community in the cabbage white butterfly larval midgut. Microb Ecol 59: 199–211.

Roeselers G, Mittge EK, Stephens WZ, Parichy DM, Cavanaugh CM, Guillemin K et al (2011). Evidence for a core gut microbiota in the zebrafish. ISME J 5: 1595–1608.

Roh S, Nam Y, Chang H, Kim K, Kim M, Ryu JH et al (2008). Phylogenetic characterization of two novel commensal bacteria related to innate immune homeostasis in Drosophila. Appl Environ Microbiol 74: 6171–6177.

Rosenthal AZ, Matson EG, Eldar A, Leadbetter JR. (2011). RNA-seq reveals cooperative metabolic interactions between two termite-gut spirochete species in co-culture. ISME J 5: 1133–1142.

Salonen A, Salojarvi J, Lahti L, de Vos WM. (2012). The adult intestinal core microbiota is determined by analysis depth and health status. Clin Microbiol Infect 18(Suppl 4): 16–20.

Savage DC. (1977). Microbial ecology of the gastrointestinal tract. Annu Rev Microbiol 31: 107–133.

Shade A, Handelsman J. (2012). Beyond the Venn diagram: the hunt for a core microbiome. Environ Microbiol 14: 4–12.

Shanbhag S, Tripathi S. (2009). Epithelial ultrastructure and cellular mechanisms of acid and base transport in the Drosophila midgut. J Exp Biol 212: 1731–1744.

Sharon G, Segal D, Ringo JM, Hefetz A, Zilber-Rosenberg I, Rosenberg E. (2010). Commensal bacteria play a role in mating preference of Drosophila melanogaster. Proc Natl Acad Sci USA 107: 20051–20056.

Shin SC, Kim SH, You H, Kim B, Kim AC, Lee KA et al (2011). Drosophila microbiome modulates host developmental and metabolic homeostasis via insulin signaling. Science 334: 670–674.

Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. (2011). Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 27: 431–432.

Stacklies W, Redestig H, Schol M, Walthe D, Selbig J. (2007). pcaMethods—a Bioconductor package providing PCA methods for incomplete data. Bioinformatics 23: 1164–1167.

Stone L, Roberts *A.* (1990). The checkerboard score and species distributions. Oecologia 85: 74–79.

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.

Sun Y, Cal Y, Huse SM, Knight R, Farmerie WG, Wang X et al (2012). A large-scale benchmark study of existing algorithms for taxonomy-independent microbial community analysis. Brief Bioinform 13: 107–121.

Tamames J, Abellan JJ, Pignatelli M, Camacho A, Moya *A.* (2010). Environmental distribution of prokaryotic taxa. BMC Microbiol 10: 85.

Tang X, Freitak D, Vogel H, Ping L, Shao Y, Cordero EA et al (2012). Compexity and variability of gut commensal microbiota in polyphagous lepidopteran larvae. PLoS One 7: 336978.

The Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. Nature 486: 207–214.

Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE et al (2008). A core gut microbiome in obese and lean twins. Nature 457: 480–484.

Walter J, Britton RA, Roos S. (2011). Host-microbial symbiosis in the vertebrate gastrointestinal tract and the Lactobacillus reuteri paradigm. Proc Natl Acad Sci USA 108(Suppl 1): 4645–4652.

Wang Y, Gilbreath TM, Kukutla P, Yan G, Xu J. (2011). Dynamic gut microbiome across life history of the malaria mosquito Anopheles gambiae in Kenya. PLoS One 6: e24767.

Werren JH, Baldo L, Clark ME. (2008). *Wolbachia*: master manipulators of invertebrate biology. Nat Rev Microbiol 6: 741–751.

Wong CN, Ng P, Douglas AE.. (2011). Low-diversity bacterial community in the gut of the fruitfly Drosophila melanogaster. Environ Microbiol 13: 1889–1900.

Wylie KM, Truty RM, Sharpton TJ, Mihindukulasuriya KA, Zhou Y, Gao H et al (2012). Novel bacterial taxa in the human microbiome. PLoS One 7: e35294.

Zhou J, Wu L, Deng Y, Zhi X, Jiang YH, Tu Q et al (2011). Reproducibility and quantitation of amplicon sequencing-based detection. ISME J 5: 1303–1313.

CHAPTER 4

THE NUTRITIONAL BENEFITS OF THE GUT MICROBIOTA IN *DROSOPHILA*[3]

Abstract

Insight into the nutritional benefits of symbiotic microorganisms to animals can be gained by analysis of the impact of experimentally eliminating the microorganisms on the fitness and nutritional status of animals reared on different diets. In this study, conventional and axenic *Drosophila melanogaster* (i.e. flies with an unmanipulated microbiota and microbe-free flies, respectively) were raised on 16 diets of systematically varied glucose and yeast content. High mortality of axenic *Drosophila* on diets of very low yeast:glucose ratio was reversed by dietary supplement of B-vitamins, suggesting that the microbiota either represent a source of these nutrients or facilitate their uptake from the diet. In axenic flies, storage of energy as lipid and carbohydrate (glycogen, trehalose and glucose) was consistently elevated relative to conventional flies, and correlated with dietary glucose:yeast ratio. The high energy storage coupled with extended development time and reduced body weight. These data suggest that the microbiota may promote host fitness by the consumption of dietary sugar and thereby balancing the availability of nutrients in yeast relative to excess dietary carbohydrates. Pyrosequencing of the 16S rRNA gene amplicons revealed that the bacterial community associated with conventional flies was dominated by *Acetobacteraceae* and *Lactobacillus*, the composition of which was not discernibly

[3] Article in preparation for journal submission by Wong, ACN., Dobson, *A*.J. and Douglas, AE. Dobson, *A*.J. conducted the statistical analyses of the performance and nutritional data.

structured by dietary yeast or glucose content. It is concluded that the microbiota contributes to the nutrition of *Drosophila* via multiple processes, and that each nutritional interaction can potentially be mediated by multiple bacterial species.

Introduction

There is increasing evidence that the resident microorganisms in animal guts have a major influence on the nutrition of their animal hosts (Flint et al., 2012; Karasov and Douglas, 2013). These microorganisms shape animal nutrition in multiple ways. They can compete with the host for ingested nutrients or provide supplementary nutrients to the host, alter animal feeding, nutrient assimilation and nutrient allocation patterns, by modulating the nutrient sensing and signaling pathways of the animal host (Backhed et al., 2004; Caricilli and Saad, 2013; Goodman et al., 2009; Vijay-Kumar et al., 2010). Multiple studies indicate that the resident microorganisms generally promote animal nutrition, although the nutritional benefit can vary with diet, composition of the microbiota and animal genotype (Benson et al., 2010; Kau et al., 2011; Smith et al., 2013). A mismatch between the microbiota and animal can result in poor host health, a condition known as dysbiosis (Nicholson et al., 2012; Stecher et al., 2013).

Much of current understanding of the nutritional significance of the gut microbiota in animals comes from comparisons between animals bearing an unmanipulated microbiota (conventional animals), animals deprived of their microbiota (axenic, microbe-free animals), and animals experimentally associated with specific gut bacteria (gnotobiotic animals) (Gordon and Pesti, 1971; Smith et al., 2007; Yi and Li, 2012). Another informative experimental strategy uses animals with mutations, especially of metabolic or immune functions, that cause correlated changes in the

composition of the microbiota and host nutritional indices (Turnbaugh et al., 2006; Vijay-Kumar et al., 2010). Most of these studies are conducted on a single diet, or two (or several) diets in which either a single nutritional component is altered, or in which the composition is fixed but at different concentrations. These approaches provide important insights into specific aspects of the nutritional interactions between animals and their microbiota, but are not designed to yield a global understanding of the nutritional significance of the microbiota.

Inclusive information on the nutritional significance of microorganisms to their animal host can be obtained from systematic dietary analysis, by subjecting conventional and axenic host to diets of systematically varied composition. For various symbioses, disproportionately low performance of microbe-free animals, relative to conventional animals, on certain diets has been attributed to a deficiency in certain nutrients that are provided by the microorganisms (Douglas 2009). Additionally or alternatively, the diet have an excess of nutrients that the microorganisms consume and, thereby, mediate amelioration of the diet. Parallel analysis of the host nutritional composition, especially densities of major macronutrients (protein, lipid etc.), provides insight into the impact of the microbiota on host nutritional allocation patterns.

We reasoned that systematic dietary experiments of this design would be especially valuable to elucidate the nutritional role of resident microorganisms in *Drosophila melanogaster*. Our choice of this species and experimental design were founded on published evidence that the microbiota is important to *Drosophila* nutrition (Broderick and Lemaitre, 2012; Erkosar et al., 2013). Elimination of the microbiota results in extended larval development time and, in some studies, depressed adult weight and total lifespan (Bakula, 1969; Brummel et al., 2004; Ridley et al., 2012; Shin et al.,

2011; Storelli et al., 2011). These performance differences between microbe-free and conventional *Drosophila* are accompanied by differences in nutritional indices, including elevated levels of free glucose and, in some studies, glycogen and triglyceride, together with reduced basal metabolic rates in microbe-free flies (Ridley et al., 2012; Shin et al., 2011). Diet composition, fly genotype and the composition of the microbiota may contribute to the among-study variation. The microbiota is generally dominated by *Acetobacteraceae* and Lactobacillales, especially *Acetobacter* and *Lactobacillus* species in long-term laboratory fly cultures (Chandler et al., 2011; Wong et al., 2011). Flies mono-associated with *A. pomorum* display elevated expression of the insulin-like peptide genes *dilp-2, -3* and *-5* relative to microbe-free flies, an effect that requires acetic acid, a metabolite produced by *A. pomorum,* and can account for the both the extended larval development time and elevated glucose and lipid levels in microbe-free flies (Shin et al., 2011). *L. plantarum* has been suggested to contribute to the protein nutrition of *Drosophila* because this bacterium promotes larval development on diets containing low concentrations of yeast (the sole dietary source of protein); and it has been proposed specifically that *L. plantarum* promotes the assimilation of dietary protein, with the downstream activation of TOR and insulin signaling pathways (Storelli et al., 2011).

The purpose of this study was to identify the key nutritional interactions between *Drosophila melanogaster* and its microbiota by a comparison of the performance and nutritional status of conventional and microbe-fee flies on diets of systematically varied composition. The 16 test diets comprised glucose and yeast (which provided protein, lipid, vitamins and minerals), at concentrations systematically varied over an eight-fold range (25-200 g l^{-1}). Our analyses revealed that the difference between

conventional and microbe-free flies was driven by the dietary concentration of yeast and yeast:glucose ratio, and that these differences could most parsimoniously be explained by two bacterial-mediated processes that are beneficial to the animal host: bacterial provisioning of B-vitamins, and bacterial consumption of excess dietary sugar.

Experimental procedures

The insects and diets

Drosophila melanogaster strain Canton S was maintained in routine culture at 25°C under a 12 h:12 h light–dark cycle on yeast-glucose medium [Y-G, comprising 100 g inactive Brewer's yeast and 100 g glucose (MP Biomedicals) l^{-1}, 12 g agar l^{-1} (Frutarom) and preservatives (0.04% phosphoric acid, 0.42% propionic acid; Sigma)]. The 16 test diets comprised yeast (Y) and glucose (G) at each combination of 25, 50, 100 or 200 $g \cdot l^{-1}$, giving Y:G ratios ranging from 1:8 to 8:1. For some experiments, casein protein (Sigma) at 78.8 g l^{-1} or 33.8 g l^{-1} final concentration was added to diet containing 25 g yeast l^{-1}, to give protein contents equivalent to diets with 200 g l^{-1} and 100 g l^{-1}, respectively (yeast comprises 45% protein, details provided by the manufacturer). The vitamin supplement to diets comprised thiamine (1.4 mg l^{-1}), riboflavin (0.7 mg l^{-1}), nicotinic acid (8.4 mg l^{-1}), pantothenate (10.8 mg l^{-1}), pyridoxine (1.7 mg l^{-1}), biotin 0.1 mg l^{-1}) and folic acid (9 mg l^{-1}), following Sang (1956) and Blatch (2010). Filter-sterilized supplements were added aseptically to autoclaved diet, and 7.5 ml volumes of the final diet were distributed to sterile vials (Corning).

Each experiment was initiated with eggs deposited overnight by mated females. The

axenic treatment was obtained by dechorionating eggs in 10% sodium hypochlorite, followed by three rinses in sterile deionized water, as described by Ridley et al. (2012), and the control eggs (giving rise to conventional flies) were generated in parallel by the same procedure, except that the hypochlorite was replaced by sterile water. Twenty five dechorionated or control eggs were transferred to each vial using aseptic technique in a laminar flow cabinet.

Drosophila performance indices

Axenic and conventional flies were raised in 5 replicate vials (one replicate of every diet formulation on 5 different days). Vials were monitored daily, and larval development time to pupation and eclosion and survival to adulthood were scored. Experiments were terminated 30 days after egg transfer.

Weight and macronutrient analyses

Each sample of 5 male flies or 5 female flies at 4-5 days post eclosion was weighed on a microbalance (Mettler MX5) to an accuracy of 1 µg. The sample was then hand-homogenized by plastic pestels in 125 µl ice-cold TE buffer (pH 7.4) comprising 10 mM Tris, 1 mM EDTA and 0.1% Triton-X-100, followed by centrifugation at 7,000 *g* at 4°C for 1 min. A portion (20 µl) of each supernatant was immediately stored at -80^{o}C for analysis of total protein, while the remaining supernatant was heat-treated at 72 ^{o}C for 20 minutes to inactivate enzymatic activity before analysis of glucose, glycogen and triglycerides. Nutritional assays were conducted in 96-well plates using commercial kits/reagents following manufacturer's instructions, as described in Ridley et al. (2012): the DC Protein Assay kit (BioRad, 500-0116), the

Triglyceride Assay kit from (Sigma, TG-5-RB), and the Glucose (GO) Assay kit (Sigma, GAGO20) for glucose and glycogen (after treatment with amyloglucosidase at 37 °C for 1 h [2 U ml^{-1}, Sigma A7420]). All colorimetric readings were obtained using a microplate spectrophotometer (BioRad xMark™).

Amino acid profiling

The free amino acid content of individual 4/5-day-old flies was assayed using the AccQ Tag derivatization kit (Waters) by UPLC with PDA detector (Waters Acquity). Ten replicate flies of each sex per treatment were homogenized in 90 µl PBS (pH 8). A 20 µl portion of the homogenate was mixed with an equal volume of 40 mM HCl, incubated on ice for 30 minutes and centrifuged at 18000 *g* at 4˚C for 10 minutes. The supernatant was then filtered through a 0.45 µm filter plate (Millipore) by centrifugation at 1500 *g* for 10 min. The filtrate (2.5 µl) was derivatized with AccQ Tag (Waters), following manufacturer's protocol, and injected into Waters Acquity UPLC with PDA detector and AccQ-Tag Ultra 2.1 x 100 mm column. The gradient was: 0-0.54 min, 99.9% A 0.1% B; 0.54-5.74 min, 90.9% A and 9.1% B; 5.74-7.74 min, 78.8% A 21.2% B; 7.74-8.04 min, 40.4% A 59.6% B; 8.04-8.64 min, 10% A 90% B; 8.05-8.64 min 10% A 90% B; 8.64-8.73 min 99.9% A 0.1% B; 8.73-9.50 min, 99.9% A 0.1% B (linear between each time point), where A is 10% AccQ-Taq Ultra Eluent A in water, and B is Accq-Taq Ultra Eluent B. Amino acids were determined by comparison to standards: 1, 5, 10, 50 and 100 pmol amino acids μl^{-1} (Waters amino acid hydrolysate standard #088122, supplemented with asparagine, tryptophan and glutamine).

Feeding assays

To measure feeding rates, 110 adult flies of each sex from each treatment were anesthetized by CO_2 and sorted into 11 vials of 10 flies to recover and starved for 2 hours. All feeding assays were conducted 6 hours after onset of the light period. 10 groups were then transferred to diet labeled with a blue dye (0.5% xylene cyanol and 0.1% bromophenol blue) and one group was transferred to dye-free diet as control. After 30 minutes, the flies were frozen at -80^{o}C, until analyzed. Samples of frozen flies were thawed for 2 minutes, rinsed gently in water, and the number of flies that had eaten, as indicated by blue dye in the abdomen was scored by examination under a dissecting microscope (7x). Each sample were then homogenized in 100 µl TE buffer (10 mM Tris, 1 mM EDTA and 0.1% Triton-X-100, pH 7.4) with 1.4 mm ceramic beads (MP Biomedicals) in FastPrep®-24 Instrument (MP Biomedicals) for 1 minute, diluted with an additional 500 µl TE buffer and centrifuged at 13,680 *g* for 3 minutes. The absorbance of the supernatant were measured at 614 nm using a microplate spectrophotometer (BioRad xMark™). Absorbance values were transformed to µg food ingested per fly, by reference to a standard curve generated with dilution series of the dye (0-200 ng dye ml^{-1}).

Bacterial complement associated with conventional Drosophila

To assess the microbiota composition of conventional flies reared on the different diets, total genomic DNA was extracted from 6 surface-sterilized whole fly bodies (3 males and 3 females) from each diet using the DNeasy Blood and Tissue Kit (Qiagen), following a protocol optimized for both Gram-positive and Gram-negative bacteria as previously described (Wong et al., 2011). A buffer-only sample was prepared as an

environmental control. 16S ribosomal RNA amplicons of the V2 region were prepared by triplicate PCR reactions (Wong *et al*., 2011), using the general 16S rRNA gene primers 27F-338R tagged with different MIDs. Equal amounts of the triplicate products per sample were mixed, purified using the QIAquick PCR purification kit (QIAGEN), and quantified by *Quant-iT*™ *PicoGreen*®. Emulsion PCR was conducted at 1.5 copies per bead using only 'A' beads for unidirectional 454 GS-FLX pyrosequencing with standard Titanium chemistry. Pyrosequencing flowgrams were analyzed by the procedure of Wong et al. (2013), using QIIME 1.4.0 virtualbox with default parameters, except that the denoising cutoff was set to remove singletons. Species identities of the OTUs were assigned by NCBI StandAlone BLAST (megablast program) using the nucleotide (nt) database (August 2012) under default settings with supplementary manual curation. OTUs with fewer reads than in the environmental sample were discarded.

Data visualisation and analysis

All data were analysed in R (version 2.15.1). Development data were analyzed using a frailty model from the coxme library, Development data were fitted to a three-way interaction of dietary glucose and dietary yeast as continuous variables, and microbiota treatment as a factor, with experimental replicate as an additional random effect. Data were plotted using the survfit function on raw development data. Other analyses were performed in parallel for males and females. Body weight was analyzed by a linear model with mass fitted to the same formula as development data. Nutritional indices were fitted to this formula, following normalization to weight. When maximal models did not yield comprehensible results, they were simplified by

stepwise removal of nonsignificant interactions or factors to obtain minimal adequate models. Data were plotted using the filled.contour function, and full 3d figures were visualised using the scatterplot3d function. Principal components analyses (PCAs) were performed using the prcomp function. Only body weight and nutritional index data were fully paired and balanced, so to obtain PCs of all datasets we took mean values of these traits, median development time, and estimated mean protein and carbohydrate ingested (assuming the dietary yeast contains 45% protein and 24% carbohydrate) for axenic and conventional males and females on each diet for which we had data (i.e. excluding 1:8 Y:G).

Results

Survival to adulthood and developmental rates

The first experiments investigated the effect of diet composition on the development time and survivorship of conventional and axenic *Drosophila* (Figure 3.1 and Table 3.1). Pre-adult mortality of conventional flies was <25% on all diets, but the axenic *Drosophila* displayed elevated mortality on low-yeast high-glucose diets, with all of the insects dying as larvae on the diet containing 25 g yeast and 200 g glucose l^{-1} (1:8 Y:G). The development of axenic flies was retarded relative to the conventional flies (Figure 3.1 - Frailty model: main effect of symbiosis; $z=8.85$, $p<0.0001$), and the response to diet also differed between the conventional and axenic flies. Extended development time was driven by low dietary yeast ($z=-9.93$, $p<0.0001$) for conventional flies, and low yeast:glucose ratio for axenic flies ($z=5.12$,$p<0.0001$).

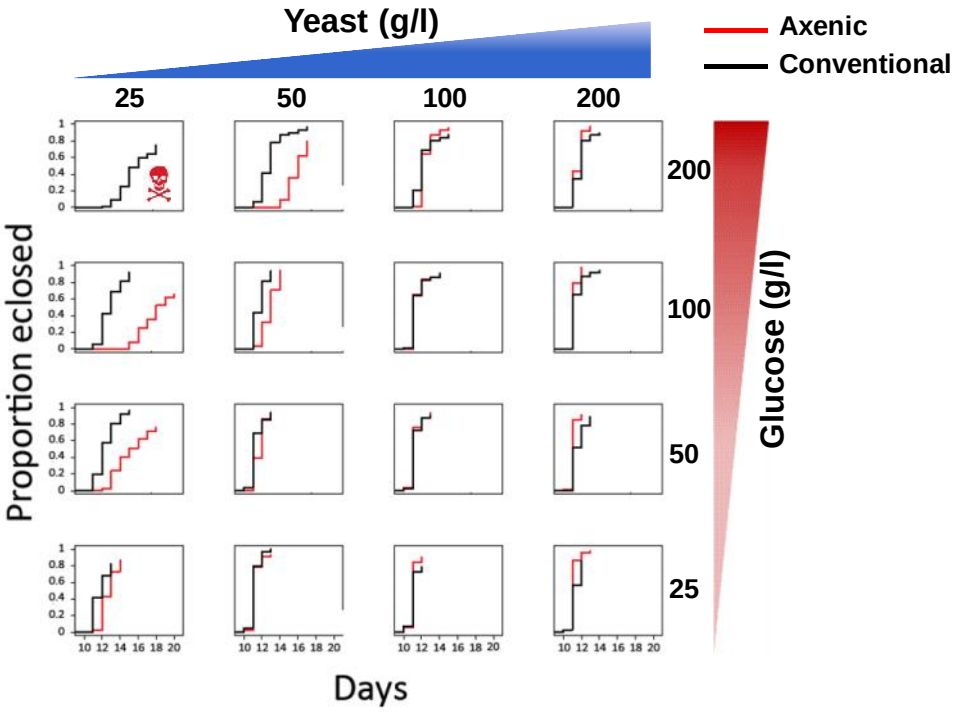


Figure 3.1 Survival and development of axenic and conventional flies across dietary space.

Table 3.1 Summary of conventional and axenic developmental indices across dietary space.

Dietary yeast (g/l)	Dietary glucose (g/l)	Median conventional development time (days)	Median axenic development time (days)	Percentage conventional eclosion	Percentage axenic eclosion
200	200	12	12	85	94
100	200	12	12	83	92
50	200	13	16	89	75
25	200	15	-*	65	-*
200	100	11	11	93	96
100	100	11	11	85	85
50	100	12	13	90	88
25	100	13	16.5	88	50
200	50	11	11	79	91
100	50	11	11	88	89
50	50	11	12	89	90
25	50	12	14	77	72
200	25	11	11	85	95
100	25	11	11	78	88
50	25	11	11	91	90
25	25	11	13	71	87

*Data not obtained due to high mortality.

The extended development time of axenic *Drosophila* on diets of low yeast content, including preadult mortality of axenic *Drosophila* on diet containing 25 g yeast and 200 g glucose l^{-1} (1:8 Y:G), suggested that the bacteria in conventional *Drosophila* may promote availability of certain yeast constituent(s). We reasoned that the bacteria may spare the insect requirement for protein or B vitamins in the yeast. To test this, the 1:8 Y:G diet was supplemented with casein protein and/or B-vitamins.

Our results demonstrated the high mortality of axenic *Drosophila* was alleviated by B-vitamins, but not casein (Figure 3.2A). Supplementary experiments with individual B-vitamin deletions revealed that riboflavin is crucially important protection against pre-adult mortality of axenic *Drosophila* (Figure 3.2B).

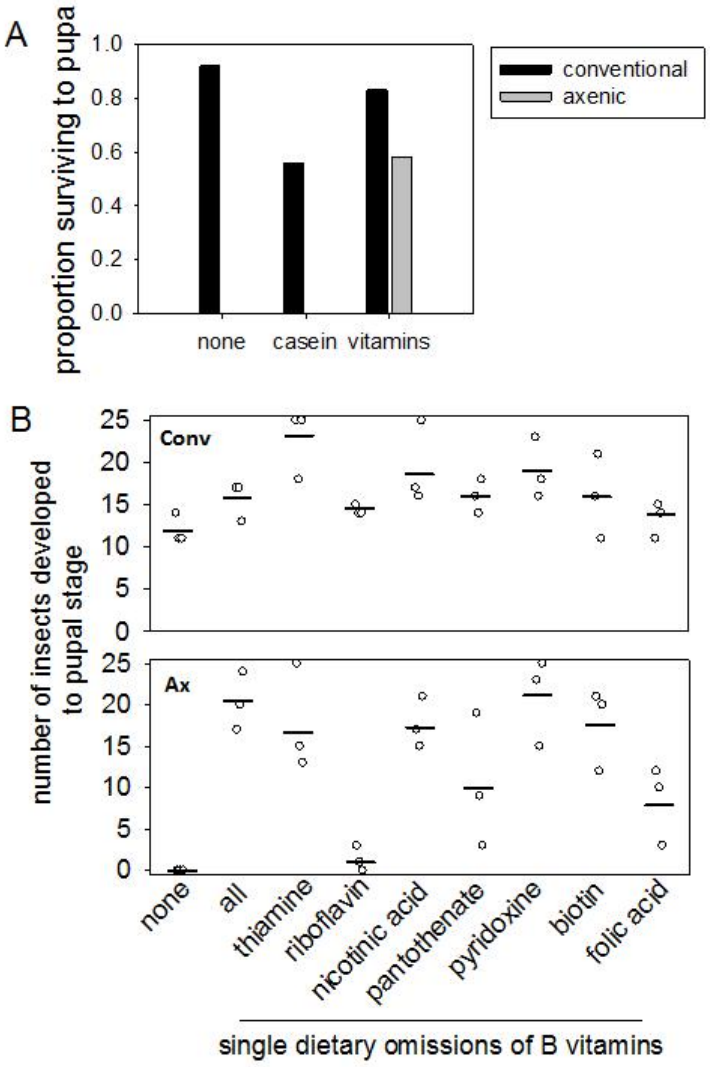


Figure 3.2 Survival of axenic flies on 25 g yeast and 200 g glucose l^{-1} (1:8 Y:G) diet supplemented with B vitamins with selective deletions.

Weight and nutritional indices of the flies

As a complementary approach to investigate further the nutritional response of conventional and axenic *Drosophila* to diet, the nutrient content of *Drosophila* was quantified. Preliminary inspection of the nutrient content of the flies revealed that the male and female flies responded very differently to diet and elimination of the

microbiota, and so the datasets for the two sexes were analyzed separately.

The weight of the flies was significantly higher in axenic flies than conventional flies, for both sexes (Figure 3.3; Figure 7.1 and 7.2 in Appendix E). It also varied with diet, generally increasing with glucose content in conventional flies, and increasing with yeast:glucose ratio in axenic flies.

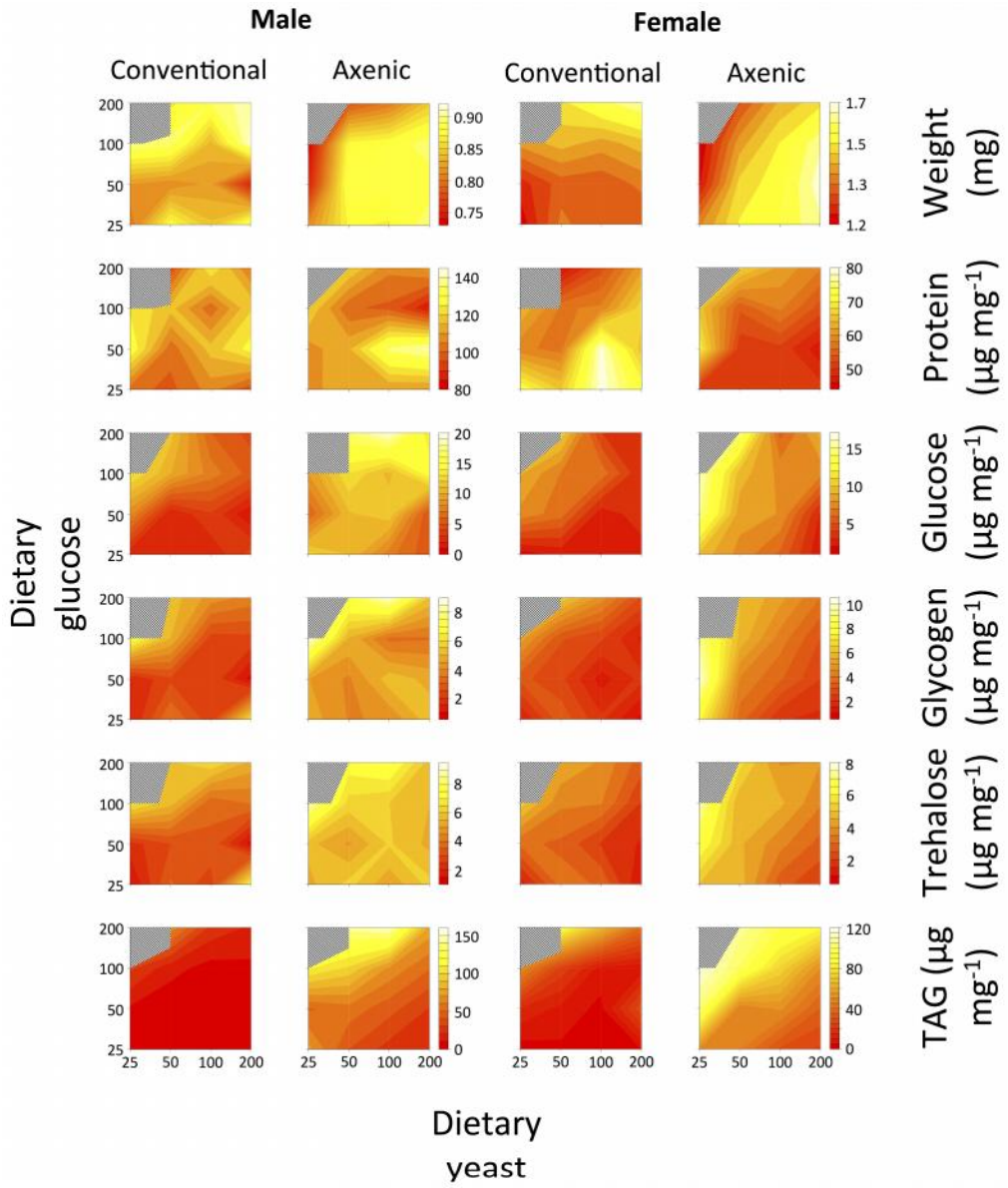


Figure 3.3 Impact of dietary yeast and glucose (25-200 g l^{-1}) on weight and nutritional indices of conventional and axenic *Drosophila* of both sexes.

The conventional flies maintained a more stable carbohydrate and lipid homeostasis than axenic flies (Figure 3.3; Figure 7.3-7.6 in Appendix E). In both sexes, axenic flies were significantly hyperglycemic, hyperlipidemic, and had elevated trehalose and glycogen levels. For males, the conventional and axenic flies displayed similar responses to diet (i.e. the 3-way interaction terms between yeast and glucose and symbiosis are not significant), and the values of all nutritional indices increased with low Y:G, apart from glucose, which increased with dietary glucose concentration. A comparable pattern was evident for the trehalose density of females, which was promoted by high glucose and depressed by low yeast in both conventional and axenic flies. For all other nutritional indices, however, the conventional and axenic females responded differently to diet. The triglyceride and glucose densities were promoted by low dietary Y:G in conventional flies and by low yeast content for axenic flies. The response to glycogen was more complex, being promoted by high dietary glucose in conventional and axenic flies, but also by low dietary yeast for axenic flies.

The protein density (μg protein mg^{-1} weight) of the male flies was not significantly affected by either diet or elimination of the microbiota. For females, protein density was reduced in axenic flies, and was responsive to diet in conventional (but not axenic) flies, specifically, increasing with high dietary yeast and low dietary glucose. To test whether specific amino acids may be limiting the protein density and development time of conventional *Drosophila* and protein density of conventional females, the free amino acid content of flies reared on diets with different yeast contents was measured. These analyses were conducted using diets with 200 g glucose l^{-1} and each of 50, 100 and 200 g yeast l^{-1}, focusing on 19 of the 20 protein-amino acids (aspartate was excluded because it was undetectable in most samples). As with

other indices, the interaction between diet and symbiosis differed between the male and female flies, and, therefore, the two sexes were analyzed separately. By MANOVA, the free amino acids (FAA) content of both males and females varied significantly with symbiosis, diet and their interaction (Table 3.2), indicating that the impact of diet on FAA composition differed between conventional and axenic flies. The basis of this difference was evident from principal components analysis (PCA) conducted on % amino acid content of conventional and axenic males and females (4 tests; figure 3.4).

Table 3.2 MANOVA results for variation in concentration of free amino acids with diet (200, 100 or 50 g yeast l^{-1}, in diet containing 200 g glucose l^{-1}) and symbiosis (conventional or axenic flies)

Factor	Test statistic	F value and significance
(a) males		
diet	Wilks' 0.021	$F_{38,70} = 10.799$, $p<0.001$
	Lawley-Hotelling 16.84	$F_{38,68} = 15.070$, $p<0.001$
	Pillai's 1.60	$F_{38,72} = 7.58$, $p<0.001$
Symbiosis	Wilks' 0.090	$F_{19,35} = 18.640$, $p<0.001$
	Lawley-Hotelling 10.12	$F_{19,35} = 18.640$, $p<0.001$
	Pillai's 18.64	$F_{19,35} = 18.640$, $p<0.001$
interaction	Wilks' 0.029	$F_{38,70} = 8.953$, $p<0.001$
	Lawley-Hotelling 9.74	$F_{38,68} = 8.717$, $p<0.001$
	Pillai's 1.66	$F_{38,72} = 9.188$, $p<0.001$
(b) females		
diet	Wilks' 0.045	$F_{38,70} = 13.449$, $p<0.001$
	Lawley-Hotelling 17.33	$F_{38,68} = 15.506$, $p<0.001$
	Pillai's 1.72	$F_{38,72} = 11.614$, $p<0.001$
Symbiosis	Wilks' 0.083	$F_{19,35} = 20.355$, $p<0.001$
	Lawley-Hotelling 11.05	$F_{19,35} = 20.355$, $p<0.001$
	Pillai's 0.92	$F_{19,35} = 20.355$, $p<0.001$
interaction	Wilks' 0.023	$F_{38,70} = 10.184$, $p<0.001$
	Lawley-Hotelling 11.31	$F_{38,68} = 10.118$, $p<0.001$
	Pillai's 1.69	$F_{38,72} = 10.241$, $p<0.001$

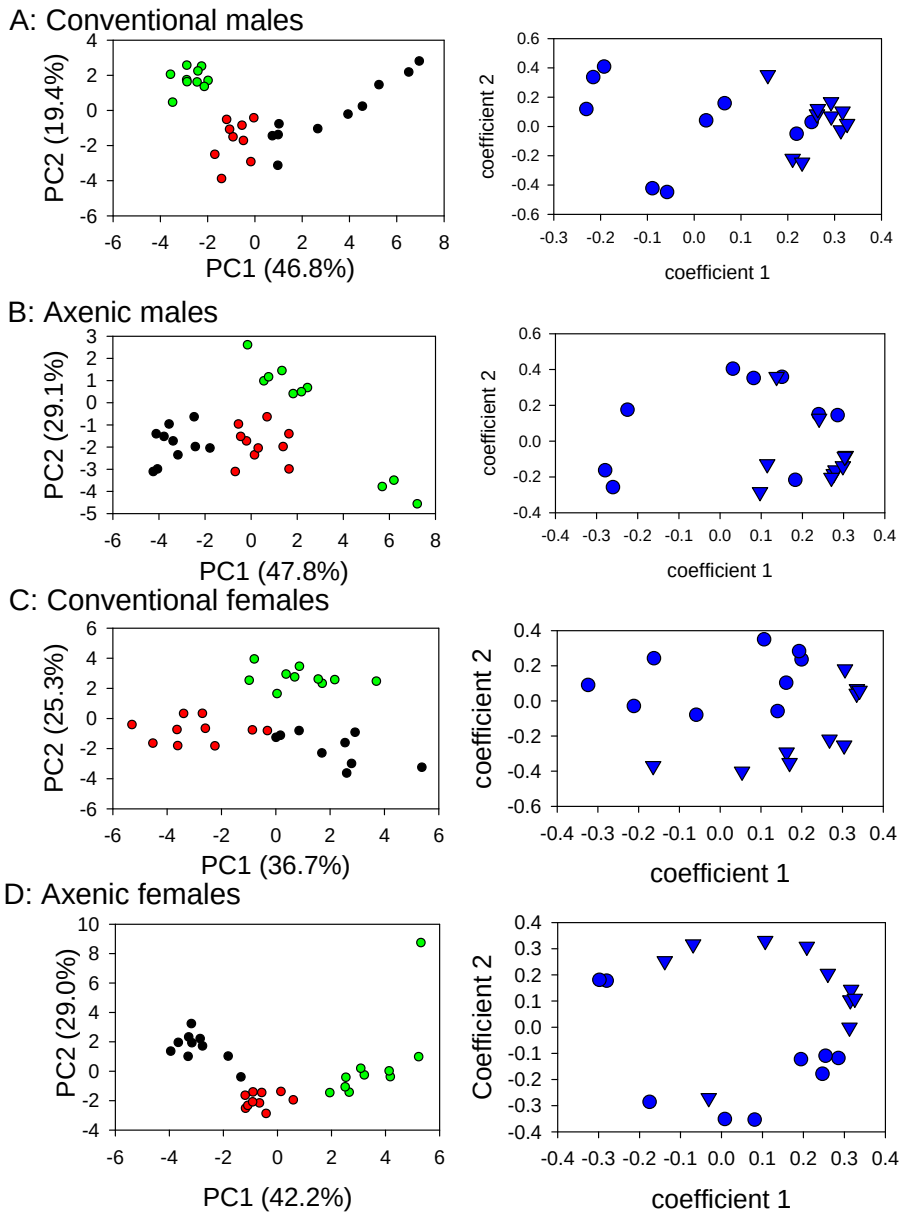


Figure 3.4. PCA of FAA content of flies. Plots of principal components (PC) 1 and 2, with % variance assigned to each axis in parentheses (left) with diets containing 200 g glucose l^{-1} and 200 (black), 100 (red) and 50 (green) g yeast l^{-1}; and loading scores of amino acids in PCA (right), with non-essential amino acids (circles) and essential amino acids (triangles).

For every analysis, the samples from the three diets could be separated on the first and second PC axes, largely on the basis of the relative amount of essential versus non-essential amino acids. For conventional flies, the contribution of essential amino acids to the FAA pool increased with increasing dietary yeast; and, for axenic flies, the association was reversed, i.e. axenic flies reared on low dietary yeast content tended to have high essential amino acid content (Figure 3.5). Consistent with this interpretation of the PCA, the interaction term ‘symbiosis x diet’ was significant in the ANOVA tests, with a significant reduction in %essential amino acid content of conventional flies, and increase of axenic flies, with diet of low yeast content (see legend to Figure 3.5).

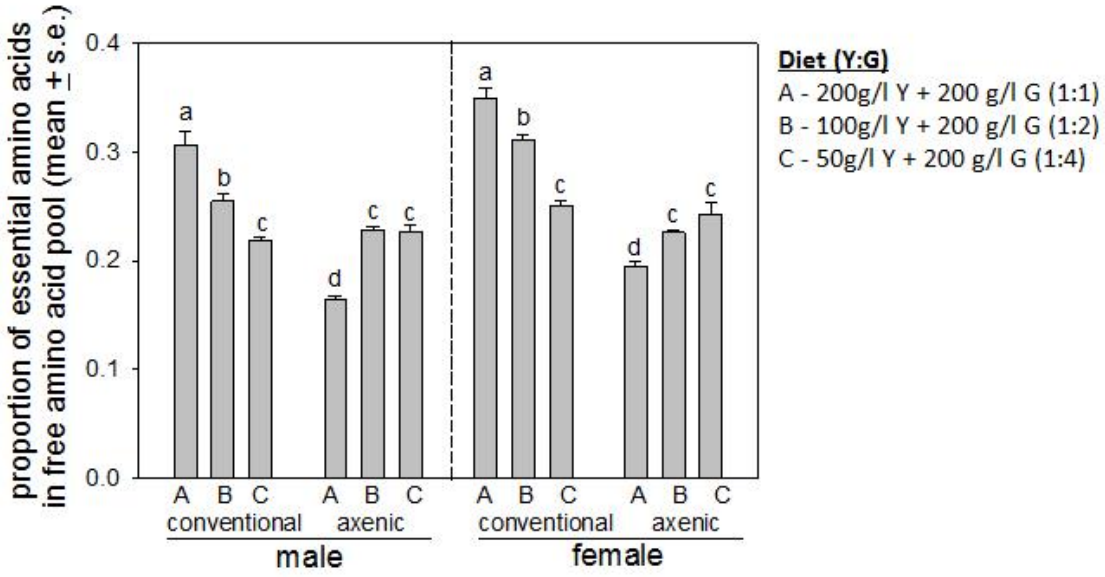
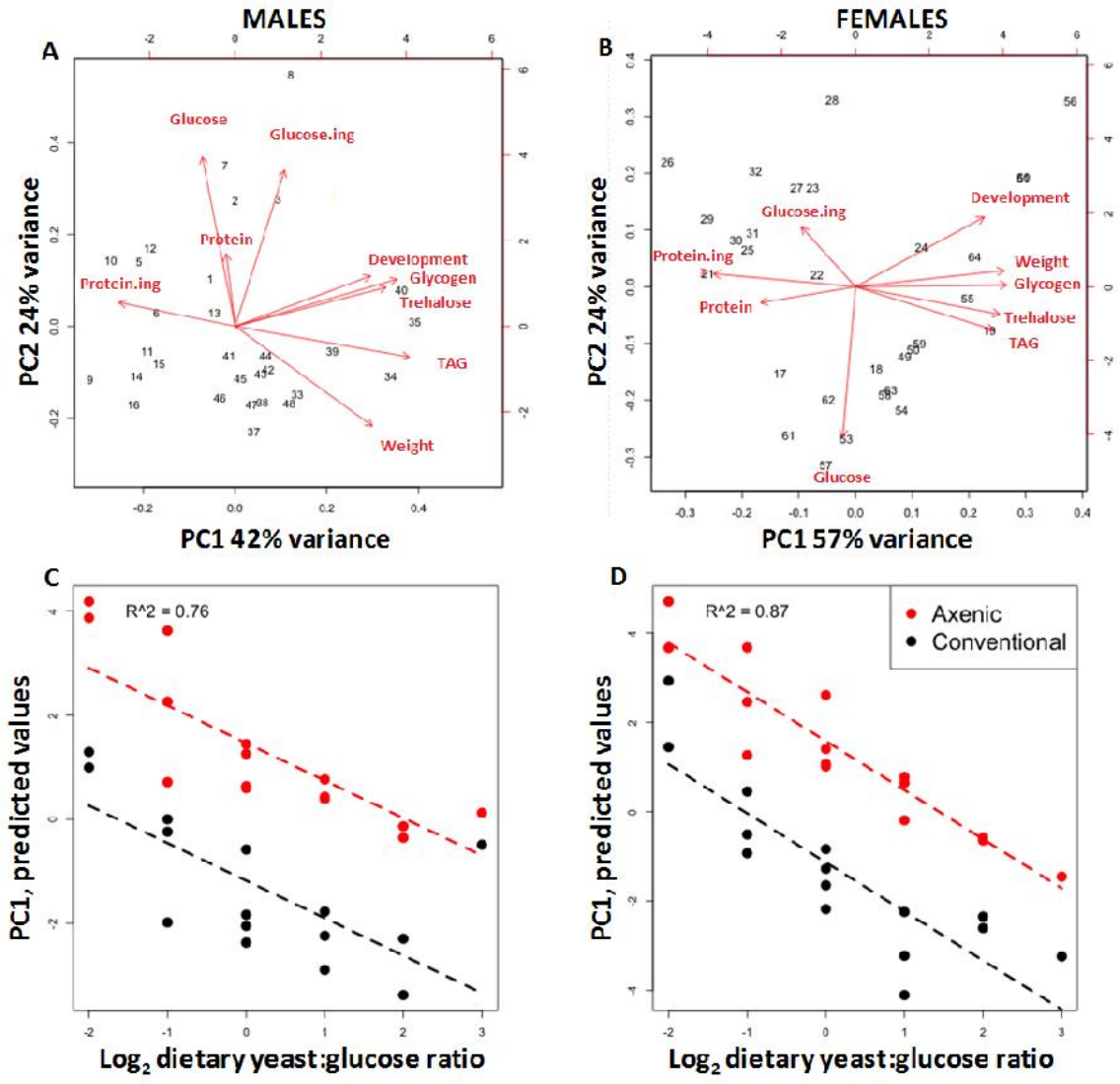


Figure 3.5 Variation in % essential amino acid content of conventional and axenic flies reared on diets A, B and C with varying dietary yeast (Y) and glucose (G). ANOVA on arcsin-squareroot transformed data: males, diet $F_{2,98} = 8.53$, $p<0.001$; symbiosis $F_{1,98} = 174.74$, $p<0.001$; interaction $F_{2,98} = 120.86$, $p<0.001$; females diet $F_{2,112} = 7.11$, $p=0.001$; symbiosis $F_{1,112} = 222.13$, $p<0.001$; interaction $F_{2,112} = 56.43$, $p<0.001$ Different letters refer to significantly different means determined by Tukey's *post hoc* test conducted on males and females separately.

Aggregate analyses of all phenotypic responses to diet and microbiota

Average values of all measured traits (means for all traits except development [medians]) obtained for flies on the 15 diets (1:8 Y:G excluded because of high mortality of axenic *Drosophila*) were compressed by separate PCAs for males and females (Figure 3.6).

Figure 3.6 Principal components analysis of average values of all measured traits per diet (means: nutritional indices and feeding rate; median deveopment time). A and B: Biplots of loadings onto PCs 1 and 2. Loadings onto principal component 1 show positive relationships between carbohydrate storage, weight and development time, which were negatively related to rates of protein ingestion. C and D: predicted PC values (PC1). For both axenic and conventional flies, principal component 1 was negatively related to the ratio of dietary yeast to glucose (ANCOVA males: $F_{1,27}$=18.19, p<0.0005; females $F_{1,27}$=110.40, p<0.0001). The values of axenics on PC 1 were signficantly greater than those of conventionals, but showed the same relationship to dietary yeast-glucose ratio (ANCOVA males $F_{1,27}$=52.19, p<0.0001; females $F_{1,27}$=66.27, p<0.0001).


MALES
A
Glucose
Glucose.ing
Protein
Protein.ing
Development
Glycogen
Trehalose
TAG
Weight
PC2 24% variance
PC1 42% variance
FEMALES
B
Development
Glucose.ing
Protein.ing
Weight
Glycogen
Protein
Trehalose
TAG
Glucose
PC2 24% variance
PC1 57% variance
C
R^2 = 0.76
PC1, predicted values
Log2 dietary yeast:glucose ratio
D
R^2 = 0.87
Axenic
Conventional
PC1, predicted values
Log2 dietary yeast:glucose ratio

PC1 accounted for 42% of the variance in males and 57% in females. For both sexes, PC1 was strongly positively loaded by development time, body mass, glycogen, trehalose and triglyceride stores, and negatively loaded with protein ingested: high values of this PC therefore correspond to an aggregate measure of high triglyceride and carbohydrate storage, slowed development, and reduced ingestion of protein. The values of PC1 have a linear negative correlation with dietary Y:G (log_2 transformed), that is significantly elevated in axenic flies, with no interaction between microbiota treatment and dietary Y:G. This result was robust to removal of protein and carbohydrate ingestion from the PCA, demonstrating that this is not an artefact of autocorrelation between dietary nutrient content and ingestion of diet. This analysis reveals that the delayed development and elevated storage of lipid and carbohydrate in *Drosophila* covary positively and in response to dietary Y:G ratio. Although the average values of these traits are elevated in response to elimination of the microbiota, the pattern of their response to diet does not differ between conventional and axenic flies. This derived result complements our more direct results from studies of individual nutritional phenotypes.

Feeding rates

One plausible explanation for the elevated carbohydrate and triglyceride levels in axenic flies was that these insects feed at higher rates and allocate the excess nutrients to storage. Contrary to expectation, axenic flies of both sexes reared on all 16 diets had lower feeding rates than conventional flies (Figure 3.7, ANOVA (males) $F=96.61$, $p<0.0001$; ANOVA (females) $F=179.66$, $p<0.0001$). The same pattern of feeding

difference between conventional and axenic flies was observed in long-term (48 hrs) Capillary Feeder (CAFE) assay (data not shown).

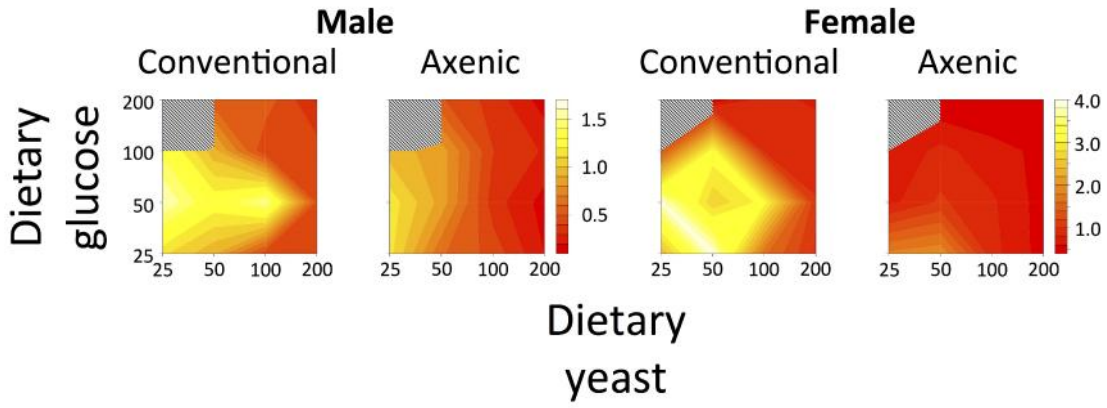


Figure 3.7 Feeding responses (μg per fly) of conventional and axenic flies to dietary space.

In conventional flies, feeding rates varied significantly with dietary yeast in males, and with dietary Y:G in females, and the response to these dietary components was significantly reduced in axenic flies for both sexes. In summary, the elevated carbohydrate and lipid levels in axenic flies is associated with reduced feeding rates and reduced responsiveness to dietary glucose/yeast, relative to conventional flies.

Bacterial complement of conventional Drosophila

The bacterial communities associated with conventionally-reared flies from the 16 tested diets were determined pyrosequencing of 16S rRNA gene amplicons. In total, 127269 reads of 16S amplicons were identified, and all the rarefaction curves tended to saturation, indicating that the OTUs were representative of the bacterial community in each sample. In all samples, the dominant bacteria were *Acetobacteraceae* and

Lactobacillaceae, as reported in previous pyrosequence analyses of *Drosophila* microbiota (Douglas, 2009; Wong et al., 2011). No systematic trends in relation to dietary nutrients were discernible, providing no evidence for diet as a major determinant of bacterial community composition (Figure 3.8).

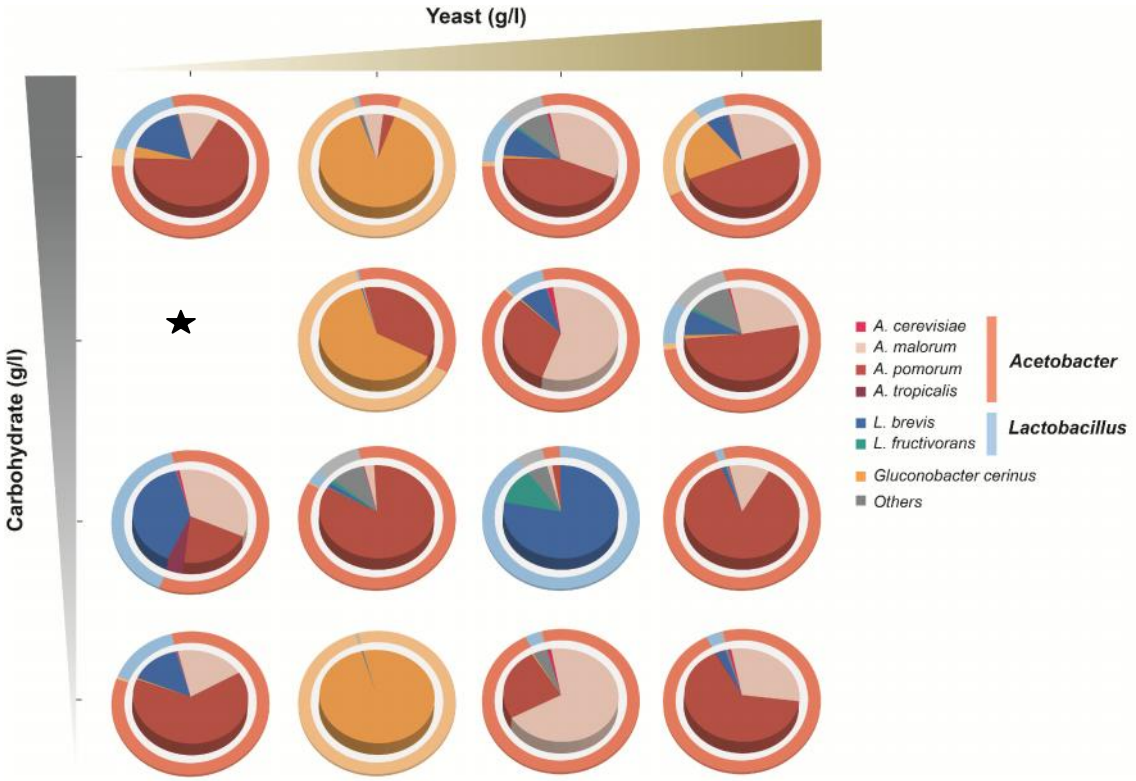


Figure 3.8 Bacterial compositions of conventional *Drosophila* across dietary space. ★ Diet not tested.

Discussion

This study demonstrates that elimination of the microbiota alters the geometry of the *Drosophila* response to diet. The differences in the effect of diet between microbe-free and conventional *Drosophila* are complex and sex-dependent, but much of the variation can be distilled into a reduced capacity of microbe-free insect to utilize diets

of low yeast:glucose ratio, associated with elevated levels of carbohydrate and lipid storage. The nutritional basis of these effects appears to be two-fold: that the microbiota represents a source of B vitamins, especially riboflavin, and reduces the deleterious effects of high dietary glucose. We will consider each of these nutritional interactions in turn.

Bacteria have been implicated in the B vitamin nutrition of various animals. In some animals, such as the insects that feed through the life cycle on vertebrate blood, a diet notoriously deficient in B vitamins, the microbial supply of these nutrients is critical. For example, the bacterial symbiont *Wigglesworthia* in tsetse flies *Glossina* has the genetic capacity to synthesize B vitamins (Akman et al., 2002). Tsetse flies are reproductively sterile when deprived of *Wigglesworthia*, and fecundity can be partially restored by adding B vitamins to the blood diet (Nogge, 1976). Similarly, elimination of the symbiotic bacteria from the blood-feeding louse *Pediculus* resulted in very high larval mortality that could be reversed by adding nicotinic acid (vitamin B3) to the diet (Puchta, 1956). Various members of the gut microbiota in humans and other mammals are capable of synthesizing B vitamins, with the potential to contribute to the vitamin nutrition of their host (Goodman et al., 2009; LeBlanc et al., 2013). In rats and other rodents, coprophagy (the ingestion of fecal pellets) is required to derive the full benefit of microbial vitamin synthesis on diets of low B vitamin content, suggesting that the relevant microbiota is resident in the distal (post-digestive) portion of the GI tract (Karasov and Douglas, 2013; Roscoe, 1931).

Our data suggest that *Drosophila* require a bacterial supply of the B vitamin riboflavin only under conditions of low dietary yeast. In other words, both yeast and bacteria can contribute to the vitamin nutrition of *Drosophila*. This nutritional role of the bacteria

may be relatively non-specific because most *Lactobacillus* and *Acetobacteraceae* are prototrophs for riboflavin and other vitamins (KEGG, http://www.genome.jp/kegg/). Nevertheless, B vitamin auxotrophy can evolve very readily (Helliwell et al., 2013), and further work is required to establish whether the various bacteria are functionally equivalent with respect to the B-vitamin nutrition of *Drosophila*. The relative importance of bacteria and yeasts in the vitamin nutrition of *Drosophila* may also vary with yeast species, in the light of evidence that yeasts vary in their capacity to support *Drosophila* growth and reproduction (Stamps et al., 2012). It is also unknown whether the *Drosophila* acquires vitamins that are released from living cells of the microorganisms, or following cell lysis in the digestive tract.

Riboflavin deficiency in microbe-free *Drosophila* is predicted to have far-reaching nutritional and metabolic consequences, contributing to explanations for some of the phenotypic traits of these flies. Riboflavin is metabolized to flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN) via riboflavin kinase and FAD synthetase, respectively, both of which enzymes are coded by the *Drosophila* genome. FAD/FMN are the defining required cofactors for all flavoproteins that play crucial roles in energy metabolism (e.g. in respiratory electron transport chain, the decarboxylation of pyruvate, and fatty acyl CoA dehydrogenase in fatty acid oxidation), redox reactions, including reduction of oxidized glutathione (GSSG) to the reduced form (GSH), and the animal-mediated synthesis of the active forms of other B-vitamins, specifically the transformation of folate (vitamin B9) to 5-methyl-tetrahydrofolate, and conversion of pyridoxal (Vitamin B6) to pyridoxal phosphate. These deficiencies may contribute, for example, to the depressed metabolic rate (Ridley et al., 2012) and increased lipid content (Shin et al., 2011; this study) of

microbe-free *Drosophila*.

The second striking discrepancy between the response of conventional and microbe-free *Drosophila* to diet was the reduced weight on microbe-free flies on diets with high glucose, especially high glucose:low yeast content. In the PCA analysis, this effect was associated with high density of lipid, glycogen, glucose and trehalose, the inverse of the relationship between these indices and weight in conventional flies, suggesting that the accumulation of lipid and carbohydrate reserves was either deleterious or a marker of other nutritional dysfunctions in the microbe-free flies. This phenotype is reminiscent of the response of conventional *Drosophila* to diets of exceptionally high sugar content (Musselman et al., 2011; Na et al., 2013). The bacteria associated with conventional *Drosophila* utilize glucose as their principal carbon and energy source, and consequently are predicted to reduce the concentration of glucose available for assimilation by the *Drosophila*. Both the laboratory diets and natural fruits utilized by *Drosophila* are sugar-rich, and bacterial consumption of sugars is unlikely to result in carbon/energy-deficiency of the *Drosophila* on most diets. Rather, this bacterial activity may function to remove excess dietary sugars. As indicated above, the riboflavin deficit in microbe-free flies, especially on low yeast diets, would tend to exacerbate these effects through the impairment of energy metabolism.

A potential caveat to the interpretation of our data is the possibility that the composition and function of the bacterial communities vary systematically with diet composition, confounding direct comparisons of the response of conventional *Drosophila* to diet. Diet composition has been identified as a major determinant of the gut microbiota in the mouse (Faith et al., 2011). Although some published data

suggest that large-scale differences in diet composition (e.g. presence/absence of corn meal, identity of sugar) may be important in structuring the microbiota of *Drosophila*, the variation in the bacterial communities identified in this study were indicative of stochastic variation with no systematic variation with diet. [Stochastic variation has been identified previously in the *Drosophila* microbiota (Wong et al., 2013).] Because the composition of the bacterial communities was not uniform across flies on the different diets, this study is anticipated to identify nutritional functions that are not specific to particular bacterial taxa. We conclude that the bacterial functions of riboflavin synthesis and sugar consumption are likely to be common to *Drosophila* in association with taxonomically-diverse communities in both the laboratory and the field. Our analysis, by definition, cannot exclude other nutritional interactions that are specific to certain bacterial taxa.

The nutritional role of the *Drosophila*-associated bacteria inferred from dietary analysis in this study provides context to previous studies. Very high dietary sugar causes dysfunction in insulin signaling in *Drosophila* (Musselman et al., 2011; Na et al., 2013). The greater functionality of insulin/TOR signaling in flies bearing bacteria than in microbe-free flies (Shin et al., 2011; Storelli et al., 2011) may reflect the high sugar nutrition of microbe-free flies, and associated insulin resistance and related metabolic pathologies. Further research is required to dissect the complex multi-way interactions between diet composition, composition of the microbiota and the microbial impacts on host nutrition and nutritional signaling.

References

Akman, *L.*, Yamashita, *A.*, Watanabe, H., Oshima, K., Shiba, T., Hattori, M., and Aksoy, S. (2002). Genome sequence of the endocellular obligate symbiont of tsetse flies, Wigglesworthia glossinidia. Nat Genet 32, 402-407.

Backhed, F., Ding, H., Wang, T., Hooper, *L.*V., Koh, G.Y., Nagy, *A.*, Semenkovich, C.F., and Gordon, J.I. (2004). The gut microbiota as an environmental factor that regulates fat storage. Proc Natl Acad Sci U S A 101, 15718-15723.

Bakula, M. (1969). The persistence of a microbial flora during postembryogenesis of Drosophila melanogaster. J Invertebr Pathol 14, 365-374.

Benson, *A.*K., Kelly, S.*A.*, Legge, R., Ma, F., Low, S.J., Kim, J., Zhang, M., Oh, P.*L.*, Nehrenberg, *D.*, Hua, K., et al. (2010). Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. Proc Natl Acad Sci U S A 107, 18933-18938.

Blatch SA, Meyer KW, Harrison JF (2010) Effects of dietary folic acid level and symbiotic folate production on fitness and development in the fruit fly *Drosophila melanogaster*. Fly 4: 312–319.

Broderick, N.*A.*, and Lemaitre, B. (2012). Gut-associated microbes of Drosophila melanogaster. Gut microbes 3, 307-321.

Brummel, T., Ching, *A.*, Seroude, *L.*, Simon, *A.*F., and Benzer, S. (2004). Drosophila lifespan enhancement by exogenous bacteria. Proc Natl Acad Sci U S A 101, 12974-12979.

Caricilli, *A.*M., and Saad, M.J. (2013). The role of gut microbiota on insulin resistance. Nutrients 5, 829-851.

Chandler, J.*A.*, Lang, J.M., Bhatnagar, S., Eisen, J.*A.*, and Kopp, *A.* (2011). Bacterial communities of diverse Drosophila species: ecological context of a host-microbe model system. PLoS Genet 7, e1002272.

Douglas, *A.*E. (2009). The microbial dimension in insect nutritional ecology. Functional Ecology 23, 38-47.

Erkosar, B., Storelli, G., Defaye, *A.*, and Leulier, F. (2013). Host-intestinal microbiota mutualism: "learning on the fly". Cell Host Microbe 13, 8-14.

Faith, J.J., McNulty, N.P., Rey, F.E., and Gordon, J.I. (2011). Predicting a human gut microbiota's response to diet in gnotobiotic mice. Science 333, 101-104.

Flint, H.J., Scott, K.P., Louis, P., and Duncan, S.H. (2012). The role of the gut microbiota in nutrition and health. Nature reviews. Gastroenterology & hepatology 9, 577-589.

Goodman, *A.L.*, McNulty, N.P., Zhao, Y., Leip, *D.*, Mitra, R.*D.*, Lozupone, C.*A.*, Knight, R., and Gordon, J.I. (2009). Identifying genetic determinants needed to establish a human gut symbiont in its habitat. Cell Host Microbe 6, 279-289.

Gordon, H.*A.*, and Pesti, *L.* (1971). The gnotobiotic animal as a tool in the study of host microbial relationships. Bacteriological Reviews 35, 390-429.

Helliwell, K.E., Wheeler, G.*L.*, and Smith, *A.*G. (2013). Widespread decay of vitamin-related pathways: coincidence or consequence? Trends in genetics : TIG.

Karasov, W.H., and Douglas, *A.*E. (2013). Comparative digestive physiology. Comprehensive Physiology 3, 741-783.

Kau, *A.L.*, Ahern, P.P., Griffin, N.W., Goodman, *A.L.*, and Gordon, J.I. (2011). Human nutrition, the gut microbiome and the immune system. Nature 474, 327-336.

LeBlanc, J.G., Milani, C., de Giori, G.S., Sesma, F., van Sinderen, *D.*, and Ventura, M. (2013). Bacteria as vitamin suppliers to their host: a gut microbiota perspective. Current opinion in biotechnology 24, 160-168.

Musselman, *L.*P., Fink, J.*L.*, Narzinski, K., Ramachandran, P.V., Hathiramani, S.S., Cagan, R.*L.*, and Baranski, T.J. (2011). A high-sugar diet produces obesity and insulin resistance in wild-type Drosophila. Disease models & mechanisms 4, 842-849.

Na, J., Musselman, *L.*P., Pendse, J., Baranski, T.J., Bodmer, R., Ocorr, K., and Cagan, R. (2013). A Drosophila model of high sugar diet-induced cardiomyopathy. PLoS Genet 9, e1003175.

Nicholson, J.K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., and Pettersson, S. (2012). Host-gut microbiota metabolic interactions. Science 336, 1262-1267.

Nogge, G. (1976). Sterility in tsetse fly (Glossina morsitans Westwood) caused by loss of symbionts. Experientia 32, 995.

Puchta, O. (1956). Zuchtungsversuche an den Symbionten von Pediculus vestimenti Burm. Nebst physiologischen und morphologischen Beobachtungen. Zeitschrift fur Morphologie und Oekologie der Tiere 44, 416-441.

Ridley, E.V., Wong, *A.*C., Westmiller, S., and Douglas, *A.*E. (2012). Impact of the resident microbiota on the nutritional phenotype of Drosophila melanogaster. PLoS One 7, e36765.

Roscoe, M.H. (1931). The effects of coprophagy in rats deprived of the vitamin B complex. Biochemistry Journal 25, 2056-2067.

Sang, J. H. (1956). The quantitative nutritional requirements of Drosophila melanogaster. J. Exp. Biol. 33. 45-72.

Shin, S.C., Kim, S.H., You, H., Kim, B., Kim, *A.*C., Lee, K.*A.*, Yoon, J.H., Ryu, J.H., and Lee, W.J. (2011). Drosophila microbiome modulates host developmental and metabolic homeostasis via insulin signaling. Science 334, 670-674.

Smith, K., McCoy, K.*D.*, and Macpherson, *A.*J. (2007). Use of axenic animals in studying the adaptation of mammals to their commensal intestinal microbiota. Seminars in immunology 19, 59-69.

Smith, M.I., Yatsunenko, T., Manary, M.J., Trehan, I., Mkakosya, R., Cheng, J., Kau, *A.L.*, Rich, S.S., Concannon, P., Mychaleckyj, J.C., et al. (2013). Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. Science 339, 548-554.

Stamps, J.*A.*, Yang, *L.*H., Morales, V.M., and Boundy-Mills, K.*L.* (2012). Drosophila regulate yeast density and increase yeast community similarity in a natural substrate. PLoS One 7, e42238.

Stecher, B., Maier, *L.*, and Hardt, W.*D.* (2013). 'Blooming' in the gut: how dysbiosis might contribute to pathogen evolution. Nature reviews. Microbiology 11, 277-284.

Storelli, G., Defaye, *A.*, Erkosar, B., Hols, P., Royet, J., and Leulier, F. (2011). Lactobacillus plantarum promotes Drosophila systemic growth by modulating hormonal signals through TOR-dependent nutrient sensing. Cell Metab 14, 403-414.

Turnbaugh, P.J., Ley, R.E., Mahowald, M.*A.*, Magrini, V., Mardis, E.R., and Gordon, J.I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. Nature 444, 1027-1031.

Vijay-Kumar, M., Aitken, J.*D.*, Carvalho, F.*A.*, Cullender, T.C., Mwangi, S., Srinivasan, S., Sitaraman, S.V., Knight, R., Ley, R.E., and Gewirtz, *A.*T. (2010). Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. Science 328, 228-231.

Wong, *A.*C., Chaston, J.M., and Douglas, *A.*E. (2013). The inconstant gut microbiota of Drosophila species revealed by 16S rRNA gene analysis. ISME J.

Wong, C.N., Ng, P., and Douglas, *A.*E. (2011). Low-diversity bacterial community in the gut of the fruitfly Drosophila melanogaster. Environ Microbiol 13, 1889-1900.

Yi, P., and Li, *L.* (2012). The germfree murine animal: an important animal model for research on the relationship between gut microbiota and the host. Veterinary microbiology 157, 1-7.

APPENDIX A

IMPACT OF THE RESIDENT MICROBIOTA ON THE NUTRITIONAL PHENOTYPE OF *DROSOPHILA MELANOGASTER*[4]

Abstract

Animals are chronically infected by benign and beneficial microorganisms that generally promote animal health through their effects on the nutrition, immune function and other physiological systems of the host. Insight into the host-microbial interactions can be obtained by comparing the traits of animals experimentally deprived of their microbiota and untreated animals. *Drosophila melanogaster* is an experimentally tractable system to study host-microbial interactions. In this study, the nutritional significance of the microbiota was investigated in *D. melanogaster* bearing unmanipulated microbiota, demonstrated by 454 sequencing of 16S rRNA amplicons to be dominated by the α-proteobacterium *Acetobacter*, and experimentally deprived of the microbiota by egg dechorionation (conventional and axenic flies, respectively). In axenic flies, larval development rate was depressed with no effect on adult size relative to conventional flies, indicating that the microbiota promotes larval growth rates. Female fecundity did not differ significantly between conventional and axenic flies, but axenic flies had significantly reduced metabolic rate and altered carbohydrate allocation, including elevated glucose levels. In summary, we have shown that

[4] Presented with minor modifications from the originally published article:
Ridley, EV., Wong, ACN., Westmiller, S. and Douglas, AE. (2012). Impact of the resident microbiota on the nutritional phenotype of *Drosophila melanogaster*. *PLoS One*, 7(5): e36765.
All supplementary materials can be found at:
http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0036765#s5

elimination of the resident microbiota extends larval development and perturbs energy homeostasis and carbohydrate allocation patterns of of *D. melanogaster*. Our results indicate that the resident microbiota promotes host nutrition and interacts with the regulation of host metabolism.

Introduction

It is increasingly recognized that all animals are chronically infected by microorganisms, and that the resident microbiota, especially the substantial microbial community in the alimentary tract, has major effects on nutrient processing, metabolic signaling and, ultimately, the health and well-being of the animal host [1], [2], [3]. There is now persuasive evidence linking the gut microbiota with energy homeostasis of rodent biomedical models and humans, including microbial-mediated promotion of nutrient acquisition and storage [4]. In particular, a causal role of the microbiota in animal energy metabolism is indicated by the elevated lipid levels and other indices of metabolic syndrome in wild-type mice infected with the microbiota from individuals that are obese as a consequence of genetic deficiencies in leptin or Toll-like receptor 5 (a component of the innate immune system that is expressed in the gut) [5], [6].

It is experimentally challenging to study the interactions between the resident microbiota and the nutrition of humans and rodent biomedical models because the microbiota of mammals includes hundreds of taxa, many of which are unculturable, with wide variation in composition among individuals [7], [8], [9]. Simple systems comprising animals bearing one or a few microbial taxa are valuable tools to investigate how resident microorganisms interact with host metabolism [10]. For example, mice experimentally infected with specific bacterial taxa have revealed the

effects of the gut microbiota on carbohydrate and energy metabolism [11], [12]. A second approach, adopted in this study, is to use insects that have a less diverse microbiota than mammals, often comprising <20 species [13], [14], [15], [16]. In particular, *Drosophila melanogaster* combines renowned genetic and experimental tractability [17] with a microbiota that is culturable, of low diversity, and uniform among individuals for a given set of conditions. The bacteria associated with *Drosophila* include *Acetobacter*, Gluconobacter, Enterococcus and *Lactobacillus* [18], [19], [20], [21], [22], [23].

Comparison between animals containing and experimentally deprived of microorganisms is a powerful strategy to investigate the interactions between animals and their resident microbiota. Here, we provide the first analysis of how the resident microbiota affects the organismal physiology of *Drosophila*, with particular emphasis on nutrition. Using insects reared on a diet that supports excellent performance of *Drosophila* with unmanipulated microbiota, we investigate the impact of eliminating the microbiota on host performance (growth, fecundity etc), nutritional status and metabolic rate. Our data suggest that, although the resident microorganisms are not essential for *Drosophila*, they have pervasive effects on the nutrition and metabolic status of their animal host.

Experimental procedures

Fly cultures and experimental design

Wolbachia-free *Drosophila melanogaster* strain Canton-S was reared at 25°C with a 12 h:12 h light–dark cycle on autoclaved medium containing 96 g glucose (Sigma), 48 g inactive dry yeast and 14 g agar (both from Genesse Scientific) l^{-1}, equivalent to 5:1

(g/g) carbohydrate:protein ratio. Experiments were initiated with eggs deposited overnight by mated females. Two egg treatments were used: dechorionated eggs (yielding axenic insects), obtained by washing in sterile deionized water, immersion in 10% sodium hypochlorite solution for 5 min, and then three rinses in sterile water; and control eggs (yielding conventional insects), for which the hypochlorite was replaced by sterile water. To initiate experiments, 10 eggs were transferred to uncrowded conditions comprising replicate vials (2 cm diam.) containing ca. 8 ml diet. All manipulations were conducted in a laminar flow cabinet with aseptic technique. In some experiments, the microbiota was depleted by rearing insects from control eggs on diet supplemented with 50 µg chloretetracycline (Sigma) ml^{-1}. To supplement the diet with *Drosophila* microbiota, adult males were cultured for 24 h on sterile medium, which was then rinsed with sterile PBS, and 50 µl of the fecal washing was added to the test diets; the fecal washings were confirmed to contain viable *Acetobacter*, by plating onto bacteriological agar.

Bacterial content of flies

DNA extractions were conducted with the DNeasy Blood and Tissue Kit (Qiagen, Valencia, California, USA) following a protocol modified from the manufacturer's instructions to ensure disruption of Gram-positive bacteria. Specifically, samples were hand-homogenised in 20 mM Tris-HCl pH 8.0, 2 mM sodium EDTA, 1.2% Triton® X-100 containing 20 mg lysozyme ml^{-1}, and the homogenates were incubated at 37°C for 1.5 h with a 5-min bead-beating in a Disruptor Genie® using 0.1 mm glass beads (Scientific Industries) at 45 min. Individual conventional and axenic flies were checked for the presence of bacteria by PCR using general 16S rRNA gene primers

16SA1: 5′- AGAGTTTGATCMTGGCTCAG-3′ and 16SB1: 5′ – TACGGYTACCTTGTTACGACTT-3′ [43], yielding ca. 1.5 kb product (27F-1522R). The PCR reactions contained 1× Taq polymerase buffer, 0.24 mM of each dNTP, 2 mM MgCl2, 0.32 µM primers, 1 µl template DNA and 0.025 U Platinum Taq in 25 µl. The cycling conditions were 5 min at 94°C, followed by one cycle of 1 min at 55°C, 72°C for 2 min and 25–30 cycles of 1 min at 94°C, 1 min at 55°C and 2 min at 72°C with a final incubation of 8 min at 72°C. All experiments included PCR reactions replacing template DNA with water, as negative control. PCR products were separated by electrophoresis in a 1% agarose gel with molecular weight markers, and visualized under ultraviolet light after staining with Sybr Safe (Invitrogen).

The diversity of bacteria associated with the flies was investigated by 454 pyrosequencing of the V2 region of the 16S rRNA gene. Three replicate PCR reactions were conducted on the experimental sample comprising DNA extracted from five pooled 7-day-old adults (3 male and 2 female) conventional flies, with a reagent-only negative controls. The primers were 27F (ACGCTCGACAAGAGTTTGATCMTGGCTCAG) and 338R (TGCTGCCTCCCGTAGGAGT), with the sample-specific 27F primer bearing a multiplex identifier (MID) sequence [MID2 (ACGCTCGACA) for the experimental sample, MID11 (TGATACGTCT) for the control sample) and all 27F and 338R primers modified with 5′-Adaptor A and 5′-Adaptor B sequences (Roche), respectively. The reactions comprised 0.6 U Platinum® Taq DNA Polymerase (Invitrogen) in 1× PCR buffer, 2 mM MgCl2, 8 pmol each primer, 0.24 mM dNTP, and 1 µl template in 25 µl final volume, at 94°C for 10 min followed by 25 cycles of 94°C for 1 min, 58°C for 1 min, and 72°C for 1 min. Following purification with the

Qiagen Qiaquick purification kit and quantification using the Quant-iT™ PicoGreen® Kit, each sample was diluted to 1×10^{7} molecules per microliter (based on 350 bp size of the products). Emulsion PCR with 1.5 copies per bead used only "A" beads for unidirectional sequencing on 454 GS-FLX pyrosequencing instrument with standard Titanium chemistry.

Pyrosequencing flowgrams were converted to sequence reads using 454 Life Science software (www.454.com). The data were then processed using Pyrotagger [44] as previously described [20], with minor modifications. In brief, reads with ambiguous nucleotides (N), <290 nucleotides after the forward primer, mismatches with the 16S rRNA gene primers, and all reads with 0.2% per-base error probability (≥3% of bases with Phred scores <27) were removed The remaining sequences were trimmed to 290 nucleotides, dereplicated and clustered into operational taxonomic units (OTUs) with 97% sequence identity (ID) threshold. The most abundant unique sequence of each OTU cluster was selected as representative, and checked for chimeras by the Mallard algorithm (Ashelford et al., 2006). Non-chimeric sequences was assigned to bacterial taxa by NCBI StandAlone BLAST (megaBLAST program) using the nucleotide (nt) database (13 August 2011) with default settings, and allocated to the experimental or control sample according to the MID sequence. The sequences of the three clusters are available at NCBI, with accession numbers provided in Table 4.1.

Insect performance indices

Vials with dechorionated or control eggs were monitored daily, and the pupation and eclosion dates of every insect surviving to adulthood was scored, from which the number surviving to adulthood and median development time per vial was determined.

Three days later, 10 females were selected at random from across each treatment for analysis of fecundity. Each insect was transferred aseptically to an individual sterile 15 ml Falcon tube containing autoclaved diet in the lid. The lid was changed daily for 7 days, the number of eggs per lid per day was scored, and the median number of eggs per day deposited by each female was determined.

Nutritional analyses

Ten replicate 7-to-10-day-old adult flies were weighed on a Mettler MX5 microbalance (1 µg accuracy). The flies were then homogenized in 80 µl ice-cold buffer comprising 10 mM Tris, 1 mM EDTA pH 8.0 and 0.1% (v/v) Triton-X-100 with hand-held homogenizer, and centrifuged at 7,000 g at 4°C for 1 min. The supernatant was used for analysis of protein, triglyceride and carbohydrates using coupled colorimetric assays with an xMark™ microplate spectrophotometer, following manufacturer's instructions (5 replicates per assay). The assay kits were the triglyceride assay kit of Sigma (catalogue number TG-5-RB); the Coomassie Brilliant Blue microassay method of BioRad (catalogue number 500-0201), with bovine serum albumin as standard (40–480 µg protein ml^{-1}) for protein; and the glucose assay kit of Sigma (catalogue number GAGO20) for glucose and, following trehalase (3.7 U/ml) and amyloglucosidase (2 U ml^{-1}) treatment, for trehalose and glycogen, respectively.

Respirometry

Respiratory oxygen consumption and carbon dioxide production by 7-to-10 day old adult flies were determined by stop-flow respirometry with air scrubbed of water vapour and carbon dioxide by silica/Ascarite columns. All experiments were

conducted at 25°C with low light conditions that minimized insect activity, and at 3–7 hours after onset of the light period, with the flies of each treatment analyzed at different times on multiple days, to avoid any confounding effects of circadian rhythm in *Drosophila* respiration rates. Each replicate of 5 flies was transferred to a respirometry chamber comprising a 5 ml syringe, and allowed to acclimate for 30 minutes prior to analysis, by which time they were quiescent. The air in the syringe was then replaced by 3.2 ml dried carbon dioxide-free air, with airflow at 57 ml min^{-1}. The carbon dioxide and oxygen content of the syringe was determined 30 minutes later by injecting 3 ml of the syringe volume into Sable Systems SS3 Gas Analyzer Sub-sampler with an FCA-10A CO2 analyzer and FC-10 O2 Analyzer (Sable systems, Nevada, USA), respectively. The gas analyzers were calibrated with 50 ppm CO2 gas and 20.9% O2 gas. Carbon dioxide and oxygen contents were analyzed using the Sable System data acquisition software (Expedata, Sable Systems, Nevada, USA). All experiments included an empty baseline chamber, as a control for drift in the baseline measures.

Results

Bacterial complement of flies

The first experiments tested for the presence of bacteria by PCR with general 16S rRNA gene primers (Figure 4.1). A PCR product of the predicted size was obtained from flies reared from eggs that had been washed in sterile water (conventional flies) but not from dechorionated eggs (axenic flies). Conventional insects reared to the pupal stage, then surface-sterilized with 10% hypochlorite, and allowed to develop on

sterile diet to 14-day-old adults, also bore bacteria.

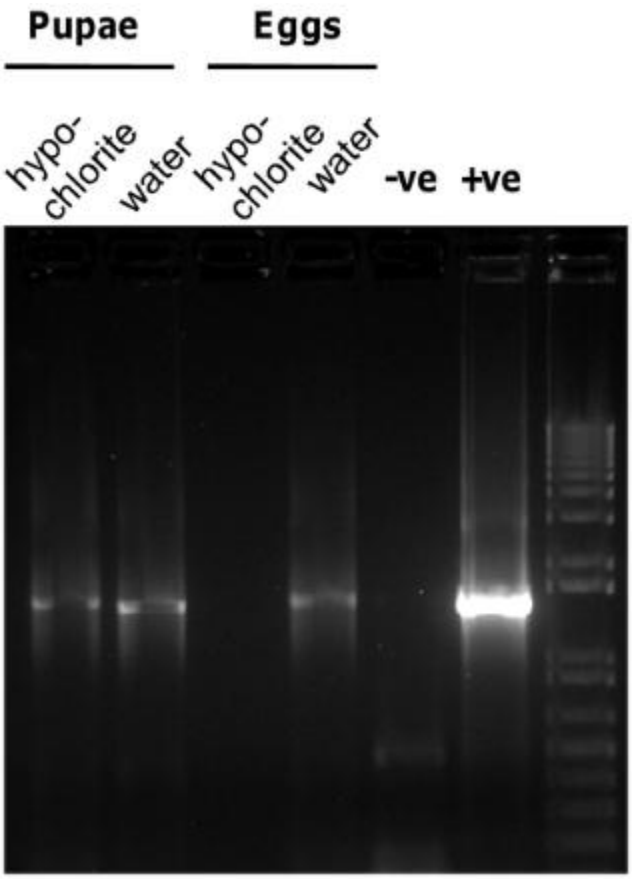


Figure 4.1 Bacterial complement of *Drosophila*. PCR assay with general 16S rRNA primers of 14-day-old adult flies, derived from pupae washed in 10% sodium hypochlorite solution or sterile water (lanes 1–2), and eggs washed in 10% sodium hypochlorite solution or sterile water (lanes 3–4). Negative and positive controls are PCR reactions with DNA from filtered water and *Drosophila* in standard culture, respectively, as template (lanes 5–6).

These data indicate that bacteria are acquired from the external environment by first-instar larvae, and persist through larval development and in internal tissues of pupae to adulthood, validating early studies [24] that quantified CFUs of culturable bacteria

without identification. In supplementary PCR assays with general 16S primers throughout the experimental study, axenic flies of all ages invariably yielded negative results, and all conventional flies bore bacteria.

The 454 pyrosequencing of 16S rRNA gene amplicons of DNA from adult flies yielded 46,752 sequence reads with an average length of 352 nucleotides (including the multiplex identifier "MID" and primer sequences), after quality filtering and removal of chimeric sequences. A single cluster with 100% sequence ID to the α-proteobacterium *Acetobacter pomorum* EW816 accounted for 98% of the reads. The remaining reads were assigned to: *Lactobacillus plantarum*, (1.9% of reads) and an uncultured γ-proteobacterium in the family Xanthomonadacae (0.1% of reads) (Table 4.1).

Table 4.1 16S rRNA gene amplicons detected by 454 pyrosequencing in 5–7-day-old adult *D.melanogaster*.

NCBI accession number of cluster (this study)	Number of sequence reads		Sequence identity		
	Experimental sample	Reagent-only control	NCBI accession number	Taxonomic identity	% sequence identity
JN592041	45682	26	EU096229.1	*Acetobacter pomorum* strain EW816	100
JN592042	873	7	AL935263.2	*Lactobadillus plantarum* strain WCFS1	100
JN592043	28	1	FJ893035	Uncultured bacterium clone nbt16f09	95.2

Insect performance

Table 4.2 displays the performance indices of conventional and axenic insects. Development time to adulthood was significantly extended by a median value of one day in axenic insects. The other fitness indices tested, survival to adulthood, adult

weight, and female fecundity over 7 days, did not differ significantly between the two treatments (Table 4.2 and Figure 4.2A).

Table 4.2 Fitness indices of conventional and axenic *Drosophila*.

Treatment	Performance indices Median (range)			
	Survival to adulthood (per 10 eggs)	Development time to adulthood (days)	Number of eggs deposited day^{-1} female^{-1}	Male lifespan (days)
Conventional	7.0 (5–8)	12 (11–13)	17.5 (5–24)	41.5 (20–62)
Axenic	7.5 (5–10)	13 (12–16)	17.0 (11.5–22)	49.0 (34–61)
Mann-Whitney U[a]	W = 107, p >0.05	W = 313, p<0.001	W = 103.5, p >0.05	W = 63, p>0.05

[a]Critical probability = 0.0125, after Bonferroni correction for 4 tests.
doi:10.1371/journal.pone.0036765.t002

The basis for the extended development time to adulthood of axenic insects was investigated. The egg dechorionation treatment used to generate axenic insects was confirmed to have no effect on survivorship or development time of the embryos: the median proportion of larvae hatching from dechorionated and control eggs was 0.9 and 0.8, respectively ($p>0.05$), and median development time to hatching was 19 h for both control and dechorionated eggs (n = 10). The development time of conventional and axenic insects from egg deposition to pupation was 7 and 8 days, respectively (Mann Whitney U, W = 874, $p<0.001$), the same difference of one day as between development time of conventional and axenic insects from egg deposition to adulthood (Table 4.2). These data indicate that larval development time was extended in axenic insects.

Two sets of supplementary experiments were conducted. In the first experiment, 10 replicate groups of 10 untreated eggs were transferred to diet containing the antibiotic chlortetracycline at 50 µg ml^{-1}, a treatment which reduces the number of culturable bacteria per fly by >90% (Ridley, unpub. data). The median development time to

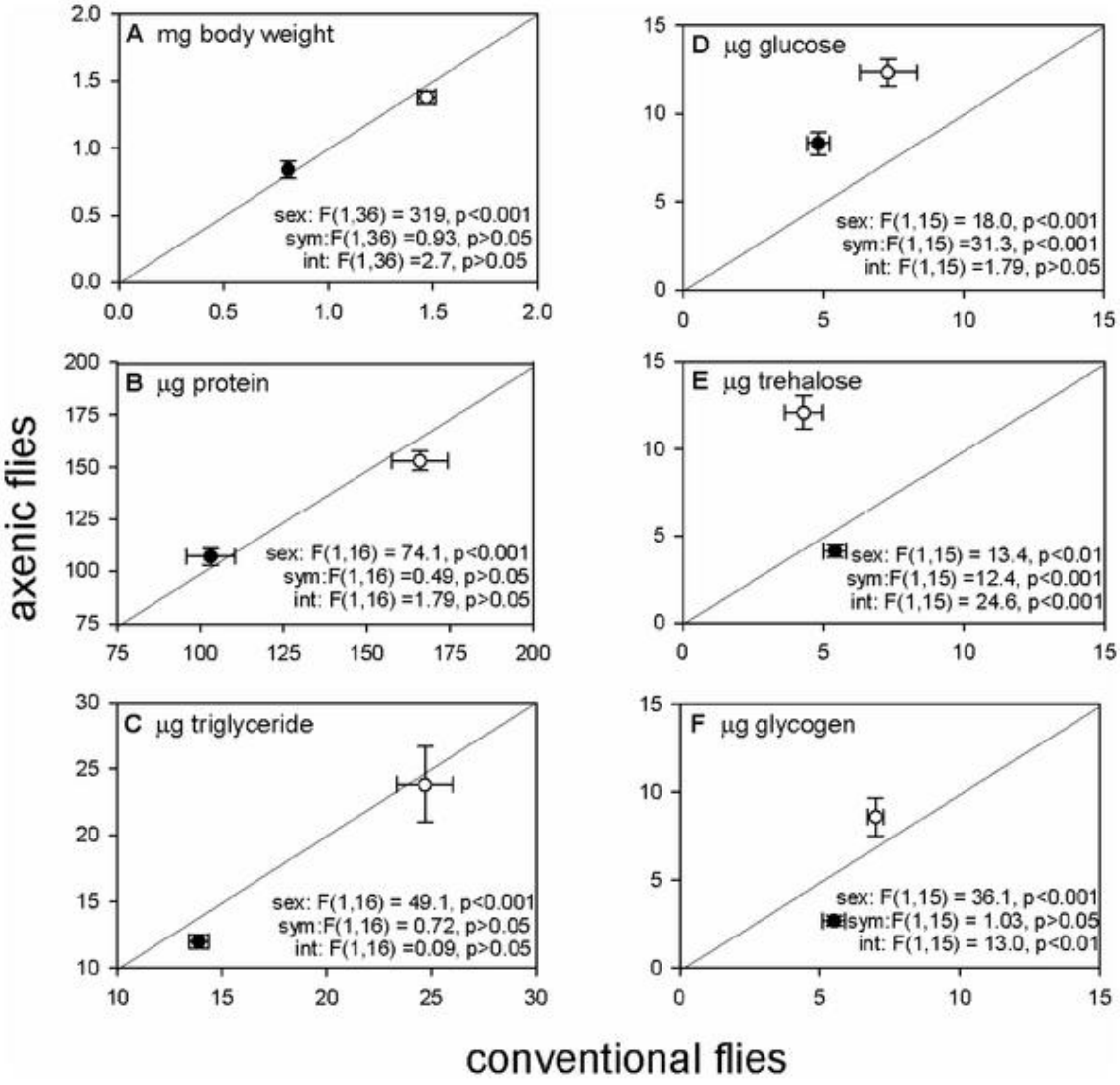
adulthood of the antibiotic-treated flies was 13.5 days (range 12–15 days, n = 10), significantly longer than for conventional flies with median development time of 12 days (range 11–13 days) (Mann Whitney W = 276.5, $p<0.001$) but not significantly different from axenic flies with median of 13 days and range 12–16 days (Mann Whitney W = 364.5, $p>0.05$). In the second set of experiments, 10 replicate groups of 10 dechorionated eggs were transferred to sterile diet and sterile diet seeded with feces collected from adult male flies, with untreated eggs as controls. The median development times of the insects reared on the fecal-seeded plates and the conventional flies were identical, at 13 days (range 12–14 days), and significantly shorter than the median development time of axenic flies (14 days, range 13–15 days) (Mann Whitney W = 55, $p<0.002$). This final analysis was conducted at a different time with a different batch of dietary yeast from the previous experiments, giving slightly different absolute values for development times but the same patterns as shown in Table 4.2. Taken together, these experiments indicate that the slow development of larvae from dechorionated eggs is caused by the absence of resident microorganisms and could not be attributed to non-specific deleterious effects of the dechorionation procedure.

Nutritional indices

The values of all nutritional indices (Figure 4.2) were significantly greater in females than males, reflecting the difference in body size between the sexes. Conventional and axenic flies did not differ significantly in protein or triglyceride contents (Figure 4.2 B–C), but did vary with respect to the three carbohydrates tested, glucose, trehalose and glycogen. The glucose content was elevated by ca. 70% in both female and male

axenic flies (Figure 4.2D), but the effect of axenic rearing on the trehalose and glycogen content differed significantly between the sexes (Figure 4.2E–F). For females, the trehalose and glycogen contents of axenic flies were elevated by 68% and 20%, respectively, relative to conventional flies; but these indices were reduced in axenic males, by 30% and 100%, respectively. The sum of glucose, trehalose and glycogen contents was significantly greater in axenic than conventional flies for females (32.1±1.75 µg versus 18.5±1.84 µg per fly, $t6 = 8.06$, $p<0.001$), but not males (15.9±1.13 versus 15.7±0.71 µg per fly, $t7 = 0.11$, $p>0.05$).

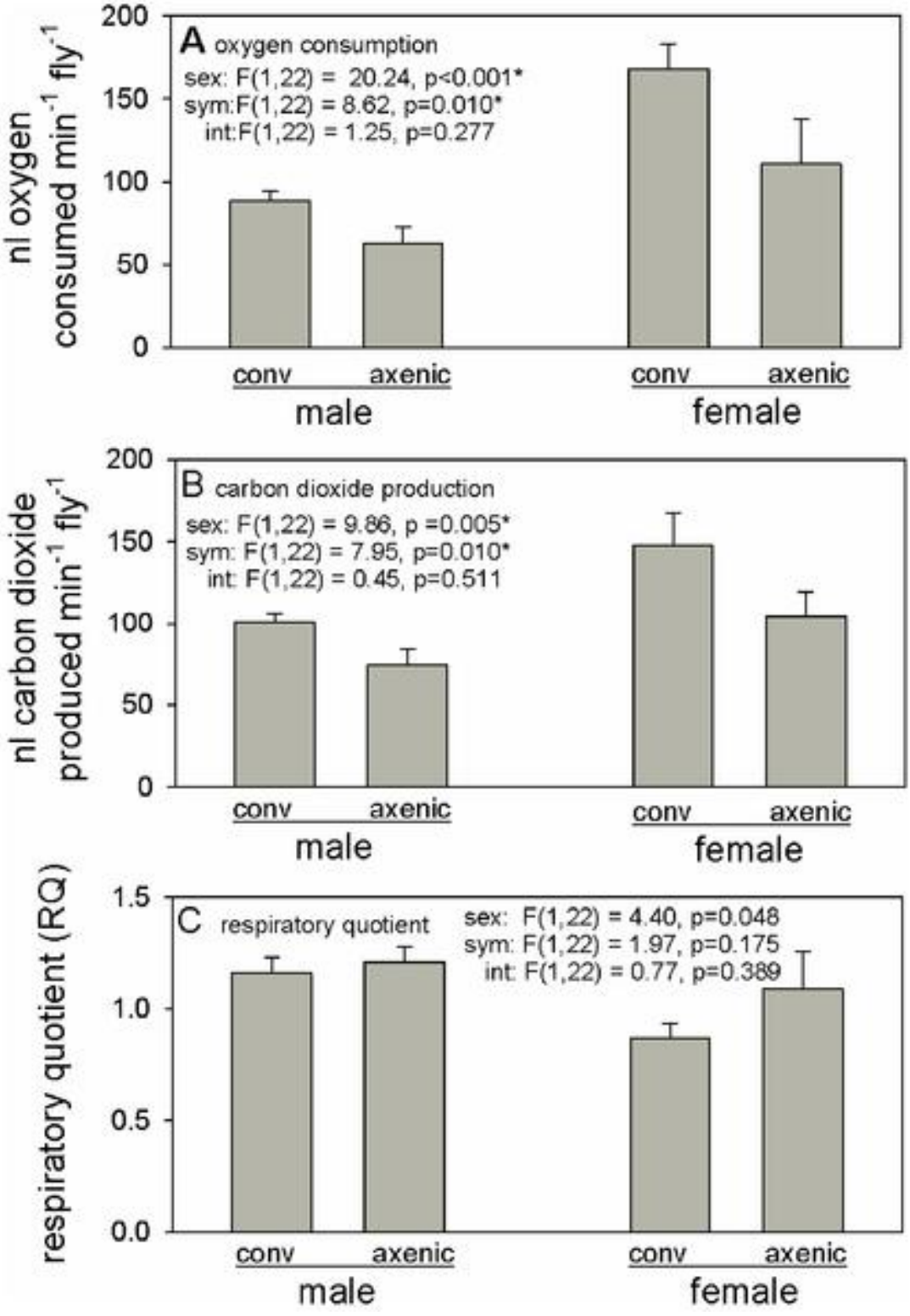
Figure 4.2 Nutritional indices of 7-to-10-day-old conventional and axenic flies, all expressed on per fly basis. Factors in ANOVA are sex (female ○, male •), sym (conventional or axenic) and int (interaction). Critical probability = 0.008 after Bonferroni correction for six tests. Data are represented as mean +/− SEM.


A mg body weight
sex: F(1,36) = 319, p<0.001
sym:F(1,36) =0.93, p>0.05
int: F(1,36) =2.7, p>0.05
B µg protein
sex: F(1,16) = 74.1, p<0.001
sym:F(1,16) =0.49, p>0.05
int: F(1,16) =1.79, p>0.05
C µg triglyceride
sex: F(1,16) = 49.1, p<0.001
sym:F(1,16) = 0.72, p>0.05
int: F(1,16) =0.09, p>0.05
D µg glucose
sex: F(1,15) = 18.0, p<0.001
sym:F(1,15) =31.3, p<0.001
int: F(1,15) =1.79, p>0.05
E µg trehalose
sex: F(1,15) = 13.4, p<0.01
sym:F(1,15) =12.4, p<0.001
int: F(1,15) = 24.6, p<0.001
F µg glycogen
sex: F(1,15) = 36.1, p<0.001
sym:F(1,15) = 1.03, p>0.05
int: F(1,15) = 13.0, p<0.01
axenic flies
conventional flies

Respiration rates

For both oxygen consumption and carbon dioxide production, exchange rates were significantly lower in males relative to females, and in axenic flies relative to conventional flies, with non-significant interaction terms in the ANOVA tests (Figure 4.3A). Bacterial respiration was calculated to contribute <2% to the difference in respiration rate between conventional and axenic flies (Text S1). The RQ was not significantly affected by either sex or treatment (Figure 4.3B), and the mean values of both males and females did not differ significantly from unity (male RQ: 1.18±0.051 ($n = 15$), $t14 = 0.229$, $p>0.05$; female RQ: 0.97±0.087 ($n = 11$), $t10 = 0.030$, $p>0.05$), indicating that the dominant respiratory fuel in all flies was glucose.

Figure 4.3 Respiratory exchange of 7-to-10-day-old conventional (conv) and axenic flies. Factors in ANOVA are sex, sym (conventional or axenic) and int (interaction). Critical probability = 0.017 after Bonferroni correction for three tests (*, statistically significant). Data are represented as mean +/− SEM.


A oxygen consumption
sex: F(1,22) = 20.24, p<0.001*
sym:F(1,22) = 8.62, p=0.010*
int:F(1,22) = 1.25, p=0.277
nl oxygen consumed min-1 fly-1
200
150
100
50
0
conv
axenic
male
female
B carbon dioxide production
sex: F(1,22) = 9.86, p =0.005*
sym: F(1,22) = 7.95, p=0.010*
int: F(1,22) = 0.45, p=0.511
nl carbon dioxide produced min-1 fly-1
C respiratory quotient
sex: F(1,22) = 4.40, p=0.048
sym: F(1,22) = 1.97, p=0.175
int: F(1,22) = 0.77, p=0.389
respiratory quotient (RQ)
1.5
1.0
0.5
0.0

Discussion

Effects of axenic cultivation on Drosophila performance

The experimental value of animals deprived of their resident microbiota to study symbiosis function depends critically on the specificity and efficacy of the methods to eliminate the microbiota, and the degree of host dependence on the microbiota. This study demonstrates that axenic *Drosophila* obtained from dechorionated eggs are ideally suited to this approach because egg dechorionation completely eliminates the microbiota (Figure 4.1), while the eggs are undamaged by the treatment, as indicated by the uniform survivorship and development time of treated and control eggs to hatching, the comparable effects of dechorionation and antibiotic treatment on development time to adulthood, and the equivalent development time of conventional insects and insects from dechorionated eggs provided with bacteria via fecal washings. The sole performance effect of eliminating the microbiota identified under the conditions tested here was extended larval development time of axenic *Drosophila*. If this effect were replicated under natural conditions, it would be beneficial for *Drosophila* because multiple eggs are deposited onto rotting fruit, such that larvae are in scramble competition for a transient resource. Individuals that develop rapidly are at a competitive advantage and more likely to pupate before exhaustion of the resource [25].

Importantly, the extended larval development time of axenic flies was not accompanied by any difference in adult body size between axenic and conventional flies (Figure 4.2) under the rearing conditions employed. Thus, axenic larvae take longer than conventional larvae to reach the critical weight at which they are committed to metamorphosis, but they are able to acquire dietary nutrients and convert

them into biomass as efficiently as conventional insects once they have passed the critical weight, i.e. during the interval to cessation of growth (ICG). This suggests that microbial effects are particularly important during larval development prior to ICG. Nevertheless, these results should be extrapolated beyond the specifics of this study with great caution. Although the literature is fragmentary, there are indications that multiple aspects of diet composition, host genotype and the identity of the resident microbiota may influence *Drosophila* performance, potentially in an interactive fashion (e.g. demonstrated by unpublished results in chapter 4). For example, elimination of the microbiota has been reported to reduce the lifespan of *Drosophila* reared on diet containing sucrose [21], but this effect was not replicated for flies reared on a diet containing glucose [18]; and the effect of sugar type on the performance of conventional *Drosophila* can vary with both sugar concentration and host genotype [26]. Performance can also vary with the composition of the microbiota, which is influenced by age and immunocompetence of the *Drosophila* [20], [27]. An indication that diet composition can also affect microbial composition comes from the comparison between the microbiota in the young adult flies studied here and a previously-published analysis of the same *Drosophila* strain reared on a diet with higher yeast content. In both studies, the young adult flies bore *Acetobacter* and *Lactobacillus*, but at ratio of 49:1 in this study (4.8% yeast diet), and 1:4 in the study using 8.6% yeast diet [20]. Further research involving systematic variation of these multiple factors is required to elucidate the multiway interactions between diet, bacterial composition, host genotype and insect performance.

Effects of axenic cultivation on the nutritional phenotype of Drosophila

A key finding of this study was the impact of the microbiota on the carbohydrate allocation pattern of the adult *Drosophila* (Figure 4.2D–F). Furthermore, the elevated female-specific body glycogen content and prolonged larval development, obtained for axenic flies on the diet used in this study [with 5:1 carbohydrate:protein ratio (5C:1P)], has also been reported for conventional flies on diets containing 10C:1P, relative to diets with more balanced C:P ratios (5C:1P and 2.5C:1P) [28]. These data suggest that the bacteria may reduce insect utilization of ingested carbohydrate. Specifically, the bacteria in the gut lumen may compete with the *Drosophila* for ingested carbohydrate. Additionally or alternatively, they may suppress insect digestion of complex dietary carbohydrates. Candidate bacterial products are acetic acid and lactic acid, which are secreted by *Acetobacter* and *Lactobacillus* species, respectively, and are known to reduce the digestibility of starch and other carbohydrates by mammals [29], [30], [31], [32]. The impact of the microbiota on the nutritional status of *Drosophila* may also arise from system-level effects on host signaling networks that regulated carbohydrate allocation patterns. In particular, *Acetobacter* and *Lactobacillus* (both resident in the flies studied here) have been implicated to promote insulin signaling in different *Drosophila* genotypes reared on diets of different formulations from this study [33], [34]. The sex-specific effect of axenic cultivation on the level of glycogen and also the disaccharide blood sugar trehalose in *Drosophila* (Figure 4.2) is consistent with the prediction that nutrient allocation to energy reserves is more responsive to diet composition in females, which have a high reproductive investment, than in males [35].

Other data suggest that the microbiota has a profound effect on energy homeostasis of

Drosophila. In particular, the significantly elevated glucose content of axenic flies can be attributed to one or both of reduced demand and increased supply of glucose. Glucose is likely the dominant respiratory fuel for both conventional and axenic flies (RQ does not differ significantly from unity), but axenic flies have a lower respiratory demand for glucose, as indicated by their lower respiration rate than conventional flies. A greater supply of glucose from ingested food for axenic than conventional flies is also predicted (see above). In particular, a contribution of bacterial-derived acetic acid in depressing the glucose content of *Drosophila* is suggested by the evidence that lowered blood glucose levels accompany the reduced digestibility of complex carbohydrates in human volunteers who include acetic acid in their diet [31]. These effects in axenic flies may be linked to reduced insulin/insulin-like growth factor signaling (IIS), which is known to promote free glucose levels [36], alter mitochondrial function resulting in reduced rates of oxygen consumption and oxidative phosphorylation [37], and depress *Drosophila* developmental rate prior to ICG [38]. The absence of any discernible effect of hyperglycemia on the weight or fecundity of axenic flies (Figure 4.2 and Table 4.2) reflects the far greater physiological tolerance of variable sugar levels in insects than in mammals [39], [40]. Both diet composition [28] and axenic cultivation (this study) had no effect on the protein density of the flies. This important result is fully consistent with previous evidence that food consumption and nutrient allocation in *Drosophila* are regulated to maintain a certain target protein content [41]. Studies involving *Drosophila* reared on diets with lower protein content and protein:carbohydrate ratio than used in this study would be required to investigate the role of the microbiota in protein nutrition.

In conclusion, this study demonstrates that the nutritional phenotype of *Drosophila* is

strongly influenced by chronic infection with microorganisms that influence energy homeostasis and carbohydrate allocation patterns. These effects are predicted to both accompany and interact with signaling interactions between the microbiota and the host that are known to underpin normal development and cellular homeostasis, especially of the *Drosophila* gut [42]. Although the detail of the relationship between animals and their resident microbiota is anticipated to vary with host and symbiont taxa and environmental circumstances, the *Drosophila* association demonstrates the generality that a comprehensive explanation of the nutritional phenotype of animals requires understanding of the animal interactions with its microbiota.

References

1. Kau AL, Ahern PP, Griffin NW, Goodman AL, Gordon JI (2011) Human nutrition, the gut microbiome and the immune system. Nature 474: 327–336.

2. Diamant M, Blaak EE, de Vos WM (2011) Do nutrient-gut-microbiotainteractions play a role in human obesity, insulin resistance and type 2 diabetes? Obes Rev 12: 272–281.

3. Musso G, Gambino R, Cassader M (2010) Gut microbiota as a regulator ofenergy homeostasis and ectopic fat deposition: mechanisms and implications for metabolic disorders. Curr Opin Lipidol 21: 76–83.

4. Venema K (2010) Role of gut microbiota in the control of energy and carbohydrate metabolism. Curr Opin Clin Nutr Metab Care 13: 432–438.

5. Ley RE, Peterson DA, Gordon JI (2006) Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell 124: 837–848.

6. Vijay-Kumar M, Aitken JD, Carvalho FA, Cullender TC, Mwangi S, et al.(2010) Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. Science 328: 228–231.

7. Dethlefsen L, McFall-Ngai M, Relman DA (2007) An ecological and evolutionary perspective on human-microbe mutualism and disease. Nature 449: 811–818.

8. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, et al. (2009) Bacterial community variation in human body habitats across space and time.Science 326: 1694–1697.

9. Nemergut DR, Costello EK, Hamady M, Lozupone C, Jiang L, et al. (2011)Global patterns in the biogeography of bacterial taxa. Environ Microbiol 13: 135–144.

10. Faith JJ, Rey FE, O'Donnell D, Karlsson M, McNulty NP, et al. (2010) Creating and characterizing communities of human gut microbes in gnotobiotic mice. ISME J 4: 1094–1098.

11. Mahowald MA, Rey FE, Seedorf H, Turnbaugh PJ, Fulton RS, et al. (2009) Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. Proc Natl Acad Sci U S A 106: 5859–5864.

12. Marco ML, Peters TH, Bongers RS, Molenaar D, van Hemert S, et al. (2009) Lifestyle of Lactobacillus plantarum in the mouse caecum. Environ Microbiol 11: 2747–2757.

13. Dillon RJ, Dillon VM (2004) The gut bacteria of insects: nonpathogenic interactions. Annu Rev Entomol 49: 71–92.14. McFall-Ngai M (2007) Adaptive immunity: care for the community. Nature 445:153.

15. Morales-Jimenez J, Zuniga G, Villa-Tanaca L, Hernandez-Rodriguez C (2009) Bacterial community and nitrogen fixation in the red turpentine beetle, Dendroctonus valens LeConte (Coleoptera: Curculionidae: Scolytinae). Microb Ecol 58: 879–891.

16. Robinson CJ, Schloss P, Ramos Y, Raffa K, Handelsman J (2010) Robustness ofthe bacterial community in the cabbage white butterfly larval midgut. Microb Ecol 59: 199–211.

17. Ashburner M (1989) Drosophila, A Laboratory Handbook Cold Spring Harbor: Cold Spring Harbor Press.

18. Ren C, Webster P, Finkel SE, Tower J (2007) Increased internal and external bacterial load during Drosophila aging without life-span trade-off. Cell Metab 6:144–152.

19. Roh SW, Nam YD, Chang HW, Kim KH, Kim MS, et al. (2008) Phylogenetic characterization of two novel commensal bacteria involved with innate immune homeostasis in Drosophila melanogaster. Appl Environ Microbiol 74: 6171–6177.

20. Wong C-N, Ng P, Douglas AE (2011) Low diversity bacterial community in the gut of the fruitfly Drosophila melanogaster. Environmental Microbiology 13:1889–1900.

21. Brummel T, Ching A, Seroude L, Simon AF, Benzer S (2004) Drosophila lifespan enhancement by exogenous bacteria. Proc Natl Acad Sci U S A 101:12974–12979.

22. Corby-Harris V, Pontaroli AC, Shimkets LJ, Bennetzen JL, Habel KE, et al.(2007) Geographical distribution and diversity of bacteria associated with natural populations of Drosophila melanogaster. Appl Environ Microbiol 73: 3470–3479.

23. Cox CR, Gilmore MS (2007) Native microbial colonization of Drosophila melanogaster and its use as a model of Enterococcus faecalis pathogenesis. Infect Immun 75: 1565–1576.

24. Bakula M (1969) The persistence of a microbial flora during postembryogenesis of Drosophila melanogaster. J Invertebr Pathol 14: 365–374.

25. Nunney L (1990) Drosophila on oranges: colonization, competition, and coexistence. Ecology 71: 1904–1915.

26. Reed LK, Williams S, Springston M, Brown J, Freeman K, et al. (2010) Genotype-by-diet interactions drive metabolic phenotype variation in Drosophila melanogaster. Genetics 185: 1009–1019.

27. Ryu JH, Ha E-M, Oh C-T, Seol J-H, Brey PT, et al. (2006) An essential complementary role of NF-kB pathway to microbicidal oxidants in Drosophila gut immunity. EMBO Journal 25: 3693–3701.

28. Matzkin LM, Johnson S, Paight C, Bozinovic G, Markow TA (2011) Dietary protein and sugar differentially affect development and metabolic pools in ecologically diverse Drosophila. J Nutr 141: 1127–1133.

29. Brighenti F, Castellani G, Benini L, Casiraghi MC, Leopardi E, et al. (1995) Effect of neutralized and native vinegar on blood glucose and acetate responses to a mixed meal in healthy subjects. Eur J Clin Nutr 49: 242–247.

30. Ogawa N, Satsu H, Watanabe H, Fukaya M, Tsukamoto Y, et al. (2000) Acetic acid suppresses the increase in disaccharidase activity that occurs during culture of caco-2 cells. J Nutr 130: 507–513.

31. Johnston CS, Steplewska I, Long CA, Harris LN, Ryals RH (2010) Examination of the antiglycemic properties of vinegar in healthy adults. Ann Nutr Metab 56:74–79.

32. Ostman EM, Nilsson M, Elmstahl HGM, Molin G, Bjorck IME (2002) On the effect of lactic acid on blood glucose and insulin responses to cereal products: mechanistic studies in healthy subjects and in vitro. Journal of Cereal Science 36: 339–346.

33. Shin SC, Kim S-H, You H, Kim B, Kim AC, et al. (2011) Drosophila microbiome modulates host developmental and metabolic homeostasis via insulin signaling. Science 334: 670–674.

34. Storelli G, Defaye A, Erkosar B, Hols P, Royet J, et al. (2011) Lactobacillus plantarum promotes Drosophila systemic growth by modulating hormonal signals through TOR-dependent nutrient sensing. Cell Metabolism 14: 403–414.

35. Boggs CL (2009) Understanding insect life histories and senescence through a resource allocation lens. Functional Ecology 23: 27–37.

36. Teleman AA (2010) Molecular mechanisms of metabolic regulation by insulin in Drosophila. Biochemical Journal 425: 13–26.

37. Stump CS, Short KR, Bigelow ML, Schimke JM, Nair KS (2003) Effect of insulin on human skeletal muscle mitochondrial ATP production, protein synthesis, and mRNA transcripts. PNAS 100: 7998–8001.

38. Edgar BA (2006) How flies get their size: genetics meets physiology. Nat Rev Genet 7: 907–916.

39. Chapman RF (1999) The Insects: Structure and Function. Cambridge, UK: Cambridge University Press.

40. Zhang H, Liu J, Li CR, Momen B, Kohanski RA, et al. (2009) Deletion of Drosophila insulin-like peptides causes growth defects and metabolic abnormalities. Proc Natl Acad Sci U S A 106: 19617–19622.

41. Lee KP, Simpson SJ, Clissold FJ, Brooks R, Ballard JW, et al. (2008) Lifespan and reproduction in Drosophila: New insights from nutritional geometry. Proc Natl Acad Sci U S A 105: 2498–2503.

42. Buchon N, Broderick NA, Chakrabarti S, Lemaitre B (2009) Invasive and indigenous microbiota impact intestinal stem cell activity through multiple pathways in Drosophila. Genes Dev 23: 2333–2344.

43. Fukatsu T, Nikoh N (1998) Two intracellular symbiotic bacteria from the mulberry psyllid Anomoneura mori (Insecta, Homoptera). Appl Environ Microbiol 64: 3599–3606.

44. Kunin V, Hugenholtz P (2010) PyroTagger: a fast, accurate pipeline for analysisof rRNA amplicon pyrosequence data. Open J 1: 1–8.

APPENDIX B

MICROBE-DEPENDENT AND NONSPECIFIC EFFECTS OF PROCEDURES TO ELIMINATE THE RESIDENT MICROBIOTA FROM *DROSOPHILA MELANOGASTER*[5]

Abstract

Comparisons of animals bearing and lacking microorganisms can offer valuable insight into the interactions between animal hosts and their resident microbiota. Most hosts are naturally infected, and therefore, these comparisons require specific procedures (e.g., antibiotic treatment or physical exclusion of microorganisms) to disrupt the microbiota, but the potential for confounding nonspecific effects of the procedure on the traits of the host exists. Microbe-dependent and nonspecific effects can be discriminated by using multiple procedures: microbe-dependent effects are evident in hosts made microbe free by different procedures, but nonspecific effects are unique to individual procedures. As a demonstration, two procedures, oral administration of chlortetracycline (50 μg ml^{-1} diet) and microbiota removal by egg dechorionation, were applied to *Drosophila melanogaster* in a 2-by-2 factorial design. Microorganisms were undetectable in flies from dechorionated eggs and reduced by >99% in chlortetracycline-treated flies. *Drosophila* flies subjected to both protocols displayed an extended preadult development time, suggesting that the microbiota promotes the development rate. Female chlortetracycline-treated flies, whether from

[5] Presented with minor modifications from the originally published article:
Ridley, EV., Wong, ACN. and Douglas, AE. (2013). Microbe-dependent and non-specific effects of procedures to eliminate the resident microbiota from *Drosophila melanogaster*. *Applied and Environmental Microbiology*. Epub Mar 8.

untreated or dechorionated eggs, displayed reduced protein content and egg fecundity, which could be attributed to the nonspecific effect of the antibiotic. We recommend that procedures used to disrupt the microbiota of animals should be selected, following systematic analysis of alternative mechanistically distinct procedures, on the basis of two criteria: those that achieve the greatest reduction (ideally, elimination) of the microbiota and those that achieve minimal nonspecific effects.

Introduction

There is now overwhelming evidence that insects, like other animals, bear a substantial resident microbiota and that multiple aspects of the insect phenotype are strongly influenced by the activities of these microorganisms (1, 2). Resident microorganisms in the gut, cells, or specialized organs contribute to the nutrition of various insect groups, e.g., termites, various xylophagous beetles, tsetse flies and other blood feeders, and plant sap feeders, such as aphids and cicadas (3). Some microorganisms contribute to insect defense against natural enemies, often by the production of specific antibiotics or stimulation of the insect immune system (4–8). Other insect traits reported to be affected by the microbiota include dispersal behavior, insecticide resistance, food choice, thermal resistance, mate choice, virus vector competence, reproductive traits (including sex ratio), and body color (9–16). Experimentally generated microbe-free insects play a pivotal role in many studies investigating microbial effects on insect traits. Multiple methods are available to disrupt the microbiota of insects, including thermal treatment, antibiotic treatment, and mechanical exclusion (17, 18). Unfortunately, all these manipulations have the potential to cause nonspecific deleterious effects on the animal host. Very commonly,

a single procedure is applied without due consideration of these nonspecific effects, and this can result in spurious claims for microbial roles in animal function.

The purpose of this paper is to recommend and illustrate an experimental approach that aids discrimination of the microbiota-dependent and nonspecific effects of procedures that disrupt the microbiota. Specifically, it is recommended that two (or more) mechanistically distinct procedures be applied, with the expectation that microbiota-dependent effects are obtained by all the procedures but nonspecific effects are unique to individual procedures. Here, we describe the application of this experimental method to the fruit fly *Drosophila melanogaster* with two treatments that have been used in previous studies: dechorionation of *Drosophila* eggs with bleach (which eliminates surface microorganisms), followed by rearing on sterile food (19–22), and feeding of the insects with food supplemented with the antibiotic chlortetracycline (CT) (12, 23). CT and other tetracyclines are broad-spectrum antibiotics that inhibit bacterial protein synthesis (24), and they are widely used to disrupt the gut microbiota in various insects and other animals (18, 25, 26).

The experiments in this study determined the impact of dietary CT and egg dechorionation on the resident microbiota and the development time, fecundity, and nutritional status (protein and free glucose contents) of *D. melanogaster*. The experiments had a 2-by-2 factorial design, with antibiotic treatment and egg dechorionation being the factors. We applied this experimental design with the aim to discriminate the specific effects of microorganisms (where the response between the antibiotic treatments differed in flies derived from untreated eggs but not in those derived from dechorionated eggs) from the nonspecific effects of either procedure (where dechorionation or antibiotic treatment affected the trait of interest, yielding a

significant main factor in the analysis). We demonstrated that some effects of dietary CT can be explained to be a consequence of the effect on the gut microbiota and others can be explained to be a direct effect of the procedure on insect function.

Experimental procedures

Insect culture and manipulations

A *Wolbachia*-free line of *Drosophila melanogaster* strain CantonS was reared in sterile Falcon tubes (BD Biosciences, San Jose, CA) at 25°C with a 12-h light and 12-h dark cycle on an autoclaved diet containing 96 g glucose (Sigma), 48 g inactive dry Saccharomyces cerevisiae, and 14 g agar (both from Genesee Scientific) per liter. To generate the CT-supplemented (+CT) diet, a filtered solution of CT (Sigma) was dispensed at a 1/100 dilution into autoclaved food at 50°C and mixed thoroughly before the food solidified. The concentration (50 μg ml^{-1}) used in the +CT diet was selected by use of the criterion of the lowest concentration yielding a >90% reduction in the number of CFU from *Drosophila* homogenates (27). For egg dechorionation, eggs deposited overnight by mated females were washed in sterile water and then immersed in 10% sodium hypochlorite solution for 5 min, followed by two rinses in sterile water, and the eggs were then transferred to an autoclaved diet. All insect manipulations were conducted in a laminar-flow cabinet with aseptic technique. The experimental design was 2-by-2 factorial, with egg treatment (dechorionation or no treatment of eggs) and diet (CT-free diet and +CT diet) being the experimental factors. Each of the four treatments comprised 10 eggs in each of 10 replicate vials containing ca. 8 ml diet. The vials were monitored daily, and the time to development to adulthood was scored. To quantify the protein and glucose contents of the flies, at 7

to 10 days of age after eclosion to adulthood, individual flies were homogenized in 80 µl ice-cold buffer comprising 10 mM Tris, 1 mM EDTA, pH 8.0, and 0.1% (vol/vol) Triton X-100 and centrifuged at 7,000 × g at 4°C for 1 min. The protein content of the supernatant was determined by the Coomassie brilliant blue microassay method (500-0201; Bio-Rad) with bovine serum albumin as the standard (40 to 480 mg protein ml^{-1}). The glucose assay kit of Sigma (GAGO20) was used for glucose assays.

To administer *Drosophila* microbiota to flies, vials (diameter, 0.9 in.) of sterile diet were pretreated with 40 adult males for 24 h. The deposited feces were washed from each vial with 500 µl sterile phosphate-buffered saline, and 50-µl fecal washings were added to each test diet. The fecal washings contained viable bacteria, including *Acetobacter* and *Lactobacillus* species, which dominate the gut microbiota (28), as revealed by plating onto nutrient agar (as below).

Identification and quantification of bacteria

The culturable bacterial load per insect was assessed by a previously described method (23). Ten replicate 7- to-10-day-old adult flies were individually hand homogenized in 250 µl sterile phosphate buffer (pH 7.4) until pieces of tissue were no longer visible. Homogenate samples (100 µl) in a 10-fold dilution series from 1× to 1/1,000× were spread onto nutrient agar plates (28 g $liter^{-1}$; Oxoid) using sterile technique, and the number of CFU was scored after 7 days at 25°C. Colonies were sampled for identification by Sanger sequencing of 16S rRNA gene sequences. Briefly, 16S rRNA gene sequences were amplified from DNA extracted from single colonies by PCR with general primers 16SA1 (5′-AGAGTTTGATCMTGGCTCAG-3′) and 16SB1 (5′-TACGGYTACCTTGTTACGACTT-3′) (29) by a previously described procedure

(30). The PCR products were purified using a QIAquick PCR purification kit (Qiagen, Valencia, CA) and sequenced on an Applied Biosystems automated 3730 DNA analyzer using BigDye Terminator chemistry and AmpliTaq-FS DNA polymerase. Sequences were trimmed using Sequencher (version 4.10.1) software and identified by NCBI nucleotide BLAST analysis.

Statistical analysis

Analysis of variance (ANOVA) was applied to data sets that conformed to normal distributions with homogeneity of variances (as determined by the Anderson-Darling test and Levene's test, respectively). Two-sample comparisons were conducted by the t test for normally distributed data sets or by the Mann-Whitney U test. Where multiple tests were conducted in parallel, the Bonferroni correction of the critical probability (P = 0.05) was applied.

Nucleotide sequence accession numbers

The sequences of the bacteria recovered from the flies were deposited in the GenBank database with accession numbers KC485818 to KC485880 (Table 5.1).

Results

Impact of CT and egg dechorionation on bacterial complement of Drosophila

The first experiments tested for the presence of bacteria in *Drosophila*. The flies derived from untreated eggs on a CT-free diet yielded 3.2 × 104 CFU per fly (median; range, 520 to 2.9 × 105 CFU per fly; n = 10). The equivalent value for flies from the

+CT diet was 118 CFU per fly (range, 5 to 9 × 103 CFU per fly; n = 10), demonstrating that, on average, >99% of the culturable bacteria were eliminated from flies reared on a +CT diet. A subset of bacterial colonies was sampled for identification by 16S rRNA gene sequencing. The most abundant bacteria were *Acetobacter* (Alphaproteobacteria), accounting for 97% and 80% of the colonies from untreated and CT-treated flies, respectively (Table 5.1). A parallel pyrosequencing analysis of PCR-generated 16S rRNA gene amplicons from flies reared on the CT-free diet yielded only *Acetobacter* species, with *Acetobacter* cerevisiae accounting for 98% of the 29,858 reads (data not shown), indicating that the bacterial community in CT-treated flies was drastically depleted and not dominated by unculturable forms.

The great majority of the culturable bacteria in flies reared on the +CT diet were susceptible to CT, as indicated by the very limited recovery of CFU from parallel fly samples reared on plates supplemented with 50 μg CT ml^{-1} (7/10 flies yielded no CFU, and the remaining 3 flies yielded 38, 260, and 420 CFU, respectively, giving a median number of CFU per fly of 0).

Every fly tested that developed from dechorionated eggs yielded no bacterial colonies on nutrient agar plates. Parallel PCR assays with general bacterial 16S rRNA gene primers also yielded no product, indicating that dechorionation eliminates all bacteria.

Table 5.1 Composition of resident microbiota in *D. melanogaster* reared on a CT-free diet and a +CT diet, determined by 16S rRNA gene sequence analysis of bacterial colonies cultured on nutrient agar.

NCBI accession number(s) of sequences	Sequence identity			Number of bacterial colonies	
	NCBI accession number	Taxonomic identity	% sequence identity	CT-free diet (n= 33)	+CT-diet (n=30)
KC485830-KC485840, KC485851-KC485862	GQ359863.1	*Acetobacter* sp. 6-C-2 16S ribosomal RNA gene, partial sequence	98-99	11	12
KC485818-KC485819	NR_025512.1	*Acetobacter cerevisiae* strain LMG 1625 16S ribosomal RNA, partial sequence	97-99	2	0
KC485872	HM218620.1	*Acetobacter malorum* strain NM156-4 16S ribosomal RNA gene, partial sequence	98-99	0	1
KC485841-KC485842, KC485871	FJ227313.1	*Acetobacter pasteurianus* strain bh12 16S ribosomal RNA gene, partial sequence	96-98	2	1
KC485843-KC485849, KC485873-KC485874	FN429068.1	*Acetobacter pasteurianus* partial 16S ribosomal RNA gene, strain SX461	96-99	7	2
KC485850	FN429074.1	*Acetobacter pasteurianus* partial 16S ribosomal RNA gene, strain ZJ362	97	1	0
KC485821-KC485829, KC485863-KC485870	EU096229.1	*Acetobacter pomorum* strain EW816 16S ribosomal RNA gene, partial sequence	97-99	9	8
KC485878	GU369767.1	*Lactobacillus brevis* strain JS-7-2 16S ribosomal RNA gene, partial sequence	99	0	1
KC485879	FJ227317.1	*Lactobacillus brevis* strain b4 16S ribosomal RNA gene, partial sequence	98	0	1
KC485875, KC485876	GU253891.1	*Lactobacillus pentosus* strain N3 16S ribosomal RNA gene, partial sequence	98	0	2
KC485880	AB494721.1	*Lactobacillus plantarum* gene for 16S ribosomal RNA, partial sequence,	99	0	1

KC485877	HM449702.1	strain: KL23 *Micrococcus luteus* strain PCSB6 16S ribosomal RNA gene, partial sequence	97	0	1
KC485820	DQ981281.1	Uncultured bacterium clone thom_c06 16S ribosomal RNA gene, partial sequence	98	1	0

Fitness indices of Drosophila

The two indices of fitness assayed yielded different patterns of response to CT and egg dechorionation. The time of insect development from oviposition to adulthood was 11 to 17 days (Figure. 5.1) and varied significantly across the four insect groups (Kruskal-Wallis test, $H = 125.77$, $P < 0.001$). As previously reported (31), dechorionation of the eggs resulted in a significantly extended time of development to adulthood relative to that for untreated eggs on the CT-free diet (Mann-Whitney test, $W = 2336.5$, $P < 0.001$). The development time on the +CT diet was also significantly prolonged relative to that on the CT-free diet for insects derived from untreated eggs (median, 12 days versus 11 days; $W = 3081$; $P < 0.001$) but not for insects derived from dechorionated eggs (median, 13 days for both treatments; $W = 5988.5$; $P = 0.75$) (Figure. 5.1). These results are consistent with the interpretation from previous studies (19, 20, 31, 32) that the microbiota increases the rate of *Drosophila* development.

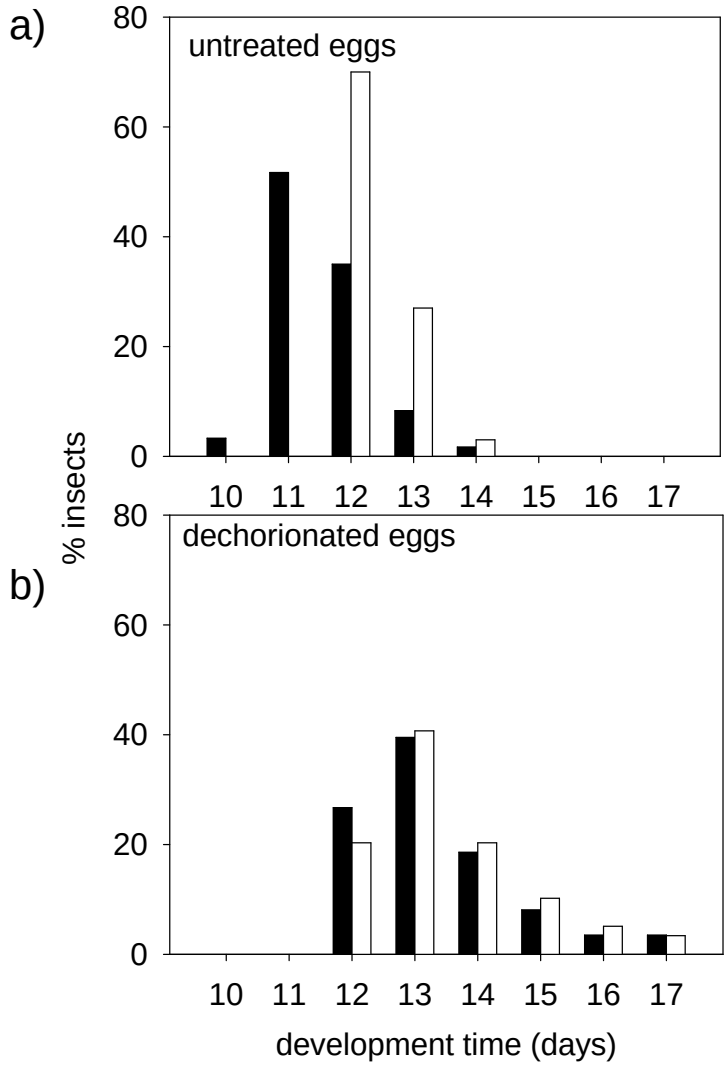


Figure 5.1 Development time of flies from oviposition to adulthood. Closed bars, CT-free diet; open bars, +CT diet. Number of replicates: 60 on CT-free diet and 100 on the +CT diet (a) and 86 on the CT-free diet and 59 on the +CT diet (b).

The fecundity of flies derived from both untreated and dechorionated eggs was negatively affected by dietary CT, with 40% fewer eggs being deposited by flies on the +CT diet than by those on the CT-free diet. In the ANOVA (Figure. 5.2), the interaction term was not statistically significant, indicating that the negative effect of

CT on fecundity cannot be explained by the elimination of microbiota and is likely a consequence of the direct effect of the antibiotic on the insect.

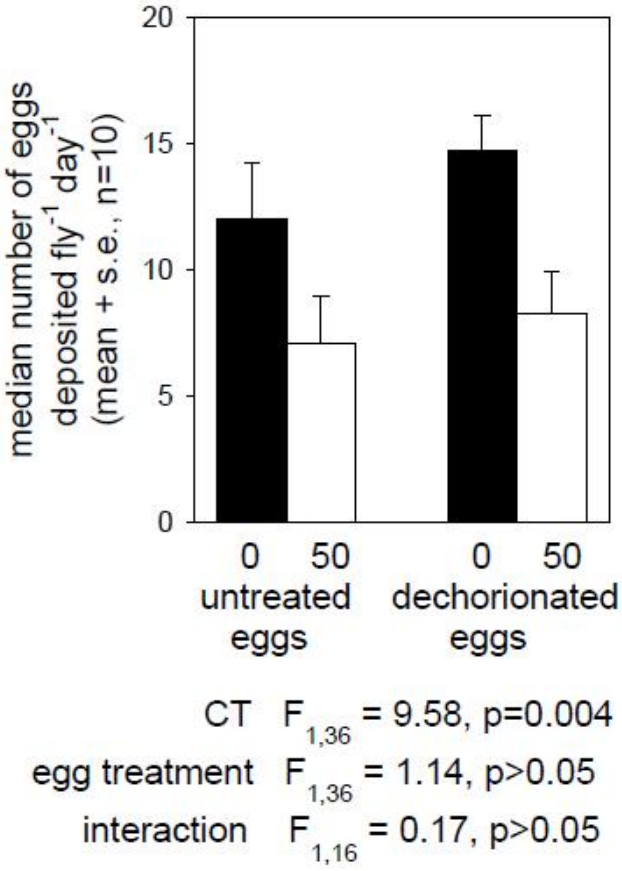


Figure 5.2 Median number of eggs deposited by 10 replicate flies over 7 days from days 3 to 10 post-eclosion. Closed bars, CT-free diet; open bars, +CT diet. s.e., standard error. ANOVA results were as follows: for CT treatment, F1,36 = 9.58, P = 0.004; for egg treatment, F1,36 = 1.14, P > 0.05; for interaction, F1,36 = 0.17, P > 0.05. The x-axis values indicate the CT concentration in μg ml^{-1}.

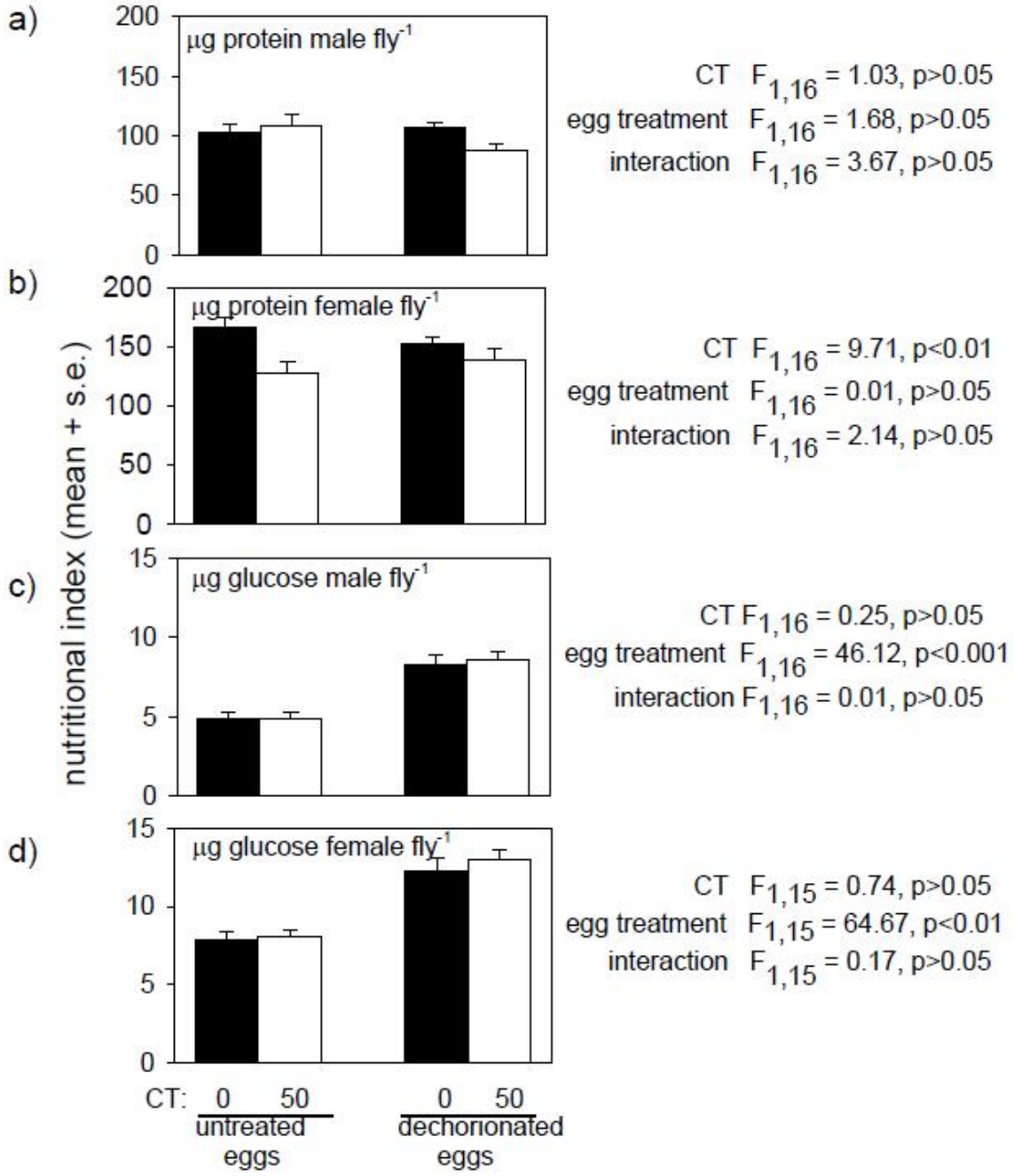
Nutritional and metabolic indices of Drosophila

It has previously been shown that *Drosophila* flies derived from dechorionated eggs have an elevated glucose content, but their protein content is comparable to that in

untreated flies (31). In this study, we investigated how these nutritional indices responded to CT treatment. (The protein and glucose contents of flies reared on a CT-free diet contributing to this analysis have been published previously [31].)

The protein content of males did not differ significantly between flies reared on the CT-free diet and those reared on the +CT diet, but that of females was reduced by 17% when they were on the +CT diet, independently of the egg treatment, and this effect was statistically significant (Figure. 5.3a and b). As with fecundity (see above), these data are indicative of a direct effect of the antibiotic on the female fly. To check whether the differential effect of CT on the protein content of the two sexes was concentration dependent, males were reared on a diet containing 300 μg CT ml^{-1}. The protein content of these flies (104 ± 6.8 μg per fly, mean ± standard error, 5 replicates) also did not differ significantly from that of flies reared on a CT-free diet (103 ± 7.3 μg per fly, 5 replicates) ($t7 = 0.887$, $P > 0.05$).

Figure 5.3 Nutritional indices of 5- to 7-day-old adult *Drosophila* flies derived from untreated and dechorionated eggs and reared on a CT-free diet (closed bars) or a +CT diet (open bars). Five replicates per treatment (except for 4 replicates for the glucose content of females from dechorionated eggs on the CT-free diet). The critical probability was 0.0125 after use of the Bonferroni correction for four tests.


a)
µg protein male fly-1
200
150
100
50
0
CT F1,16 = 1.03, p>0.05
egg treatment F1,16 = 1.68, p>0.05
interaction F1,16 = 3.67, p>0.05
b)
µg protein female fly-1
200
150
100
50
0
CT F1,16 = 9.71, p<0.01
egg treatment F1,16 = 0.01, p>0.05
interaction F1,16 = 2.14, p>0.05
nutritional index (mean + s.e.)
c)
µg glucose male fly-1
15
10
5
0
CT F1,16 = 0.25, p>0.05
egg treatment F1,16 = 46.12, p<0.001
interaction F1,16 = 0.01, p>0.05
d)
µg glucose female fly-1
15
10
5
0
CT F1,15 = 0.74, p>0.05
egg treatment F1,15 = 64.67, p<0.01
interaction F1,15 = 0.17, p>0.05
CT: 0 50 0 50
untreated eggs
dechorionated eggs

The glucose content of both male and female flies reared from untreated eggs on the +CT diet did not differ significantly from that of the equivalent flies reared on the CT-free diet and was significantly lower than that of flies derived from dechorionated eggs (Figure. 5.3c and d). These data are open to two alternative interpretations: (i) the elevated glucose content of flies from dechorionated eggs is a nonspecific effect of the egg treatment, or (ii) the small numbers of bacteria associated with the flies on the +CT diet is sufficient to reduce the glucose content to values comparable to those for conventionally reared flies. To discriminate between these possibilities, dechorionated eggs were transferred to a sterile diet supplemented with *Drosophila* feces, which contain live bacteria. This treatment resulted in a significant reduction in the glucose content of the flies for both males and females to levels that did not differ significantly from those for flies derived from untreated eggs (Figure. 5.4).

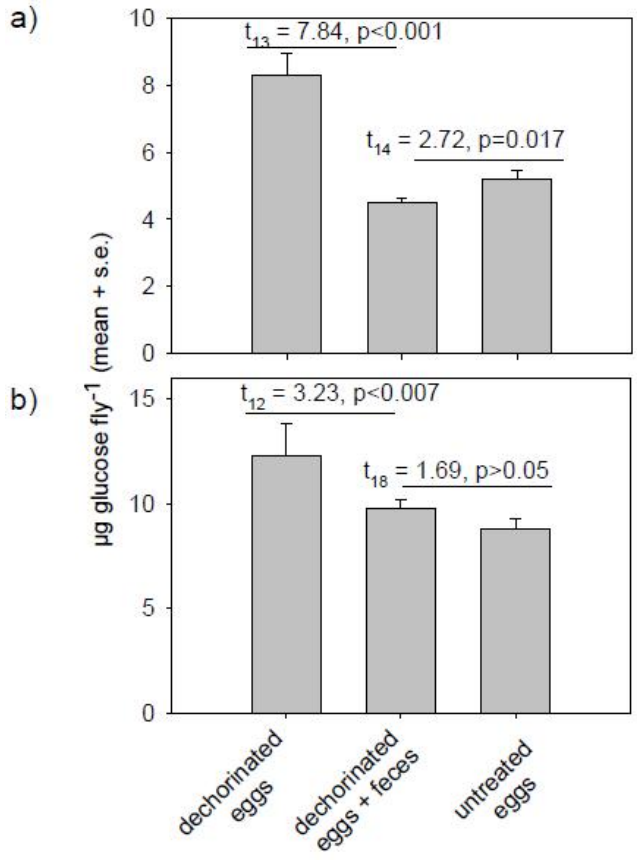


Figure 5.4 Glucose content of *Drosophila* flies derived from dechorionated and untreated eggs. (Top) Males; (bottom) females. Ten replicates were used for untreated eggs and dechorionated eggs plus feces; 5 replicates were used for dechorionated eggs. The critical probability was 0.0125 after use of the Bonferroni correction for four t tests.

Discussion

Any intervention to disrupt the resident microbiota of insects has the potential to cause nonspecific effects, and interpretation of results is critically dependent on discrimination between these nonspecific effects and effects attributable to the microbiota. This study demonstrates how the application of two mechanistically

different procedures in a factorial design can be useful to make this discrimination. Insect eggs can be sensitive to physical manipulations, especially removal of the egg shell (17). Nevertheless, multiple lines of evidence indicate that dechorionation of *Drosophila* eggs has no discernible nonspecific effect on preadult development. Specifically, untreated and dechorionated eggs develop to hatching at the same rates (31); the extended development time of insects derived from dechorionated eggs was also displayed by flies reared from untreated eggs on a +CT diet, in which the microbiota was greatly depleted (Figure. 5.1), and administration of bacteria to insects derived from dechorionated eggs rescued both the low glucose content (this study) and the rapid development rate (31) of untreated insects. Intriguingly, the glucose levels, but not the development rates, in CT-treated flies with a much-depleted bacterial content were comparable to those in untreated flies, suggesting that these two indices differ in their responsiveness to the abundance of bacteria. Further research is required to investigate the basis of this effect.

The application of two mechanistically different methods to disrupt the microbiota has revealed that CT treatment is a less satisfactory method than egg dechorionation for elimination of the microbiota of *Drosophila*. Importantly, direct effects of the antibiotic treatment on the insect are obtained at a concentration (50 μg ml^{-1}) that fails to eliminate the bacteria, indicating that no CT concentration would achieve bacterial elimination without side effects. Although greater bacterial depletion can be achieved by the use of CT concentrations higher than those used in this study (27), treatment with tetracycline antibiotics at concentrations of ≥100 μg CT ml^{-1} diet is well-known to have substantial and transgenerational effects on mitochondria (33), particularly affecting systems strongly dependent on mitochondrial function, e.g., embryo

development (17) and sperm viability (34). For this reason, the use of protocols using antibiotics to eliminate specific bacteria, e.g., *Wolbachia*, delays the ability to perform experiments on insects for multiple generations after antibiotic treatment (33). It has been assumed that the nonspecific deleterious effects of tetracyclines and other antibiotics are insignificant at concentrations of ≤100 μg ml^{-1}, and 50 μg ml^{-1} is widely used to remove bacteria from insects used for study within a single generation (18, 35). This study demonstrates that, for *Drosophila*, this supposition is invalid. Both protein content and fecundity are significantly depressed in female *Drosophila* flies feeding on a +CT diet (50 μg ml^{-1}) relative to that in flies derived from untreated and dechorionated eggs on a CT-free diet.

A related issue is the physiological condition of the residual bacteria in the CT-treated flies. The culturable bacteria were largely CT susceptible, as revealed by the minimal growth of bacteria from CT-treated flies on CT-supplemented medium. Taken with other data indicating that the bacteria associated with *Drosophila* are generally culturable (23), these data suggest that bacterial protein synthesis and linked processes (metabolism, growth, division, etc.) are largely inactive in flies on a +CT diet. They contrast with data for some insects which are known to bear antibiotic-resistant bacteria (36–39), whose interactions with the insect host would presumably be unaffected by the antibiotic treatment.

These considerations lead to two methodological recommendations. First, dechorionation is preferable over CT treatment for elimination of the gut microbiota of *Drosophila*, because CT has microbe-independent deleterious effects on *Drosophila* function at concentrations that are insufficient to achieve complete bacterial elimination. Second, control experiments with insects derived from dechorionated

eggs should be conducted in studies where other antibiotics, e.g., erythromycin and rifampin (40, 41), are used to manipulate the microbiota of *Drosophila*. The microbe-independent effects of tetracycline and possibly other antibiotics used to disrupt the microbiota likely apply to other insects, and the tolerance of insect eggs to dechorionation procedures may vary among insect species. The factorial design (e.g., antibiotic and dechorionation) used in this study has general value to tease apart the microbe-dependent and microbe-independent effects of treatments to eliminate the microbiota.

An important caveat to these considerations relates to *Wolbachia*, present in 20 to 70% of all insect species (21, 42), including many laboratory lines and field isolates of *D. melanogaster* (43–46); the strain of *D. melanogaster* used in this study was *Wolbachia* free. As well as being a reproductive parasite (16), *Wolbachia* can confer virus resistance and nutritional benefits (47–49). *Wolbachia* can be eliminated by antibiotics but not surface sterilization/dechorionation of eggs, because it is transmitted vertically in the egg cytoplasm (16). Although the interactions between the gut microbiota and *Wolbachia* have received little study, elimination of the gut microbiota and the resultant changes in the insect signaling networks and immune function (5, 20, 32) could lead to changes to the population size, tissue tropism, and activities of *Wolbachia*. Consequently, differences between *Wolbachia*-positive insects derived from untreated and dechorionated eggs may be caused by gut microbe-dependent effects on both the insect and *Wolbachia*.

In conclusion, experimentally generated insects in which the microbiota is depleted or eliminated offer a vitally important tool to investigate insect-microbe interactions. Access to alternative methods with different modes of action is particularly valuable,

to provide independent confirmation of proposed interactions, to identify microbe-independent effects (e.g., depressed protein content of CT-treated *Drosophila* flies), and to select the most appropriate method to manipulate the microbiota for different purposes.

References

1. Ferrari J, Vavre F. 2011. Bacterial symbionts in insects or the story of communities affecting communities. Philos. Trans. R. Soc. Lond. B Biol. Sci. 366:1389–1400.

2. Feldhaar H. 2011. Bacterial symbionts as mediators of ecologically important traits of insect hosts. Ecol. Entomol. 36:533–543.

3. Douglas AE. 2009. The microbial dimension in insect nutritional ecology. Funct. Ecol. 23:38–47.

4. Haine ER. 2008. Symbiont-mediated protection. Proc. Biol. Sci. 275:353–361.

5. Jaenike J, Unckless R, Cockburn SN, Boelio LM, Perlman SJ. 2010. Adaptation via symbiosis: recent spread of a Drosophila defensive symbiont. Science 329:212–215.

6. Oliver KM, Moran NA, Hunter MS. 2005. Variation in resistance to parasitism in aphids is due to symbionts not host genotype. Proc. Natl. Acad. Sci. U. S. *A.* 102:12795–12800.

7. Kroiss J, Kaltenpoth M, Schneider B, Schwinger MG, Hertweck C, Muaddula RK, Strohm E, Svatos *A.* 2010. Symbiotic Streptomycetes provide antibiotic combination prophylaxis for wasp offspring. Nat. Chem. Biol. 6:261–263.

8. Giordano R, Weber E, Waite J, Bencivenga N, Krogh PH. 2010. Effect of a high dose of three antibiotics on the reproduction of a parthenogenetic strain of Folsomia candida (Isotomidae: Collembola). Env. Entomol. 39:1170–1177.

9. Hosokawa T, Kikuchi Y, Shimada M, Fukatsu T. 2007. Obligate symbiont involved in pest status of host insect. Proc. Biol. Sci. 274:1979–1984.

10. Tsuchida T, Koga R, Horikawa M, Tsunoda T, Maoka T, Matsumoto S, Simon JC, Fukatsu T. 2010. Symbiotic bacterium modifies aphid body color. Science 330:1102–1104.

11. Kikuchi Y, Hayatsu M, Hosokawa T, Nagayama A, Tago K, Fukatsu T. 2012. Symbiont-mediated insecticide resistance. Proc. Natl. Acad. Sci. U. S. *A.* 109:8618–8622.

12. Sharon G, Segal D, Ringo JM, 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.

13. Goodacre SL, Martin OY, Bonte D, Hutchings L, Woolley C, Ibrahim K, Thomas C, Hewitt GM. 2009. Microbial modification of host long-distance dispersal capacity. BMC Biol. 7:32. doi:10.1186/1741-7007-7-32.

14. Gottlieb Y, Zchori-Fein E, Mozes-Daube N, Kontsedalov S, Skaljac M, Brumin M, Sobol I, Czosnek H, Vavre F, Fleury F, Ghanim M. 2010. The transmission

efficiency of tomato yellow leaf curl virus by the whitefly Bemisia tabaci is correlated with the presence of a specific symbiotic bacterium species. J. Virol. 84:9310–9317.

15. Wernegreen JJ. 2012. Mutualism meltdown in insects: bacteria constrain thermal adaptation. Curr. Opin. Microbiol. 15:255–262.

16. O'Neill SL, Hoffmann AA, Werren JH (ed). 1997. Influential passengers—inherited microorganisms and arthropod reproduction. Oxford University Press, Oxford, United Kingdom.

17. Douglas AE. 1989. Mycetocyte symbiosis in insects. Biol. Rev. 64:409–434.

18. Wilkinson T*L*. 1998. The elimination of intracellular microorganisms from insects: an analysis of antibiotic treatment in the pea aphid (Acyrthosiphon pisum). Comp. Biochem. Physiol. 119A:871–881.

19. Bakula M. 1969. The persistence of a microbial flora during postembryogenesis of Drosophila melanogaster. J. Invertebr. Pathol. 14:365–374.

20. Shin SC, Kim SH, You H, Kim B, Kim AC, Lee KA, Yoon JH, Ryu JH, Lee WJ. 2011. Drosophila microbiome modulates host developmental and metabolic homeostasis via insulin signaling. Science 334:670–674.

21. Werren JH, Windsor DM, Guo *L*. 1995. Distribution of *Wolbachia* among neotropical arthropods. Proc. Biol. Sci. 262:197–204.

22. Brummel T, Ching A, Seroude L, Simon AF, Benzer S. 2004. Drosophila lifespan enhancement by exogenous bacteria. Proc. Natl. Acad. Sci. U. S. *A.* 101:12974–12979.

23. Ren C, Webster P, Finkel SE, Tower J. 2007. Increased internal and external bacterial load during Drosophila aging without life-span trade-off. Cell Metab. 6:144–152.

24. Chopra I, Roberts M. 2001. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. Microbiol. Mol. Biol. Rev. 65:232–260.

25. Barr KL, Hearne LB, Briesacher S, Clark TL, Davis GE. 2010. Microbial symbionts in insects influence down-regulation of defense genes in maize. PLoS One 5:e11339. doi:10.1371/journal.pone.0011339.

26. Douglas AE. 1998. Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria Buchnera. Annu. Rev. Entomol. 43:17–37.

27. Ridley EV. 2011. The impact of chlortetracycline on Drosophila melanogaster and Aedes aegypti. Ph.*D.* thesis. University of York, York, United Kingdom.

28. Wong C-N, Ng P, Douglas AE. 2011. Low diversity bacterial community in the

gut of the fruitfly Drosophila melanogaster. Environ. Microbiol. 13:1889–1900.

29. Fukatsu T, Nikoh N. 1998. Two intracellular symbiotic bacteria from the mulberry psyllid Anomoneura mori (Insecta, Homoptera). Appl. Environ. Microbiol. 64:3599–3606.

30. Douglas AE, Francois CLMJ, Minto LB. 2006. Facultative 'secondary' bacterial symbionts and the nutrition of the pea aphid, Acyrthosiphon pisum. Physiol. Entomol. 31:262–269.

31. Ridley EV, Wong AC, Westmiller S, Douglas AE. 2012. Impact of the resident microbiota on the nutritional phenotype of Drosophila melanogaster. PLoS One 7:e36765. doi:10.1371/journal.pone.0036765.

32. 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.

33. Ballard JW, Melvin RG. 2007. Tetracycline treatment influences mitochondrial metabolism and mtDNA density two generations after treatment in Drosophila. Insect Mol. Biol. 16:799–802.

34. Zeh JA, Bonilla MM, Adrian AJ, Mesfin S, Zeh DW. 2012. From father to son: transgenerational effect of tetracycline on sperm viability. Sci. Rep. 2:375. doi:10.1038/srep00375.

35. Ahmed MZ, Ren SX, Xue X, Li XX, Jin GH, Qiu B*L*. 2010. Prevalence of endosymbionts in Bemisia tabaci populations and their in vivo sensitivity to antibiotics. Curr. Microbiol. 61:322–328.

36. Channaiah LH, Subramanyam B, McKinney LJ, Zurek *L*. 2010. Stored-product insects carry antibiotic-resistant and potentially virulent enterococci. FEMS Microbiol. Ecol. 74:464–471.

37. Macovei L, Zurek *L*. 2006. Ecology of antibiotic resistance genes: characterization of enterococci from houseflies collected in food settings. Appl. Environ. Microbiol. 72:4028–4035.

38. Allen HK, Cloud-Hansen KA, Wolinski JM, Guan C, Greene S, Lu S, Beyink M, Broderick NA, Raffa KF, Handelsman J. 2009. Resident microbiota of the gypsy moth midgut harbors antibiotic resistance determinants. DNA Cell Biol. 28:109–117.

39. Kadavy DR, Hornby JM, Haverkost T, Nickerson KW. 2000. Natural antibiotic resistance of bacteria isolated from larvae of the oil fly, Helaeomyia petrolei. Appl. Environ. Microbiol. 66:4615–4619.

40. Blatch SA, Meyer KW, Harrison JF. 2010. Effects of dietary folic acid level and symbiotic folate production on fitness and development in the fruit fly Drosophila

melanogaster. Fly 4:1–8.

41. Cox CR, Gilmore MS. 2007. Native microbial colonization of Drosophila melanogaster and its use as a model of Enterococcus faecalis pathogenesis. Infect. Immun. 75:1565–1576.

42. Jeyaprakash A, Hoy M*A.* 2000. Long PCR improves *Wolbachia* DNA amplification: wsp sequences found in 76% of sixty-three arthropod species. Insect Mol. Biol. 9:393–405.

43. Werren JH, Baldo L, Clark ME. 2008. *Wolbachia*: master manipulators of invertebrate biology. Nat. Rev. Microbiol. 6:741–751.

44. Corby-Harris V, Pontaroli AC, Shimkets LJ, Bennetzen JL, Habel KE, Promislow DE. 2007. Geographical distribution and diversity of bacteria associated with natural populations of Drosophila melanogaster. Appl. Environ. Microbiol. 73:3470–3479.

45. Clark ME, Anderson CL, Cande J, Karr T*L*. 2005. Widespread prevalence of *Wolbachia* in laboratory stocks and the implications for Drosophila research. Genetics 170:1667–1675.

46. Chandler SM, Wilkinson TL, Douglas AE. 2008. Impact of plant nutrients on the relationship between a herbivorous insect and its symbiotic bacteria. Proc. Biol. Sci. 275:565–570.

47. Hedges LM, Brownlie JC, O'Neill SL, Johnson KN. 2008. *Wolbachia* and virus protection in insects. Science 322:702.

48. Brownlie JC, Cass BN, Riegler M, Witsenburg JJ, Iturbe-Ormaetxe I, McGraw EA, O'Neill S*L*. 2009. Evidence for metabolic provisioning by a common invertebrate endosymbiont, *Wolbachia* pipientis, during periods of nutritional stress. PLoS Pathog. 5:e1000368. doi:10.1371/journal.ppat.1000368.

49. Bian G, Xu Y, Lu P, Xie Y, Xi Z. 2010. The endosymbiotic bacterium *Wolbachia* induces resistance to dengue virus in Aedes aegypti. PLoS Pathog. 6:e1000833. doi:10.1371/journal.ppat.1000833.

APPENDIX C

RESIDENT MICROBIOTA OF EXCEPTIONALLY LOW DIVERSITY IN A PLANT SAP FEEDING INSECT[6]

Abstract

The resident microbiota of animals represents an important contribution to the global microbial diversity, but it is poorly studied in most animals other than humans and a few invertebrates. This study investigated the bacterial communities in 7 species of the whitefly *Bemisia tabaci* complex by pyrosequencing bacterial 16S rRNA gene amplicons. Representatives of just 9 bacterial genera were identified, with 7 previously described vertically-transmitted taxa accounting for >99.9-100% of the reads in each sample. The number of OTUs tended to saturation in every biological sample and, after correction for sequencing error, each sample was concluded to bear 3-5 bacterial taxa, with low diversity (Shannon's index: 0.36-1.46, Simpson's index: 0.17-0.74). The abundance of the bacterial symbionts in *B. tabaci* MEAM1 was elevated in insect cultures bearing begomoviruses (tomato mosaic virus and tomato yellow leaf curl virus) relative to a non-viruliferous culture, and varied significantly with rearing plant. Generally, the abundance of the different symbionts varied in concert, suggesting that they were regulated by common or linked insect mechanism(s). These host controls over bacterial abundance, together with the frequencies of vertical and horizontal transmission and the fitness of insects with different bacterial complements, were identified as candidate factors contributing to

[6] Article revised for resubmission to *Molecular Ecology* by Jing, XF*., Wong, ACN*., Colvin, J., Mckenzie, C. and Douglas, AE. *co-first authors.

Supplementary materials available upon request.

the variation in composition and diversity of the bacterial communities in *B. tabaci*. The demonstration of a very low bacterial diversity in this insect contributes to an emerging pattern of lower bacterial diversity in many invertebrates than in vertebrate animals, although the factors shaping this pattern remain to be established.

Introduction

All macroorganisms, including animals and plants, are constantly associated with microorganisms; but only a restricted range of microorganisms exploit these hosts. For example, among the bacteria, members of just 11 of the ca. 52 phyla of Eubacteria are known to be associated with hosts (Sachs *et al*. 2011), and the Archaea include just a few genera of mutualists (Gill & Brinkman 2011). Furthermore, many host-associated microorganisms are rare or unknown in the free-living condition, i.e. apart from a host (Ley *et al*. 2008; Tamames *et al*. 2010) and can exhibit high rates of evolutionary diversification, linked to their coevolutionary interactions with their hosts (Brucker & Bordenstein 2012; Dethlefsen *et al*. 2007; Walter *et al*. 2011). As a result, animals and plants are increasingly recognized as habitats for an important fraction of the global microbial diversity.

Our appreciation of the diversity of host-associated microorganisms, especially in animals, has been transformed by high throughput sequencing technologies which facilitate the identification and enumeration of the many uncultured microbial taxa. To date, most research has focused on bacteria, revealing different patterns of diversity between vertebrates and most invertebrate animals. In particular, the guts of vertebrates are generally dominated by members of two phyla, the Bacteroidetes and Firmicutes, but include hundreds of bacterial taxa at the level of $\leq$97% sequence

identity of specific variable (V) regions of the rRNA gene sequences (Costello *et al*. 2010; Hong *et al*. 2011; Human Microbiome Project 2012; Kohl 2012; McDonald *et al*. 2012; Rawls *et al*. 2006) but the guts or whole bodies of many invertebrates, especially some insects, harbor just tens of taxa, often comprising members of multiple phyla (Li *et al*. 2012; Martinson *et al*. 2011; Robinson *et al*. 2010; Wong *et al*. 2011), with certain taxa apparently restricted to particular invertebrate groups (Fieseler *et al*. 2004). Although possible factors contributing to this striking pattern of diversity have been proposed (Ley *et al*. 2008; McFall-Ngai 2007), a comprehensive understanding of the underlying processes remains as a major challenge for the discipline of microbial ecology.

Explanations of the diversity of animal-associated microorganisms are only as good as the datasets on which they are based. Unfortunately, the number and scale of analyses of host-associated microorganisms are far smaller in invertebrates than vertebrates, especially humans and mice, with the implication that some (perhaps many) unstudied invertebrate taxa may not fit to the pattern identified to date. It is, therefore, an important priority to obtain detailed data on the microbiota in key invertebrate taxa, including those believed to bear a microbiota of very low diversities.

This study concerns the bacterial diversity in an invertebrate group anecdotally cited to bear a microbiota of exceptionally low diversity: the phloem sap feeders. The capacity to utilize this diet through the life cycle has evolved multiple times in the order Hemiptera, but is otherwise unknown across the entire animal kingdom (Dolling 1991; Douglas 2003). Furthermore, this feeding trait is correlated absolutely with the possession of microorganisms, usually bacteria that are localized to specialized insect cells and are obligately vertically transmitted (Buchner 1965). These bacteria are

known as primary symbionts. In addition to the primary symbiont, many phloem feeding insects bear one to several other bacteria, informally called secondary symbionts, which may be localized to the bacteriocytes, other insect cells or the body cavity, and are capable of both vertical and horizontal transmission (Buchner 1965; Degnan *et al*. 2010; Ferrari &Vavre 2011). All primary and most secondary symbionts have reduced genomes and are nutritionally fastidious, and most have not been brought into culture (Burke & Moran 2011; McCutcheon & Moran 2012). The widespread, informal belief that the bacterial diversity in these insects is dominated by these primary and secondary symbionts is based on the finding that these insects generally harbor few or no bacteria culturable on routine bacteriological media (Davidson *et al*. 2000; Grenier *et al*. 1994); and 16S rRNA genes amplicons obtained by low-resolution methods (e.g PCR cloning, TRFLP) can be assigned to these few taxa (Ferrari *et al*. 2012; Haynes *et al*. 2003; Singh *et al*. 2012). Exceptionally, the aphid *Brevicoryne brassicae* was reported to bear multiple bacteria, including members of the Pseudomonales and Enterobacteriales (Clark *et al*. 2012). To our knowledge, next generation sequencing methods have not been applied to test the supposed low bacterial diversity in any phloem feeding insect.

The purpose of this study is to quantify the bacterial diversity in *Bemisia tabaci* whiteflies, by pyrosequencing 16S rRNA gene fragments amplified with general bacterial primers. *B. tabaci* is a morphological species of cosmopolitan distribution that comprises a complex of >30 partially or completely reproductively isolated candidate species (De Barro *et al*. 2011). [We refer to these candidate species as "*B. tabaci* Asia1", "*B. tabaci* MEAM1" *etc.*, following Tay *et al*. (2012).] As for whiteflies generally, the primary symbiont in all *B. tabaci* species is the γ-

proteobacterium Candidatus *Portiera aleyrodidarum* (Baumann *et al*. 2004). A total of 7 different secondary symbionts have been identified in *B. tabaci* species: *Arsenophonus* (Thao & Baumann 2004a), *Cardinium* (Weeks *et al*. 2003), *Fritschea* (Everett *et al*. 2005), *Hamiltonella* (Thao & Baumann 2004a), *Hemipteriphilus* (allied to *Orientia*) (Bing *et al*. 2013), *Rickettsia* (Gottlieb *et al*. 2006) and *Wolbachia* (Nirgianaki *et al*. 2003). Most current research on the bacterial diversity in *B. tabaci* has focused on the prevalence of these bacteria in different species and from different geographical locations. These studies are based on end-point PCR using refined diagnostic primers and template DNA obtained by gentle extraction methods unsuitable for Gram-positive bacteria, with no estimation of the detection limits of the assay (Ahmed *et al*. 2013; Chiel *et al*. 2007; Chu *et al*. 2011; Gueguen *et al*. 2010; Thierry *et al*. 2011). This widespread practice could grossly underestimate the bacterial diversity in *B. tabaci* because the template DNA may be representative of a subset of bacteria in the original samples, previously unreported taxa are excluded from diagnostic PCR assays, and target bacteria at abundances below the assay detection limit may be erroneously scored as absent. Circumstantial evidence that *B. tabaci* may include bacteria additional to the 8 bacterial taxa described comes from the cultivation of various bacteria, including the Gram-positive *Bacillus* and *Staphylococcus*, from *B. tabaci* MEAM1 (Davidson *et al*. 2000) and recovery of *Bacillus*, *Enterobacter*, *Paracoccus*, *Acinetobacter* and other bacterial sequences from a metatranscriptomic analysis of *B. tabaci* Asia I and Asia II (Singh *et al*. 2012). This study has two specific aims. The first was to quantify the bacterial diversity in 7 *B. tabaci* species, including two highly invasive species, *B. tabaci* MEAM1 and *B. tabaci* MED (also known as biotypes B and Q, respectively). We used material from

the *B. tabaci* culture repository at Natural Resources Institute, UK for this analysis, so that the insect material is available for future investigations, including as controls for analyses of bacterial diversity in field populations. The second aim was to compare the bacterial diversity in long-term *B. tabaci* MEAM1 cultures reared on different plants or with/without plant begomoviruses vectored by this species. Secondary symbionts have been implicated in the capacity of *B. tabaci* insects to vector begomoviruses (Gottlieb *et al*. 2010; Rana *et al*. 2012), and the prevalence and abundance of some secondary symbionts are reported to vary with rearing plant for *B. tabaci* and other phloem feeders (Chandler *et al*. 2008; Chen *et al*. 2000; Chiel *et al*. 2007; Tsuchida *et al*. 2004). Interpretation of these associations, however, depends critically on the effects of virus carriage and rearing plant on the overall bacterial complement of the insects. This first comprehensive analysis of the bacterial diversity in any phloem feeding insect demonstrates a remarkably low diversity, raising fundamental questions about the processes that shape the variation in diversity of host-associated microbiota in different animals.

Experimental procedures

The experimental material

The analyses were conducted on adult *Bemisia tabaci* in laboratory culture at three locations (Table 6.1). The 7 *B. tabaci* species in set-1 were reared at 27 °C with a 14L:10D photoperiod at The Natural Resources Institute, UK. Set-2 were *B. tabaci* MEAM1 derived from three separate field collections, one nonviruliferous, one infected with tomato yellow leaf curl virus (TYLCV) and one infected with tomato mosaic virus (ToMoV) (Sinisterra *et al*. 2005), and maintained at 25 ± 1°C with

16L:8D photoperiod at the U.S. Horticultural Research Laboratory, Ft. Pierce, F*L.* Set-3 comprised *B. tabaci* MEAM1, from a culture established in 1989 from a field population by Dr John Sanderson at Cornell University, Ithaca, NY, and reared on poinsettia cv. Freedom Red, *Vicia faba* cv. Windsor, *Nicotinia tabacum* cv. Xanthi and *Solanum tuberosum* cv. Desiree at 24/20 °C (L/D) and 14L:10D photoperiod. The identity of every isolate was confirmed by sequencing the mitochondria cytochrome c oxidase I (*mtCOI*) gene (Shatters *et al*. 2009). Samples of set-1 and set-2 were preserved in 90% ethanol prior to molecular analysis, and set-3 comprised fresh material.

Table 6.1 Complement of bacterial symbionts in *Bemisia tabaci* species

Insect culture	Rearing plant	Collection site	Symbionts*					
			A	C	H	P	R	W
Set-1: *Bemisia tabaci* species reared at Natural Resources Institute, University of Greenwich, UK								
Asia1	Egg plant	India	+	-	-	+	+	+
AsiaII-5	Cassava	India	+	-	-	+	-	+
AsiaII-7	Cotton	China	+	+	-	+	+	+
Australia	Wild poinsettia	Australia	+	-	-	+	-	+
China1	Cotton	China	-	-	-	+	+	+
MEAM1	Cotton/Brassica	Peru	-	-	+	+	+	-
Mediterranean	Pepper	France	-	+	+	+	-	-
Set-2: *Bemisia tabaci* MEAM1 reared at U.S. Horticultural Research Laboratory, Fort Pierce, FL, USA†								
Nonviruliferous	Tomato cv. Florida Lanai	Florida, USA	-	-	+	+	+	-
Nonviruliferous	Tobacco cv. Xanthi		-	-	+	+	+	-
ToMoV-infected	Tomato cv. Florida Lanai	Florida, USA	-	-	+	+	+	-
TYLCV-infected	Tomato cv. Florida Lanai	Florida, USA	-	-	+	+	+	-
Set 3: *Bemisia tabaci* MEAM1 strain XF13/01 reared at Cornell University, Ithaca, NY								
	Fava bean cv. Windsor	New York, USA	-	-	+	+	+	-
	Poinsettia cv. Freedom Red		-	-	+	+	+	-
	Potato cv. Desiree		-	-	+	+	+	-
	Tobacco cv. Xanthi		-	-	+	+	+	-

* *Arsenophonus* (A), *Cardinium* (C), *Hamiltonella* (H), *Portiera* (P), *Rickettsia* (R) and *Wolbachia* (W)

† The nonviruliferous cultures and cultures infected with each virus were derived from different field populations, see Sinisterra *et al*. (2005).

DNA isolation

The insect samples, each comprising 30-40 adult whiteflies, were rinsed three times in extraction buffer (20 mM Tris-HCl pH 8.0, 2 mM sodium EDTA, 1.2% Triton® X-100) and then hand- homogenized in extraction buffer containing 20 mg lysozyme ml^{-1}, and incubated at 37 °C for 1.5 h to achieve DNA extraction from both Gram-positive and Gram-negative bacteria. The DNA in the samples was then extracted using the DNeasy Blood and Tissue Kit (Qiagen, Valencia, California, USA) following manufacturer's instructions for Gram-positive bacteria. The quantity and quality of the DNA were measured by Nanodrop 2000 (Thermo Scientific).

Multiplex 454 pyrosequencing of 16S rRNA gene sequences

16S ribosomal RNA amplicons of the V6-V7 region were prepared using general 16S rRNA gene primers 907Fmod (5'-AAACTCAAADGAATTGACGG-3') modified from (Sundquist *et al*. 2007) and 1237R (5'-GTAGYACGYGTGTWGCCC-3') (Turner *et al*. 1999). The 907Fmod primer comprised the 907F primer of (Sundquist *et al*. 2007) with the degenerated nucleotide D (G or T) in place of G at nt10, following preliminary analysis indicated poor predicted amplification of *Portiera* 16S rDNA with the 907F primer. Each sample-specific 907Fmod primer bears a multiplex identifier (MID) sequence (Supplementary Table 1a). PCR reactions were conducted as previously described (Wong *et al*. 2011). Briefly, equal amounts (ng) of the products of three PCR reactions per sample were mixed, purified using the QIAquick PCR purification kit (QIAGEN), and quantified by the *Quant-iT™ PicoGreen*® Kit. Emulsion PCR was conducted at 1.5 copies per bead using only 'A' beads for unidirectional 454 GS-FLX pyrosequencing with standard Titanium chemistry.

Pyrosequencing flowgrams were converted to sequence reads using 454 Life Science software (www.454.com). Reads with ambiguous nucleotides (N) and < 270 nucleotides after the forward primer, and mismatches with the 16S rRNA gene primers were excluded in the initial filtering. The QIIME 1.4.0 virtualbox package was used to split the multiplexed sequences, discard chimeras, denoise the data, bin sequences at 97% sequence identity, and make taxonomy calls to genus level (Caporaso *et al*. 2010). Default parameters were used except that the denoising cutoff was set to retain doubletons, and the RDP classifier was applied using a custom Greengenes database to assign class through genus designations. Bacterial species identity of each OTU was assigned by NCBI StandAlone BLAST (megablast program) using the nucleotide (nt) database (August 2012) under default settings with supplementary manual curation in April 2013. OTUs with either single reads or fewer reads than in the negative controls were excluded.

The analysis yielded several instances of multiple OTUs with the same bacterial species as the BLAST top hit (i.e. species identity), including many more reads in one OTU (which we call the "major" OTU) than the other OTUs (the "minor" OTUs). To assess whether the minor OTUs are an artifact arising from sequencing error, the Poisson probabilities (P_{Pois}) for a single artifactual read at different %ID were calculated for each biological sample, using the sequencing error rate of 1.07% (Gilles *et al*. 2011) (Supplementary Table 2). The probability of each minor OTU arising by sequencing error was determined by multiplying the P_{Pois} with the number of reads in a given biological sample.

Bacterial co-occurrence analyses

The patterns of bacterial species co-occurrence among whitefly samples were analyzed by the *C*-score test (Stone & Roberts 1990). The pyrosequence datasets (Supplementary Table 3) were transformed into presence–absence matrices with the five bacterial species (*Arsenophonus, Cardinium, Hamiltonella, Rickettsia* and *Wolbachia*) as rows and individual whitefly samples as columns. *C*-score test calculates the mean number of instances where two bacterial species co-occur. The computed *C*-score is significantly greater than the null distribution if the bacteria co-occur less frequently than predicted by chance (segregation), and less than the null distribution for positive co-occurrence (aggregation). The most appropriate null model for these data, in which the presence/absence of each bacterial species in the whitefly populations is known, is the "fixed-fixed" null model [SIM9 of Gotelli (2000)]. The observed data matrices were compared to 5000 randomly generated matrices using EcoSim 7.72 (Gotelli & Entsminger 2012).

Quantitative real-time PCR (qRT-PCR) analysis

The abundance of *Portiera, Hamiltonella* and *Rickettsia* in *B. tabaci* MEAM1 was determined by qRT-PCR of 16S rRNA gene amplicons, normalized to the *B. tabaci* gene tubulin alpha-1 chain (*αTUB*: NCBI_EE598061), using specific primers covering ca.100 bp of the V2 region of the 16S rRNA gene (Supplementary Table 1b). The reactions were conducted in C1000TM Thermal cycler (Bio-Rad) with 10 µl Power SYBR Green PCR Master Mix (Applied Biosystems), 2.4 µl 2.5 µM primers and ca. 25 ng DNA template in 20 µl volume, with reagent-only samples as negative controls. The thermal profile was 95°C for 10 min, 35 amplification cycles of 95°C for 10 s,

60°C for 30 s and dissociation cycle of 95°C for 10 s, 60°C for 5 s then brought back to 95°C. The dissociation curve confirmed that every reaction yielded a single PCR product with the predicted Tm. The experiments comprised three biological samples per treatment, for each of which the mean Ct of three technical replicates was calculated. The fold-difference between the abundance of each bacterial16S gene and the whitefly *αTUB* gene was determined by the ΔΔCt method (Schmittgen & Livak 2008).

Statistical analysis

The variation in abundance of bacterial 16S rRNA gene amplicons was analyzed by ANOVA, following confirmation that the datasets were normally distributed (Anderson Darling test) with homogeneous variance (Bartlett's and Levene's tests) after $\log_2$ transformation. The significance of individual pair-wise differences of biological interest was tested by *post hoc* test. Paired t-tests were applied to test for difference in bacterial diversity indices and number of *Portiera* reads between male and female insects.

Results

Pyrosequencing data

The pyrosequencing of the 16S rRNA gene amplicons from the whitefly samples yielded 229,457 reads, after quality filtering and removal of chimeric sequences. Between 2,119 and 15,197 reads were obtained per insect sample, and the reads could be assigned to 40 operational taxonomic units (OTUs) at 97% sequence identity (Supplementary Table 3; Table 6.2). The rarefaction curve for every sample tended to

saturation (Supplementary Figure. 1), indicating that the OTUs detected were representative of the total bacterial community in each sample.

Table 6.2 Pyrosequencing analysis of 16S rRNA gene amplicons with diversity indices

Sample	Males				Females			
	No. reads	No. OTUs*	Diversity indices		No. reads	No. OTUs*	Diversity indices	
			Shannon index	Simpson index			Shannon index	Simpson index
Set-1 (different *B. tabaci* species)								
Australia	2566	12/4	1.34/0.69	0.62/0.42	5328	15/3	1.18/0.61	0.53/0.36
Asia 1	5562	17/5	1.60/1.14	0.75/0.65	8834	18/4	1.49/0.92	0.70/0.53
China 1	2119	14/3	1.26/0.97	0.62/0.58	4419	17/5	1.21/0.90	0.57/0.53
Asia II-5	2820	12/3	0.74/0.36	0.29/0.17	3202	15/4	0.88/0.50	0.37/0.25
Asia II-7	4866	16/5	1.66/1.39	0.77/0.73	11152	22/8	1.79/1.46	0.80/0.74
MEAM 1	6696	14/3	0.77/0.54	0.33/0.28	13982	19/6	1.02/0.81	0.50/0.46
Mediterranean	5380	15/3	1.38/1.02	0.67/0.61	15197	18/3	1.24/0.83	0.56/0.48
Set-2 (*B. tabaci* MEAM 1 reared at U.S. Horticultural Research Laboratory, Fort Pierce, FL, USA †)								
Nonviruliferous on tomato	3688	14/3	1.21/0.97	0.62/0.58	8984	18/3	1.29/1.02	0.65/0.61
Nonviruliferous on tobacco	8989	17/3	1.09/0.85	0.53/0.49	6969	20/3	1.25/0.97	0.64/0.59
ToMoV-infected	11298	19/4	1.37/1.07	0.68/0.65	11022	18/4	1.20/0.88	0.57/0.51
TYLCV-infected	10584	20/3	1.30/1.06	0.68/0.64	13218	21/4	1.18/0.83	0.55/0.47
Set-3 (*B. tabaci* MEAM 1, strain XJ13/01 cultured on different plants)								
Bean	3900	20/4	1.35/1.09	0.69/0.66	7098	20/3	1.36/1.10	0.70/0.67
Poinsettia	6531	17/4	1.22/1.00	0.63/0.60	5690	20/5	1.32/1.04	0.66/0.62
Potato	11949	18/4	1.23/0.99	0.62/0.59	15913	20/5	1.30/1.07	0.67/0.65
Tobacco	5625	18/4	1.32/1.04	0.67/0.63	5839	19/5	1.35/1.05	0.67/0.63

* OTUs (operational taxonomic units) were defined with pair-wise 97% sequence identity

† Values provided are before/after correction for sequence error

Thirty-eight (95%) of the 40 OTUs could be assigned to the primary symbiont *Portiera* and six secondary symbionts, *Arsenophonus*, *Cardinium*, *Hamiltonella*, *Hemipteriphilus*, *Rickettsia* and *Wolbachia* (*Fritschea* was not detected). The remaining two OTUs were α-Proteobacteria (Supplementary Table 3). OTU38 is allied with an un-named bacterium (NCBI accession HM256949) previously detected in human skin samples (Kong *et al*. 2012) and was detected in *B. tabaci* Australia males [2 (0.08%) of 2,568 reads], and was probably a handling contaminant. OTU20 has high sequence identity with the 16S rRNA gene sequence (NCBI accession JX560791) of *Methylobacterium oryzae* endophyte of pineapple plants, and was detected in *B. tabaci* Asia1 males [2 (0.04%) of 5,564 reads] in set-1 and *B. tabaci* MEAM1 females on potato plants in set-3 [2 (0.01%) of 15, 915 reads]. OTU20 may be an insect culture contaminant or occasional associate of *B. tabaci*.

The reads assigned to *Portiera* and five of the secondary symbionts comprised multiple OTUs, with many more reads in one major OTU than the other minor OTUs (*Hemipteriphilus* comprised a single OTU, OTU34) (Supplementary Table 4). For example, *Portiera* was represented by a total of 12 OTUs, of which one (OTU31) accounted for >90% of the reads in every sample. All insect samples bore the same major OTU of each symbiont type, with the exception of *Rickettsia*, which included two major OTUs (Supplementary Table 4; Figure. 6.1). *Rickettsia* OTU26 (with related minor OTUs 1, 21, 22 40 and 46) was detected in all samples of *B. tabaci* MEAM 1, and had >97% sequence identity to *Rickettsia* sp. described previously in *B. tabaci* MEAM 1 (Gottlieb *et al*. 2006). The second major *Rickettsia* OTU, OTU11 (with related minor OTUs 3, 21 and 44), was detected in *B. tabaci* China 1, Asia 1 and Asia II-7, originating from China and India. This *Rickettsia* is allied to *Rickettsia*

strain RI1 identified in *B. tabaci* Asia II-India, collected from India (Singh *et al*. 2012).

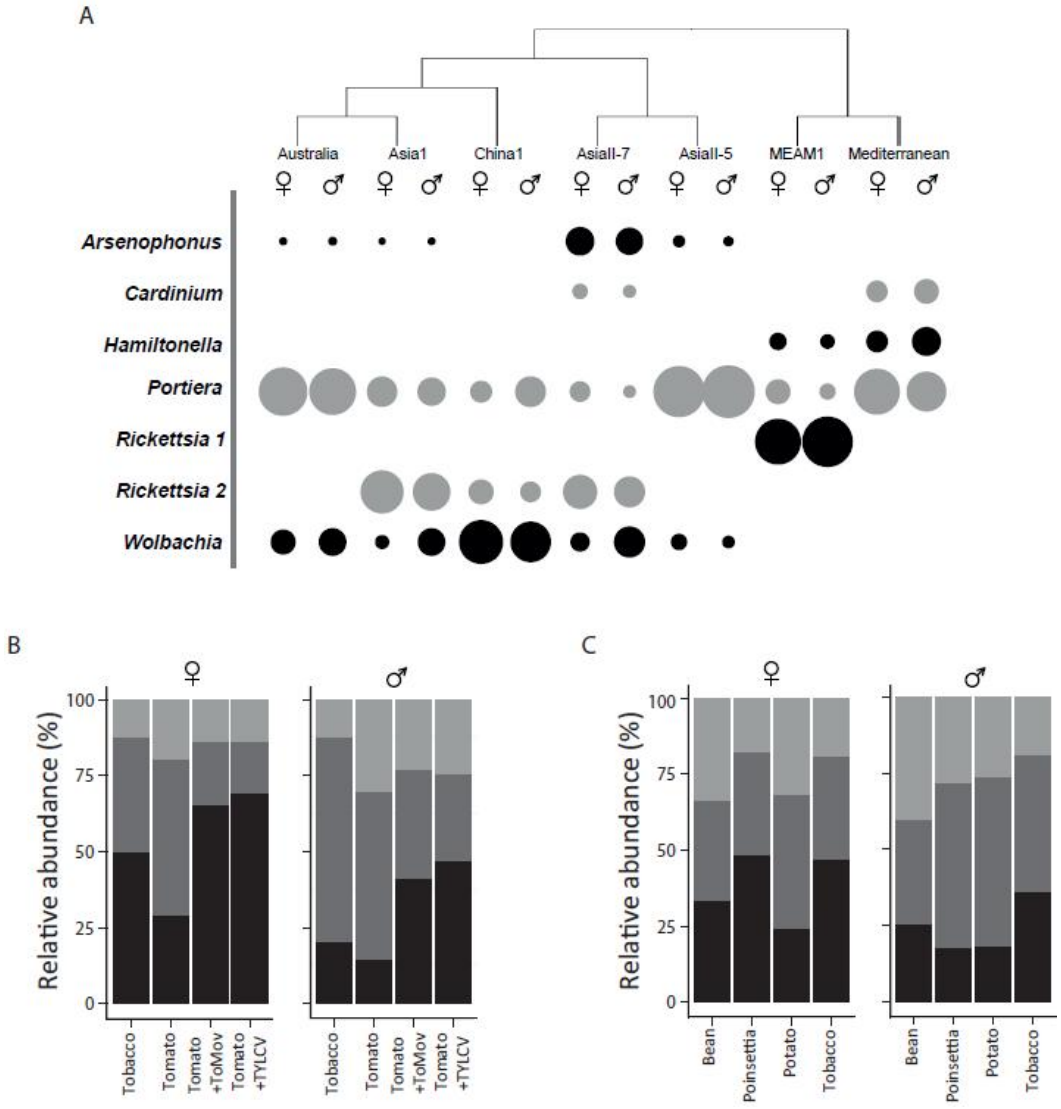


Figure 6.1 Relative abundance of pyrosequence reads assigned to bacterial symbionts of *B. tabaci*. *A.* Different species of set-1 (circle area corresponds to relative abundance of bacteria in each sample). *Rickettsia*-1 and *Rickettsia*-2 have high sequence ID with *Rickettsia* described by Gottlieb *et al*. (2006) and Singh *et al*. (2012), respectively. *Hemipterophilus*, OTU20 and OTU38 are not included in this figure because they account for <1% of the reads. B. *B. tabaci* MEAM1 set-2. C. *B.*

tabaci MEAM1 set-3.Key for Figure. 1 B&C: ■ *PortieraI,* ■ *Rickettsia,* ■ *Hamiltonella.*

To investigate whether the minor OTUs were likely the result of sequencing error, the Poisson probability distribution (Supplementary Table 2) was applied to the data, with a critical probability of 0.05. Sequencing error could account for the % sequence difference between each major OTU and all the related minor OTUs except *Portiera* OTU25 and *Arsenophonus* OTU36 (Supplementary Table 4). These two variant OTUs are each represented by just two reads in a single insect sample, accounting for <0.003% of the total reads. For subsequent analysis, the reads for all other minor OTUs were assigned to the related major OTU.

Bacterial communities in different species of B. tabaci

The bacterial communities in *B. tabaci* are of low diversity, as indicated by the low values for the Shannon index (0.36-1.46) and Simpson index (0.17-0.74) (Table 6.2). The indices did not differ between males and females (Shannon index: t=0.42, p=0.681; Simpson index: t=0.47, p=0.646).

The primary symbiont *Portiera* was detected in every sample of *B. tabaci* tested, and accounted for between 5% (in AsiaII-7 males) and 86% (in AsiaII-5 females) of the total reads per sample (Figure. 6.1A, Supplementary Table 3). For 13 of the 15 *B. tabaci* cultures tested, *Portiera* contributed a higher percentage of reads in the sample of females than male insects (Figure. 6.2) and, overall, the difference between the two sexes was statistically significant (t = 3.60, p=0.003).

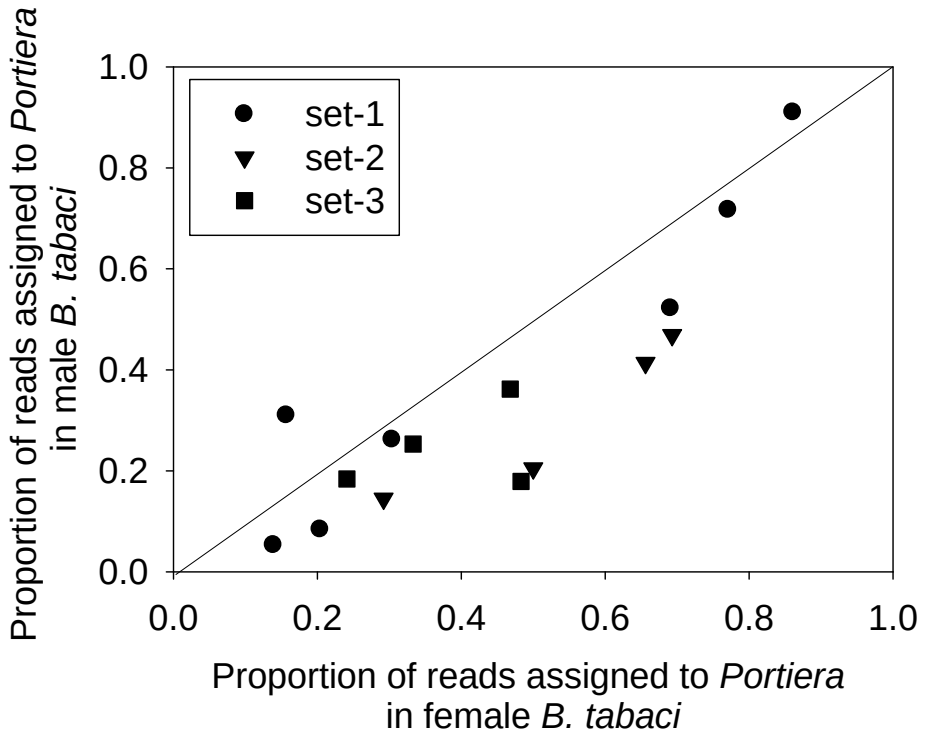


Figure 6.2 Relative abundance of *Portiera* pyrosequence reads in male and female *B. tabaci.*

The male and female samples of all 7 species of *B. tabaci* in set-1 bore between two and four secondary symbionts, in addition to *Portiera*; and the *B. tabaci* MEAM1 in set-2 and set-3 had the same secondary symbionts (*Hamiltonella* and *Rickettsia*) as for the MEAM1 culture in set-1 (Figure. 6.1; Supplementary Table 3). The complement of secondary symbionts was unique to each species, apart from AsiaII-5 and Australia, both of which harbored *Arsenophonus* and *Wolbachia*. Visual inspection revealed that *Wolbachia* and *Hamiltonella* did not co-occur in any samples, and that *Arsenophonus* was scored only in samples that also bore *Wolbachia*. To assess whether the incidence of co-occurrence of secondary symbionts across the 7 *B. tabaci* species differed from

that predicted by chance, the *C*-scores test was applied to the data. The *C*-score for the full dataset, 15.2, was significantly higher than expected by chance (13.3, $p < 0.001$), indicating that, overall, the bacteria were less likely to co-occur than predicted by chance.

The bacterial complement of viruliferous and non-viruliferous B. tabaci

The pyrosequence data were used to compare the symbiont complement of viruliferous and non-viruliferous *B. tabaci* MEAM1. In the context of evidence that the secondary symbiont *Hamiltonella* may promote begomovirus transmission (Gottlieb *et al*. 2010; Rana *et al*. 2012), we hypothesized that the viruliferous cultures (infected with ToMoV and TYLCV) would have elevated titers of *Hamiltonella*. Contrary to this expectation, the % contribution of *Hamiltonella* reads was marginally lower in the cultures bearing ToMoV and TYLCV than in the non-viruliferous culture, but the relative abundance of *Portiera* was >two-fold greater in viruliferous compared to non-viriliferous adults (Figure. 6.1B).

To test for among-culture variation in the absolute density of the bacteria, the abundance of 16S rRNA gene sequences of *Portiera*, *Rickettsia* and *Hamiltonella* was determined by qPCR (Figure. 6.3). By ANOVA, bacterial abundance varied significantly across the three cultures, but the interaction term "symbiont x insect culture" was not significant (see legend to Figure. 6.3). These results indicate that, although the overall bacterial density varied among the cultures (highest in the ToMoV-infected culture and lowest in the non-viruliferous insects), there was no significant variation in the relative abundance of the different symbionts across the three insect cultures. In summary, this analysis is suggestive of an association between

begomovirus infection and high bacterial titers, but no association specifically with elevated *Hamiltonella*.

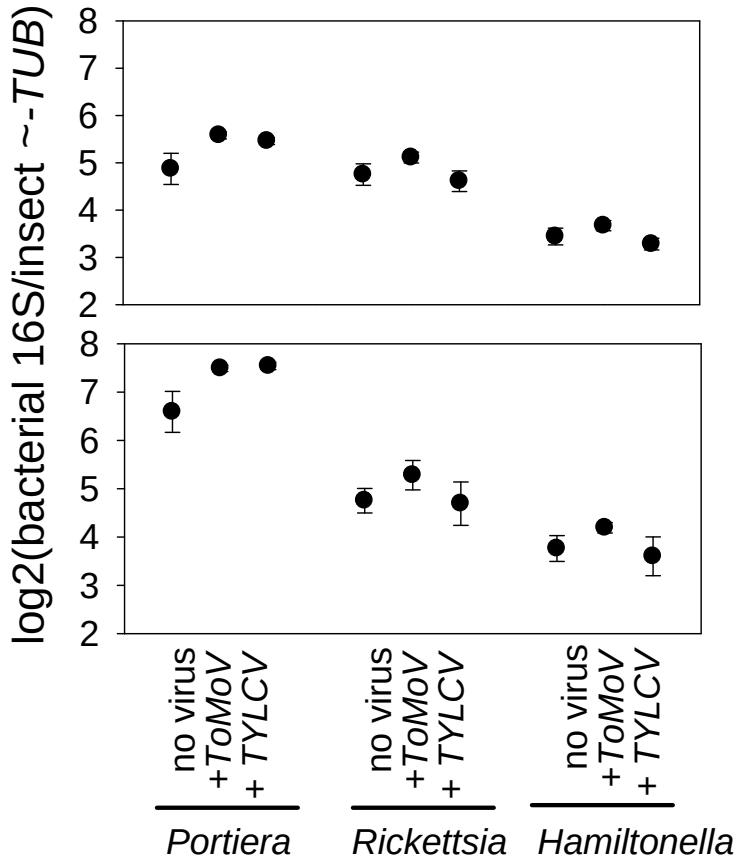
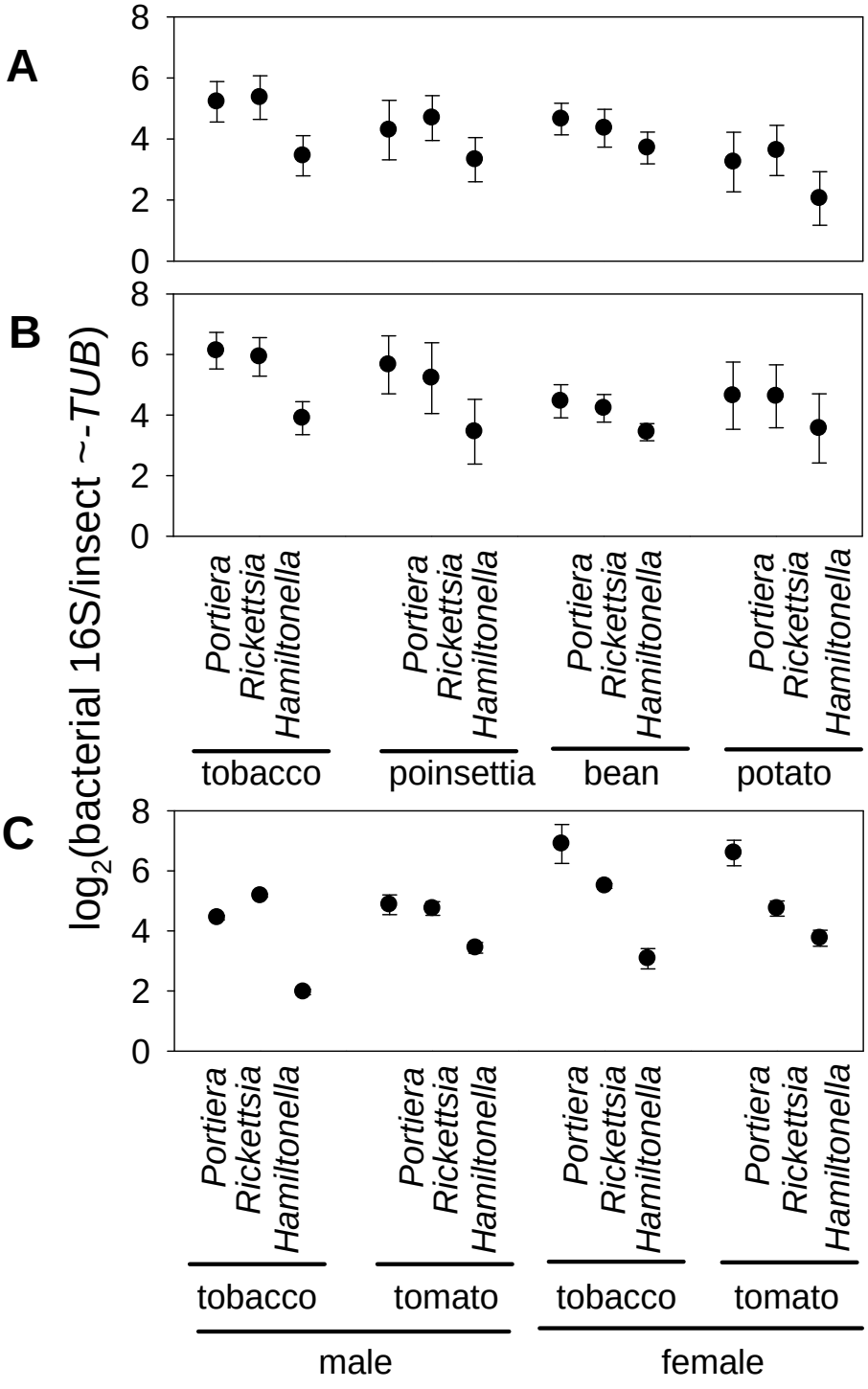


Figure 6.3 qPCR analysis of relative copy number of bacterial 16S rRNA genes and insect αTubulin gene in *B. tabaci* MEAM 1 cultures that are non-viruliferous or infected with ToMoV or TYLCV. *A.* Males B. females. ["no virus" data are also presented as "tomato" dataset in Figure. 4C]. ANOVA: insect culture $F_{1,36} = 7.21$, $p=0.002$; symbiont type: $F_{2,36} = 169$, $p<0.001$; sex: $F_{1,36} = 47.31$, $p<0.001$; insect culture x symbiont type: $F_{4,36} = 2.27$, $p>0.05$; insect culture x sex: $F_{2,36} = 0.25$, $p>0.05$; symbiont x sex: $F_{4,36} = 0.10$, $p>0.05$; 3-way interaction: $F_{4,36} = 0.10$, $p>0.05$

Bacterial communities in B. tabaci reared on different plants

The pyrosequence reads obtained for non-viruliferous *B. tabaci* MEAM 1 reared on different plants in set-2 and set-3 suggested that the relative abundance of the three symbionts (*Portiera*, *Hamiltonella* and *Rickettsia*) varied subtly between the cultures on different plants (Figure. 6.1B & C). For set-3, symbiont abundance varied significantly with rearing plant, attaining a significantly higher abundance in insects on tobacco than potato ($p<0.05$) (Figure. 6.4A & B, with statistical analysis in legend), but the relative abundance of different symbionts did not vary across the plants (the ANOVA interaction term "symbiont x plant" was not significant). In the parallel analysis of set-2 (Figure. 6.4C), this interaction term was significant, and the *post hoc* test revealed that *Hamiltonella* (but not *Portiera* or *Rickettsia*) was significantly reduced in the insect culture reared on tobacco relative to tomato.

Figure. 6.4 qPCR analysis of relative copy number of bacterial 16S rRNA genes and insect αTubulin gene in *B. tabaci* MEAM 1 reared on different plants. *A.* Males of set-3. B. females of set-3. C. Non-viruliferous insects of set-2. ANOVA for set-3 data (displayed as Figure. 4A & B): rearing plant $F_{3,48} = 3.06$, $p=0.037$; symbiont type: $F_{2,48} = 15.91$, $p<0.001$; sex: $F_{1,48} = 3.44$, $p>0.05$; rearing plant x symbiont type: $F_{3,48} = 0.27$, $p>0.05$; rearing plant x sex: $F_{3,48} = 0.88$, $p>0.05$; symbiont type x sex: $F_{2,48} = 0.16$, $p>0.05$; 3-way interaction: $F_{6,48} = 0.08$, $p>0.05$. ANOVA for set-2 data (displayed in Figure. 4C): rearing plant: $F_{1,24} = 1.12$, $p>0.05$; symbiont type: $F_{2,24} = 85.22$, $p<0.001$; sex: $F_{1,24} = 33.09$, $p<0.001$; rearing plant x symbiont type: $F_{2,24} = 7.93$, $p=0.002$; rearing plant x sex: $F_{1,24} = 3.17$, $p>0.05$; symbiont type x sex: $F_{2,24} = 11.15$, $p<0.001$; 3-way interaction: $F_{2,24} = 0.17$, $p>0.05$.


A
B
C
log2(bacterial 16S/insect ~-TUB)
0
2
4
6
8
Portiera
Rickettsia
Hamiltonella
tobacco
poinsettia
bean
potato
tomato
male
female

Discussion

Our appreciation of microbial diversity has been transformed by the advent of culture-independent methods to identify and enumerate taxa. A majority of environmental microbes are not readily cultivated on standard media and, although microbiologists are increasingly responding to the challenge to develop new culture strategies (Carini *et al*. 2013; Goodman *et al*. 2011; Singh *et al*. 2013), sequence-based methods remain the method of choice for establishing microbial diversity, with PCR-generated 16S rRNA gene amplicons the mainstay for bacteria. However, reliable estimates of bacterial diversity require close attention to methods. Pitfalls include unsuitable methods to extract DNA (see Introduction) and poor sequence identity between the primers and the some taxa in the sample, leading to biased amplification. Two further difficulties can arise with the interpretation of sequence data: underestimation of diversity because some low abundance taxa may remain undetected even with a sequencing depth of $\geq$10,000 reads; and the artifactual inflation of OTUs through misinterpretation of PCR and sequencing errors as novel sequences (Huse *et al*. 2010; Kunin *et al*. 2010; Sun *et al*. 2012). In this study on the bacterial diversity in *B. tabaci*, we sought to minimize these sources of error. Specifically, we used extraction methods designed to obtain DNA from Gram-positive bacteria; and we increased the degeneracy of the primer sequences to avoid exclusion of bacteria, including *Portiera*, in the amplicon set. Furthermore, the number of reads in all samples yielded saturation of the rarefaction curves (Supplementary Figure. 1), indicating that our sampling was exhaustive, and we determined the probability that OTUs with similar sequence could be attributed to sequence error.

With these multiple precautions in place, the pyrosequencing revealed a total of just 9 genera of bacteria across the 7 species and 15 samples of *B. tabaci*, with >99.99% of the sequence reads assigned to 7 genera described previously as *B. tabaci* symbionts. These bacteria are localized to cells (specifically bacteriocytes for *Portiera*), hemolymph (blood) and other tissues of their whitefly hosts (Brumin *et al*. 2012; Gottlieb *et al*. 2006; Gottlieb *et al*. 2008). Our results are in sharp contrast to the report of the Gram-positive bacteria *Bacillus* and *Enterobacter*, and well as other bacteria of unknown taxonomic position, in two independent studies of *B. tabaci*, one on *B. tabaci* MEAM1 in laboratory culture (Davidson *et al*. 2000) and the other on multiple *B. tabaci* species collected from various geographical locations in India (Singh *et al*. 2012). Perhaps these bacteria are surface contaminants, and the different studies varied in the efficacy of washing methods to remove them. Alternatively (and we consider less likely), these additional bacteria are carried internally, but their distribution varies such that, by chance, they were entirely absent from the *B. tabaci* cultures derived from 11 independent collections and maintained in three different locations (New York, Florida and United Kingdom) in this study.

The resident microbiota in humans and other mammals includes a high diversity of bacteria at the species and strain level relative to higher level taxonomic units (Dethlefsen *et al*. 2007; Human Microbiome Project 2012). Unlike mammals, the low diversity at the level of genus is also displayed at lower taxonomic levels in *B. tabaci*. Each *B. tabaci* sample bore just one major OTU for each bacterial genus, apart from the two *Rickettsia* taxa (one in *B. tabaci* MEAM1 collected from the Americas, and the other in three *B. tabaci* species collected from China, Figure. 6.1A). Furthermore, almost all of the fine-scale sequence variation was within the range of variation

predicted from sequencing error, and the pattern of variation suggests that the probability analysis did not attribute biologically-distinct OTUs erroneously to sequencing error. Specifically, every *B. tabaci* species bore the same major OTU of each symbiont-type; and *Portiera*, predicted to have especially low within-sample biological diversity because of its long-term obligate vertical transmission and very small effective population size (Moran & Wernegreen 2000), included more minor OTUs than any of the secondary symbionts.

The likely reason for the low diversity of the resident microbiota in *B. tabaci* is the lack of a gut microbiota, as previously reported by (Davidson *et al*. 2000), whose detailed electron microscopical analysis revealed no bacteria in the gut lumen of *B. tabaci*, unless the insects were fed on high densities of bacteria. Many bacteria cannot gain access to the whitefly gut because the natural diet of plant phloem sap is generally microbe-free (although it can bear some highly specialized pathogens, notably the phytoplasmas) and the very narrow whitefly stylets exclude any particles of diameter $\geq$ 0.5 µm (Davidson *et al*. 2000). Phloem-sap feeding aphids also support a resident microbiota of exceptionally low diversity, but this trait is not universal among plant sap feeding insects. For example, the leafhopper *Euscelis incisus* and sharp-shooter *Homalodisca vitripennis* bear a diverse gut microbiota (Douglas 1988; Hail *et al*. 2011). Further research is required to explain the among-insect variation in the extent of microbial colonization of the gut.

Barriers to gut colonization are not, however, the sole determinants of the bacterial diversity in *B. tabaci*. The incidence of a bacterial taxon in a host population is shaped by the balance between gain and loss of bacterial cells in each insect and the selective differential between insects with different bacterial communities. The sole route by

which *B. tabaci* acquires its primary symbiont is by vertical transmission via the egg, such that the phylogenies of whiteflies and *Portiera* are congruent (Thao & Baumann, 2004b); and whiteflies are believed to be nutritionally dependent on *Portiera* (Sloan & Moran 2012), such that any *Portiera*-free insects would be eliminated rapidly from the population. Secondary symbionts are also maternally inherited via the egg, apparently with high fidelity in the laboratory. The multiple reports of intermediate prevalence of most secondary symbionts in wild populations, e.g. Bing *et al*. (2013), Chu *et al*. (2011), Thierry *et al*. (2011), suggests that the more variable environmental conditions in the field may reduce vertical transmission efficiency; and the incongruence of host/secondary symbiont phylogenies is indicative of horizontal transmission, potentially via feeding, sexual contact and aborted parasitoid attack (Caspi-Fluger *et al*. 2012; Chiel *et al*. 2009; Moran & Dunbar 2006). Evidence that selective differentials can drive rapid changes in the prevalence of a secondary symbiont in *B. tabaci* comes from the study of Himler *et al*. (2011), who showed that *Rickettsia* significantly promotes the fecundity and survivorship of *B. tabaci* MEAM 1, as well distorting the sex ratio to favor female offspring (which carry *Rickettsia*). These processes were sufficient to drive an increase in prevalence of *Rickettsia* from $\leq$1% to $\geq$95% of the population in SW USA in less than 80 host generations.

Diversity indices are sensitive to the relative abundance of different taxa. Intriguingly, the abundance of the different symbionts in *B. tabaci* MEAM1 tended to vary in concert, both among insect cultures with different virus infections and reared on different plants (Figure. 6.3 & 4). (Exceptionally, the abundance of *Hamiltonella* 16S was specifically reduced in *B. tabaci* set-2 reared on tobacco relative to tomato.) The simplest interpretation of these data is that the multiple symbionts are maintained by a

common host regulatory mechanism, with the implication that net rates of proliferation are controlled by the same or linked sets of conditions and resources in the host insect. The systemic immune system may be involved because the symbionts have different tissue tropisms, with *Portiera* and *Hamiltonella* located in bacteriocytes and *Rickettsia* widely distributed throughout the insect body (Brumin *et al*. 2012; Gottlieb *et al*. 2008). Similarly, the elevated abundance of bacterial 16S in the viruliferous cultures of *B. tabaci* MEAM1 (set-2) may be indicative of virus manipulation of the immune status of the insect, with correlated consequences for insect regulation of symbionts.

In conclusion, the exhaustive sampling of the bacterial diversity in *B. tabaci* made possible by this first study using high throughput sequencing methods has revealed a resident microbiota of exceptionally low diversity that is structured predominantly by vertical transmission from mother to offspring. These data raise important questions about the function of the gastro-intestinal tract that is not naturally heavily colonized by bacteria, the role of the insect immune system in maintaining the relative abundance of the microbiota, and the relative importance of horizontal transmission of secondary symbionts and selection on insects with different bacterial communities in shaping the composition of the microbiota in individuals, populations and species of *B. tabaci.* More generally, this study provides a valuable “low diversity” datapoint for the broader ecological questions posed by the remarkable variation in diversity of resident microbiota across different animals.

References

Ahmed MZ, De Barro PJ, Ren SX, Greeff JM, Qiu BL (2013) Evidence for horizontal transmission of secondary endosymbionts in the Bemisia tabaci cryptic species complex. PLoS One 8, e53084.

Baumann L, Thao ML, Funk CJ, Falk BW, Ng JC, Baumann P (2004) Sequence analysis of DNA fragments from the genome of the primary endosymbiont of the whitefly Bemisia tabaci. Current Microbiology 48, 77-81.

Bing XL, Yang J, Zchori-Fein E, Wang XW, Liu SS (2013) Characterization of a newly discovered symbiont of the whitefly Bemisia tabaci (Hemiptera: Aleyrodidae). Applied and Environmental Microbiology 79, 569-575.

Brucker RM, Bordenstein SR (2012) Speciation by symbiosis. Trends in Ecology and Evolution 27, 443-451.

Brumin M, Levy M, Ghanim M (2012) Transovarial transmission of Rickettsia spp. and organ-specific infection of the whitefly Bemisia tabaci. Applied and Environmental Microbiology 78, 5565-5574.

Buchner P (1965) Endosymbioses of Animals with Plant Microorganisms (Chichester, UK: John Wiley and Sons).

Burke GR, Moran NA (2011) Massive genomic decay in Serratia symbiotica, a recently evolved symbiont of aphids. Genome Biology and Evolution 3, 195-208.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI et al. (2010) QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7, 335-336.

Carini P, Steindler L, Beszteri S, Giovannoni SJ (2013) Nutrient requirements for growth of the extreme oligotroph 'Candidatus Pelagibacter ubique' HTCC1062 on a defined medium. The ISME Journal 7, 592-602.

Caspi-Fluger A, Inbar M, Mozes-Daube N, Katzir N, Portnoy V, Belausov E, Hunter MS, Zhori-Fein E (2012) Horizontal transmission of the insect symbiont Rickettsia is plant-mediated. Proceedings of the Royal Society of London B 279, 1791-6.

Chandler SM, Wilkinson TL, Douglas AE (2008) Impact of plant nutrients on the relationship between a herbivorous insect and its symbiotic bacteria. Proceedings of the Royal Society of London B 275, 565-570.

Chen DQ, Montllor CB, Purcell AH (2000) Fitness effects of two facultative endosymbiotic bacteria on the pea aphid, Acyrthosiphon pisum, and the blue alfalfa aphid, *A.* kondoi. Entomologia Experimentalis et Applicata 95, 315-323.

Chiel E, Gottlieb Y, Zchori-Fein E, Mozes-Daube N, Katzir N, Inbar M, Ghanim M

(2007) Biotype-dependent secondary symbiont communities in sympatric populations of Bemisia tabaci. Bulletin of Entomological Research 97, 407-413.

Chiel E, Zchori-Fein E, Inbar M, Gottlieb Y, Adachi-Hagimori T, et al. (2009) Almost there: transmission routes of bacterial symbionts between trophic levels. PLoS ONE 4, e4767.

Chu D, Gao CS, De Barro P, Zhang YJ, Wan FH, Khan IA (2011) Further insights into the strange role of bacterial endosymbionts in whitefly, Bemisia tabaci: comparison of secondary symbionts from biotypes B and Q in China. Bulletin of Entomological Research 101, 477-486.

Clark EL, Daniell TJ, Wishart J, Hubbard SF, Karley AJ (2012) How conserved are the bacterial communities associated with aphids? A detailed assessment of the Brevicoryne brassicae (Hemiptera: Aphididae) using 16S rDN*A*. Environmental Entomology 41, 1386-1397.

Costello EK, Gordon JI, Secor SM, Knight R (2010) Postprandial remodeling of the gut microbiota in Burmese pythons. The ISME Journal 4, 1375-1385.

Davidson EW, Rosell RC, Hendrix D (2000) Culturable bacteria associated with the whitefly, Bemisia anrgentifolii (Homoptera: Aleyrodidae). The Florida Entomologist 83, 159-171.

De Barro PJ, Liu SS, Boykin LM, Dinsdale AB (2011) Bemisia tabaci: a statement of species status. Annual Reviews of Entomology 56, 1-19.

Degnan PH, Leonardo TE, Cass BN, Hurwitz B, Stern D, Gibbs RA, Richards S, Moran NA (2010) Dynamics of genome evolution in facultative symbionts of aphids. Environmental Microbiology 12, 2060-2069.

Dethlefsen L, McFall-Ngai M, Relman DA (2007) An ecological and evolutionary perspective on human-microbe mutualism and disease. Nature 449, 811-818.

Dolling WR (1991) The Hemiptera (Oxford: Oxford University Press).

Douglas AE (1988) Experimental studies on the mycetome symbiosis in the leafhopper Euscelis incisus. Journal of Insect Physiology 34, 1043-1053.

Douglas AE (2003) The nutritional physiology of aphids. Advances in Insect Physiology 31, 73-140.

Everett KD, Thao M, Horn M, Dyszynski GE, Baumann P (2005) Novel chlamydiae in whiteflies and scale insects: endosymbionts 'Candidatus Fritschea bemisiae' strain Falk and 'Candidatus Fritschea eriococci' strain Elm. International Journal of Systematic and Evolutionary Microbiology 55, 1581-1587.

Ferrari J, Vavre F (2011) Bacterial symbionts in insects or the story of communities

affecting communities. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 366, 1389-1400.

Ferrari J, West JA, Via S, Godfray HC (2012) Population genetic structure and secondary symbionts in host-associated populations of the pea aphid complex. Evolution 66, 375-390.

Fieseler L, Horn M, Wagner M, Hentschel U (2004) Discovery of the novel candidate phylum "Poribacteria" in marine sponges. Applied and Environmental Microbiology 70, 3724-3732.

Gill EE, Brinkman FS (2011) The proportional lack of archaeal pathogens: Do viruses/phages hold the key? Bioessays 33, 248-254.

Gilles A, Meglecz E, Pech N, Ferreira S, Malausa T, Martin JF (2011).Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. BMC Genomics 12, 245.

Goodman AL, Kallstrom G, Faith JJ, Reyes A, Moore A, Dantas G, Gordon JI (2011) Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. Proceedings of the National Academy of Sciences USA 108, 6252-6257.

Gotelli NJ (2000) Null model analysis of species co-occurrence patterns. Ecology 81, 2606-2621.

Gotelli NJ, Entsminger G*L*. (2012) EcoSim 7.72.Acquired Intelligence, Inc.

http://www.uvm.edu/~ngotelli/EcoSim/EcoSim.html.

Gottlieb Y, Ghanim M, Chiel E, Gerling D, Portnoy V, Steinberg S, Tzuri G, Horowitz AR, Belausov E, Mozes-Daube N, et al. (2006) Identification and localization of a Rickettsia sp. in Bemisia tabaci (Homoptera: Aleyrodidae). Applied and Environmental Microbiology 72, 3646-3652.

Gottlieb Y, Ghanim M, Gueguen G, Kontsedalov S, Vavre F, Fleury F, Zchori-Fein E (2008) Inherited intracellular ecosystem: symbiotic bacteria share bacteriocytes in whiteflies. FASEB Journal 22, 2591-2599.

Gottlieb Y, Zchori-Fein E, Mozes-Daube N, Kontsedalov S, Skaljac M, Brumin M, Sobol I, Czosnek H, Vavre F, Fleury F, Ghanim M. (2010) The transmission efficiency of tomato yellow leaf curl virus by the whitefly Bemisia tabaci is correlated with the presence of a specific symbiotic bacterium species. Journal of Virology 84, 9310-9317.

Grenier AM, Nardon C, Rahbe Y (1994) Observations on the micro-organisms occurring in the gut of the pea aphid Acyithosiphon pisum. Entomologia Experimentalis et Applicata 70, 91-96.

Gueguen G, Vavre F, Gnankine O, Peterschmitt M, Charif D, Chiel E, Gottlieb Y, Ghanim M, Zchori-Fein E, Fleury F (2010) Endosymbiont metacommunities, mtDNA diversity and the evolution of the Bemisia tabaci (Hemiptera: Aleyrodidae) species complex. Molecular Ecology

Hail D, Lauziere I, Dowd SE, Bextine B (2011) Culture independent survey of the microbiota of the glassy-winged sharpshooter (Homalodisca vitripennis) using 454 pyrosequencing. Environmental Entomology 40, 23-29.

Haynes S, Darby AC, Daniell TJ, Webster G, Van Veen FJ, Godfray HC, Prosser JI, Douglas AE (2003) Diversity of bacteria associated with natural aphid populations. Applied and Environmental Microbiology 69, 7216-7223.

Himler AG, Adachi-Hagimori T, Bergen JE, Kozuch A, Kelly SE, Tabashnik BE, Chiel E, Duckworth VE, Dennehy TJ, Zchori-Fein E, Hunter MS (2011) Rapid spread of a bacterial symbiont in an invasive whitefly is driven by fitness benefits and female bias. Science 332, 254-256.

Hong PY, Wheeler E, Cann IK, Mackie RI (2011) Phylogenetic analysis of the fecal microbial community in herbivorous land and marine iguanas of the Galapagos Islands using 16S rRNA-based pyrosequencing. The ISME Journal 5, 1461-1470.

Human Microbiome Project (2012) Structure, function and diversity of the healthy human microbiome. Nature 486, 207-214.

Huse SM, Welch DM, Morrison HG, Sogin ML (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environmental Microbiology 12, 1889-1898.

Kohl KD (2012) Diversity and function of the avian gut microbiota. Journal of Comparative Physiology B 182, 591-602.

Kong HH, Oh J, Deming C, Conlan S, Grice EA, Beatson MA, Nomicos E, Polley EC, Komarow HD, Program NCS et al. (2012) Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. Genome Research 22, 850-859.

Kunin V, Engelbrektson A, Ochman H, Hugenholtz P (2010) Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. Environmental Microbiology 12, 118-123.

Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI (2008) Worlds within worlds: evolution of the vertebrate gut microbiota. Nature Reviews Microbiology 6, 776-788.

Li X, Nan X, Wei C, He H (2012) The gut bacteria associated with Camponotus japonicus Mayr with culture-dependent and DGGE methods. Current Microbiology 65, 610-616.

Martinson VG, Danforth BN, Minckley RL, Rueppell O, Tingek S, Moran NA (2011) A simple and distinctive microbiota associated with honey bees and bumble bees. Molecular Ecology 20, 619-628.

McCutcheon JP, Moran NA (2012) Extreme genome reduction in symbiotic bacteria. Nature Reviews Microbiology 10, 13-26.

McDonald R, Schreier HJ, Watts JE (2012) Phylogenetic analysis of microbial communities in different regions of the gastrointestinal tract in Panaque nigrolineatus, a wood-eating fish. PLoS One 7, e48018.

McFall-Ngai M (2007) Adaptive immunity: care for the community. Nature 445, 153.

Moran NA, Wernegreen JJ (2000) Lifestyle evolution in symbiotic bacteria: insights from genomics. Trends in Ecology and Evolution 15, 321-326.

Moran NA, Dunbar HE (2006) Sexual acquisition of beneficial symbionts in aphids. Proceedings of the National Academy of Sciences 103, 12803-6

Nirgianaki A, Banks GK, Frohlich DR, Veneti Z, Braig HR, Miller TA, Bedford ID, Markham PG, Savakis C, Bourtzis K (2003) *Wolbachia* infections of the whitefly Bemisia tabaci. Current Microbiology 47, 93-101.

Rana VS, Singh ST, Priya NG, Kumar J, Rajagopal R (2012) Arsenophonus GroEL interacts with CLCuV and is localized in midgut and salivary gland of whitefly B. tabaci. PLoS One 7, e42168.

Rawls JF, Mahowald MA, Ley RE, Gordon JI (2006) Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. Cell 127, 423-433.

Robinson CJ, Schloss P, Ramos Y, Raffa K, Handelsman J (2010) Robustness of the bacterial community in the cabbage white butterfly larval midgut. Microbial Ecology 59, 199-211.

Sachs JL, Skophammer RG, Regus JU (2011) Evolutionary transitions in bacterial symbiosis. Proceedings of the National Academy of Sciences USA 108 Suppl 2, 10800-10807.

Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative C(T) method. Nature Protocols 3, 1101-1108.

Shatters RG, Powell CA, Boykin LM, Liansheng H, McKenzie CL (2009) Improved DNA barcoding method for Bemisia tabaci and related Aleyrodidae: development of universal and Bemisia tabaci biotype-specific mitochondrial cytochrome c oxidase I polymerase chain reaction primers. Journal of Economic Entomology 102, 750-758.

Singh S, Eldin C, Kowalczewska M, Raoult D (2013) Axenic culture of fastidious and

intracellular bacteria. Trends in Microbiology 21, 92-99.

Singh ST, Priya NG, Kumar J, Rana VS, Ellango R, Joshi A, Priyadarshini G, Asokan R, Rajagopal R (2012) Diversity and phylogenetic analysis of endosymbiotic bacteria from field caught Bemisia tabaci from different locations of North India based on 16S rDNA library screening. Infection, Genetics and Evolution 12, 411-419.

Sinisterra XH, McKenzie CL, Hunter WB, Powell CA, Shatters RG (2005) Differential transcriptional activity of plant-pathogenic begomoviruses in their whitefly vector (Bemisia tabaci, Gennadius: Hemiptera Aleyrodidae). Journal of General Virology 86, 1525-1532.

Sloan DB, Moran NA (2012) Endosymbiotic bacteria as a source of carotenoids in whiteflies. Biology Letters 8, 986-989.

Stone L, Roberts A (1990) The checkerboard score and species distributions. Oecologia 85, 74-79.

Sun Y, Cai Y, Huse SM, Knight R, Farmerie WG, Wang X, Mai V (2012) A large-scale benchmark study of existing algorithms for taxonomy-independent microbial community analysis. Briefings in Bioinformatics 13, 107-121.

Sundquist A, Bigdeli S, Jalili R, Druzin ML, Waller S, Pullen KM, El-Sayed YY, Taslimi MM, Batzoglou S, Ronaghi M (2007) Bacterial flora-typing with targeted, chip-based pyrosequencing. BMC Microbiology 7, 108.

Tamames J, Abellan JJ, Pignatelli M, Camacho A, Moya A (2010) Environmental distribution of prokaryotic taxa. BMC Microbiology 10, 85.

Tay WT, Evans GA, Boykin LM, De Barro PJ (2012) Will the real Bemisia tabaci please stand up? PLoS One 7, e50550.

Thao ML, Baumann P (2004a). Evidence for multiple acquisition of Arsenophonus by whitefly species (Sternorrhyncha: Aleyrodidae). Current Microbiology 48, 140-144.

Thao ML, Baumann P (2004b). Evolutionary relationships of primary prokaryotic endosymbionts of whiteflies and their hosts. Applied and Environmental Microbiology 70, 3401-3406.

Thierry M, Becker N, Hajri A, Reynaud B, Lett JM, Delatte H (2011) Symbiont diversity and non-random hybridization among indigenous (Ms) and invasive (B) biotypes of Bemisia tabaci. Molecular Ecology 20, 2172-2187.

Tsuchida T, Koga R, Fukatsu T (2004) Host plant specialization governed by facultative symbiont. Science 303, 1989.

Turner S, Pryer KM, Miao VP, Palmer JD (1999) Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence

analysis. Journal of Eukaryotic Microbiology 46, 327-338.

Walter J, Britton RA, Roos S (2011) Host-microbial symbiosis in the vertebrate gastrointestinal tract and the Lactobacillus reuteri paradigm. Proceedings of the National Academy of Sciences USA 108 Suppl 1, 4645-4652.

Weeks AR, Velten R, Stouthamer R (2003) Incidence of a new sex-ratio-distorting endosymbiotic bacterium among arthropods. Proceedings of the Royal Society of London B 270, 1857-1865.

Wong CN, Ng P, Douglas AE (2011) Low-diversity bacterial community in the gut of the fruitfly Drosophila melanogaster. Environmental Microbiology 13, 1889-1900.

APPENDIX D

16SpeB: DEFINING BACTERIAL SPECIES BOUNDARIES BY MINIMUM INTRA-SPECIES 16S RRNA GENE SEQUENCE IDENTITY[7]

Abstract

Summary: 16SpeB (16S rRNA-based Species Boundary) is a package of Perl programs that evaluates total sequence variation of a bacterial species at the levels of the whole 16S rRNA gene sequences or single hypervariable (V) regions, using publicly-available sequences. The 16SpeB pipelines filter sequences from duplicated strains and of low quality, extracts a V region of interest using general primer sequences, and calculates sequence percentage identity (%ID) through all possible pairwise alignments. *Results*: The minimum %ID of 16S rRNA gene sequences for 15 clinically-important bacterial species, as determined by 16SpeB, ranged from 82.6% to 99.8%. The relationship between minimum %ID of V2/V6 regions and full-gene sequences varied among species, indicating that %ID species limits should be resolved independently for each region of the 16S rRNA gene and bacterial species. *Availability*: 16SpeB and user manual are freely available for download from http://www.angeladouglaslab.com/content/view/16SpeB.html. It currently supports the Linux operation system.

[7] Article in preparation for journal submission by Wong, ACN., Ng, P. and Douglas, AE.
Ng, P contributed to the scripting of the 16SpeB.
Supplementary materials available upon request.

Introduction

16S rRNA-based identification of bacteria from clinical and environmental samples can be a challenging task. Although 97% sequence identity (%ID) of the full length 16S gene is widely used as the threshold defining bacterial species (Clarridge, 2004; Drancourt and Raoult, 2005; Janda and Abbott, 2007; Petti, 2007), this criterion has been validated for very few species of bacteria (Drancourt, *et al.*, 2004; Ueda, *et al.*, 1999). Compounding the uncertainty about the validity of the 97% ID threshold for species identification, there is increasing demand for taxonomic identification of bacteria from relatively short sequence reads (<400 bp) of single hypervariable (V) regions of the 16S gene, usually V2 or V6 (Bowen, *et al.*, 2011; De Filippo, *et al.*, 2010; Guss, *et al.*, 2011; Kirchman, *et al.*, 2010; Ravussin, *et al.*, 2011; Wu, *et al.*, 2011).

16SpeB is an analytical tool designed to identify the range of 16S %ID encompassed by individual bacterial species based on known 16S rRNA gene sequence variation. This tool provides accurate taxonomic identification of bacteria, using both (near)-full 16S sequences and short reads obtained by 454 or Illumina sequencing of 16S rRNA gene amplicons.

Usage

16S rRNA sequences from three 16S rRNA databases can be downloaded from *Greengenes* (DeSantis, *et al.*, 2006) *Ribosomal Database Project* (Cole, *et al.*, 2007) and *Silva* (Pruesse, *et al.*, 2007). 16SpeB allows users to trim the (near-)full 16S rRNA sequences to their preferred length. It can also extract the sequences of the V2 and V6 regions, which are widely used in 454 sequencing studies, by reference to the general

primer sets 27F-338R and 784F-1061R, respectively. Sequences that fail to satisfy the two following conditions are removed: (1) <2 bp mismatches with the general 16S primers (i.e. conserved regions of the 16S gene), and (2) relative coordinates of matched primers are within +/- 50 bp from the relative coordinates of the literature. The V2 region is trimmed to 270 bp upstream of the 338R primer. 16SpeB conducts all possible pairwise sequence comparisons by aligning all pairwise sequences using Needleman-Wunsch alignment algorithm with match/mismatch score of 1/-2 and affine gap penalty open/extension of -5/-2. The minimum and 95% quantile %ID are computed for each species, providing a measure of the total known sequence variation that defines the species.

Application of 16SpeB

16SpeB was initially developed to identify species limits of *Acetobacter* and *Lactobacillus* in a pyroseqeuncing analysis of the gut microbiota of *Drosophila melanogaster* (Wong, *et al.*, 2011). Here we extend the application of 16SpeB to determine the %ID of (near-)full 16S rRNA genes that defines the species boundary of 15 clinically-important bacterial species (listed in Supplementary Data Set 1); and to determine the %ID of the V2 and V6 regions widely used in pyrosequencing studies that correlate with this species boundary. The 15 bacterial species were selected on the criteria that a broad range of publicly-available sequences (3 to 454) and phylogenetic diversity (including representatives of Actinobacteria, Bacteroidetes, Chlamydiae, Firmicutes and Proteobacteria) were represented. In total, 1,296 sequences were analyzed. The minimum %ID of (near-) full 16S sequences varied from 99.8% (*Neisseria gonorrhoeae*) to 82.6% (*Staphylococcus aureus*) (Table 7.1). Just two

(13%) of the 15 species had minimum %ID close to predicted 97% threshold for species boundary (*Neisseria meningitidis* 97.0%, and *Listeria monocytogenes* 97.1%); and 11 (73%) species deviated from 97% by more than one percentage point. Values of the 95% quantile are provided in Table 7.1 and may prove to be more useful than minimum %ID for some species, e.g. *Staphylococcus aureus*, where the minimum %ID is suspected to be artefactually low (possibly through mis-identification).

Table 7.1 The minimum and 95% quantile %ID of the (near-)full 16S rRNA gene, and the V2 and V6 regions of the 15 clinically-important bacteria.

Species	Number of sequences (pairs)	Minimum %ID			95% quantile %ID		
		(near)-full 16S	V2	V6	(near)-full 16S	V2	V6
Bacteroides fragilis	345 (59340)	0.928	0.899	0.928	0.978	0.959	0.973
Clostridium bifermentans	58 (1653)	0.926	0.928	0.787	0.967	0.967	0.893
Chlamydia trachomati	15 (105)	0.950	0.921	0.954	0.958	0.942	0.959
Corynebacterium diphtheriae	10 (45)	0.942	0.920	0.912	0.946	0.934	0.918
Haemophilus influenzae	92 (4186)	0.901	0.831	0.891	0.951	0.925	0.907
Helicobacter pylori	59 (1711)	0.949	0.895	0.939	0.977	0.960	0.961
Listeria monocytogenes	26 (325)	0.971	0.939	0.966	0.974	0.953	0.969
Mycobacterium leprae	4 (6)	0.984	0.967	0.992	0.984	0.967	0.992
Mycobacterium tuberculosis	10 (45)	0.984	0.982	0.988	0.989	0.985	0.992
Mycoplasma hominis	6 (15)	0.899	0.949	0.681	0.899	0.949	0.681
Neisseria gonorrhoeae	3 (3)	0.998	1.000	1.000	0.999	1.000	1.000
Neisseria meningitidis	133 (8778)	0.970	0.927	0.962	0.990	0.978	0.981
Staphylococcus aureus	454 (102831)	0.826	0.604	0.843	0.980	0.981	0.973
Streptococcus pneumoniae	47 (1081)	0.980	0.938	0.977	0.986	0.963	0.985
Yersinia pestis	34 (561)	0.979	0.960	0.966	0.986	0.967	0.977

As anticipated, the minimum %ID of both the V2 and V6 regions varied positively with %minimum ID of the (near-) full sequence of the 16S genes (Supplementary Figure 1). The relationships were not, however, tight indicating that the rates of sequence evolution of individual V regions are not closely correlated to each other or to other regions of the 16S gene. Our data suggest the 97% threshold is not a reliable index of the taxonomic species limit, and there is no simple linear relationship linking the minimum %ID of the V2 or V6 sequences to the (near-) full 16S sequence across multiple bacterial species.

We conclude that the %ID species limits should be resolved independently for each region of the 16S rRNA gene and each bacterial species. Therefore, 16SpeB can serve as an important tool that facilitates accurate taxonomic identification and proper interpretation of 16S rRNA gene pyrosequencing data.

References

Bowen, J.*L.*, et al. (2011) Microbial community composition in sediments resists perturbation by nutrient enrichment, ISME J, 5, 1540-1548.

Clarridge, J.E., 3rd (2004) Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases, Clin Microbiol Rev, 17, 840-862, table of contents.

Cole, J.R., et al. (2007) The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data, Nucleic Acids Res, 35, D169-172.

De Filippo, C., et al. (2010) Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa, Proc Natl Acad Sci U S A, 107, 14691-14696.

DeSantis, T.Z., et al. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB, Appl Environ Microbiol, 72, 5069-5072.

Drancourt, M., Berger, P. and Raoult, *D.* (2004) Systematic 16S rRNA gene sequencing of atypical clinical isolates identified 27 new bacterial species associated with humans, J Clin Microbiol, 42, 2197-2202.

Drancourt, M. and Raoult, *D.* (2005) Sequence-based identification of new bacteria: a proposition for creation of an orphan bacterium repository, J Clin Microbiol, 43, 4311-4315.

Guss, *A.*M., et al. (2011) Phylogenetic and metabolic diversity of bacteria associated with cystic fibrosis, ISME J, 5, 20-29.

Janda, J.M. and Abbott, S.*L.* (2007) 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls, J Clin Microbiol, 45, 2761-2764.

Kirchman, *D.L.*, Cottrell, M.T. and Lovejoy, C. (2010) The structure of bacterial communities in the western Arctic Ocean as revealed by pyrosequencing of 16S rRNA genes, Environ Microbiol, 12, 1132-1143.

Petti, C.*A.* (2007) Detection and identification of microorganisms by gene amplification and sequencing, Clin Infect Dis, 44, 1108-1114.

Pruesse, E., et al. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB, Nucleic Acids Res, 35, 7188-7196.

Ravussin, Y., et al. (2011) Responses of Gut Microbiota to Diet Composition and Weight Loss in Lean and Obese Mice, Obesity (Silver Spring).

Ueda, K., et al. (1999) Two distinct mechanisms cause heterogeneity of 16S rRNA, J

Bacteriol, 181, 78-82.

Wong, C.N., Ng, P. and Douglas, *A*.E. (2011) Low-diversity bacterial community in the gut of the fruitfly Drosophila melanogaster, Environ Microbiol, 13, 1889-1900.

Wu, G.*D*., et al. (2011) Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes, Science. 334, (6052), 105-108.

APPENDIX E

SUPPLEMENTARY NUTRITIONAL INDICES FIGURES FOR CHAPTER 4

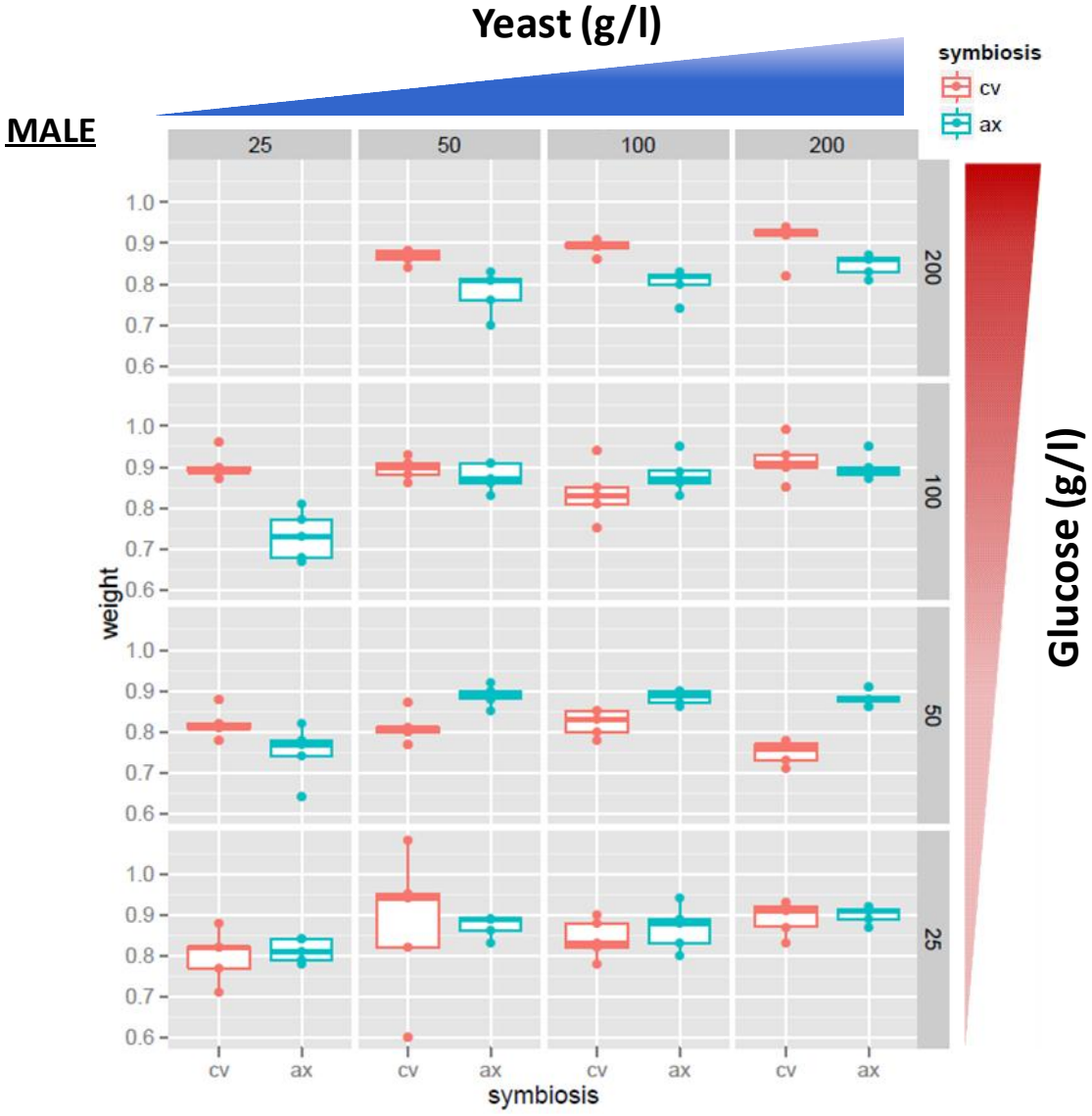


Figure 7.1 Impact of dietary yeast and glucose (25-200 g l^{-1}) on weight (mg) of conventional and axenic male *Drosophila*.

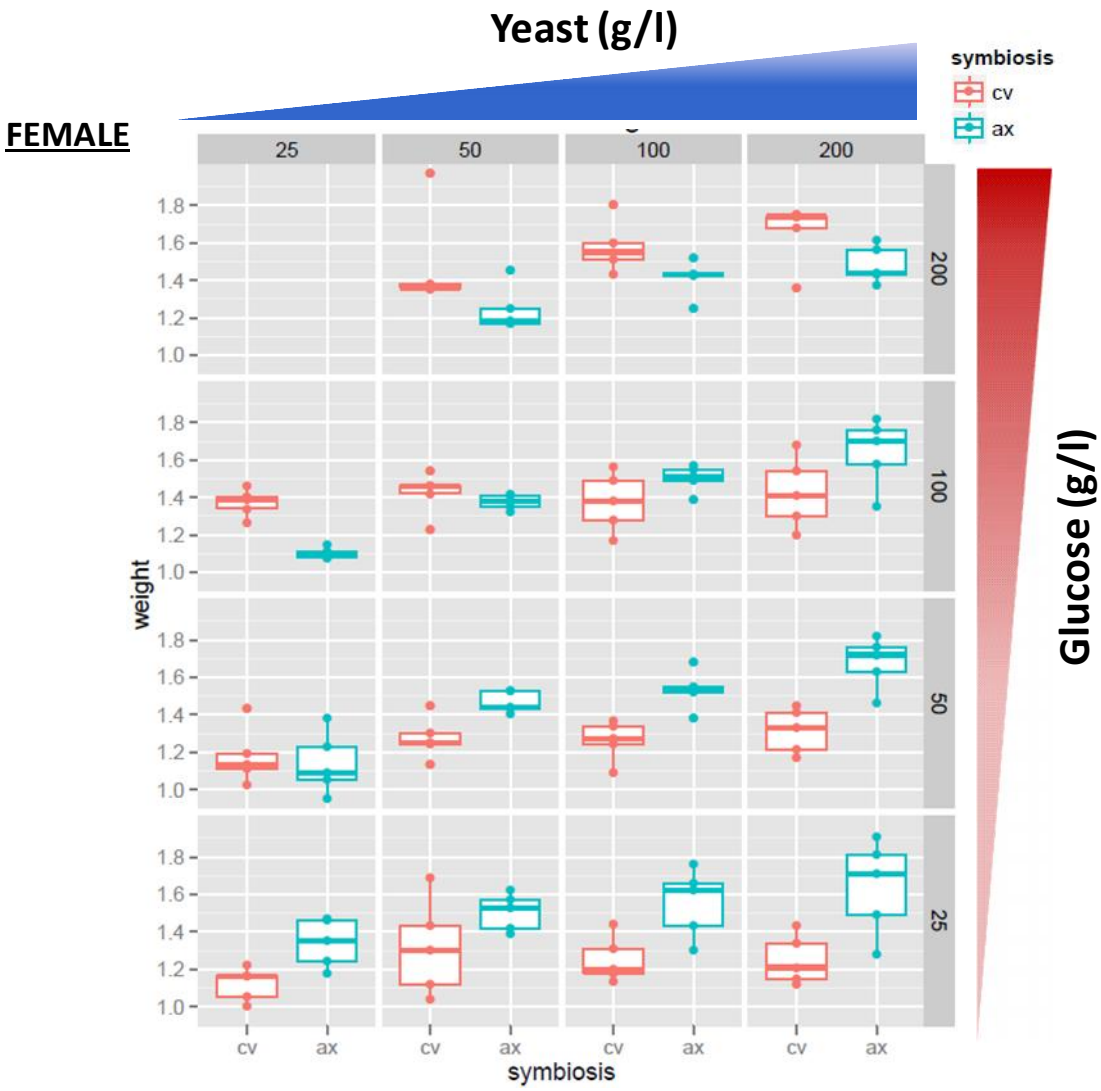


Figure 7.2 Impact of dietary yeast and glucose (25-200 g l^{-1}) on weight (mg) of conventional and axenic female *Drosophila*.

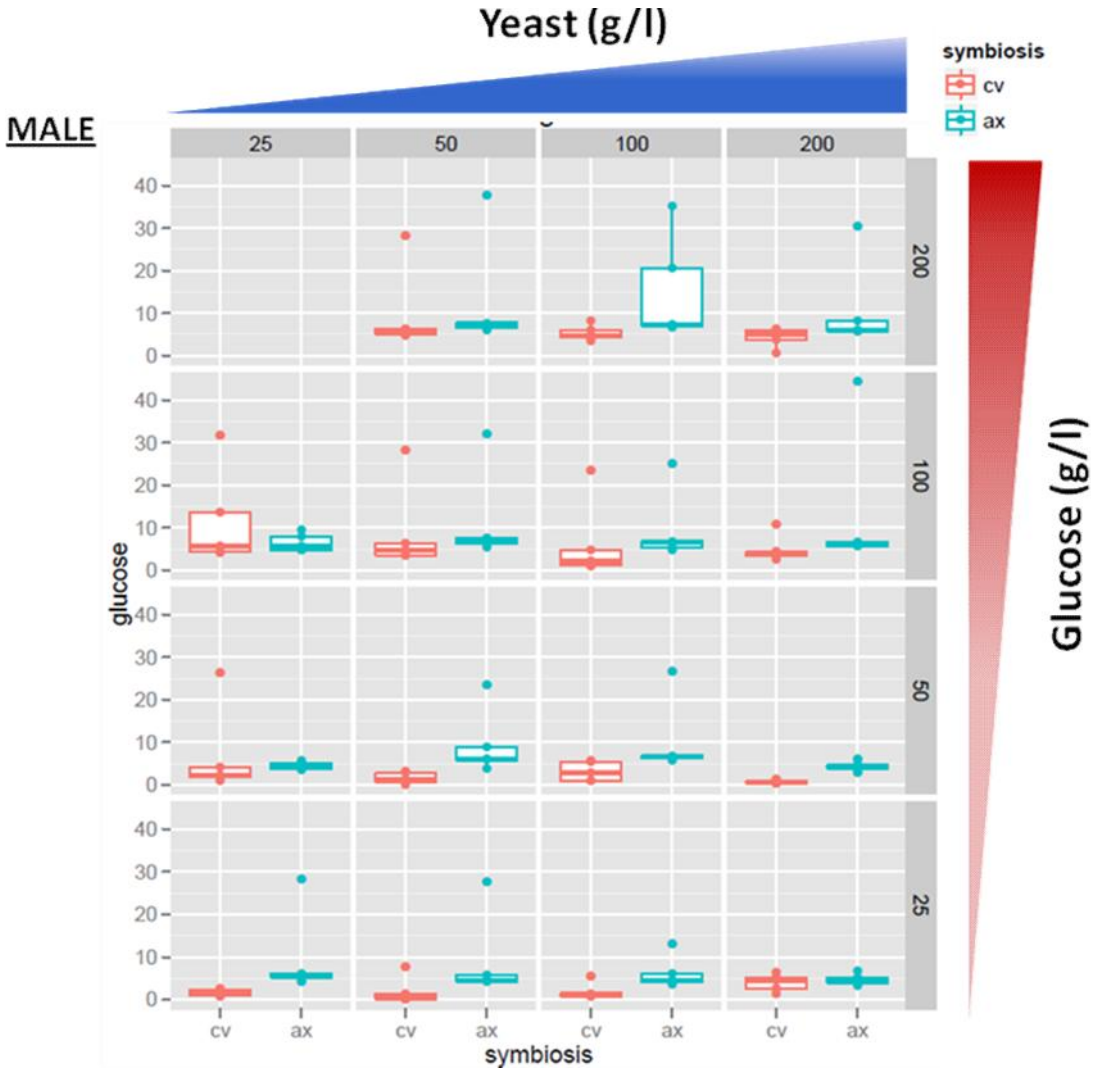


Figure 7.3 Impact of dietary yeast and glucose (25-200 g l^{-1}) on glucose content ($\mu g mg^{-1}$) of conventional and axenic male *Drosophila*.

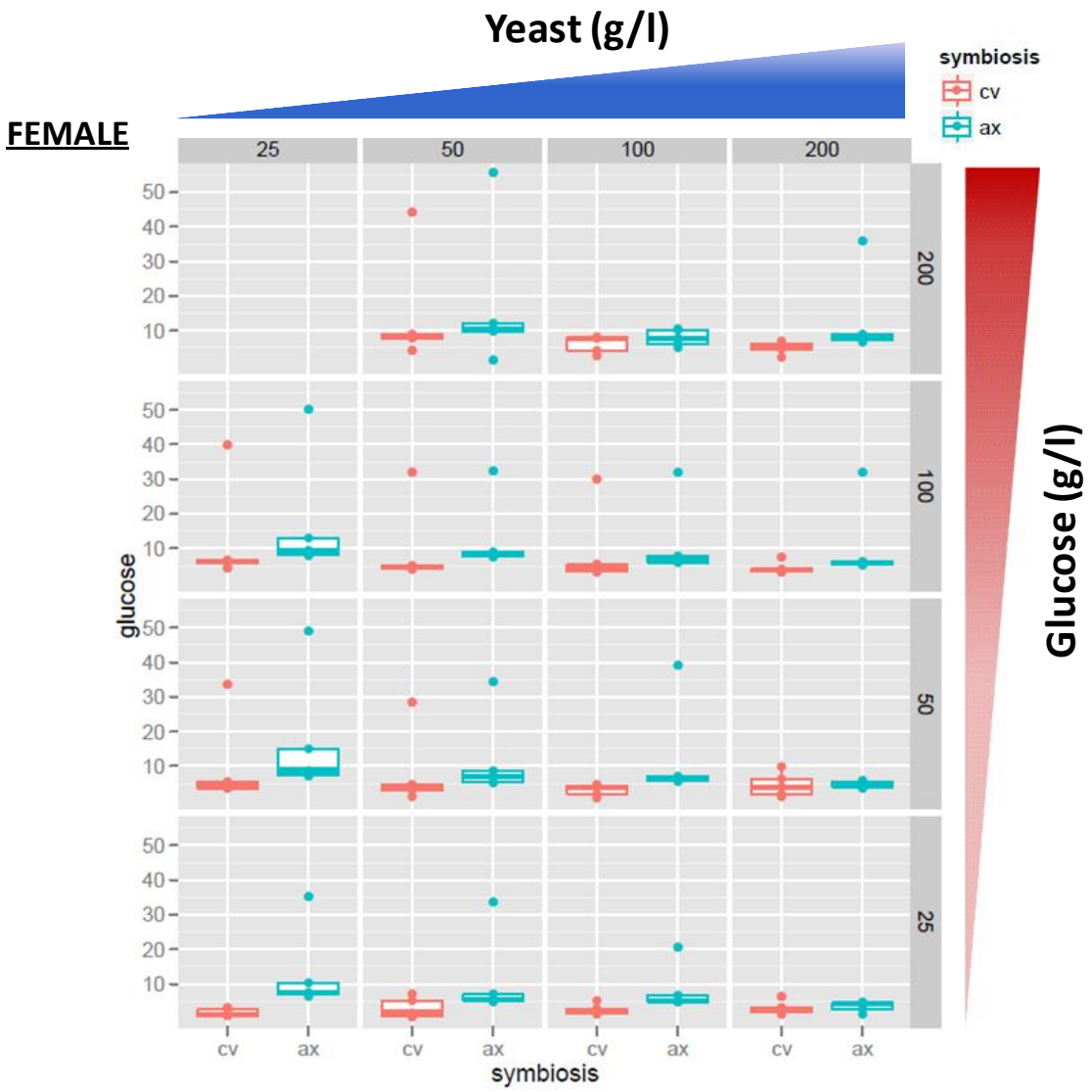


Figure 7.4 Impact of dietary yeast and glucose (25-200 g l^{-1}) on glucose content ($\mu g mg^{-1}$) of conventional and axenic female *Drosophila*.

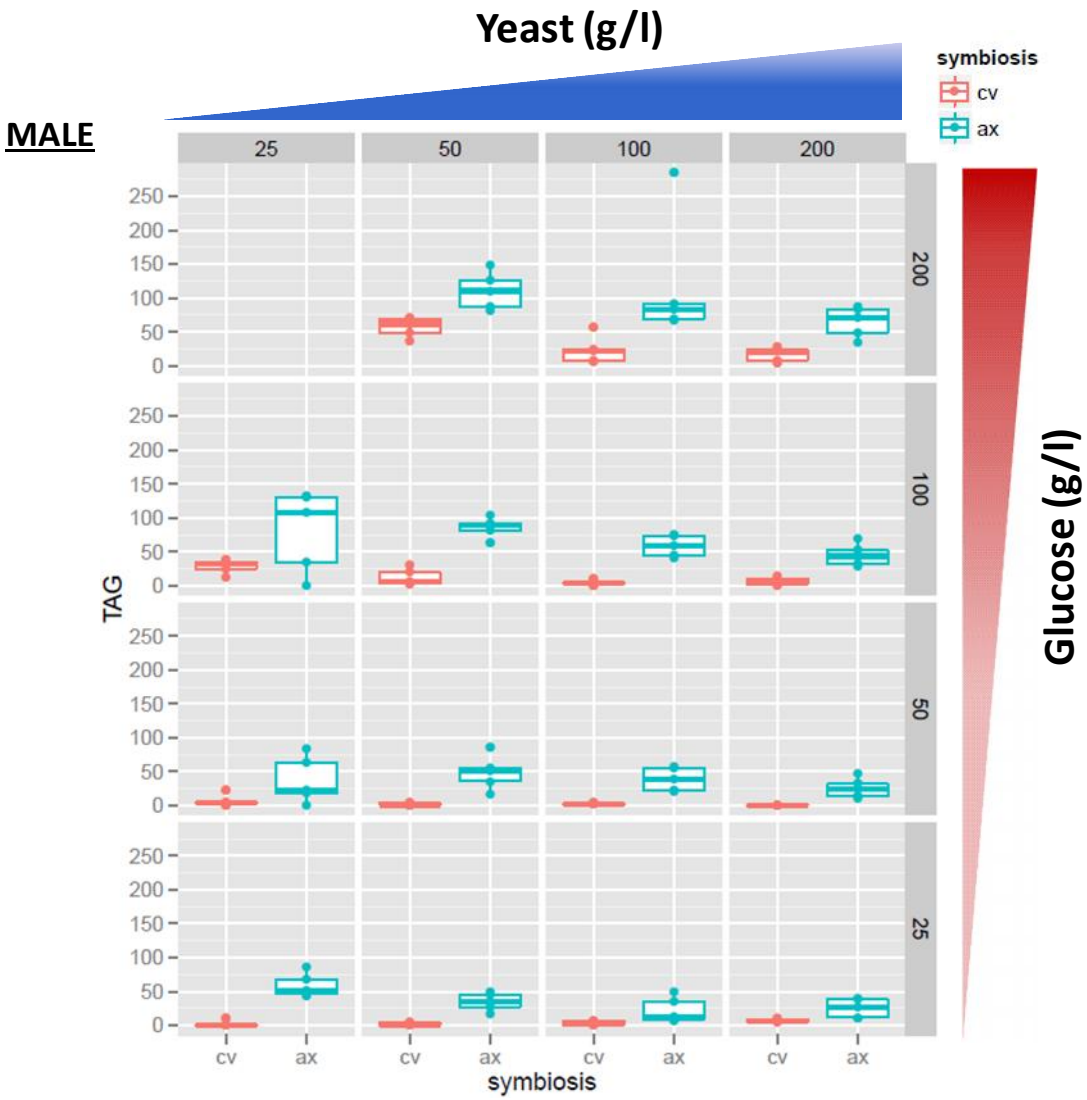


Figure 7.5 Impact of dietary yeast and glucose (25-200 g l^{-1}) on triglycerides (TAG) content ($\mu g mg^{-1}$) of conventional and axenic male *Drosophila*.

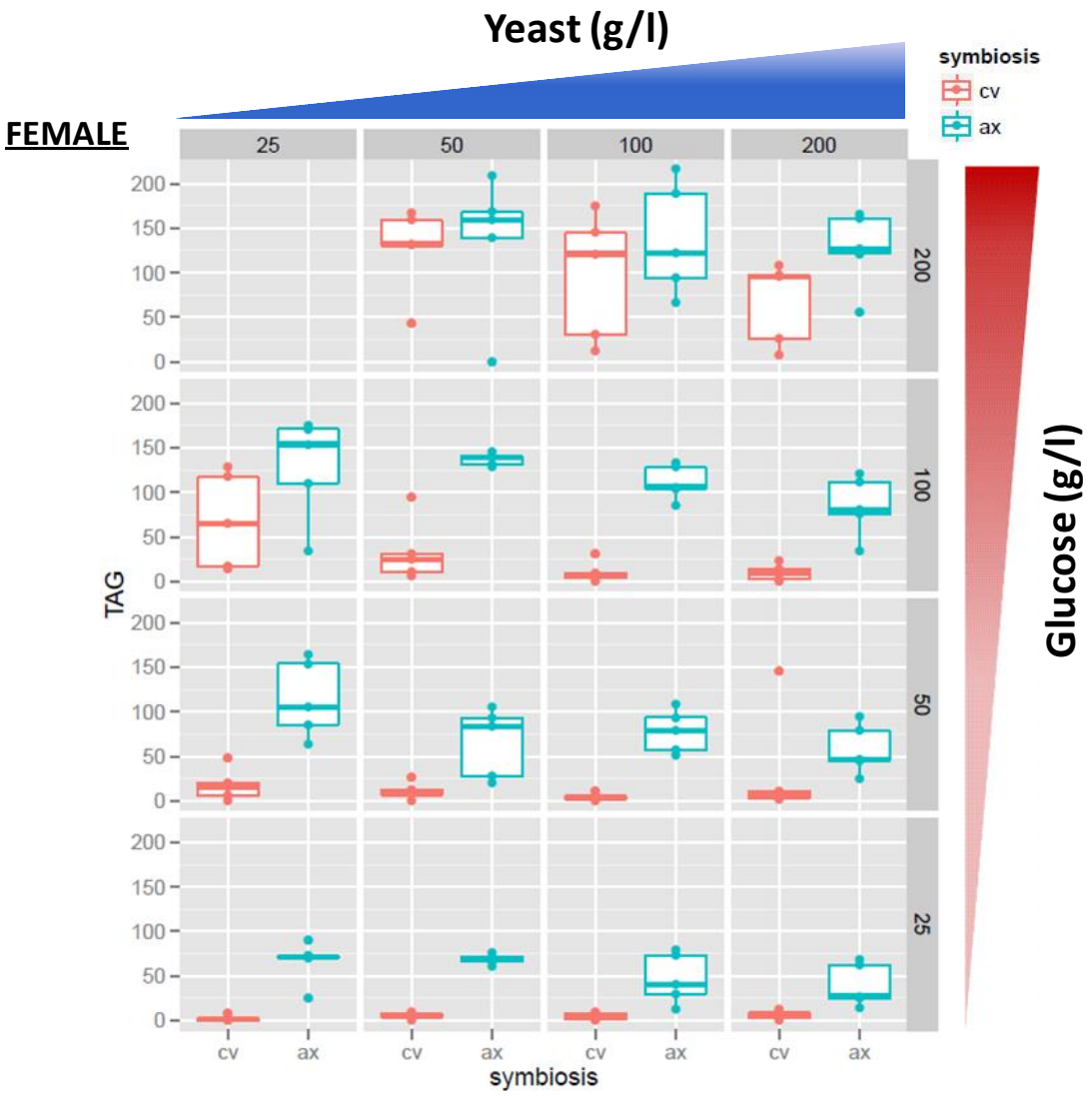


Figure 7.6 Impact of dietary yeast and glucose (25-200 g l^{-1}) on triglycerides (TAG) content ($\mu g mg^{-1}$) of conventional and axenic female *Drosophila*.

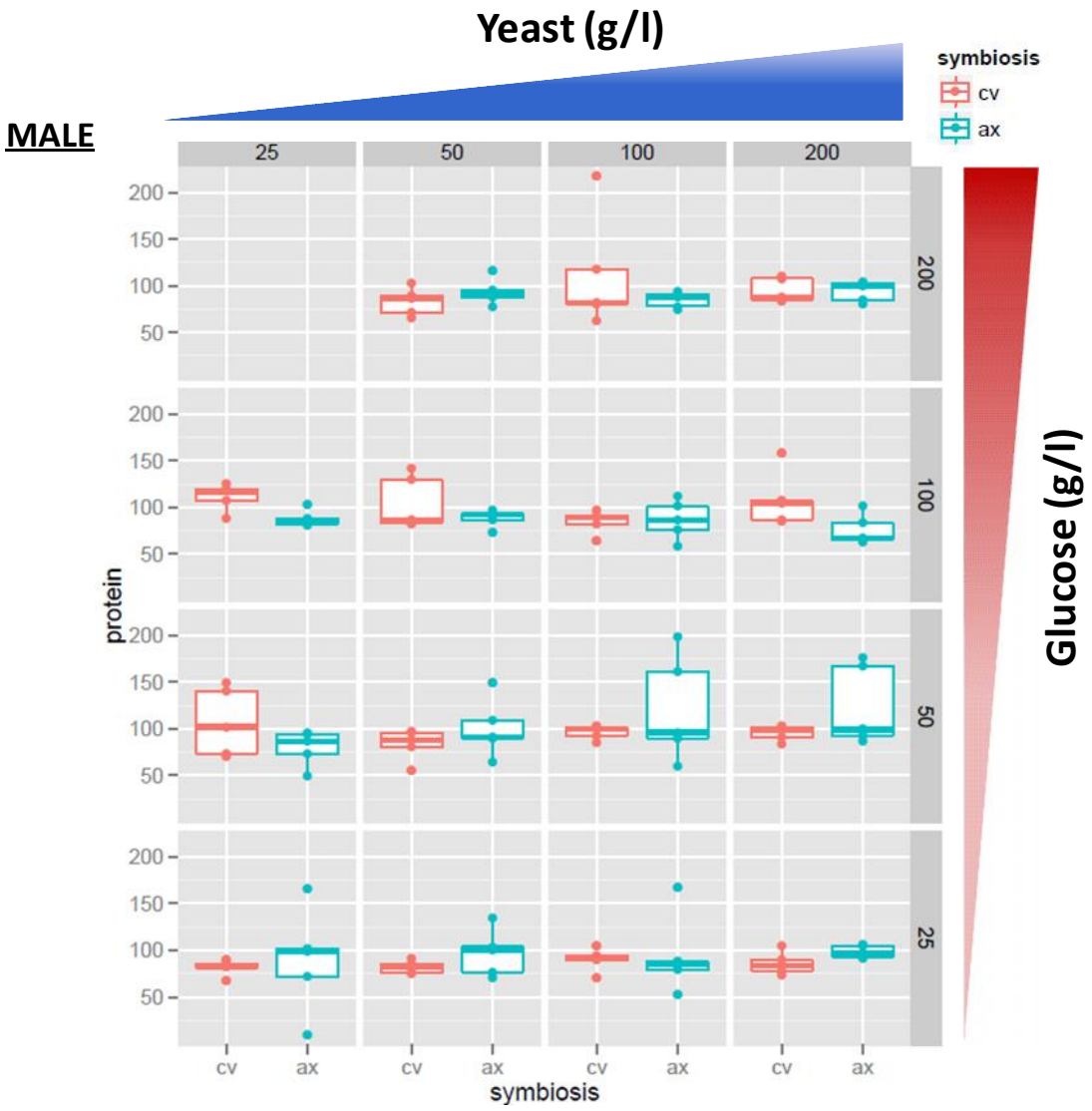


Figure 7.7 Impact of dietary yeast and glucose (25-200 g l^{-1}) on protein content ($\mu g mg^{-1}$) of conventional and axenic male *Drosophila*.

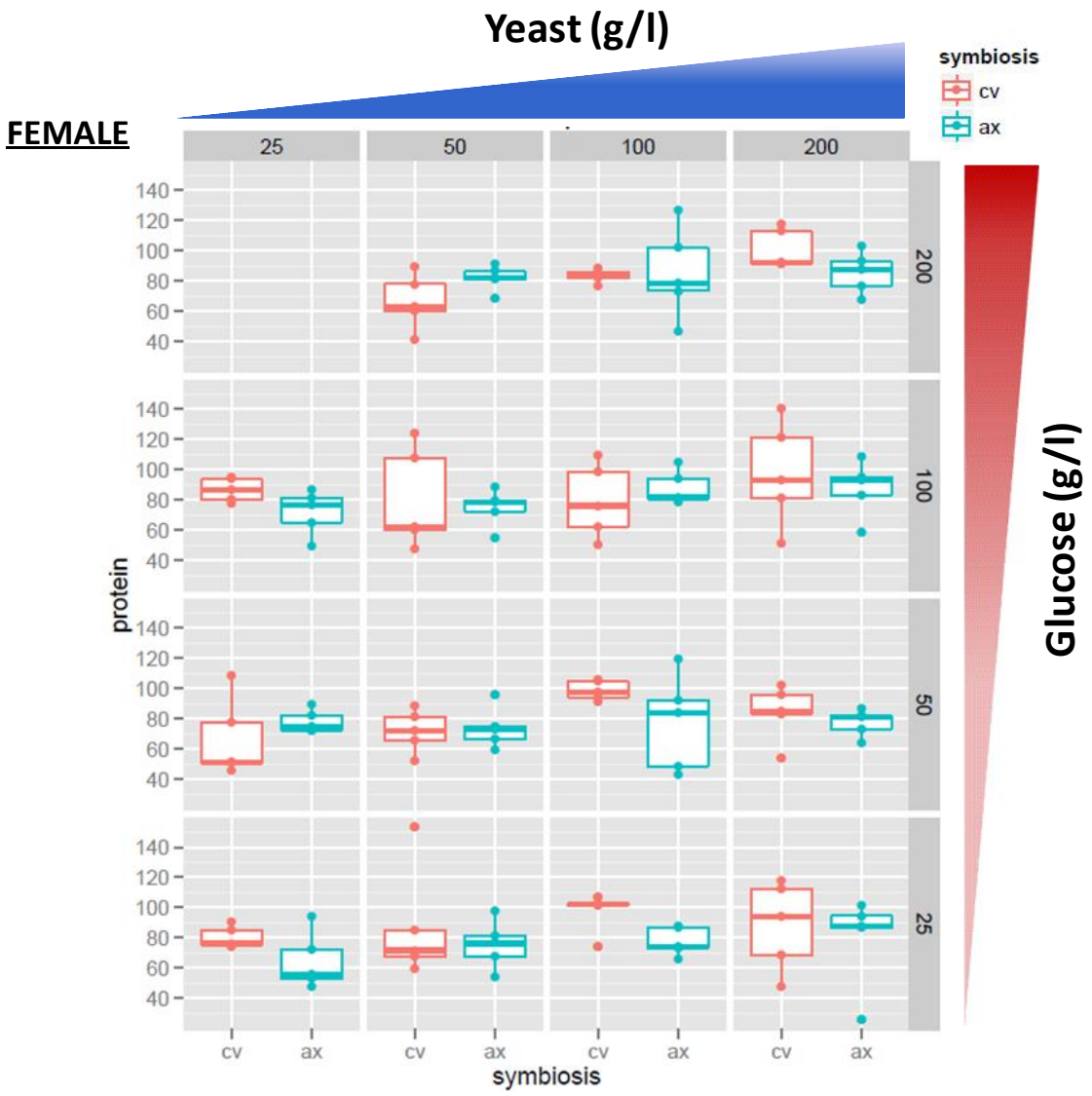


Figure 7.8 Impact of dietary yeast and glucose (25-200 g l^{-1}) on protein content ($\mu g mg^{-1}$) of conventional and axenic female *Drosophila*.