Towards an integrated understanding of gut microbiota using insects as model systems

Mathieu Pernice \textsuperscript{a,b}, Stephen J Simpson\textsuperscript{a,c}, Fleur Ponton\textsuperscript{a,c,*}

\textsuperscript{a}School of Biological Sciences, The University of Sydney, NSW 2006, Australia
\textsuperscript{b}Plant Functional Biology and Climate Change Cluster, University of Technology Sydney, New South Wales 2007, Australia.
\textsuperscript{c}Charles Perkins Centre, The University of Sydney, NSW 2006, Australia

*Corresponding author

Doctor Fleur Ponton
School of Biological Sciences
Heydon-Laurence Building, A08
The University of Sydney
New South Wales 2006
AUSTRALIA
ABSTRACT

Metazoans form symbioses with microorganisms that synthesize essential nutritional compounds and increase their efficiency to digest and absorb nutrients. Despite the growing awareness that microbes play key roles on metabolism, health and development of metazoans, symbiotic relationships within the gut are far from fully understood. Perhaps the most important obstacle to understanding these symbiotic relationships resides in the high diversity of bacterial communities living within the gut of most vertebrates. In this regard, insects, which generally harbor a lower microbial diversity within their gut, offer an interesting alternative to vertebrates and have recently emerged as potential model systems to study these interactions. In this review, we give a brief overview of the characteristics of the gut microbiota in insects in terms of low diversity but high variability at intra- and interspecific levels and we investigate some of the ecological and methodological factors that might explain such variability. We then emphasize how studies integrating a vast array of techniques and disciplines have the potential to provide a groundbreaking understanding of the biology of this micro eco-system.
Extracting essential nutritious components from food can be challenging. Metazoans have partly faced this challenge by forming symbioses with microorganisms that both synthesize essential nutritional compounds and increase the efficiency of nutrient digestion and absorption (Fraune and Bosch, 2010; Moran, 2007). In insects, nutritional symbioses can be split into two main categories: (i) intracellular associations, that are generally found in arthropods with restricted diets such as blood and plant sap and involve only few types of symbionts, and (ii) extracellular associations, that are more common among metazoan and involve a complex community of symbionts that generally live within the gut lumen. Symbionts can serve a wide range of nutritional functions, from mobilizing stored nitrogen to contributing essential amino acids (Brune and Ohkuma, 2011; Douglas, 2009; Feldhaar, 2011; Kaufman and Klug, 1991), and hosts often rely on symbiotic microorganisms to supply nutrients required for viability and fertility (Dillon and Dillon, 2003; Douglas, 2010; Moran and Baumann, 2000).

In recent decades, numerous investigations have been devoted to understanding the metabolic roles of associated microorganisms (Douglas, 2009; Moran, 2007) with a particular emphasizes on the bacteria that compose the gut microbiota (Dillon and Dillon, 2003; Nicholson et al., 2012; Ryu et al., 2008; Storelli et al., 2011). In this context, it has been shown, for instance, that the gut microbiota can contribute up to 70% of a vertebrate’s energy needs (Flint et al., 2008). However, despite the growing awareness that microbes play key roles on metabolism, health and development of metazoans (Fraune and Bosch, 2010; Lee and Brey, 2013; Maslowski and Mackay, 2011), symbiotic relationships within the gut are far from fully understood (Engel and Moran, 2013).
Perhaps the most important obstacle to understanding these symbiotic relationships resides in the high diversity of bacterial communities living within the gut of most vertebrates. In this regard, insects, which generally harbour a lower microbial diversity within their gut, offer an interesting alternative to vertebrates and have recently emerged as potential model systems to study these interactions. Insects are not only tractable and easy to manipulate, they also offer substantial genetic resources allowing investigations of conversed metabolic and immune pathways. However, capturing the properties of insect gut microbiota has been challenging so far due to a high variability in composition between individuals and closely related species. Here, we give a brief overview of the characteristics of the gut microbiota in insects in terms of low diversity but high variability at intra- and interspecific levels and we investigate some of the ecological and methodological factors that might explain such variability. We then emphasize how studies integrating the latest technological advances from molecular biology and stable isotopes based-technics can improve our understanding of host/symbiont interactions.
1. Gut microbiota in insects

1.1. A low diversity

In contrast to mammals, the bacterial diversity in insect digestive tracts is generally low and rarely exceeds a few tens of species (Colman et al., 2012). In *Drosophila*, the gut only contains 2 to 20 bacterial species in natural and/or field conditions (Apidianakis and Rahme, 2011; Bae et al., 2010; Blum et al., 2013; Chandler et al., 2011; Corby-Harris et al., 2007; Cox and Gilmore, 2007; Ren et al., 2007; Ryu et al., 2008; Storelli et al., 2011; Wong et al., 2011), while in humans the gut microbiota diversity generally exceed 1000 bacterial species (Ley et al., 2008b). Several immunological, physiological and morphological hypotheses have been proposed (see (Broderick and Lemaitre, 2012; Engel and Moran, 2013). The lack of adaptive immune function in invertebrates might partly explain this low diversity. Indeed, the innate immune system may only be capable of managing the simple communities of resident bacteria typically present in the invertebrate gut, while the adaptive immune system of higher metazoans might have facilitated association with a greater number of different microbes (McFall-Ngai, 2007). Invertebrates have developed physical barriers to separate the microbes from the host and this might explain why they rely primarily on innate immunity. This can be done either by the formation of a specific organ that will host the bacteria or by the existence of a specific tissue that effectively separates the microbes from the host tissues. For instance, the peritrophic membrane in insect digestive tract does not allow the passage of microorganisms and contains them in the gut lumen. This could also limit the diversity of microbes that would settle in the gut.
Profiling of the commensal community members in *Drosophila* has nevertheless revealed different bacterial constituents with high taxonomic diversity at the species level. *Lactobacillus* and *Acetobacter* are the most abundant and common genera in the gut of *D. melanogaster*. Most of the commensal bacteria of the fruit fly are cultivable *in vitro*, which facilitates experimental manipulations of gut microbial communities and microbial genetic analysis. Also *Drosophila* is one of the major insect model systems in the study of innate immunity (Hultmark, 2003; Lemaitre and Hoffmann, 2007), aging (Bjedov et al., 2010), metabolism (Bharucha, 2009; Birse et al., 2010), intestinal stem cells homeostasis (Apidianakis and Rahme, 2011; Buchon et al., 2010; Casali and Batlle, 2009), large-scale dietary studies (e.g. Lee et al., 2008) and offers substantial molecular genetic resources. There is then strong expectations that laboratory experiments on *Drosophila* will define future research in biomedical systems (mammals and humans), which currently lacks a framework to better understand the relationships between nutrition, immunity and gut microbial ecology at different stages of life and in distinct environments. (see Bae et al., 2010; Broderick and Lemaitre, 2012; Charroux and Royet, 2012; Erkosar et al., 2013; Kau et al., 2011; Lee and Brey, 2013; Ponton et al., 2011, 2013; Ryu et al., 2010).

1.2. A versatile gut microbiota

One of the first steps to understanding the symbiotic relationship between gut microbes and their host is to characterize the baseline healthy microbiota and the differences that are associated with metabolic perturbations and disease. Once the healthy composition and functional states of gut microbiota are understood, the features that, when disrupted, are
associated with disease can be determined. Recent studies have however shown that defining this “healthy” composition is challenging. Indeed, while it has been shown that in some insects that have a restricted diet, some gut bacterial strains are specifically associated and can be maternally inherited (see for instance ), in lots of other insect species the composition of the gut microbiota between individuals of a same species or closely related species varies not only in total size but also in composition (Colman et al., 2012; Lozupone et al., 2012; Staubach et al., 2013). Wong et al. (2013) explored the gut microbiota composition of drosophilid flies by first investigating the prevalence of five strains of bacteria usually found in the gut of individual flies for 21 strains in 10 \textit{Drosophila} species; and, in a second analysis, they investigated the gut microbiota of 11 species of \textit{Drosophila}. Their results have shown that the five bacterial strains were not systematically found in all individuals, without any evidence of a core gut microbiota for the different species of \textit{Drosophila}. In a recent review paper, Broderick and Lemaitre (2012) summarized the results of different studies that also investigated the composition of gut microbiota in \textit{Drosophila} for laboratory stocks and wild-caught flies (see also Erkosar et al., 2013). In laboratory reared flies, only two strains of bacteria seem to be consistently associated with \textit{Drosophila}: \textit{Lactobacillus plantarum} and \textit{Acetobacter pomorum/pasteurianus}. In wild-caught flies, even if the diversity of bacteria present in the digestive tract was greater, the two same bacterial genera, \textit{Acetobacter/Gluconobacter} and \textit{Lactobacillus}, were the only symbionts to remain consistently present. In mosquitoes, Osei-Poku et al. (2012) catalogued the inter-individual bacterial diversity in the guts of eight species collected from the coastal region of Kenya. Extensive variation in gut microbiota has also been found between individuals of the same species. Better understanding this variation in the composition of the gut microbiota is
important because (i) it will give new insights into the factors that modulate the gut microbiota composition; (ii) it will allow to identify relevant bacterial diversity information and to target species or functions that are key; and (iii) it will stimulate the development of functional analyses that do not account for taxonomic diversity.

1.3. **Diet, a key driver of gut microbiota composition**

Several biological and ecological factors such as age, genetics and environment have been proposed to explain gut microbiota composition. However, diet seems to be one of the main factors driving variation in the composition of the gut microbiota in vertebrates and invertebrates (Lozupone et al., 2012). In insects, effects of diet have been particularly investigated and diet composition has been shown to influence the bacterial community in the midgut of different species such as larval gypsy moths (Broderick et al., 2004) and cotton bollworms (Xiang et al., 2006). Also, investigations in flies have shown that diet shapes the microbiota composition regardless of taxonomy and geography (Chandler et al., 2011). Indeed, whereas taxonomically- and geographically-distant fly populations collected from various food sources have had very different microbiota compositions, when maintained on the same type of food, they showed similar communities of bacteria in their gut (Chandler et al., 2011). Similarly, Staubach et al (2013) have found no evidence for host species effects in lab-reared *D. melanogaster* and *D. simulans*; instead the lab of origin has had pronounced effects reflecting the importance of the culture conditions. The *Drosophila*–associated microbiota appears thus to be predominantly shaped by food substrate with an additional but smaller effect of host species identity. Also, one major difference between the human and insect microbiota is that all
insect bacteria seem to be aerobic and therefore capable of living on (and "digesting") the food outside the fly. This may contribute to explaining why diet is such a key driver of microbiota composition (erkosar??).

Food composition is then a major determinant of gut bacterial community (see also Ley et al., 2008a; Ravussin et al., 2012; Wu et al., 2011). This might be explained by the fact that diet composition promotes specific bacterial strains by providing them with appropriate nutritional conditions. It has also been suggested that diet may influence the physical and chemical milieu of the gut (Clissold et al., 2010; Flint et al., 2008; Ley et al., 2008b; Sorensen et al., 2010) and will constrain the type of bacterial strains that can survive in the gut ecosystem. In mosquitoes, for instance, blood meals have been associated with massive proliferation of bacteria residing in the digestive tract (Kumar et al., 2010; Oliveira et al., 2011; Wang et al., 2011). Blood meals decrease the levels of reactive oxygen species (ROS) in the digestive tract, creating a favorable environment for bacterial growth (Oliveira et al., 2011). Although intestinal microbes are more abundant after a blood meal, the diversity of the gut microbiota seems to be affected with enteric bacteria being favored due to their capacity to cope with oxidative and nitrosative stresses (Wang et al., 2011). When not ingesting blood, mosquitoes can feed on nectars and these meals might also influence the composition of the bacterial communities found in the gut (Lefèvre et al., 2013).

Extracellular symbionts in insects are not usually inherited from the mother or transmitted from host-to-host, but they mainly come opportunistically from the environment (see for instance Storelli et al, 2011 and Blum et al, 2013; but see also Kikuchi et al, 2009). Food
can be itself a vector of commensals, and different diets will provide microbial inoculates of
different community compositions (Broderick and Lemaitre, 2012; Gendrin and Christophides,
2013). In a recent study, Blum et al. (2013) investigated the role of food in supplying bacteria to
flies and allowing insect to maintain a gut microbiota. They have shown that flies establish and
maintain their gut microbiota by frequently consuming food containing bacteria. Hence, food
can be considered as a bacterial reservoir for flies and the establishment of *Drosophila*
microbiome is only possible if flies consume exogenous bacteria. Food sources such as living
animals can also be vectors of bacteria for blood feeding insects that become infected when
feeding on the host (Gendrin and Christophides, 2013). Unravelling the relationships between
nutrition, food composition and, the composition and function of symbiont populations is
fundamental to predicting the outcome of parasitic infections. This is particularly significant
when the application of microbial symbionts is considered to reduce vector competence and to
control the spread of arthropod-transmitted pathogens (Weiss and Aksoy, 2011).

2. Towards an integrated understanding of gut microbiota: from genes to function

Molecular techniques that have been employed to analyse gut microbiota might also be
sources of variability, particularly for taxonomic identification. The integration of studies
combining different methods might allow a better understanding of individual and communal
roles of bacteria in the physiology and ecology of the host

2.1. History and limitation of 16S rRNA as a universal marker
For decades, the study of microbial diversity has been hampered by the resistance of the vast majority of bacteria to cultivation under artificial conditions. In the late 1970s, the work of Carl Woese brought a new perspective on microbial diversity by providing the first bacterial phylogeny based on ribosomal RNA sequence (16S rRNA) (Woese and Fox, 1977). The 16S rRNA is a molecular marker that is universally present in bacteria and has highly conserved domains flanking hypervariable sequences, which can easily be used for gene amplification using PCR and further sequencing and analysis, enabling phylogenetic identification of bacteria without any cultivating steps (Dave et al., 2012). The seminal work of Woese and Fox led to a breakthrough in the classification of the “uncultivated majority of microbes”, which are estimated to account for over 80% of human gut microbes (Eckburg et al., 2005). The emergence of next-generation sequencing technologies, which became commercially available in 2005, has further increased the speed of phylogenetic coverage and decreased the cost through massively parallel sequencing methods. As a result, over 3.8 million 16S rRNA sequences are now available in databases such as Silva (QUAST et al. 2013) (http://www.arb-silva.de/). Despite these enormous advances, the microbial community of the gut remains unclear, hiding behind its diversity and unexpected variability between individuals (Lozupone et al., 2012).

There have been an increasing number of studies targeting gut microbial communities in insects based on the sequencing and analysis of 16S rRNA (see Broderick and Lemaitre, 2012; Colman et al., 2012 for review). However, determining and comparing the taxonomic composition of the bacterial communities based on sequencing and analysis of highly conserved genes should be considered with caution particularly when the identifications
were done in separate studies. Not only are protocols usually different in terms of tissues and methods used for extraction, but also taxonomic analyses can be different and use different taxonomic levels to distinguish the bacterial groups. Modern sequencing technologies define microbial taxonomic groups by 16S rRNA sequence similarity (OTUs, or operational taxonomic units), the species-level OTU being generally defined by a percentage of similarity greater than 97% between two sequences (Gevers et al., 2005). The intriguing question arises as to whether the level of diversity and variability found in the gut microbiota communities of different individuals may reflect the level of taxonomic classification (i.e. percentage of similarity) chosen (also discussed in Wong et al., 2013). It is very likely that at higher-order taxonomic levels (e.g., phylum), the gut microbiota communities between individuals begin to resemble one another more closely. It is still unclear whether studies should consider strain-, species-, genus-, or higher order-level to assess the diversity of gut microbiota and search for differences among individuals or conditions. Future work evaluating critically the appropriateness of taxonomic level used for assessing the diversity of microbial communities in the gut of insects, and more generally in metazoans, is therefore required. It is important to note that the differences in gut microbiota communities can also be measured as changes in the proportional representation of OTUs or Phyla. For example, the relative abundance of the bacterial phyla Firmicutes, Actinobacteria, and Bacteroidetes in the gut has been shown to be consistently associated with obesity in both humans and mice (Turnbaugh et al., 2006; Turnbaugh et al., 2009). Also, De Filippo et al. (2010) demonstrated a significant increase in the abundance of Bacteroidetes relative to Firmicutes in the gut microbial communities of children characterized by a rural diet compared to a modern western diet. Despite the considerable progresses realized in this field,
the general consensus is that microbial studies based on 16S rRNA enable little by way of functional conclusions and give little information on the metabolic capabilities of the different bacterial groups.

2.2. Molecular approaches to investigate potential functions of bacteria

Recently, some studies based on 16S rRNA gene have suggested that it may be possible to predict the functional potential of microbes from phylogeny (Langille et al., 2013). PICRUSt, an automated method based on evolutionary modeling uses phylogenetic information contained in 16S marker gene sequences in relation to existing reference genomes to predict the function of microbial communities (Langille et al., 2013). Although PICRUSt does not infer function for viruses or eukaryotes and its accuracy is affected by phylogenetic dissimilarity among environmental organisms and sequenced genomes, it can predict and compare probable functions for bacteria across many samples from a wide range of habitats at a limited financial cost. This approach has the potential to provide the first functional glimpses into the vast amount of existing samples for which only 16S data are available and can be seen as an important step for future cost-effective studies including two steps: (i) integration of completed genome sequences and 16S rRNA gene studies to approximate functional information and then (ii) design of more costly metagenomics or functional studies to assess precisely the metabolic role of gut microbiota communities.

Metagenomics allows direct sequencing of genomes contained within a community providing access to the functional gene composition of microbes and therefore to a much broader description than phylogenetic surveys based on the diversity of 16S rRNA. Despite the
fact that deep, and therefore expensive, metagenomic sequencing is generally required to
access rare organisms and genes, more than 100 metagenomics sequencing projects have been
completed so far (Liolios et al., 2010). Some of these projects have been responsible for
substantial advances in the study of microbiomes over the past 10 years, enabling the
characterization of microbial assemblages at a functional level. Suen et al. (2010) did the first
functional characterization of the fungus garden microbiome of leaf-cutter ants using
metagenomics and whole-genome sequencing. They have shown that the microbial community
within the fungus gardens of leaf-cutter ants contains not only the fungal cultivar, but also a
diverse assembly of bacteria. Using metagenomics analysis of a carbohydrate-active enzyme,
they further showed that these bacteria are likely to participate in the symbiotic degradation of
plant biomass in the fungus garden while previous studies suggested that the fungal cultivar
was solely responsible for this process. Metagenomics therefore provides important
opportunities for the emerging field of eco-systems biology (i.e. considering molecular systems
biology at the ecosystem level) which unites molecular microbiology and ecology to develop an
understanding of community function (Raes and Bork, 2008). Also, looking at the functional
capabilities of communities reveals conservation of metabolism even when species composition
varies. One intriguing result of this emerging field is that completely different microbial
communities found in different individuals or habitats appear to converge on the same
functional gene repertoire (Dinsdale et al., 2008). For instance, using the same set of samples
from lean and obese twins, different species assemblages appear to lead to very similar
functional profiles, demonstrating the presence of a shared core of metabolic capabilities in the
microbiome (Turnbaugh et al., 2009). Discovering the relationships between these consistent
functional signatures and the sum of all metabolic processes (i.e. the nutrients and energy cycling) occurring within the gut will be an especially important step to better understand the biology and functioning of this micro eco-system (Figure 1). Further down the track, transcriptomic approaches, which allow investigating change in gene expression in both the insect host and the microbial symbionts, are powerful to understand the interplay between the different actors within this micro-ecosystem. Although the vast progress of transcriptomics approaches are well described elsewhere (REF) and beyond the scope of this review, it is important to note that combined metagenomic and metatranscriptomic datasets further allow examining the link between gene expression level and sequence conservation, revealing broad evolutionary patterns across taxonomically and functionally diverse communities (Stewart et al, 2011 http://genomebiology.com/content/12/3/R26).

2.3. Stable isotopes based technics to quantify metabolic activity of bacteria in situ

By allowing the researchers to simultaneously investigate genes, their functions and the bacteria that exert them, molecular approaches such as metagenomics have greatly improved our understanding of microbial communities and their metabolic potential. However, genetic information is not always well correlated to the metabolic activity of specific bacteria in situ. In this context, the development of single cell approaches using radioactive or stable isotope-based techniques such as microautoradiography, Raman microspectroscopy and Secondary Ion Mass Spectrometry (SIMS) has been a revolutionary step in modern microbial ecology, revealing individual microorganisms that are metabolically active in their natural habitat and
within complex communities (for review on these techniques and their application in microbiology (see Musat et al., 2012; Orphan and House, 2009; Wagner, 2009). Among these single-cell approaches, the latest version of high spatial resolution SIMS instrument (NanoSIMS) can provide direct imaging and precise quantification of up to 7 different isotopes at a micrometer or submicrometer scale (up to ~50 nm) (Hoppe et al., 2013). Recently, Carpenter and coworkers (Carpenter et al., 2013) applied, for the first time, this technique to insect gut microbiota studies, combining stable isotope labeling using 13C-cellulose with NanoSIMS analysis to investigate nutrients flow within the gut microbiota of the desert dampwood termite *Paraneotermes simplicicornis*. Their results suggested an unexpected tripartite nutritional interaction, the protist *Oxymonas dimorpha* degrading wood fragments ingested by *P. simplicicornis* and subsequently transferring carbon derived products to bacterial symbionts. Given its spatial resolution and detection limit (parts per million), NanoSIMS allows the quantification of the relative metabolic contribution of different partners (i.e. metabolic rates of individual host and symbiont cells) within an intact symbiosis (Pernice et al., 2012). This technique can also be used in concert with *in situ* hybridizations to simultaneously identify individual microbial cells and quantify their substrate uptake (Behrens et al., 2008; Musat et al., 2008). Using this combination, Berry et al (2013) elegantly demonstrated that within the complex communities of the mouse gut microbiota, two bacterial species, *Bacteroides acidifaciens* and *Akkermansia muciniphila*, are important host protein foragers. In this context, the use of stable isotope approaches in aposymbiotic animals that can be infected selectively has a great potential to experimentally test hypotheses about nutritional function of a specific bacterial species (Salem et al, 2012). Despite the fact that the application of NanoSIMS to study
insect gut microbiota is still in its infancy, there is no doubt that future studies integrating this powerful analytical technique can dramatically improve our understanding of the nutritional interactions that lie at the very heart of insect gut microbiota. Together, these investigations are revolutionizing our understanding of the gut microbiome.

CONCLUSION

There is a growing interest in using model systems to provide a comprehensive and integrated understanding of the functioning of the gut microbiota and its interactions with the host metabolism and immunity. Insects are relevant systems in many ways (Apidianakis and Rahme, 2011; Charroux and Royet, 2012; Erkosar et al., 2013) and numerous studies already provide an abundance of data on the taxonomic diversity of gut microbiota, but also the ecological factors that might influence this diversity and the in situ roles of bacteria residing in the digestive tract. However, readability of data and harmonization of the data collection and analyses could be the largest obstacle to providing reliable and repeatable results that might be translated to mammalian models, thereby achieving the full potential of this field. There is therefore a need for a consortium where researcher working on gut microbiota share technical details (see also Broderick and Lemaitre, 2012) and produce some standard guidelines that can be followed by the researcher community when publishing gut microbiota data. These guidelines will allow an adequate, more transparent and comprehensive reporting of experimental details. This has already been successfully applied for studies based on quantitative real-time PCR for instance (see Bustin et al., 2009). Also, the successful development of research on gut microbiota will result only with the integration of a vast array of techniques and disciplines allowing a deeper
understanding of the functioning of this micro eco-system. Technological advances in combination with ecological and evolutionary approaches will soon provide groundbreaking understanding of the function of microbiomes and comprehensive knowledge of the effects of gut microbiomes on the host biology, physiology and immunity; and might provide new efficient treatments against infection and metabolic diseases.
Figure 1: Gut eco-system biology: from genes to function.

An overview of the different methods that can be used for understanding the function of the gut eco-system by integrating different levels (Diversity, Gene function and Metabolic activity). Recently developed methods (listed below a figure of representative output) can be used to assess changes in the different levels of response to different factors (such as change in microbial diversity and host diet). Arrows between the different levels indicate interdependencies (for example, metabolic activity depends on gene function and ultimately on microbial diversity). The representative output for isotope-based technics and 16S rRNA analysis are modified from (Pernice et al., 2007; Pernice et al., 2012).
References


Birse, R.T., Choi, J., Reardon, K., Rodriguez, J., Graham, S., Diop, S., Ocorr, K., Bodmer, R., Oldham, S., 2010. High-Fat-Diet-Induced Obesity and Heart Dysfunction Are Regulated by the TOR Pathway in *Drosophila*. Cell Metabolism 12, 533-544.


Figure 1: Gut eco-system biology: from genes to function