Insect science

Insect Science (2010) 17, 199–219, DOI 10.1111/j.1744-7917.2010.01340.x

REVIEW

Molecular approaches to study the insect gut symbiotic microbiota at the 'omics' age

Weibing Shi[1,2], Ryan Syrenne[1], Jian-Zhong Sun[3] and Joshua S. Yuan[1,2,3,4]

[1]*Department of Plant Pathology and Microbiology and* [2]*Institute for Plant Genomics and Biotechnology, Texas A&M University, Texas, USA,* [3]*Biofuels Institute, School of the Environment, Jiangsu University, Zhenjiang, Jiangsu Province, China, and* [4]*Advanced Research Institute for Sustainable Energy (ARISE), Texas A&M University, Texas, USA*

Abstract Insect gut symbiotic microbiota play essential roles in the growth, development, pathogenesis and environmental adaptation of host insects. The molecular and systems level analysis of insect gut symbiotic microbial community will allow us to discover novel biocatalysts for biomass deconstruction and to develop innovative strategies for pest management. We hereby review the various molecular biology techniques as applied to insect gut symbiont analysis. This review aims to serve as an informative resource for experimental design and research strategy development in the field. We first discuss various strategies for sample preparation and their pros and cons. The traditional molecular techniques like DGGE, RFLP and FISH are covered with respect to how they are applied to study the composition, diversity and dynamics of insect gut symbiotic microbiota. We then focus on the various 'omics' techniques. The metagenome analysis together with the recent advancements in next-generation sequencing will provide enormous sequencing information, allowing in-depth microbial diversity analysis and modeling of pathways for biological processes such as biomass degradation. The metagenome sequencing will also enable the study of system dynamics and gene expression with metatranscriptome and metaproteome methods. The integration of different 'omics' level data will allow us to understand how insect gut works as a system to carry out its functions. The molecular and systems-level understanding will also guide the reverse design of next-generation biorefinery.

Key words DGGE, insect gut, metagenomics, metaproteomics, symbiotic microbiota, systems biology

Introduction – Why study insect gut symbionts?

Insects are one of the most diverse groups of living organisms on earth (Chapman, 2006; Erwin, 1982). Due to their diverse behaviors and feeding habits, almost no terrestrial food source can escape the consumption by one or more insect species. Despite the diversity, the highly interdependent and well-regulated symbiotic interactions with micro-organisms seem to be an important common property for different insect species (Breznak, 2004).

Correspondence: Joshua S. Yuan, Department of Plant Pathology and Microbiology, Texas A&M University, College Station, Texas, 77843, USA. Tel: 979 845 3016; email: syuan@tamu.edu

The definition and importance of symbiosis

Symbiosis often refers to the long-term and mutually beneficial interactions among different species. Symbiotic microbes living inside the host species are referred to as endosymbionts, and the symbiotic microbes living upon or outside an insect's body are often defined as ectosymbionts (Breznak, 2004). Based on previous studies, endosymbionts are prevalent in a variety of insect species such as scarab beetles, cockroaches, termites and so on

(Brune, 2003; Dasch *et al.*, 1984; Kane & Pierce, 1994b; Kaufman *et al.*, 2000). Overall, it is estimated that a majority of members of the Insecta are involved in some type of symbiosis (Moran, 2002; Moran, 2007; Moran & Telang, 1998). Considering that Insecta is the largest group of invertebrates, it is important to study symbiosis in various insect species to understand the evolutionary and ecological significance of the predominant phenomena. In particular, we need to better understand what roles the symbiotic microbiota plays in plant–insect interaction in terms of host selection and co-evolution of host–insect relationships. From an application perspective, the study of insect symbionts will help to discover novel biocatalysts for biomass deconstruction and develop innovative strategies for pest management.

Function of insect symbiotic microbiota

The herbivore insect gut microbiota has been well-established for at least two aspects of the function: the nutrient biosynthesis and the biomass deconstruction. The nutritional function of the insect endosymbiotic microbes have been well studied by feeding experiments with unbalanced or poor diets lacking essential nutrients such as amino acids and vitamins (Douglas, 1998). Some feeding experiments demonstrated that the insect endosymbiont can help to produce nutrients that do not exist in the food (Khachane *et al.*, 2007; Tamas *et al.*, 2002; Tamas *et al.*, 2008; van Ham *et al.*, 2003). The genome sequence of an obligate symbiont *Wigglesworthia glossinidia* revealed many genes for nutrient biosynthesis and transport (Akman *et al.*, 2002). The phenomena are typical for symbiotic microbes, which often dedicate part of their genomes for the benefit of the hosts (Moran, 2001; Ochman & Moran, 2001). A recent metagenome project also revealed that the viruses affecting the symbionts of the honeybee will lead to detrimental effects on honeybee growth and development and could be a major cause for CCD (colony collapse disease) (Cox-Foster *et al.*, 2007).

The second well-characterized function for insect symbiotic microbiota is the biomass deconstruction and digestion function. Both herbivore insects and symbiotic microbes can secrete cellulytic enzymes for biomass deconstruction and hydrolysis (Ohkuma, 2003; Tokuda & Watanabe, 2007; Warnecke *et al.*, 2007; Sun & Zhou, 2009). It has been controversial about which plays a more important role for biomass deconstruction, the symbionts or insect host itself. Despite the controversy, the importance of symbiotic microbes for biomass deconstruction has recently been established by various genome-level studies. For example, symbiotic microbiota can help termites to deconstruct lignocellulosic biomass with high efficiency (Ohkuma, 2003). The termite gut has actually been referred to as the smallest bioreactor in the world (Brune, 1998). The recent sequencing of the symbiotic microbiota of the higher termite revealed many glycosyl hydrolase enzymes with activities for degrading cell wall components such as cellulose and hemicellulose (Warnecke *et al.*, 2007). In addition, the recent completion of the genome sequence of a prokaryotic symbiont of cellulolytic protozoa *Pseudotrichonympha grassi* has also unveiled its ability to fix nitrogen and to recycle putative host nitrogen wastes for the biosynthesis of diverse amino acids and cofactors (Hongoh *et al.*, 2008b). The protozoa contains up to 70% of the bacterial cells in the gut of the termite *Coptotermes formosanus* and is an important component of the termite gut symbiont.

Both nutrient production and biomass deconstruction functions of the insect gut symbiotic microbiota can be exploited for biotechnology purposes. On one side, it might be possible to develop various strategies for pest management through the control of insect gut symbiotic microbes. On the other side, the insect gut symbiotic microbiota can be exploited for novel biocatalysts and microbe strain discovery. Combined with functional validation, these new biocatalysts and microbe strains could greatly improve the design and efficiency of the next-generation biorefinery. The thorough understanding of the insect gut as a natural biocatalyst system with various molecular techniques will also enable the reverse design of next-generation biorefinery. Regardless of the goal of analysis, the first task for analyzing insect gut microbiota is to prepare the samples that well represent the microbe community in the insect guts.

Sample preparation for insect gut symbiotic microbial study

At the 'omics' age, DNA, RNA, protein and metabolite samples can be prepared from insect gut symbionts. We hereby focus on metagenomic DNA sample preparation and then briefly discuss the sample preparation for metaproteomics.

Insect gut metagenomic DNA extraction

Metagenomics can be defined as the study of the metagenome, the whole genetic material of the microbial community existing in certain eco-environments (Sleator *et al.*, 2008). The ultimate goal of metagenomics is to acquire a global view of the composition and function of the microbial community (Guazzaroni *et al.*, 2009). The proper methods for DNA extraction remain keys

Table 1 Commercial kits for metagenome DNA extraction and their application in insect gut systems.

Company	Target product	Website	Application in insect gut symbiota
MP Biomedicals	FastDNA SPIN Kit for Soil	http://www.mpbio.com	Zhang & Jackson, 2008 Dillon *et al.*, 2008 Shinzato *et al.*, 2005
Sigma-Aldrich	GenElute bacterial Genomic DNA Kit	http://www.sigmaaldrich.com/	Guan *et al.*, 2007
QIAGEN	Qiagen DNeasy Tissue Kit	http://www1.qiagen.com/	Fisher *et al.*, 2007
QIAGEN	QIAamp DNA Mini Kit	http://www1.qiagen.com/	Hosokawa *et al.*, 2006
Promega	Wizard™ Genomic DNA Purification Kit	http://www.promega.com/Default.asp	Wei *et al.*, 2006
Mo Bio Laboratories	PowerSoil™ DNA Isolation Kit	http://www.mobio.com/index.php	Pittman *et al.*, 2008b

to reaching a comprehensive and unbiased evaluation of metagenomes of the community, particularly for the unculturable micro-organisms (Cowan *et al.*, 2005). In order to reach such a goal, there are three aspects to consider during the sample preparation (Schmeisser *et al.*, 2007). The first aspect is the coverage. Metagenomic DNA should cover as many microbial species as possible. The second aspect is the integrity of the DNA sample. Shearing should be avoided to obtain high molecular weight and high quality metagenomic DNA. The third aspect is purity. The metagenomic DNA should be free of contaminants interfering with downstream DNA processing such as enzyme digestion, polymerase chain reaction (PCR) and vector ligation (Schmeisser *et al.*, 2007).

Many of the insect gut microbial DNA isolation protocols were derived from those for soil microbial community analysis and the first paper on the extraction of DNA from soil was published more than three decades ago (Torsvik, 1980). Two strategies have been popular for metagenomic DNA isolation, and they are the cell recovery method and the direct lysis method (Roose-Amsaleg *et al.*, 2001). The cell recovery method isolates intact organisms from the gut content prior to cell lysis, and the cell isolation is achieved either by repeated homogenization and differential centrifugation (Holben *et al.*, 1988; Hopkins *et al.*, 1991) or by gradient centrifugation in media such as sucrose, Nycodenz®, Percoll® or metrizamide (Pillai *et al.*, 1991; Robe *et al.*, 2003). Some commercial kits have recently become available and these kits greatly simplified many cultivation-independent analysis methods (Smalla, 2004). The commercial kits used for DNA extraction from insect gut systems are shown in Table 1. For instance, Schloss *et al.* (2006) used FastDNA SPIN kit for soil (MP Biomedical, Solon, OH, US) to isolate the metagenomic DNA from wood-boring beetle gut after the sonication and centrifugation separation of bacterial cells from insect gut wall. The DNA isolation involves mechanical lysis by bead beating followed by purification of DNA on a silica matrix (Schloss *et al.*, 2006). The same kit has also been used widely for metagenomic DNA extraction from the gut systems of grass grub (Zhang & Jackson, 2008), feral locusts, grasshoppers (Dillon *et al.*, 2008) and termites (Shinzato *et al.*, 2005). Other commercial kits used for insect gut symbiotic microbial metagenomic DNA isolation includes the GenElute bacterial genomic DNA kit (Sigma-Aldrich Corp., St. Louis, MO, US) (Guan *et al.*, 2007), Qiagen DNeasy Tissue kit (Fisher *et al.*, 2007), QIAamp DNA Mini Kit (Hosokawa *et al.*, 2006), Wizard™ Genomic DNA Purification Kit from Promega (Wei *et al.*, 2006), and PowerSoil™ DNA isolation kit (Pittman *et al.*, 2008b). Despite the available commercial kits, one has to realize that the metagenomic DNA preparation protocol has to be optimized because most of these kits are not designed for metagenomic DNA isolation from insect gut (Broderick *et al.*, 2004; Warnecke *et al.*, 2007). For example, we have recently modified an indirect DNA extraction method for various insect gut symbiont metagenomic DNA extractions (Shi *et al.*, 2009, unpublished data).

Besides the cell separation approaches, another approach is based on direct or *in situ* lysis of microbial cells in the presence of the environmental matrix (e.g., soil, sediments or plant material), followed by the separation of nucleic acids from matrix components and cell debris (Ogram *et al.*, 1987). The strategy generally yields more DNA and is believed to provide a better representation of environmental biodiversity (More *et al.*, 1994). However, the largest disadvantage of direct lysis methods is the co-recovery of contaminants like humic and fulvic acids with environmental DNA, and these contaminants are visible as a dark color in the DNA sample. The contaminants have been demonstrated to be inhibitors for DNA

hybridization, digestion and PCR (polymerase chain reactions) (Jackson *et al.*, 1997; Miller *et al.*, 1999; Tebbe & Vahjen, 1993). The removal of co-extracted humic acids is critical for the direct lysis method. Lilburn *et al.* (1999) used direct lysis method for phylogenetic diversity study of termite gut spirochaetes (Lilburn *et al.*, 1999). Despite the advantages of the direct lysis method, much fewer studies used the method to study the insect gut symbiont, probably because of the concerns over contamination of the host DNA. Overall, cell recovery method has been much more popular in the insect gut metagenomic analysis and various commercial kits and modified protocols are available for the analysis. The cell recovery method can also be modified to isolate RNA from the symbiotic microbiota.

Protein for 'omics' analysis

Besides the metagenomics, metaproteomics are also important perspectives for analyzing insect gut microbe communities. Metaproteome describes the proteins expressed in the environmental samples and provides the real-time dynamics of the system (Handelsman *et al.*, 1998). Among the various proteomic techniques, mass spectrometry (MS)-based shot-gun proteomics has emerged as the primary method for the identification and quantification of protein expression (Cravatt *et al.*, 2007). As for metagenome analysis, sample preparation is also crucial for metaproteomics. The challenges come from requirements from both the environmental samples and the ESI (electrospray ionization) MS analysis. On one side, ESI is highly sensitive to detergent and requires the sample to be relatively pure. The extra purification step is often involved for sample preparation for shot-gun proteomics and the use of detergent like sodium dodecyl sulfate (SDS) should be avoided. On the other side, the sample preparation from insect guts needs to be comprehensive and contamination from the host tissue needs to be avoided. Several protocols were developed based on the previous metaproteomics analysis of environmental samples. Ogunseitan developed and evaluated two methods for extracting proteins from water, sediments and soil samples (Ogunseitan, 1993, 1997). One is the boiling method, which recovered high concentrations of proteins from waste water but not from soil and sediments. The other one is the freeze–thaw method, which worked better for soils and sediments (Ogunseitan, 1993, 1997). After the pioneering work, different extracting methods were developed for various purposes (Schulze *et al.*, 2005; Singleton *et al.*, 2003). As compared to the environmental samples like soil and sediment, the insect gut samples are normally very limited and need specific modification of the protocols for efficient and comprehensive extraction of proteome for LC-MS/MS (liquid chromatography-mass spectrometry/mass spectrometry) analysis. In addition, the extraction of total microbial protein and the extraction of free proteins in the gut content will be different. Warnecke and colleagues employed metaproteomics approaches to study the free proteins extracted from wood-feeding higher termite hindgut (Warnecke *et al.*, 2007). The sample preparation involves high-speed centrifugation of luminal contents in saline buffer to remove the insoluble fraction. The soluble proteins were then denatured, reduced, alkylated, and digested with trypsin for the LC-MS/MS-based shot-gun proteomics analysis. The analysis allowed measurement of soluble proteins in the gut contents. However, analysis of total microbial protein will have to follow a protocol similar to the cell recovery metagenomic DNA extraction method, where the microbial cells will be first separated and then total protein will be extracted. We have recently developed such a protocol for cattle rumen metaproteomics analysis, which can also be used for insect gut analysis.

Traditional molecular techniques to investigate insect gut microbiota

Traditional molecular techniques played an important role in furthering our understanding of the composition and function of insect gut symbionts. These techniques continue to provide solutions for insect gut microbial community analysis at the 'omics' age. Over the past two decades, the study of insect gut samples with molecular methods has revealed a large discrepancy between the relatively few culturable micro-organisms and the significant diversity present in insect gut (Head *et al.*, 1998; Pace, 1997). Due to the limitation of cultivation-based methods, it was expected that most of the diversity in insect gut microbiomes were still unknown (Stokes *et al.*, 2001). In order to study the diversity of insect gut microbial communities, three major molecular approaches have been employed to discover new genes and investigate the composition of gut microbial communities. These three approaches include gene targeting PCR, molecular fingerprinting techniques such as DGGE (denaturing gradient gel electrophoresis), and oligonucleotide probe-based hybridization techniques such as FISH (fluorescent *in situ* hybridization) (Stokes *et al.*, 2001).

Gene targeting: gene-specific PCR

Gene targeting techniques employ gene-specific primers to specifically amplify target genes, including

conserved *16S rRNA* gene or a gene of specific functional interest from the metagenomic DNA of insect gut symbionts. This approach has been widely applied to insect gut symbiotic microbiota analysis and has revealed substantial bacterial diversity and groups of unculturable microbes (Brauman *et al.*, 2001; Paster *et al.*, 1996; Spiteller *et al.*, 2000). Kane and Pierce (1994a) were among the first to use PCR-based ribosomal DNA sequencing to study insect gut microbial communities. Later on, Mckillip and colleagues analyzed the composition of the microbiome in the midgut of *Pandemis pyrusana* Kearfott by both PCR and culturing techniques (McKillip *et al.*, 1997). Lilburn and colleagues sequenced 98 clones of near-full-length *16S rDNA* genes of *Spirochaetes* in the gut of termite species *Reticulitermes flavipes*. The research revealed substantial phylogenetic diversity in the termite gut (Lilburn *et al.*, 1999). Phylogenetic analysis of *16S rRNA* genes recovered from the hindgut of soil-feeding termites also revealed an enormous diversity of bacteria in the different gut compartments (Schmitt-Wagner *et al.*, 2003b). Based on the PCR targeting of *16S rRNA*, it has also been shown that most of the gut microbial *16S rRNA*s from termite *Reticulitermes speratus* were unknown (Ohkuma & Kudo, 1996). Most of the early *16S rRNA* gene targeting analyses revealed a significant number of unknown bacterial species at the time.

Besides *16S rRNA*, gene-specific PCR has also been widely used to discover genes of interest and survey metabolic pathways. This approach has been particularly useful in cell wall degrading enzyme discovery for bioenergy purposes. A number of cellulases belonging to glycosyl hydrolase family 45 were cloned by gene targeting from the flagellates *Koruga bonita* and *Deltotrichonympha nana*, both of which were cultured from termite gut (Li *et al.*, 2003). In addition, Inoue and colleagues identified a cellulase gene from lower termite hindgut using PCR with gene-specific primers and *in situ* hybridization (Inoue *et al.*, 2005).

In addition to gene-targeting PCR of DNA samples, reverse transcriptase PCR (RT-PCR) from RNA has also been employed to clone genes from environmental samples (Manefield *et al.*, 2002). By combining the RT-PCR with immune-blotting, Casu and colleagues identified a major excretory/secretory protease from *Lucilia cuprina* larvae (Casu *et al.*, 1996). Noda and colleagues also amplified a nitrogen fixation gene from microbial RNA in the gut of the termite *Neotermes koshunensis* by RT-PCR (Noda *et al.*, 1999). RT-PCR experiments also revealed that five GHF9 EG (Glycosyl Hydrolase Family 9 Endoglucanase) homologs were expressed in the salivary glands and the midgut of termites (Nakashima *et al.*, 2002). Other examples employing the RT-PCR technique for gene discovery in insect guts includes studies in *Ancylostoma caninum* hookworms (Jones & Hotez, 2002), *Creontiades dilutus* (Colebatch *et al.*, 2002), *Protaetia brevitarsis* (Yoon *et al.*, 2003), *Aedes aegypti* (Pootanakit *et al.*, 2003), *Helicoverpa armigera* (Chougule *et al.*, 2005), and *Manduca sexta* (Brinkmann *et al.*, 2008; Hogenkamp *et al.*, 2005).

Even though gene-specific PCR was proven to be effective for gene discovery and microbial diversity analysis, two major limitations have restricted the application of the technique (Cowan *et al.*, 2005). First, the gene-targeting techniques depend on existing sequence information to design primers for PCR amplification, which greatly limited the application of the technique. Second, normally only partial sequence of the genes can be cloned. The cloning of full-length genes will have to involve further PCR-based chromosome walking (Cowan *et al.*, 2005). The available next-generation sequencing techniques and the metagenomic strategies will certainly revolutionize both gene discovery and biodiversity analysis for the insect gut symbiotic microbiota. In addition to traditional gene-targeting PCR-based techniques, PCR can also be used for various molecular fingerprinting techniques to study microbial diversity.

Molecular fingerprinting techniques

Besides the library-based gene targeting PCR, several other PCR-based techniques have also been widely used to study microbial diversity in various environmental samples. These molecular fingerprinting techniques include denaturing or temperature gradient gel electrophoresis (DGGE or TGGE) (Muyzer *et al.*, 1993; Muyzer & Smalla, 1998), restriction fragment length polymorphisms (RFLP) (Liu *et al.*, 1997; Osborn *et al.*, 2000), single strand conformation polymorphism (SSCP) (Lee *et al.*, 1996; Schwieger & Tebbe, 1998), and random amplified polymorphic DNA (RAPD) (Kauppinen *et al.*, 1999). For microbial diversity analysis, these techniques are usually used to analyze the sequence of *16s rRNA* from different microbial species, where both molecular fingerprints and phylogenetic affiliation of microbial species can be generated (Smalla, 2004). These techniques have been proven to be helpful in providing an overview of microbial diversity in certain insect gut symbiotic microbiota. We hereby review the previous application of these techniques in insect gut microbial diversity analysis.

Among the different aforementioned genetic fingerprinting techniques, DGGE is perhaps the most commonly used. Recent application of the technique to study insect gut microbial diversity has led to a much more

comprehensive understanding of insect symbionts (da Mota *et al.*, 2005; Schabereiter-Gurtner *et al.*, 2003; Smalla *et al.*, 2007; Webster *et al.*, 2003). The DGGE profiling of wasp larval *Vespula germanica* revealed a diverse group of micro-organisms in the gut and indicated that the wasp larva are not dependent on one particular type of mutualist (Reeson *et al.*, 2003). Behar and colleagues analyzed Mediterranean fruit fly gut bacterial communities using both culture-dependent and culture-independent approaches such as DGGE and revealed that the family Enterobacteriaceae was the most dominant species in the fruit fly gut (Behar *et al.*, 2005). Recently, DGGE was employed to explore microbial diversity in herbivore insects to study the potential mechanisms for biomass degradation. Enterobacterial repetitive intergenic consensus PCR (ERIC-PCR) and DGGE were combined to compare the diversity of lactic acid bacteria communities in wood- and soil-feeding termites (Bauer *et al.*, 2000). The DGGE method was also used to survey and screen for gut micro-organisms in wood-feeding termites (Hayashi *et al.*, 2007), soil-feeding termites, and their mounds (Fall *et al.*, 2007). In addition to termites, the symbiotic microbiota in the hindguts of scarab beetle larvae were also explored with metagenomic approaches mainly based on DGGE (Pittman *et al.*, 2008b; Vasanthakumar *et al.*, 2006). Moreover, Dillon and colleagues surveyed microbial diversity from four species of feral locusts and grasshoppers by DGGE analysis of bacterial *16S* gene fragments and revealed that *Gammaproteobacteria* from the family Enterobacteriaceae is the most predominant species in grasshopper and locust guts (Dillon *et al.*, 2008). Recently, we revealed the diversity of gut bacteria from different insect species by DGGE and found significant microbial diversity differences among wood-feeding, grass-feeding and leaf-feeding insects (Shi *et al.*, 2009, unpublished data). DGGE has also been used to study symbiotic microbiota in a variety of insect species such as *Dermolepida albohirturn* (Pittman *et al.*, 2008a; Pittman *et al.*, 2008b), *Gadus morhua* L. (McIntosh *et al.*, 2008), diamondback moth (Raymond *et al.*, 2008), *Anopheles gambiae* (Lindh *et al.*, 2008), *Hippoglossus hippoglossus L.* (Bjornsdottir *et al.*, 2009), and *Artemia franciscana* (Orozco-Medina *et al.*, 2009).

Restriction fragment length polymorphism (RFLP) analysis differentiates homologous DNA sequences based on the distinct DNA fragment patterns resulting from the sequence specificity toward restriction enzymes (Esumi *et al.*, 1982). In 1993, Harada and Ishikawa used RFLP to analyze *16S rRNA* from the group of prokaryote microbes in the gut of the pea aphid. The result suggested that gut microbes have a close relationship with aphid intracellular symbionts (Harada & Ishikawa, 1993). Despite this analysis, the application of traditional RFLP in microbial diversity studies is very limited due to the inherent technical limitations of the technology. Domingo used RFLP of *16S rRNA* to study cricket hindgut microbial communities and suggested that community RFLP methods did not have sufficient resolution or specificity required to study the effect of diets on cricket hindgut microbial community dynamics (Domingo, 1998). Due to the limitations of traditional RFLP, terminal restriction fragment length polymorphism (T-RFLP) has been employed to study microbial diversity in insect gut (Shinzato *et al.*, 2005). Different from RFLP, T-RFLP will separate homologous DNA based on the length and sequence of the end sequence generated from restriction enzyme digestion of 16S rRNA, which makes it much more efficient in revealing microbial diversity. T-RFLP was used to analyze the bacterial *16S rRNA* genes in the midguts of individual European cockchafer (*Melolontha melolontha*) larvae and revealed a simple but variable community structure (Egert *et al.*, 2005). In addition, T-RFLP has been used for gut symbiotic microbial community research of various termites such as soil-feeding termites (Donovan *et al.*, 2004; Friedrich *et al.*, 2001; Kohler *et al.*, 2008; Schmitt-Wagner *et al.*, 2003a), wood-feeding lower termites (Miyata *et al.*, 2007; Stingl & Brune, 2003), and fungus-growing termites (Hongoh *et al.*, 2006; Mackenzie *et al.*, 2007; Shinzato *et al.*, 2007). These studies helped to reveal the composition and dynamics of termite gut microbial communities and led to some speculations on how symbiotic microbes could contribute to biomass degradation.

Another traditional molecular fingerprinting technique is random amplified polymorphic DNA (RAPD). The analysis is based on amplification of genomic DNA using random primers. RAPD-PCR was carried out to compare microbiota composition between different generations of western flower thrips *Frankliniella occidentalis* and revealed a surprising result that some bacteria in the thrips can be passed from generation to generation for up to 50 generations (de Vries *et al.*, 2001a, b). The discovery highlighted that symbiotic microbiota can be indigenous instead of exogenous from the food material (de Vries *et al.*, 2001a, b). The application of RAPD is also very limited due to technical complexity and low reproducibility of the technique.

Single-strand conformation polymorphism (SSCP) is a technique that uses electrophoresis to separate single-strand DNA to differentiate the homologous sequences (Yandell, 1991). SSCP was introduced to insect gut microbiota analysis very recently and has not been widely used. Mohr and Tebbe used SSCP to study the diversity

and phylogenetic consistency of bacteria in the guts of three bee species at the same oilseed rape field (Mohr & Tebbe, 2006). In a recent study, PCR-SSCP, RT-PCR-SSCP and stable isotope probing (SIP) were combined to study partial bacterial *16S rRNA* genes to survey the diversity of metabolically active bacteria in the larval gut of *Manduca sexta* (Brinkmann *et al.*, 2008).

Even though these different molecular fingerprinting techniques have revealed significant microbial diversity in the guts of various insect species, all of them are rather limited in providing comprehensive and detailed analysis of microbial diversity. The techniques are particularly limited if we want to survey the dynamics of microbial communities during biomass deconstruction. The recently developed metagenomics platforms are rapidly replacing these molecular fingerprinting techniques.

Fluorescent in situ *hybridization*

Fluorescent *in situ* hybridization (FISH) is commonly used in microbial ecology studies to visualize symbiotic bacteria in the gut (Aminov *et al.*, 2006; Cheung *et al.*, 1977). The application of FISH in insect gut microbial studies often involves fluorescently labeled probes targeting 16s rRNA with sequences specific for a bacterial species or genus (Turroni *et al.*, 2008). FISH has been used to detect, visualize and characterize the intracellular symbiotic bacteria of aphids (Fukatsu *et al.*, 1998), crickets (Domingo *et al.*, 1998), termites (Berchtold *et al.*, 1999) and some others. For biomass degradation-related studies, Berchtold and colleagues examined the abundance and spatial distribution of major phylogenetic groups of bacteria in the hindguts of the Australian lower termite *Mastotermes darwiniensis* using FISH with group-specific, fluorescently labeled, rRNA-targeted oligonucleotide probes. The approach has been shown to be particularly useful in studying uncultivated microbes to observe the dynamics of microbiota (Santo Domingo *et al.*, 1998). However, when complex bacterial communities from environmental samples are analyzed by FISH with rRNA-targeted probes, several technical problems and potential artifacts might occur and the detailed composition of the microbiota cannot be revealed. In addition, bacteria in less nutrient-rich environments have low ribosome content, which will affect the sensitivity of detection (Smalla, 2004). In complement to FISH, DAPI (4′,6-diamidino-2-phenylindole) and GFP (green fluorescent protein) have also been used to visualize microbial communities. DAPI staining of bacterial cells highlighted the significant differences in the number of bacterial cells among different insect species when reared under the same conditions (Cazemier *et al.*, 1997a, b). GFP can be used to track target microbial species in the host. It has been used to show that the colonization of bacterium *Serratia entomophila* in the gut of the host *Costelytra zealandica* was not confined to a specific site in the gut (Hurst & Jackson, 2002).

Overall, the various molecular techniques have greatly advanced our understanding of insect gut microbial communities, and many of these techniques will continue to be important to further our understanding of insect gut symbionts today. However, due to the inherent limitations of these techniques, they cannot provide detailed information regarding the gene and pathway for different biological processes and a comprehensive coverage of microbial taxonomy in the gut. In order to understand the biological processes involved in biomass degradation, we have to reach a detailed understanding of the biocatalysts, pathways and compositions of insect gut symbionts. The recently available different 'omics' platforms enabled such studies.

Techniques for "meta-omics" analysis of insect gut symbionts

The recent advances in 'omics' technologies enabled us to explore micro-organism communities in an unprecedented way (Allen & Banfield, 2005; Tyson *et al.*, 2004). The high-throughput metagenome, metatranscriptome and metaproteome analysis of micro-organism populations will allow molecular, organism and population-level investigation of how chemical and biological processes have enabled, controlled and evolved (Allen & Banfield, 2005). The complementary data annotation and high-throughput functional screening will allow the identification of novel catalysts and strains for bioremediation, biomass processing, bioproduct synthesis and so on (Hongoh *et al.*, 2008a; Lorenz & Eck, 2005; Warnecke *et al.*, 2007). The so-called 'metagenomics' often involves sequencing genomic DNA extracted from a microbe population in a certain eco-environmental setting (Handelsman, 2004). It often involves sequence-based, compositional and/or functional analyses of the combined microbial genomes contained within an environmental sample such as the insect gut (Handelsman *et al.*, 1998). Metatranscriptomics refers to sequencing analysis of mRNA from a microbial population. Metaproteomics refers to the quantification and identification of all the proteins in a microbial community.

The different 'meta-omics' techniques have been broadly used to explore the function and dynamics of diverse microbe populations in various eco-environmental systems (Green *et al.*, 2008; Keller & Zengler, 2004; Strom, 2008). From the human intestine to the depths of

the ocean, metagenomes from microbe communities have been sequenced and analyzed for evolutionary, pathological, physiological, environmental and ecological studies (Allen & Banfield, 2005; Tyson *et al.*, 2004). The diversity, composition and dynamics of a microbial community largely defines its effectiveness, specificity and reactivity for a certain function related to life, biogeochemical cycles and environmental mitigation (Allen & Banfield, 2005; Backhed *et al.*, 2005; Falkowski *et al.*, 2008; Green *et al.*, 2008; Keller & Zengler, 2004; Lorenz & Eck, 2005; Tyson *et al.*, 2004). In the past two decades, much effort has been dedicated to exploring the components of microbial communities from different niches at the molecular, organism and ecological level to discover novel enzymes, pathways and organisms for various applications (Green *et al.*, 2008; Roussel *et al.*, 2008). For example, metagenome and metatranscriptome sequencing have also become important approaches for exploring biomass degrading mechanisms in wood-feeding insects. Several studies have been carried out to study symbionts in the hindgut and midgut of wood-feeding higher termites (Warnecke *et al.*, 2007) and lower termites (Todaka *et al.*, 2007; Hongoh *et al.*, 2008a, b). The termite is believed to recycle up to 30% of the total carbon on earth, and the highly efficient lignocellulosic biomass deconstruction has made the termite a potential source for novel biocatalysts for biomass deconstruction (Hongoh *et al.*, 2008a; Warnecke *et al.*, 2007). Recent studies have indicated that symbiotic bacteria and protozoa in the hindgut of the termite play an important role in the hydrolysis of cellulose and hemicellulose (Nakashima *et al.*, 2002; Tokuda & Watanabe, 2007; Warnecke & Hugenholtz, 2007; Warnecke *et al.*, 2007; Wheeler *et al.*, 2007; Zhou *et al.*, 2007). These analyses not only revealed a diverse group of bacteria covering 12 phyla and 216 phylotypes, but also led to more than 100 candidate glycoside hydrolases. Moreover, the study also indicated other important functions of symbiotic microbiota, including hydrogen metabolism, carbon dioxide-reductive acetogenesis, and nitrogen fixation (Warnecke *et al.*, 2007). Overall, the development of metagenomics, metatranscripomics and metaproteomics over the past decades has been focused on the better understanding of microbial diversity and function in the eco-environment, and has been driven by increasing demands for biocatalysts and biomolecules for applications such as biorefinery (Schmeisser *et al.*, 2007). We hereby review the application of these 'omics' platforms to study insect gut symbiotic microbiota from several perspectives, including the overview of metagenome analysis of microbial communities, next-generation sequencing and metagenome sequencing, functional metagenomics, metatranscriptomics and metaproteomics.

Metagenome sequencing and next-generation sequencing

There are two principal metagenomic strategies for metagenomics, the sequence-based metagenomics approach and functional metagenomics (Fig. 1). Sequence-based metagenomics involves metagenome sequencing and downstream data analysis. Functional metagenomics involves screening of DNA or cDNA library for gene discovery. Sequence-based analysis of metagenomic DNA from insect gut symbionts has been well-established during the past decade. Metagenomics was first carried out with the conventional Sanger sequencing techniques (Smalla, 2004). Sanger sequencing is more used toward the 16s rRNA library or metagenomic DNA library (Smalla, 2004). The aforementioned metagenomic analysis of termite hindgut symbiotic microbiota involves Sanger sequencing of the metagenomic DNA library. Total metagenomic DNA from pooled P3 luminal contents was purified, cloned and sequenced (Warnecke *et al.*, 2007). Approximately 71 million base pairs of sequence data were generated and assembled. The assembled sequences are highly fragmented. In order to better understand the shot-gun data, 15 fosmids were selected for further sequencing and training of the dataset. The data have led to a comprehensive coverage and quantification of the microbial composition in termite gut symbionts. In addition, more than 700 glycoside hydrolase (GH) catalytic domains corresponding to 45 different CAZy families were identified through the analysis. The study highlighted how metagenome sequencing can help to identify natural biocatalysts, including different cellulases and hemicellulases (Warnecke *et al.*, 2007). Another successful metagenome analysis is from the study of aphid symbionts showing that heat tolerance of the host aphid species can be conferred by gene mutation in their symbiotic microbes, which confers an evolutionary advantage for the host in the field (Harmon *et al.*, 2009).

The recent development of next-generation sequencing has offered the potential to revolutionize metagenome analysis (Marusina, 2006). When next-generation sequencing is used, the approach can be the direct shot-gun sequencing of metagenomic DNA. Up to now, four major next-generation sequencing platforms have been available. 454 sequencing technology is the first available next generation sequencing technique and the platform is based on 'pyrosequencing' and emulsion PCR amplification (Margulies *et al.*, 2005). The sequence read length for 454 sequencing can be up to 400 bases and the throughput is relatively lower at 400 million bases per run. The advantage of the 454 sequencing is the read length, which makes it easier for the sequence assembly in *de novo*

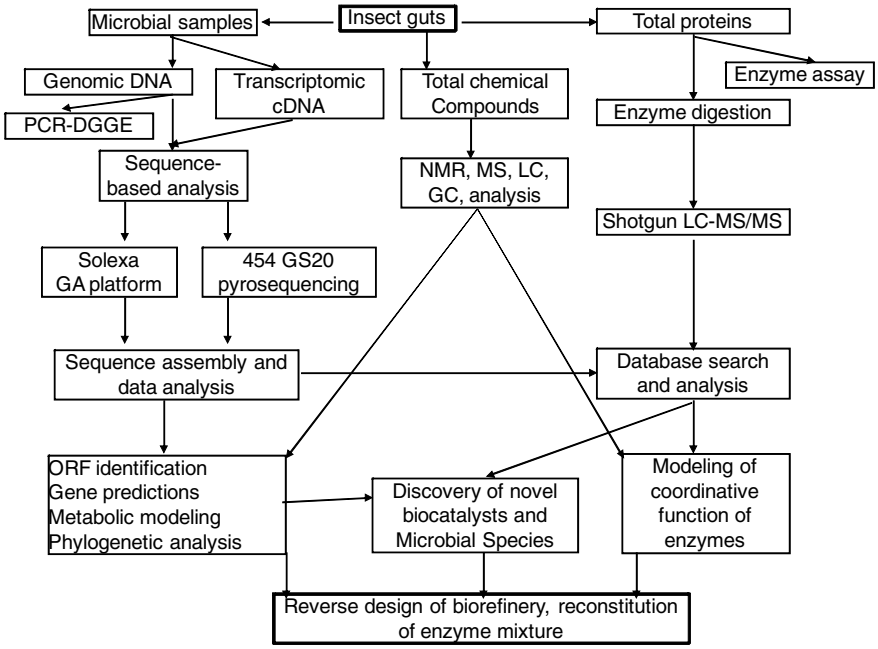

Fig. 1 'Omics' analysis of insect gut as a natural biocatalyst system.

sequencing (Shendure & Ji, 2008; Yuan *et al.*, 2008). Illumina genome analyzer, formerly known as Solexa, is based on the concept of 'sequencing by syntheis' (SBS) (Adams *et al.*, 2009; Mardis, 2008). With the latest development of the technology, Illumina genome analyzer can generate pairwise sequencing of 100 base pairs and 40 gigabase sequences per run. Another two platforms are ABSOLiD and Helocus, both of which have similar sequencing throughput and less sequence read-length (Mardis, 2008). For this reason, 454 and Illumina have been the major approaches for metagenome sequencing. The advantage of 454 is the longer read length, while the strength of Illumina is the sequence throughput (Stangier, 2009). It is expected companies like Pacific Biosciences will soon have the next-next-generation sequencing techniques available. The accuracy and coverage of the metagenome analysis highly depends on the sequence coverage depth. The capacity of the next-generation sequencing technique has enabled a deeper coverage of the metagenomes and allows better annotation of more genes.

Considering the pros and cons for Solexa and 454 sequencing technology, some recent studies have combined the analysis with the two platforms to allow both better assembly of the sequence, and the deeper coverage of the genome (Ansorge, 2009; Shendure & Ji, 2008). Despite the limitations of next-generation sequencing techniques, they have been broadly used for metagenome sequencing of environmental microbial communities from different niches, including soil (Blaha *et al.*, 2007; Tringe *et al.*, 2005; Voget *et al.*, 2003), the human gastrointestinal tract (Gill *et al.*, 2006), human feces (Breitbart *et al.*, 2003), the oceans (Culley *et al.*, 2006; Venter *et al.*, 2004), the rumen (Brulc *et al.*, 2009), acid-mine drains (Tyson *et al.*, 2004) and Zodletone Spring, OK, US (Elshahed *et al.*, 2005). However, more limited efforts have been employed in insect gut symbionts. Very recently, the next-generation-based metagenomic analysis of the grasshopper (Orthoptera) and cutworm (Lepidoptera) gut symbiotic microbiota were carried out to compare the differences in community structure as related to feeding habits and to discover novel genes for biomass degradation (W.B. Shi, X. Zhou, L.T. Liu, P. Gao, X.Y. Chen, N. Kyprides, E.G. No, S.Y. Dai and J.S. Yuan, unpubl. data). The analysis has led to the discovery of numerous novel biocatalysts.

Functional metagenomics

Functional metagenomics involves screening for target genes in a library built with metagenomic DNA or RNA (Allen *et al.*, 2009). Traditionally, metagenomic DNA can be stored stably as a DNA library for further investigation. In a similar way, RNA can be extracted to build a cDNA library. The information held within a DNA or cDNA library can be used to determine community diversity and

search for the enzymes with a particular activity (Steele & Streit, 2005). For the DNA library, the basic steps of library construction include the extraction of metagenomic DNA as aforementioned, the generation of suitably sized DNA fragments, and the cloning of these fragments into an appropriate vector (Cowan *et al.*, 2005). For the cDNA library, total RNA will be extracted and cDNA will be synthesized for building into a proper vector. Both types of libraries can be screened for genes of interest via DNA hybridization using the probes of target genes or homolog genes (Demaneche *et al.*, 2009). The approach has been used to search for various genes from insect guts. For example, Shen and Jacobs Lorena reported the cloning and characterization of a novel chitinase gene expressed specifically in the midgut of adult *Anopheles gambiae* females (Shen & Jacobs Lorena, 1997). They cloned the chitinase gene from a cDNA library via screening and further confirmed by Northern blot that the chitinase is expressed exclusively in the guts of adult females.

One of the major limitations of the traditional screening strategy is the need for probes specific to a certain gene. The sensitivity and reproducibility often also depends on the probe design. The combination of library screening with gene expression and/or enzyme activity assay has been developed to overcome such limitations. The method has been successfully applied to discover new genes and enzymes with different activities. A cDNA clone encoding carboxypeptidase was isolated from a larval gut library of *Helicoverpa armigera*, and the complete cDNA sequence was expressed in insect cells using the baculovirus system to verify carboxypetidase activity (Bown *et al.*, 1998). Girard and Jouanin isolated a cDNA encoding chitinase of *Pheadon cochleariae* from a larval gut library (Girard & Jouanin, 1999). For bioenergy research, novel xylanases with distinct domains have been discovered using metagenomic libraries of microbiota in several insects belonging to *Isoptera* (termites) and *Lepidoptera* (moths) (Brennan *et al.*, 2004). Considering that this strategy does not require the homolog sequences for genes of interest, it has the potential to identify entirely new classes of genes of new or known function (Handelsman, 2004). However, the heterologous gene expression also has some limitations, including low gene expression level and wrong post-translational modification (Handelsman *et al.*, 2002).

A recent development of functional metagenomics is the use of biosensor technology in gene discovery from insect symbioints. Guan and colleagues at the University of Wisconsin constructed a metagenomic library consisting of DNA extracted directly from gypsy moth midgut microbiota, and analyzed it using an intracellular screen designated as METREX (Guan *et al.*, 2007). In this method, the biosensor detects compounds that induce the expression of GFP from a bacterial quorum promoter by fluorescence microscopy or fluorescence-activated cell sorting (Williamson *et al.*, 2005). The authors identified an active metagenomic clone encoding a mono-oxygenase homologue that mediates a pathway of indole oxidation. It was the first to identify a new structural class of quorum-sensing inducer from uncultured bacteria.

The functional metagenomics based on the cDNA library allows us to identify novel enzymes and genes for a particular application; however, the analysis is limited by the available probes for cDNA library screening and the assay used for protein function determination (Chaves *et al.*, 2009; Moran *et al.*, 2008). A more comprehensive approach is to sequence the metatranscriptome of microbial communities and annotate the metatranscriptome to discover the novel genes.

Metatranscriptomics

Metatranscriptome involves the analysis of RNA in a microbial community. RNA is converted to cDNA for the analysis. The random sequencing of cDNA thus may lead to a high percentage of rRNA signals. Different strategies have been developed to remove rRNA to improve the coverage of mRNA. In addition, the available next generation sequencing technique has greatly enhanced the capacity to carry out metatranscriptome analysis.

Cox-Foster and colleagues (Cox-Foster *et al.*, 2007) used an unbiased metatranscriptomic approach to characterize microflora associated with honeybee *Apis mellifera* in a search for the cause of colony collapse disorder (CCD). In this study, total RNA was extracted to capture RNA viruses in presumed CCD-positive and negative bees for 454 sequencing. The raw sequencing reads were trimmed and assembled into contigs, and then analyzed using BLASTN and BLASTX for function annotation. This analysis revealed the presence of bacteria, fungi, parasites, metazoans and viruses in the bee gut content. For example, sequences homologous to bacterial 16S ribosomal RNA were assembled into 48 contigs. Eighty-one distinct fungal 18S rRNA sequences were recovered from the pooled samples. More importantly, the RNA profiling indicated that CCD may be caused by the virus disruption of microbial community structure in the bee gut system (Cox-Foster *et al.*, 2007). More recently, a parallel metatranscriptome analyses was used to identify host and symbiont contributions in collaborative lignocellulose digestion by termites (Tartar *et al.*, 2009). In this study, over 10 000 expressed sequence tags (ESTs) were sequenced from host and symbiont libraries that aligned into 6 555 putative transcripts, including 171

putative lignocelluase genes. They found that cellulases were contributed by host plus symbiont genomes, whereas hemicellulases were contributed exclusively by symbiont genomes. However, ligninase, antioxidant and detoxification enzymes were identified exclusively from the host library.

These researches highlighted the importance of the insect symbionts for host health and showed how the metatranscriptome can be applied to study insect gut systems. The advantage for metatranscriptome sequencing is that it can better reflect the dynamics and function of the insect gut symbionts.

Metaprotoeomics techniques for insect gut symbiont studies

Another way to explore systems dynamics is to study the metaproteomics of insect gut symbionts. Like any genome sequencing project, metagenome sequencing is only the first step toward a comprehensive understanding of composition, dynamics and function of insect gut symbiotic microbiota. The sequence itself won't allow us to understand the protein activity and the dynamic changes of the system (Nelson, 2008). Post-genomic molecular approaches such as proteomics will allow us to study the ultimate functional products of genes/genomes and derive the function and dynamics of insect gut system. The collective study of all proteins in microbial communities, such as those in insect gut, is referred as 'metaproteomics', to distinguish from the proteomics study of single species (Nelson, 2008). Metaproteomics allows the measurement of gene expression from the perspective of presence and abundance of translated proteins (Blackstock & Weir, 1999; Wilmes & Bond, 2004). The proteomics platform can be generally classified as gel-free or gel-based systems (Kan *et al.*, 2005). The traditional approach is to analyze the protein sample with two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) at first and then further cut the spot for MS-based protein identification. The MS techniques that can be used for protein identification include both matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI). MALDI is often coupled with time-of-flight (TOF) mass analyzer, while ESI can be coupled with a variety of mass analyzers. The earliest approach for protein identification of gel spot is through peptide fingerprinting, where the peptides from protease-digested protein will be measured by MALDI-TOF for the m/z value. The pattern of peptide distribution will be searched against a database of candidate proteins for identification. Even though the method was successfully applied for protein identification in gel-based proteomics, the accuracy and reproducibility of the method is often inconsistent. In particular, the post-translational modification of the protein will severely distort the m/z value for the protein identification. For this reason, peptide fingerprinting has been gradually replaced with tandem MS (MS/MS) analysis, where individual peptides will be subject to two rounds of MS analyses. The first round of MS analysis will render the m/z value of the peptide, and the peptide will be further broken into fragment ions by electron or chemical dissociation for the second round of measurement. According to the fragment ion pattern, a protein sequence can be identified based on the search for fragment patterns against the database with protein sequences. The tandem MS method has become the most popular approach for protein identification.

Even though gel-based proteomics was the golden standard for proteomics, the 2D-gel-based methods have numerous inherent limitations including low sensitivity, low coverage of proteome and difficulties in quantification. For all these reasons, gel-based proteomics has been gradually replaced with the gel-free proteomics, which mainly relies on LC-MS/MS platform. The most popular approach for gel-free proteomics is MudPIT (multidimensional protein identification technology)-based shot-gun proteomics (Delahunty & Yates, 2007; Lohrig & Wolters, 2009). In this approach, the total protein from a sample is first digested by protease into a peptide mixture and the peptide mixture is further separated by multidimensional LC. The separated peptides are further analyzed by MS/MS for protein identification as aforementioned. MudPIT can be combined with the different labeling techniques like ICAT (isotope coded affinity tags), ICPL (isotope coded protein labels), or iTRAQ (isobaric tag for relative and absolute quantification) for protein quantification (Delahunty & Yates, 2007). MudPIT can also be used as a label-free platform, where peptide quantification can be based on total ion counts and numbers of peptides (Delahunty & Yates, 2007). Despite the broad application of proteomics techniques in various studies, the use of proteomics in the analysis of insect gut symbiotic microbiota is still very limited. In the aforementioned termite gut metagenomics analysis, the authors carried out a proteomics analysis of total gut protein to examine which enzymes are expressed (Warnecke *et al.*, 2007). The total proteins were first extracted from P3 luminal contents of wood-feeding higher termites as aforementioned. The digested peptides were then subject to three-dimensional LC-MS/MS analysis for protein expression analysis. The fragment ion patterns from metaproteomics were searched against a sequence database derived from metagenome sequencing for protein identification. The study has revealed that expression

of glycosyl hydrolases are regulated at the protein level, and enzymes in the metagenome were not expressed at the same time and same level (Warnecke *et al.*, 2007). Further study of the metaproteome in the natural biocatalyst systems such as termite gut will allow us to understand how enzymes coordinate to degrade plant cell walls. Metaproteomics analysis will be based on the metagenome sequencing data and will help to further understanding of insect gut symbiotic microbes to the proteome level.

Looking into the future

The study of insect symbiotic microbiota is important for insect physiology, pest management, evolutionary study and discovery of various biocatalysts for different applications, including pest management and biorefinery development. In particular, the gut systems of many herbivore insects can be considered as effective bioreactors, where biomass material can be deconstructed for the synthesis of various bioproducts important for insect growth and development (Breznak, 2004). The coordinative function of host and symbiont enzymes plays important roles in biomass processing and degradation. The study of insect gut symbiotic microbiota at the systems level will enable us to reverse-design the next-generation biorefinery.

The techniques to study insect gut symbionts have experienced dramatic changes during the past two decades. The initial studies of insect gut symbionts were based on microbial culture-dependent platforms, which provided very limited information for the diversity and functions of insect gut symbiotic microbiota (Amann *et al.*, 1995; Dillon & Dillon, 2004). The culture-dependent technique only allows us to access to a small portion of the microbe community in insect guts (Oliver, 2000). The culture-dependent analysis was quickly replaced and complemented by molecular biotechniques independent of microbial culturing. Methods like DGGE, SSCP, RFLP and FISH allowed us to better explore the complexity of natural microbial communities. These techniques provided some speculations of microbial community composition, dynamics and function. However, traditional molecular techniques still cannot provide a comprehensive view of the composition and dynamics of insect symbiotic microbial communities. The recently developed metagenome sequencing techniques enabled us to reach much deeper sequencing and better coverage of the metagenome (Mardis, 2008). In particular, the advancements in next-generation sequencing techniques allowed us to explore the metagenomes from insect gut symbiotic microbiota to an unprecedented depth and comprehensiveness (Adams *et al.*, 2009; Stangier, 2009). In addition, functional analysis, metatranscriptomics, metaproteomics and metabolite profiling are all providing important information regarding the function of insect hosts and symbionts from different perspectives. The integration of information will lead to a systems-level understanding of insect gut as the system for biomass deconstruction, nutrient biosynthesis and so on. Despite significant progresses, several aspects of research need to be emphasized to better exploit insect gut systems for various biotechnology applications.

First, more insect gut systems need to be studied with various 'omics' techniques. Current research mainly focuses on the termite gut as the model system for biomass degradation. Comprehensive metagenomics and metatrascriptomics were carried out to study termite gut systems (Tartar *et al.*, 2009; Warnecke *et al.*, 2007). However, there are many other insect species with strong capacities to degrade lignocellulosic biomass (Sun & Zhou, 2009). The cellullolytic enzyme activity in grasshopper gut is actually comparable to that of the termite gut (Shi *et al.*, 2010). The comparative analyses of different insect gut systems will allow us to identify common and unique features for degrading different lignocellulosic biomasses in various insect gut systems. Such studies will also help to understand the co-evolution of insect hosts and symbionts toward different food sources.

Second, bioinformatics challenges for the assembly of next-generation sequencing data need to be better addressed. Despite the potential of next-generation sequencing in increasing the sequencing coverage of metagenome, sequence assembly for metagenome is much more challenging than single species, in particular for complex systems. The more microbe species in a community, the more complexity and overall genome size there will be for insect gut symbiotic microbiota. Illumina genome analyzer has the most potential for increasing sequence coverage due to higher sequencing throughput and lower per base cost. However, short sequence read length together with large overall genome size from this technology make it extremely challenging to assemble metagenome sequences. The recent development of several assemblers for short sequences like SSAKE, VEVELT, ABySS and Euler have provided solutions for the assembly of short sequence reads of genome sequencing (Scheibye-Alsing *et al.*, 2009). However, the conditions used for single genome assembly are not suitable for metagenome sequencing. On one side, we need to find the optimized parameters and criteria for the assembly of metagenomes; one the other side, these software packages need to be further improved for metagenome sequencing.

Third, lignocellulose digestion models of insects consider both host and symbiont. In particular, enzymes

secreted by host insect species play particularly important roles in lignin degradation (Tartar *et al.*, 2009). In previous studies, the metabolism of monoaromatic model compounds by termites and their gut microflora were studied; the results indicated that microbial degradation of plant aromatic compounds can occur in termite guts and may contribute to the carbon and energy requirement of the host (Brune *et al.*, 1995). The recent metagenome and metatrascriptome sequencing of gut symbionts for termite, grasshopper and cutworm has led to the finding of very few lignin-degrading laccases, peroxidases or esterases (Tartar *et al.*, 2009; W.B. Shi, X. Zhou, L.T. Liu, P. Gao, X.Y. Chen, N. Kyprides, E.G. No, S.Y. Dai and J.S. Yuan, unpubl. data). Metaproteomics will provide a powerful solution toward the observation of how biocatalysts from the host and microbes work together to degrade biomass. However, more sequencing information needs to be available to enable such analysis. The study of coordinative function of host and symbiotic microbial biocatalysts will help to guide the reverse-design of biorefineries and the reconstitution of effective enzyme mixtures for biomass degradation.

Fourth, the integration of different 'omics' data into systems-level understanding of insect guts will be important for the reverse-design of artificial reactors mimicking natural biocatalyst systems. Systems biology enables the observation of biological systems and processes at an integrated view (Rachlin *et al.*, 2006). The interaction, dynamics and network of multiple components in a system will be modeled based on genome, proteome, metabolome and transcriptome analyses (Rachlin *et al.*, 2006; Vieites *et al.*, 2009). The accumulation of different 'omics' data regarding insect gut systems will allow us to investigate how different components and biocatalysts work together to fulfill various functions, including biomass degradation.

Overall, we are at a golden age of addressing basic and applied questions involved in insect gut systems. In particular, the recently available 'omics' techniques will revolutionize the field with enormous data to enable unprecedented understanding of insect gut symbiotic microbiota and their interactions with hosts. The systems-level integration of this tremendous information will enable in-depth understanding of natural biocatalyst systems, like insect guts, toward providing novel solutions for next-generation biorefineries.

Acknowledgments

This research was supported by the Texas Agrilife Research Bioenergy Research Initiative, Sungrant and a special research initiative for Bioenergy sponsored by Jiangsu University, China.

References

Adams, I.P., Glover, R.H., Monger, W.A., Mumford, R., Jackeviciene, E., Navalinskiene, M., Samuitiene, M. and Boonham, N. (2009) Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. *Molecular Plant Pathology*, 10, 537–545.

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*. *Nature Genetics*, 32, 402–407.

Allen, E.E. and Banfield, J.F. (2005) Community genomics in microbial ecology and evolution. *Nature Reviews Microbiology*, 3, 489–498.

Allen, H.K., Moe, L.A., Rodbumrer, J., Gaarder, A. and Handelsman, J. (2009) Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. *ISME Journal*, 3, 243–251.

Amann, R.I., Ludwig, W. and Schleifer, K.H. (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews*, 59, 143–169.

Aminov, R.I., Walker, A.W., Duncan, S.H., Harmsen, H.J., Welling, G.W. and Flint, H.J. (2006) Molecular diversity, cultivation, and improved detection by fluorescent in situ hybridization of a dominant group of human gut bacteria related to *Roseburia* spp. or *Eubacterium rectale*. *Applied and Environmental Microbiology*, 72, 6371–6376.

Ansorge, W.J. (2009) Next-generation DNA sequencing techniques. *New Biotechnology*, 25, 195–203.

Backhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A. and Gordon, J.I. (2005) Host-bacterial mutualism in the human intestine. *Science*, 307, 1915–1920.

Bauer, S., Tholen, A., Overmann, J. and Brune, A. (2000) Characterization of abundance and diversity of lactic acid bacteria in the hindgut of wood- and soil-feeding termites by molecular and culture-dependent techniques. *Archives of Microbiology*, 173, 126–137.

Behar, A., Yuval, B. and Jurkevitch E. (2005) Enterobacteria-mediated nitrogen fixation in natural populations of the fruit fly *Ceratitis capitata*. *Molecular Ecology*, 14, 2637–2643.

Berchtold, M., Chatzinotas, A., Schönhuber, W., Brune, A., Amann, R., Hahn, D. and König, H. (1999) Differential enumeration and in situ localization of microorganisms in the hindgut of the lower termite *Mastotermes darwiniensis* by hybridization with rRNA-targeted probes. *Archives of Microbiology*, 172, 407–416.

Bjornsdottir, R., Johannsdottir, J., Coe, J., Smaradottir, H., Agustsson, T., Sigurgisladottir, S. and Gudmundsdottir, B.K. (2009) Survival and quality of halibut larvae (*Hippoglossus hippoglossus* L.) in intensive farming: Possible impact of the intestinal bacterial community. *Aquaculture*, 286, 53–63.

Blackstock, W.P. and Weir, M.P. (1999) Proteomics: quantitative and physical mapping of cellular proteins. *Trends In Biotechnology*, 17, 121–127.

Blaha, C.A.G., Santos, I.L.V.D., Pereira, M.S., Carvalho, C.M., da Silva, C.L.V., de Melo, J.T.A., Silva, A.L.D., de Medeiros, S.R. and Lima, L.F.A. (2007) Metagenomic evaluation of indigenous microbial diversity in tropical mangrove and semi-arid soils from petroliferous basin. *Journal of Biotechnology*, 131, S239.

Bown, D.P., Wilkinson, H.S. and Gatehouse, J.A. (1998) Midgut carboxypeptidase from *Helicoverpa armigera* (Lepidoptera: Noctuidae) larvae: enzyme characterisation, cDNA cloning and expression. *Insect Biochemistry and Molecular Biology*, 28, 739–749.

Brauman, A., Dore, J., Eggleton, P., Bignell, D., Breznak, J.A. and Kane, M.D. (2001) Molecular phylogenetic profiling of prokaryotic communities in guts of termites with different feeding habits. *FEMS Microbiology Ecology*, 35, 27–36.

Breitbart, M., Hewson, I., Felts, B., Mahaffy, J.M., Nulton, J., Salamon, P. and Rohwer, F. (2003) Metagenomic analyses of an uncultured viral community from human feces. *Journal of Bacteriology*, 185, 6220–6223.

Brennan, Y., Callen, W.N., Christoffersen, L., Dupree, P., Goubet, F., Healey, S., Hernandez, M., Keller, M., Li, K., Palackal, N., Sittenfeld, A., Tamayo, G., Wells, S., Hazlewood, G.P., Mathur, E.J., Short, J.M., Robertson, D.E. and Steer, B.A. (2004) Unusual microbial xylanases from insect guts. *Applied and Environmental Microbiology*, 70, 3609–3617.

Breznak, J.A. (2004) Invertebrates-Insects. *Microbial Diversity and Bioprospecting* (ed. A.T. Bull), pp. 191–203. ASM Press, Washington, D.C.

Brinkmann, N., Martens, R. and Tebbc, C.C. (2008) Origin and diversity of metabolically active gut bacteria from laboratory-bred larvae of *Manduca sexta* (Sphingidae, Lepidoptera, Insecta). *Applied and Environmental Microbiology*, 74, 7189–7196.

Broderick, N.A., Raffa, K.F., Goodman, R.M. and Handelsman, J. (2004) Census of the bacterial community of the gypsy moth larval midgut by using culturing and culture-independent methods. *Applied and Environmental Microbiology*, 70, 293–300.

Brulc, J.M., Antonopoulos, D.A., Miller, M.E.B., Wilson, M.K., Yannarell, A.C., Dinsdale, E.A., Edwards, R.E., Frank, E.D., Emerson, J.B., Wacklin, P., Coutinho, P.M., Henrissat, B., Nelson, K.E. and White, B.A. (2009) Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 1948–1953.

Brune, A. (1998) Termite guts: the world's smallest bioreactors. *Trends in Biotechnology*, 16, 16–21.

Brune, A., Miambi, E. and Breznak, J.A. (1995) Roles of oxygen and the intestinal microflora in the metabolism of lignin-derived phenylpropanoids and other monoaromatic compounds by termites. *Applied and Environmental Microbiology*, 61(7), 2688–2695.

Brune, A. (2003) Symbionts aiding digestion. *Encyclopedia of Insects* (eds. R.T. Cardé & V.H. Resh), pp. 1102–1107. Academic Press, Inc., New York.

Casu, R.E., Eisemann, C.H., Vuocolo, T. and Tellam, R.L. (1996) The major excretory/secretory protease from *Lucilia cuprina* larvae is also a gut digestive protease. *International Journal for Parasitology*, 26, 623–628.

Cazemier, A.E., Hackstein, J.H.P., Op den Camp, H.J.M., Rosenberg, J. and van der Drift, C. (1997a) Bacteria in the intestinal tract of different species of arthropods. *Microbial Ecology*, 33, 189–197.

Cazemier, A.E., Op den Camp, H.J.M., Hackstein, J.H.P. and Vogels, G.D. (1997b) Fibre digestion in arthropods. *Comparative Biochemistry and Physiology A: Physiology*, 118, 101–109.

Chapman, A.D. (2006) *Numbers of Living Species in Australia and the World*. Australian Biological Resources Study. Canberra, Australia. 61 pp.

Chaves, S., Neto, M. and Tenreiro, R. (2009) Insect-symbiont systems: From complex relationships to biotechnological applications. *Biotechnology Journal*, 4(12), 1753–1765.

Cheung, S.W., Tishler, P.V., Atkins, L., Sengupta, S.K., Modest, E.J. and Forget, B.G. (1977) Gene mapping by fluorescent in situ hybridization. *Cell Biology International Report*, 1, 255–262.

Chougule, N.P., Giri, A.P., Sainani, M.N. and Gupta, V.S. (2005) Gene expression patterns of *Helicoverpa armigera* gut proteases. *Insect Biochemistry and Molecular Biology*, 35, 355–367.

Colebatch, G., Cooper, P. and East, P. (2002) cDNA cloning of a salivary chymotrypsin-like protease and the identification of six additional cDNAs encoding putative digestive proteases from the green mirid, *Creontiades dilutus* (Hemiptera: Miridae). *Insect Biochemistry and Molecular Biology*, 32, 1065–1075.

Cowan, D., Meyer, Q., Stafford, W., Muyanga, S., Cameron, R. and Wittwer, P. (2005) Metagenomic gene discovery: past, present and future. *Trends in Biotechnology*, 23, 321–329.

Cox-Foster, D.L., Conlan, S., Holmes, E.C., Palacios, G., Evans, J.D., Moran, N.A., Quan, P.L., Briese, T., Hornig, M., Geiser, D.M., Martinson, V., vanEngelsdorp, D., Kalkstein, A.L., Drysdale, A., Hui, J., Zhai, J.H., Cui, L.W., Hutchison, S.K., Simons, J.F., Egholm, M., Pettis, J.S. and Lipkin, W.I. (2007)

A metagenomic survey of microbes in honey bee colony collapse disorder. *Science*, 318, 283–287.

Cravatt, B.F., Simon, G.M. and Yates, J.R. (2007) The biological impact of mass-spectrometry-based proteomics. *Nature*, 450, 991–1000.

Culley, A.I., Lang, A.S. and Suttle, C.A. (2006) Metagenomic analysis of coastal RNA virus communities. *Science*, 312, 1795–1798.

da Mota, F.F., Gomes, E.A., Paiva, E. and Seldin, L. (2005) Assessment of the diversity of *Paenibacillus* species in environmental samples by a novel rpoB-based PCR-DGGE method. *FEMS Microbiology Ecology*, 53, 317–328.

Dasch, G.A., Weiss, E.W. and Chang, K.P. (1984) Endosymbionts of insects. *Bergey's Manual of Systematic Bacteriology* (eds. N.R. Krieg & N.R. Krieg), pp. 811–833. Williams and Wilkins Co., Baltimore, MD.

de Vries, E.J., Breeuwer, J.A.J., Jacobs, G. and Mollema, C. (2001a) The association of western flower thrips, *Frankliniella occidentalis*, with a near *Erwinia* species gut bacteria: Transient or permanent? *Journal of Invertebrate Pathology*, 77, 120–128.

de Vries, E.J., Jacobs, G. and Breeuwer, J.A.J. (2001b) Growth and transmission of gut bacteria in the western flower thrips, *Frankliniella occidentalis*. *Journal of Invertebrate Pathology*, 77, 129–137.

Delahunty, C.M. and Yates, J.R. (2007) MudPIT: multidimensional protein identification technology. *Biotechniques*, 43, 563, 565, 567 passim.

Demaneche, S., David, M.M., Navarro, E., Simonet, P. and Vogel, T.M. (2009) Evaluation of functional gene enrichment in a soil metagenomic clone library. *Journal of Microbiological Methods*, 76, 105–107.

Dillon, R.J. and Dillon, V.M. (2004) The gut bacteria of insects: Nonpathogenic interactions. *Annual Review of Entomology*, 49, 71–92.

Dillon, R.J., Webster, G., Weightman, A.J., Dillon, V. M., Blanford, S. and Charnley, A.K. (2008) Composition of Acridid gut bacterial communities as revealed by 16S rRNA gene analysis. *Journal of Invertebrate Pathology*, 97, 265–272.

Domingo, J.W.S. (1998) Use of 16S rDNA community fingerprints to study cricket hindgut microbial communities. *Journal of General and Applied Microbiology*, 44, 119–127.

Domingo, J.W.S., Kaufman, M.G., Klug, M.J. and Tiedje, J.M. (1998) Characterization of the cricket hindgut microbiota with fluorescently labeled rRNA-targeted oligonucleotide probes. *Applied and Environmental Microbiology*, 64, 752–755.

Donovan, S.E., Purdy, K.J., Kane, M.D. and Eggleton, P. (2004) Comparison of *Euryarchaea* strains in the guts and food-soil of the soil-feeding termite *Cubitermes fungifaber* across different soil types. *Applied and Environmental Microbiology*, 70, 3884–3892.

Douglas, A.E. (1998) Nutritional interactions in insect-microbial symbioses: Aphids and their symbiotic bacteria *Buchnera*. *Annual Review of Entomology*, 43, 17–37.

Egert, M., Stingl, U., Bruun, L.D., Pommerenke, B., Brune, A. and Friedrich, M.W. (2005) Structure and topology of microbial communities in the major gut compartments of *Melolontha melolontha* larvae (Coleoptera: Scarabaeidae). *Applied and Environmental Microbiology*, 71, 4556–4566.

Elshahed, M.S., Najar, F.Z., Aycock, M., Qu, C.N., Roe, B.A. and Krumhoz, L.R. (2005) Metagenomic analysis of the microbial community at Zodletone Spring (Oklahoma): Insights into the genome of a member of the novel candidate division OD1. *Applied and Environmental Microbiology*, 71, 7598–7602.

Erwin, T.L. (1982) Tropical forests: Their richness in Coleoptera and other Arthropod species. *The Coleopterists Bulletin*, 36, 74–75.

Esumi, H., Takahashi, Y., Sato, S. and Sugimura, T. (1982) Restriction fragment length polymorphism of the rat albumin gene in Sprague-Dawley rats and its application in genetic study of analbuminemia. *Nucleic Acids Research*, 10, 4247–4257.

Falkowski, P.G., Fenchel, T. and Delong, E.F. (2008) The microbial engines that drive Earth's biogeochemical cycles. *Science*, 320, 1034–1039.

Fall, S., Hamelin, J., Ndiaye, F., Assigbetse, K., Aragno, M., Chotte, J.L. and Brauman, A. (2007) Differences between bacterial communities in the gut of a soil-feeding termite (*Cubitermes niokoloensis*) and its mounds. *Applied and Environmental Microbiology*, 73, 5199–5208.

Fisher, M., Miller, D., Brewster, C., Husseneder, C. and Dickerman, A. (2007) Diversity of gut bacteria of *Reticulitermes flavipes* as examined by 16S rRNA gene sequencing and amplified rDNA restriction analysis. *Current Microbiology*, 55, 254–259.

Friedrich, M.W., Schmitt-Wagner, D., Lueders, T. and Brune, A. (2001) Axial differences in community structure of *Crenarchaeota* and *Euryarchaeota* in the highly compartmentalized gut of the soil-feeding termite *Cubitermes orthognathus*. *Applied and Environmental Microbiology*, 67, 4880–4890.

Fukatsu, T., Watanable, K. and Sekiguchi, Y. (1998) Specific detection of intracellular symbiotic bacteria of aphids by oligonucleotide-probed in situ hybridization. *Applied Entomology and Zoology*, 33, 461–472.

Gill, S.R., Pop, M., DeBoy, R.T., Eckburg, P.B., Turnbaugh, P.J., Samuel, B.S., Gordon, J.I., Relman, D.A., Fraser-Liggett, C.M. and Nelson, K.E. (2006) Metagenomic analysis of the human distal gut microbiome. *Science*, 312, 1355–1359.

Girard, C. and Jouanin, L. (1999) Molecular cloning of a gut-specific chitinase cDNA from the beetle *Phaedon cochleariae*. *Insect Biochemistry and Molecular Biology*, 29, 549–556.

Green, J.L., Bohannan, B.J. and Whitaker, R.J. (2008) Microbial biogeography: From taxonomy to traits. *Science*, 320, 1039–1043.

Guan, C.H., Ju, J.J., Borlee, B.R., Williamson, L.L., Shen, B., Raffa, K.F. and Handelsman, J. (2007) Signal mimics derived from a metagenomic analysis of the gypsy moth gut microbiota. *Applied and Environmental Microbiology*, 73, 3669–3676.

Guazzaroni, M.E., Beloqui, A., Golyshin, P.N. and Ferrer, M. (2009) Metagenomics as a new technological tool to gain scientific knowledge. *World Journal of Microbiology and Biotechnology*, 25, 945–954.

Handelsman, J. (2004) Metagenomics: application of genomics to uncultured microorganisms. *Microbiology and Molecular Biology Reviews*, 68, 669–685.

Handelsman, J., Liles, M., Mann, D., Riesenfeld, C. and Goodman, R.M. (2002) Cloning the metagenome: Culture-independent access to the diversity and functions of the uncultivated microbial world. *Functional Microbial Genomics* (eds. B. Wren & N. Dorrell), pp. 241–255, Academic Press Inc, San Diego.

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J. and Goodman, R.M. (1998) Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. *Chemistry & Biology*, 5, R245–R249.

Harada, H. and Ishikawa, H. (1993) Gut microbe of aphid closely-related to its intracellular symbiont. *Biosystems*, 31, 185–191.

Harmon, J.P., Moran, N.A. and Ives, A.R. (2009) Species response to environmental change: Impacts of food web interactions and evolution. *Science*, 323, 1347–1350.

Hayashi, A., Aoyagi, H., Yoshimura, T. and Tanaka, H. (2007) Development of novel method for screening microorganisms using symbiotic association between insect (*Coptotermes formosanus* Shiraki) and intestinal microorganisms. *Journal of Bioscience and Bioengineering*, 103, 358–367.

Head, I.M., Saunders, J.R. and Pickup, R.W. (1998) Microbial evolution, diversity, and ecology: A decade of ribosomal RNA analysis of uncultivated microorganisms. *Microbial Ecology*, 35, 1–21.

Hogenkamp, D.G., Arakane, Y., Zimoch, L., Merzendorfer, H., Kramer, K.J., Beeman, R.W., Kanost, M.R., Specht, C.A. and Muthukrishnan, S. (2005) Chitin synthase genes in *Manduca sexta*: characterization of a gut-specific transcript and differential tissue expression of alternately spliced mRNAs during development. *Insect Biochemistry and Molecular Biology*, 35, 529–540.

Holben, W.E., Jansson, J.K., Chemlm, B.K. and Tiedje, J.M. (1988) DNA probe method for the detection of specific microorganisms in the soil bacterial community. *Applied and Environmental Microbiology*, 54, 703–711.

Hongoh, Y., Ekpornprasit, L., Inoue, T., Moriya, S., Trakulnaleamsai, S., Ohkuma, M. Noparatnaraporn, N. and Kudo, T. (2006) Intracolony variation of bacterial gut microbiota among castes and ages in the fungus-growing termite *Macrotermes gilvus*. *Molecular Ecology*, 15, 505–516.

Hongoh, Y., Sharma, V. K., Prakash, T., Noda, S., Taylor, T.D., Kudo, T., Sakaki, Y., Toyoda, A., Hattori, M. and Ohkuma, M. (2008a) Complete genome of the uncultured termite group 1 bacteria in a single host protist cell. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 5555–5560.

Hongoh, Y., Sharma, V.K., Prakash, T., Noda, S., Toh, H., Taylor, T.D., Kudo, T., Sakaki, Y., Toyoda, A., Hattori, M. and Ohkuma, M. (2008b) Genome of an endosymbiont coupling N-2 fixation to cellulolysis within protist cells in termite gut. *Science*, 322, 1108–1109.

Hopkins, D.W., Macnaughton, S.J. and Odonnell, A.G. (1991) A dispersion and differential centrifugation technique for representatively sampling microorganisms from soil. *Soil Biology & Biochemistry*, 23, 217–225.

Hosokawa, T., Kikuchi, Y., Nikoh, N., Shimada, M. and Fukatsu, T. (2006) Strict host-symbiont cospeciation and reductive genome evolution in insect gut bacteria. *PLoS Biology*, 4, 1841–1851.

Hurst, M.R.H. and Jackson, T.A. (2002) Use of the green fluorescent protein to monitor the fate of *Serratia entomophila* causing amber disease in the New Zealand grass grub, *Costelytra zealandica*. *Journal of Microbiological Methods*, 50, 1–8.

Inoue, T., Moriya, S., Ohkuma, M. and Kudo, T. (2005) Molecular cloning and characterization of a cellulase gene from a symbiotic protist of the lower termite, *Coptotermes formosanus*. *Gene*, 349, 67–75.

Jackson, C.R., Harper, J.P., Willoughby, D., Roden, E.E. and Churchill, P.F. (1997) A simple, efficient method for the separation of humic substances and DNA from environmental samples. *Applied and Environmental Microbiology*, 63, 4993–4995.

Jones, B.F. and Hotez, P.J. (2002) Molecular cloning and characterization of Ac-mep-1, a developmentally regulated gut luminal metalloendopeptidase from adult *Ancylostoma caninum* hookworms. *Molecular and Biochemical Parasitology*, 119, 107–116.

Kan, J., Hanson, T.E., Ginter, J.M., Wang, K. and Chen, F. (2005) Metaproteomic analysis of Chesapeake Bay microbial communities. *Saline Systems*, 1, 7.

Kane, M.D. and Pierce, N.E. (1994a) Diversity within diversity: molecular approaches to studying microbial interactions with insects. *EXS*, 69, 509–524.

Kane, M.D. and Pierce, N.E. (1994b) Diversity within diversity: molecular approaches to studying microbial interactions with insects. *Molecular Methods in Ecology and Evolution* (eds.

B. Schierwater, B. Streit, G.P. Wagner & R. DeSalle), pp. 509–524. Birkhauser Verlag, Berlin, Germany.

Kaufman, M.G., Walker, E.D., Odelson, D.A. and Klug, M.J. (2000) Microbial community ecology & insect nutrition. *American Entomologist*, 46, 173–185.

Kauppinen, J., Mantyjarvi, R. and Katila, M.L. (1999) Mycobacterium malmoense-specific nested PCR based on a conserved sequence detected in random amplified polymorphic DNA fingerprints. *Journal of Clinical Microbiology*, 37, 1454–1458.

Keller, M. and Zengler, K. (2004) Tapping into microbial diversity. *Nature Reviews Microbiology*, 2, 141–150.

Khachane, A.N., Timmis, K.N. and Martins dos Santos, V.A. (2007) Dynamics of reductive genome evolution in mitochondria and obligate intracellular microbes. *Molecular Biology and Evolution*, 24, 449–456.

Kohler, T., Stingl, U., Meuser, K. and Brune, A. (2008) Novel lineages of Planctomycetes densely colonize the alkaline gut of soil-feeding termites (*Cubitermes* spp.). *Environmental Microbiology*, 10, 1260–1270.

Lee, D.H., Zo, Y.G. and Kin, S.J. (1996) Nonradioactive method to study genetic profiles of natural bacterial communities by PCR-single-strand-conformation polymorphism. *Applied and Environmental Microbiology*, 62, 3112–3120.

Li, L., Frohlich, J., Pfeiffer, P. and König, H. (2003) Termite gut symbiotic archaezoa are becoming living metabolic fossils. *Eukaryotic Cell*, 2, 1091–1098.

Lilburn, T.G., Schmidt, T.M. and Breznak, J.A. (1999) Phylogenetic diversity of termite gut spirochaetes. *Environmental Microbiology*, 1, 331–345.

Lindh, J.M., Borg-Karlson, A.K. and Faye, I. (2008) Transstadial and horizontal transfer of bacteria within a colony of *Anopheles gambiae* (Diptera: Culicidae) and oviposition response to bacteria-containing water. *Acta Tropica*, 107, 242–250.

Liu, W.T., Marsh, T.L., Cheng, H. and Forney, L.J. (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Applied and Environmental Microbiology*, 63, 4516–4522.

Lohrig, K. and Wolters, D. (2009) Multidimensional protein identification technology. *Methods in Molecular Biology*, 564, 143–153.

Lorenz, P. and Eck, J. (2005) Metagenomics and industrial applications. *Nature Reviews Microbiology*, 3, 510–516.

Mackenzie, L.M., Muigai, A.T., Osir, E.O., Lwande, W., Keller, M., Toledo, G. and Boga, H.I. (2007) Bacterial diversity in the intestinal tract of the fungus-cultivating termite *Macrotermes michaelseni* (Sjostedt). *African Journal of Biotechnology*, 6, 658–667.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J. M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F. and Rothberg, J.M. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057), 376–380.

Manefield, M., Whiteley, A.S., Griffiths, R.I. and Bailey, M.J. (2002) RNA stable isotope probing, a novel means of linking microbial community function to phylogeny. *Applied and Environmental Microbiology*, 68, 5367–5373.

Mardis, E.R. (2008) The impact of next-generation sequencing technology on genetics. *Trends of Genetics*, 24(3), 133–141.

Moran, N.A, McCutcheon, J.P. and Nakabachi, A. (2008) Genomics and evolution of heritable bacterial symbionts. *Annual Review of Genetics*, 42, 165–190.

Marusina, K. (2006) The next generation of DNA sequencing. *Genetic Engineering News*, 26, 28–31.

McIntosh, D., Ji, B., Forward, B.S., Puvanendran, V., Boyce, D. and Ritchie, R. (2008) Culture-independent characterization of the bacterial populations associated with cod (*Gadus morhua* L.) and live feed at an experimental hatchery facility using denaturing gradient gel electrophoresis. *Aquaculture*, 275, 42–50.

McKillip, J.L., Small, C.L., Brown, J.L., Brunner, J.F. and Spence, K.D. (1997) Sporogenous midgut bacteria of the leafroller, *Pandemis pyrusana* (Lepidoptera: Tortricidae). *Environmental Entomology*, 26, 1475–1481.

Miller, D.N., Bryant, J.E., Madsen, E.L. and Ghiorse, W.C. (1999) Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. *Applied and Environmental Microbiology*, 65, 4715–4724.

Miyata, R., Noda, N., Tamaki, H., Kinjyo, K., Aoyagi, H., Uchiyama, H. and Tanaka, H. (2007) Influence of feed components on symbiotic bacterial community structure in the gut of the wood-feeding higher termite *Nasutitermes takasagoensis*. *Bioscience Biotechnology and Biochemistry*, 71, 1244–1251.

Mohr, K.I. and Tebbe, C.C. (2006) Diversity and phylotype consistency of bacteria in the guts of three bee species (*Apoidea*) at an oilseed rape field. *Environmental Microbiology*, 8, 258–272.

Moran, N.A. (2001) Bacterial menageries inside insects. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 1338–1340.

Moran, N.A. (2002) Microbial minimalism: Genome reduction in bacterial pathogens. *Cell*, 108, 583–586.

Moran, N.A. (2007) Symbiosis as an adaptive process and source of phenotypic complexity. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 8627–8633.

Moran, N.A. and Telang, A. (1998) Bacteriocyte-associated symbionts of insects – A variety of insect groups harbor ancient prokaryotic endosymbionts. *Bioscience*, 48, 295–304.

More, M.I., Herrick, J.B., Silva, M.C., Ghiorse, W.C. and Madsen, E.L. (1994) Quantitative cell-Lysis of indigenous microorganisms and rapid extraction of microbial DNA from sediment. *Applied and Environmental Microbiology*, 60, 1572–1580.

Muyzer, G. and Smalla, K. (1998) Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonie van Leeuwenhoek*, 73, 127–141.

Muyzer, G., de Waal, E.C. and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by denaturing gradient gel-electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16s rRNA. *Applied and Environmental Microbiology*, 59, 695–700.

Nakashima, K., Watanabe, H., Saitoh, H., Tokuda, G. and Azuma, J.I. (2002) Dual cellulose-digesting system of the wood-feeding termite, *Coptotermes formosanus* Shiraki. *Insect Biochemistry and Molecular Biology*, 32, 777–784.

Nelson, K.E. (2008) Metagenomics as a tool to study biodiversity. *Accessing Uncultivated Microorganisms: From the Environment to Organisms and Genomes and Back* (eds. K. Zengler), pp. 153–170. ASM Press, Washington, DC.

Noda, S., Ohkuma, M., Usami, R., Horikoshi, K. and Kudo, T. (1999) Culture-independent characterization of a gene responsible for nitrogen fixation in the symbiotic microbial community in the gut of the termite *Neotermes koshunensis*. *Applied and Environmental Microbiology*, 65, 4935–4942.

Ochman, H. and Moran, N.A. (2001) Genes lost and genes found: Evolution of bacterial pathogenesis and symbiosis. *Science*, 292, 1096–1098.

Ogram, A., Sayler, G.S. and Barkay, T. (1987) The extraction and purification of microbial DNA from sediments. *Journal of Microbiological Methods*, 7, 57–66.

Ogunseitan, O.A. (1993) Direct extraction of proteins from environmental-samples. *Journal of Microbiological Methods*, 17, 273–281.

Ogunseitan, O.A. (1997) Direct extraction of catalytic proteins from natural microbial communities. *Journal of Microbiological Methods*, 28, 55–63.

Ohkuma, M. (2003) Termite symbiotic systems: efficient bio-recycling of lignocellulose. *Applied Microbiology and Biotechnology*, 61, 1–9.

Ohkuma, M. and Kudo, T. (1996) Phylogenetic diversity of the intestinal bacterial community in the termite *Reticulitermes speratus*. *Applied and Environmental Microbiology*, 62, 461–468.

Oliver, J.D. (2000) Problems in detecting dormant (VBNC) cells, and the role of DNA elements in this response. *Tracking Genetically-Engineered Microorganisms* (eds. J.K. Jansson, J.D. van Elsas & M. Bailey), pp. 1–15. Landes Bioscience, Georgetown, Texas.

Orozco-Medina, C., Lopez-Cortes, A. and Maeda-Martinez, A.M. (2009) Aerobic Gram-positive heterotrophic bacteria *Exiguobacterium mexicanum* and *Microbacterium* sp. in the gut lumen of *Artemia franciscana* larvae under gnotobiotic conditions. *Current Science*, 96, 120–129.

Osborn, A.M., Moor, E.R. and Timmis, K.N. (2000) An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics. *Environmental Microbiology*, 2, 39–50.

Pace, N.R. (1997) A molecular view of microbial diversity and the biosphere. *Science*, 276, 734–740.

Paster, B.J., Dewhirst, F.E., Cooke, S.M., Fussing, V., Poulsen, L.K. and Breznak, J.A. (1996) Phylogeny of not-yet-cultured spirochetes from termite guts. *Applied and Environmental Microbiology*, 62, 347–352.

Pillai, S.D., Josephson, K.L., Bailey, R.L., Gerba, C.P. and Pepper, I.L. (1991) Rapid method for processing soil samples for polymerase chain reaction amplification of specific gene sequences. *Applied and Environmental Microbiology*, 57, 2283–2286.

Pittman, G.W., Brumbley, S.M., Allsopp, P.G. and O'Neill, S.L. (2008a) Assessment of gut bacteria for a paratransgenic approach to control *Dermolepida albohirtum* larvae. *Applied and Environmental Microbiology*, 74, 4036–4043.

Pittman, G.W., Brumbley, S.M., Allsopp, P.G. and O'Neill, S.L. (2008b) "Endomicrobia" and other bacteria associated with the hindgut of *Dermolepida albohirtum* larvae. *Applied and Environmental Microbiology*, 74, 762–767.

Pootanakit, K., Angsuthanasombat, C. and Panyim, S. (2003) Identification of two isoforms of aminopeptidase N in *Aedes aegypti* larval midgut. *Journal of Biochemistry and Molecular Biology*, 36, 508–513.

Rachlin, J., Cohen, D.D., Cantor, C. and Kasif, S. (2006) Biological context networks: a mosaic view of the interactome. *Molecular Systems Biology*, 2, 66.

Raymond, B., Lijek, R.S., Griffiths, R.I. and Bonsall, M.B. (2008) Ecological consequences of ingestion of *Bacillus cereus* on *Bacillus thuringiensis* infections and on the gut flora of a lepidopteran host. *Journal of Invertebrate Pathology*, 99, 103–111.

Reeson, A.F., Jankovic, T., Kasper, M.L., Rogers, S. and Austin, A.D. (2003) Application of 16S rDNA-DGGE to examine the microbial ecology associated with a social wasp *Vespula germanica*. *Insect Molecular Biology*, 12, 85–91.

Robe, P., Nalin, R., Capellano, C., Vogel, T.A. and Simonet, P. (2003) Extraction of DNA from soil. *European Journal of Soil Biology*, 39, 183–190.

Roose-Amsaleg, C.L., Garnier-Sillam, E. and Harry, M. (2001) Extraction and purification of microbial DNA from soil and sediment samples. *Applied Soil Ecology*, 18, 47–60.

Roussel, E.G., Bonavita, M.A.C., Querellou, J., Cragg, B.A., Webster, G., Prieur, D. and Parkes, R.J. (2008) Extending the sub-sea-floor biosphere. *Science*, 320, 1046.

Santo Domingo, J.W., Kaufman, M.G., Klug, M.J., Holben, W.E., Harris, D. and Tiedje, J.M. (1998) Influence of diet on the structure and function of the bacterial hindgut community of crickets. *Molecular Ecology*, 7, 761–767.

Schabereiter-Gurtner, C., Lubitz, W. and Rolleke, S. (2003) Application of broad-range 16S rRNA PCR amplification and DGGE fingerprinting for detection of tick-infecting bacteria. *Journal of Microbiological Methods*, 52, 251–260.

Scheibye-Alsing, K., Hoffmann, S., Frankel, A., Jensen, P., Stadler, P.F., Mang, Y., Tommerup, N., Gilchrist, M.J., Nygard, A.B., Cirera, S., Jorgensen, C.B., Fredholm, M. and Gorodkin, J. (2009) Sequence assembly. *Computational Biology and Chemistry*, 33, 121–136.

Schloss, P.D., Delalibera, I., Handelsman, J. and Raffa, K.F. (2006) Bacteria associated with the guts of two wood-boring beetles: *Anoplophora glabripennis* and *Saperda vestita* (Cerambycidae). *Environmental Entomology*, 35, 625–629.

Schmeisser, C., Steele, H. and Streit, W.R. (2007) Metagenomics, biotechnology with non-culturable microbes. *Applied Microbiology and Biotechnology*, 75, 955–962.

Schmitt-Wagner, D., Friedrich, M.W., Wagner, B. and Brune, A. (2003a) Axial dynamics, stability, and interspecies similarity of bacterial community structure in the highly compartmentalized gut of soil-feeding termites (*Cubitermes* spp.). *Applied and Environmental Microbiology*, 69, 6018–6024.

Schmitt-Wagner, D., Friedrich, M.W., Wagner, B. and Brune, A. (2003b) Phylogenetic diversity, abundance, and axial distribution of bacteria in the intestinal tract of two soil-feeding termites (*Cubitermes* spp.). *Applied and Environmental Microbiology*, 69, 6007–6017.

Schulze, W.X., Gleixner, G., Kaiser, K., Guggenberger, G., Mann, M., Schulze, E.D. (2005) A proteomic fingerprint of dissolved organic carbon and of soil particles. *Oecologia*, 142, 335–343.

Schwieger, F. and Tebbe, C.C. (1998) A new approach to utilize PCR-single-strand-conformation polymorphism for 16s rRNA gene-based microbial community analysis. *Applied and Environmental Microbiology*, 64, 4870–4876.

Shen, Z. and Jacobs-Lorena, M. (1997) Characterization of a novel gut-specific chitinase gene from the human malaria vector *Anopheles gambiae*. *Journal of Biological Chemistry*, 272, 28895–28900.

Shendure, J. and Ji, H. (2008) Next-generation DNA sequencing. *Nature Biotechnology*, 26, 1135–1145.

Shi, W.B., Ding, S.Y. and Yuan, J.S. (2010) Comparison of insect gut cellulase and xylanase activity across different insect species with distinct food sources. *BioEnergy Research*. (in press)

Shinzato, N., Muramatsu, M., Matsui, T. and Watanabe, Y. (2007) Phylogenetic analysis of the gut bacterial microflora of the fungus-growing termite *Odontotermes formosanus*. *Bioscience, Biotechnology, and Biochemistry*, 71, 906–915.

Shinzato, N., Muramatsu, M., Matsui, T. and Watanabe, Y. (2005) Molecular phylogenetic diversity of the bacterial community in the gut of the termite *Coptotermes formosanus*. *Bioscience, Biotechnology, and Biochemistry*, 69, 1145–1155.

Singleton, I., Merrington, G., Colvan, S. and Delahunty, J.S. (2003) The potential of soil protein-based methods to indicate metal contamination. *Applied Soil Ecology*, 23, 25–32.

Sleator, R.D., Shortall, C. and Hill, C. (2008) Metagenomics. *Letters in Applied Microbiology*, 47, 361–366.

Smalla, K. (2004) Culture-independent microbiology. *Microbial Diversity and Bioprospecting* (ed. A.T. Bull), pp. 88–99. ASM Press, Washington, D.C.

Smalla, K., Oros-Sichler, M., Milling, A., Heuer, H., Baumgarte, S., Becker, R., Neuber, G., Kropf, S., Ulrich, A. and Tebbe, C.C. (2007) Bacterial diversity of soils assessed by DGGE, T-RFLP and SSCP fingerprints of PCR-amplified 16S rRNA gene fragments: Do the different methods provide similar results? *Journal of Microbiological Methods*, 69, 470–479.

Spiteller, D., Dettner, K. and Bolan, W. (2000) Gut bacteria may be involved in interactions between plants, herbivores and their predators: Microbial biosynthesis of N-acylglutamine surfactants as elicitors of plant volatiles. *Biological Chemistry*, 381, 755–762.

Stangier, K.A. (2009) Next-generation sequencing: a short comparison. *LaboratoriumsMedizin-Journal of Laboratory Medicine*, 33, 153–157.

Steele, H.L. and Streit, W.R. (2005) Metagenomics: Advances in ecology and biotechnology. *FEMs Microbiology Letters*, 247, 105–111.

Stingl, U. and Brune, A. (2003) Phylogenetic diversity and whole-cell hybridization of oxymonad flagellates from the hindgut of the wood-feeding lower termite *Reticulitermes flavipes*. *Protist*, 154, 147–155.

Stokes, H.W., Holmes, A.J., Nield, B.S., Holley, M.P., Nevalainen, K.M.H., Mabbutt, B.C. and Gillings, M.R. (2001) Gene cassette PCR: Sequence-independent recovery of entire genes from environmental DNA. *Applied and Environmental Microbiology*, 67, 5240–5246.

Strom, S.L. (2008) Microbial ecology of ocean biogeochemistry: A community perspective. *Science*, 320, 1043–1045.

Sun, J.Z. and Zhou, X.G. (2009) Utilization of lignocellulose-feeding insects for viable biofuels: an emerging and promising area of entomological science. *Recent Advances in Entomological Research* (eds. T.-X. Liu & L. Kang), pp. 251–291. Higher Education Press, Beijing, China (in press).

Todaka, N., Moriya, S., Saita, K., Hondo, T., Kiuchi, I., Takasu, H., Ohkuma, M., Piero, C., Hayashizaki, Y. and Kudo, T. (2007) Environmental cDNA analysis of the genes involved in lignocellulose digestion in the symbiotic protist community of *Reticulitermes speratus*. *Fems Microbiology Ecology*, 59(3), 592–599.

Tamas, I., Klasson, L., Canback, B., Naslund, A.K., Eriksson, A.S., Wernegreen, J.J., Sandstrom, J.P., Moran, N.A. and Andersson, S.G.E. (2002) 50 million years of genomic stasis in endosymbiotic bacteria. *Science*, 296, 2376–2379.

Tamas, I., Wernegreen, J.J., Nystedt, B., Kauppinen, S.N., Darby, A.C., Gomez-Valero, L., Lundin, D., Poole, A.M. and Andersson, S.G.E. (2008) Endosymbiont gene functions impaired and rescued by polymerase infidelity at poly(A) tracts. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 14934–14939.

Tartar, A., Wheeler, M.M., Zhou, X.G., Coy, M.R., Boucias, D.G. and Scharf, M.E. (2009) Parallel metatranscriptome analyses of host and symbiont gene expression in the gut of the termite *Reticulitermes flavipes*. *Biotechnology for Biofuels*, 2, 25.

Tebbe, C.C. and Vahjen, W. (1993) Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant-DNA from bacteria and a yeast. *Applied and Environmental Microbiology*, 59, 2657–2665.

Tokuda, G. and Watanabe, H. (2007) Hidden cellulases in termites: revision of an old hypothesis. *Biology Letters*, 3, 336–339.

Torsvik, V.L. (1980) Isolation of bacterial-DNA from soil. *Soil Biology & Biochemistry*, 12, 15–21.

Tringe, S.G., von Mering, C., Kobayashi, A., Salamov, A.A., Chen, K., Chang, H.W., Podar, M., Short, J.M., Mathur, E.J., Detter, J.C., Bork, P., Hugenholtz, P. and Rubin, E.M. (2005) Comparative metagenomics of microbial communities. *Science*, 308, 554–557.

Turroni, F., Ribbera, A., Foroni, E., van Sinderen, D. and Ventura, M. (2008) Human gut microbiota and bifidobacteria: from composition to functionality. *Antonie van Leeuwenhoek*, 94, 35–50.

Tyson, G.W., Chapman, J., Hugenholtz, P., Allen, E.E., Ram, R.J., Richardson, P.M., Solovyev, V.V., Rubin, E.M., Rokhsar, D.S. and Banfield, J.F. (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature*, 428, 37–43.

van Ham, R., Kamerbeek, J., Palacios, C., Rausell, C., Abascal, F., Bastolla, U., Fernandez, J.M., Jimenez, L., Postigo, M., Silva, F.J., Tamames, J., Viguera, E., Latorre, A., Valencia, A., Moran, F. and Moya, A. (2003) Reductive genome evolution in *Buchnera aphidicola*. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 581–586.

Vasanthakumar, A., Delalibera, I., Handelsman, J., Klepzig, K.D., Schloss, P.D. and Raffa, K.F. (2006) Characterization of gut-associated bacteria in larvae and adults of the southern pine beetle, *Dendroctonus frontalis* Zimmermann. *Environmental Entomology*, 35, 1710–1717.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D.Y., Paulsen, I., Nelson, K.E., Nelson, W., Fouts, D.E., Levy, S., Knap, A.H., Lomas, M.W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannkoch, C., Rogers, Y.H. and Smith, H.O. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science*, 304, 66–74.

Vieites, J.M., Guazzaroni, M.E., Beloqui, A., Golyshin, P.N. and Ferrer, M. (2009) Metagenomics approaches in systems microbiology. *FEMS Microbiol Reviews*, 33, 236–255.

Voget, S., Leggewie, C., Uesbeck, A., Raasch, C., Jaeger, K.E. and Streit, W.R. (2003) Prospecting for novel biocatalysts in a soil metagenome. *Applied and Environmental Microbiology*, 69, 6235–6242.

Warnecke, F. and Hugenholtz, P. (2007) Building on basic metagenomics with complementary technologies. *Genome Biology*, 8, 231.

Warnecke, F., Luginbuhl, P., Ivanova, N., Ghassemian, M., Richardson, T.H., Stege, J.T., Cayouette, M., McHardy, A.C., Djordjevic, G., Aboushadi, N., Sorek, R., Tringe, S.G., Podar, M., Martin, H.G., Kunin, V., Dalevi, D., Madejska, J., Kirton, E., Platt, D., Szeto, E., Salamov, A., Barry, K., Mikhailova, N., Kyrpides, N.C., Matson, E.G., Ottesen, E.A., Zhang, X., Hernandez, M., Murillo, C., Acosta, L.G., Rigoutsos, I., Tamayo, G., Green, B.D., Chang, C., Rubin, E.M., Mathur, E.J., Robertson, D.E., Hugenholtz, P. and Leadbetter, J.R. (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature*, 450, 560–565.

Webster, G., Newberry, C.J., Fry, J.C. and Weightman, A.J. (2003) Assessment of bacterial community structure in the deep sub-seafloor biosphere by 16S rDNA-based techniques: a cautionary tale. *Journal of Microbiological Methods*, 55, 155–164.

Wei, Y.D., Lee, K.S., Gui, Z.Z., Yoon, H.J., Kim, I., Zhang, G.Z., Guo, X., Sohn, H.D. and Jin, B.R. (2006) Molecular cloning, expression, and enzymatic activity of a novel endogenous cellulase from the mulberry longicorn beetle, *Apriona germari*. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology*, 145, 220–229.

Wheeler, M.M., Zhou, X.G., Scharf, M.E. and Oi, F.M. (2007) Molecular and biochemical markers for monitoring dynamic shifts of cellulolytic protozoa in *Reticulitermes flavipes*.

Insect Biochemistry and Molecular Biology, 37(12), 1366–1374.

Williamson, L.L., Borlee, B.R., Schloss, P.D., Guan, C.H., Allen, H.K. and Handelsman, J. (2005) Intracellular screen to identify metagenomic clones that induce or inhibit a quorum-sensing biosensor. *Applied and Environmental Microbiology*, 71, 6335–6344.

Wilmes, P. and Bond, P.L. (2004) The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms. *Environmental Microbiology*, 6, 911–920.

Yandell, D.W. (1991) SSCP (single strand conformation polymorphism) analysis of PCR fragments. *Japanese Journal of Cancer Research*, 82, 1468–1469.

Yoon, H.S., Lee, C.S., Lee, S.Y., Choi, C.S., Lee, I.H., Yeo, S.M. and Kim, H.R. (2003) Purification and cDNA cloning of inducible antibacterial peptides from *Protaetia brevitarsis* (Coleoptera). *Archives of Insect Biochemistry and Physiology*, 52, 92–103.

Yuan, J.S., Galbraith, D.W., Dai, S.Y., Griffin, P. and Stewart, C.N. (2008) Plant systems biology comes of age. *Trends in Plant Science*, 13, 165–171.

Zhang, H. and Jackson, T.A. (2008) Autochthonous bacterial flora indicated by PCR-DGGE of 16S rRNA gene fragments from the alimentary tract of *Costelytra zealandica* (Coleoptera: Scarabaeidae). *Journal of Applied Microbiology*, 105, 1277–1285.

Zhou, X.G., Smith, J.A., Oi, F.M., Koehler, P.G., Bennett, G.W. and Scharf, M.E. (2007) Correlation of cellulase gene expression and cellulolytic activity throughout the gut of the termite *Reticulitermes flavipes*. *Gene*, 395(1–2), 29–39.

Accepted February 22, 2010