Molecular Approaches to Study the Insect Gut Symbiotic Microbiota at the 'Omics' Age[①]

Weibing Shi[1,2], Ryan Syrenne[1], Jianzhong Sun[3], and Joshua S. Yuan[1,2,3,4]*

1, Department of Plant Pathology and Microbiology, Texas A&M University, College Station, TX, 77843, USA

2, Institute for Plant Genomics and Biotechnology, Texas A&M University, College Station, TX, 77843, USA

3, Biofuels Institute, School of the Environments, Jiangsu University, Zhenjiang, Jiangsu Province, 212013, China

4, Advanced Research Institute for Sustainable Energy (ARISE), Texas A&M University, College Station, TX, 77843, 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 the insect gut symbiont analysis. The review aims to serve as an informative resource for experimental design and research strategy development in the field. We first discussed various strategies for sample preparation and their pros and cons. The traditional molecular techniques like DGGE, RFLP, and FISH were covered with respect to how they are applied to study the composition, diversity, and dynamics of insect gut symbiotic microbiota. We then focused on the various 'omics' techniques. The metagenome analysis together with the recent advancements in next generation sequencing will provide enormous sequencing information allowing the in-depth microbial diversity analysis and the modeling of pathways for biological processes like 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 the functions. The molecular and systems level understanding will also

① 原文载于 *Insect Science* (2010) 17, 199 - 219

guide the reverse design of next generation biorefinery.

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

I Introduction-Why to 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 microorganisms seem to be an important common property for different insect species (Breznak, 2004).

I.1 The Definition and Importance of the Symbiosis

Symbiosis often refers to the long-term and mutually beneficial interactions among different species. Symbiotic microbes living inside of the host species are referred as endosymbionts, and the symbiotic microbes living upon or outside an insect's body are often defined as ectosymbionts (Breznak, 2004). Based on the previous studies, endosymbionts are prevalent in a variety of insect species such as scarab beetles, cockroaches, and termites et al. (Brune, 2003; Dasch et al., 1984; Kane and 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 and Telang, 1998). Considering that Insecta is the largest group of invertebrates, it is important to study the 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 the plant-insect interaction in terms of host selection and co-evolution of host-insect relationships. From the application perspective, the study of insect symbionts will help to discover novel biocatalysts for biomass deconstruction and develop innovative strategies for pest management.

I.2 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 the 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 the 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 (A*km*an et al., 2002). The phenomena are typical for the symbiotic microbes, which often dedicate part of their genomes for the benefit of the hosts (Moran, 2001; Ochman and Moran, 2001). Recent metagenome project also revealed that the viruses affecting the symbionts of honeybee will lead to detrimental effects on the 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 the insect symbiotic microbiota is the biomass deconstruction and digestion functions. Both herbivore insects and symbiotic microbes can secrete cellulytic enzymes for biomass deconstruction and hydrolysis (Ohkuma, 2003; Tokuda and Watanabe, 2007; Warnecke et al., 2007; Sun and Zhou 2009). It has been controversial about which plays a more important roles for biomass deconstruction, the symbionts or insect host itself. Despite the controversy, the importance of the symbiotic microbes for biomass deconstruction has recently been established by the various genome-level studies. For example, the symbiotic microbiota can help termites to deconstruct the lignocellulosic biomass with a high efficiency (Ohkuma, 2003). Termite gut has actually been referred 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 like 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 the novel biocatalysts and microbe strain discovery. Combining with the 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.

II 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.

II.1 Insect Gut Metagenomic DNA Extraction

Metagenomics can be defined as the study of metagenome, the whole genetic material of the microbial community existing in certain eco-environment (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 key to reaching a comprehensive and unbiased evaluation of metagenomes of the community, particularly for the unculturable microorganisms (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 the high molecular weight and high quality metagenomic DNA. The third aspect is the purity. The metagenomic DNA should be free of contaminants interfering with downstream DNA processing such as enzyme digestion, 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 were published more than three decades ago (Torsvik, 1980). Two strategies have been popular for metagenomic DNA isolation, and they are cell recovery method and direct lysis method (Roose-Amsaleg, 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 medium 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 as shown in Table 1. For instance, Schloss et al. used FastDNA SPIN kit for soil (MP Biomedical, USA) to isolate the metagenomic DNA from the wood-boring beetle's 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 and Jackson, 2008), feral locusts, grasshoppers (Dillon et al., 2008), and termite (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) (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 gutsymbiont metagenomic DNA extraction (Shi et al., 2009b).

Table1 Commercial kits for metagenome DNA extraction and their application in insect gut system

Company	Target product	Website	Application in insect gut symbiota
MP Biomedicals	FastDNA SPIN kit for soil	http://www.mpbio.com	(Zhang and Jackson,2008)Dillon et al., 2008 (Shinzato et al., 2005a)
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

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 the 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 the environmental DNA, and these contaminants are visible as a dark color in the DNA sample. The contaminants

have been demonstrated to be the inhibitors for DNA hybridization, digestion and PCR (polymerase chain reactions) (Jackson et al., 1997; Miller et al., 1999; Tebbe and 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.

II. 2 Protein for 'Omics' Analysis

Besides the metagenomics, metaproteomics are also important perspectives for analyzing the 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 the protein expression (Cravatt et al., 2007). As for the metagenome analysis, sample preparation is also crucial for the metaproteomics. The challenges come from the requirements from both the environmental samples and the ESI (electrospray ionization) mass spectrometry 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 SDS should be avoided. On the other side, the sample preparation from insect guts needs to be comprehensive and the 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; Ogunseitan, 1997). One is the boiling method, which recovered high concentrations of proteins from wastewater but not from soil and sediments. The other one is the freeze-thaw method, which worked better for soils and sediments (Ogunseitan, 1993; Ogunseitan, 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 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 ex-

tracted from wood-feeding higher termite hindgut (Warnecke et al., 2007). The sample preparation involves the 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 the measurement of the soluble proteins in the gut contents. However, the 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 the 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.

III Traditional Molecular Techniques to Investigate the Insect Gut Microbiota

Traditional molecular techniques played an important role in furthering our understanding of the composition and function of the 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 microorganisms and the significant diversity present in insect gut (Head et al., 1998; Pace, 1997). Due to the limitation of the cultivation-based methods, it was expected that most of the diversity in the insect gut microbiomes were still unknown (Stokes et al., 2001). In order to study the diversity of insect gut microbial community, three major molecular approaches have been employed to discover new genes and investigate the composition of the gut microbial community. These three approaches includes 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).

III.1 Gene Targeting: Gene Specific PCR

Gene targeting techniques employs the gene-specific primers to specifically amplify the 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 (Kane and Pierce, 1994a) were one of the first to use PCR-based ribosomal DNA sequencing to study insect gut microbial community. Later on, Mckillip and his colleagues analyzed the composition of the mi-

crobiome 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 termite also revealed an enormous diversity of bacteria in the different gut compartments (Schmitt-Wagner et al., 2003b). Based on the PCR targeting 16S rRNA, it has also been shown that most of the gut microbial 16S rRNAs from termite *Reticulitermes speratus* were unknown (Ohkuma and Kudo, 1996). Most of the early 16S rRNA gene targeting analysis revealed a significant number of unknown bacterial species at the time.

Besides the 16S rRNA, gene specific PCR has also been widely used to discover genes of interest and survey the metabolic pathways. The 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 the 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 his 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 the gene targeting PCR of DNA sample, 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 his 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 the studies in *Ancylostoma caninum* hookworms (Jones and 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 the 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 the gene discovery and biodiversity analysis for the insect gut symbiotic microbiota. In addition to traditional gene-targeting PCR-based techniques, PCR can be also used for various molecular fingerprinting techniques to study the microbial diversity.

III. 2 Molecular fingerprinting techniques

Besides the library-based gene targeting PCR, several other PCR-based techniques have also been widely used to study the 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 and 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 and Tebbe, 1998), and random amplified polymorphic DNA (RAPD) (Kauppinen et al., 1999). For the 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 to provide 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, denaturing gradient gel electrophoresis (DGGE) is perhaps the most commonly used one. The recent application of the technique to study the 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 microorganisms in the gut and indicated that the wasp larva are not dependent on one particular type of mutualists (Reeson et al., 2003). Behar and his colleagues analyzed the Mediterranean fruit fly gut bacterial communities using both culture-dependent and culture-independent approaches like DGGE and revealed that family *Enterobacteriaceae* was the most dominant species in the fruit fly gut (Behar et al., 2005). Recently, DGGE was employed to explore the microbial diversity in herbivore insects to study the potential mechanisms for biomass degradation. ERIC-PCR (enterobacterial repetitive intergenic consensus) and DGGE were combined to compare the diversity of the lactic acid bacteria communities in the wood- and soil-feeding termites (Bauer et al., 2000). DGGE method was also used to survey and screen for the gut microorganisms in wood-feeding termite

(Hayashi et al., 2007), soil-feeding termite, and its mounds (Fall et al., 2007). In addition to termite, the symbiotic microbiota in the hindguts of scarab beetle larvae were also explored with metagenomic approaches mainly based on the DGGE (Pittman et al., 2008b; Vasanthakumar et al., 2006). Moreover, Dillon and his 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 the grasshopper and locust guts (Dillon et al., 2008). Recently, we revealed the diversity of gut bacteria from different insect species by denaturing gradient gel electrophoresis (DGGE) and found significant microbial diversity differences among the wood-feeding, grass-feeding, and leaf-feeding insects (Shi et al., 2009b). DGGE has also been used to study the sympbiotic microbiota in a variety of insect species like *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).

RFLP (restriction fragment length polymorphism) analysis differentiates homologous DNA sequences based on the distinct DNA fragment patterns resulted from the sequence specificity toward restriction enzymes (Esumi et al., 1982). In 1993, Harada and Ishikawa used RFLP to analyze the 16S rRNA from the group of prokaryote microbes in the gut of Pea aphid. The result suggested that gut microbes have a close relationship with aphid intracellular symbionts (Harada and Ishikawa, 1993). Despite the analysis, the application of traditional RFLP in microbial diversity study is very limited due to the inherent technical limitations of the technology. Domingo used RFLP of 16S rRNA to study the cricket hindgut microbial communities and suggested that community RFLP methods did not have the sufficient resolution or specificity required to study the effect of diets on the 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 the 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 communities research of various termites such as soil-feeding termite (Donovan et al., 2004; Friedrich et al., 2001; Kohler et al., 2008; Schmitt-Wagner et al., 2003a), wood-feeding lower termite (Miyata et al., 2007; Stingl and

Brune, 2003), and fungus-growing termite (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 community and led to some speculations of how the symbiotic microbes could contribute to the 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 the 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; de Vries et al., 2001b). The discovery highlighted that the symbiotic microbiota can be indigenous instead of exogenous from the food material (de Vries et al., 2001a; de Vries et al., 2001b). 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 is 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 and Tebbe, 2006). In a recent study, PCR-SSCP, RT-PCR-SSCP and stable isotope probing (SIP) were combined to study the partial bacterial 16S rRNA genes to survey the diversity of metabolically active bacteria in the larval gut of *Manduca sexta* (Brin*km*ann et al., 2008).

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

III. 3 **Fluorescent** in situ *hybridization* (*FISH*)

Fluorescent in situ hybridization (FISH) is commonly used in microbial ecology studies to visualize the symbiotic bacteria in the gut (Aminov et al., 2006; Cheung et al., 1977). The application of FISH in insect gut microbial study often involves fluorescently labeled probes targeting the 16s rRNA with sequence 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 the

biomass degradation related studies, Berchtold and his 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 to study the 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 complimentary to FISH, DAPI (4′,6-diamidino-2-phenylindole) and GFP (Green Fluorescent Protein) also used to visualize microbial community. DAPI staining of bacterial cells highlighted the significant differences in the number of bacterial cells among different insect species when reared in the same condition (Cazemier et al., 1997a; Cazemier et al., 1997b). Green fluorescent protein (GFP) can be used to track the 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 and Jackson, 2002).

Overall, the various molecular techniques have greatly advanced our understanding of insect gut microbial community, 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 recent available different 'omics' platforms enabled such study.

IV Techniques for "Meta-omics" Analysis of Insect Gut Symbionts

The recent advances of 'omics' technologies enabled us to explore the microorganism communities in an unprecedented way (Allen and Banfield, 2005; Tyson et al., 2004). The high-throughput metagenome, metatranscriptome, and metaproteome analysis of microorganism populations will allow the molecular, organism, and population level investigation of how the chemical and biological processes were enabled, controlled, and evolved (Allen and Banfield, 2005). The complimentary data annotation and high-throughput functional screening will allow the identification of novel catalysts

and strains for bioremediation, biomass processing, bioproduct synthesis, and such (Hongoh et al., 2008a; Lorenz and Eck, 2005; Warnecke et al., 2007). The so-called 'metagenomics' often involves sequencing the 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 like insect gut (Handelsman et al., 1998). Metatranscriptomics refers to sequenceing analysis of the 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 the diverse microbe populations in various eco-environmental systems (Green et al., 2008; Keller and 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 and 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 and Banfield, 2005; Backhed et al., 2005; Falkowski et al., 2008; Green et al., 2008; Keller and Zengler, 2004; Lorenz and Eck, 2005; Tyson et al., 2004). In the past two decades, much effort has been dedicated to exploring the components of the microbial community 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 an important approach for exploring the biomass degrading mechanisms in wood feeding insects. Several studies have been carried out to study the symbionts in the hindgut and midgut of the wood-feeding higher termite (Warnecke et al., 2007) and lower termites (Todaka et al., 2007; Hongoh et al., 2008a,b). Termite is believed to recycle up to 30% of the total carbon on earth, and the highly efficient lignocellulosic biomass deconstruction has made termite a potential source for novel biocatalysts for biomass deconstruction (Hongoh et al., 2008a; Warnecke et al., 2007). Recent studies have indicated that the 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 and Watanabe, 2007; Warnecke and Hugenholtz, 2007; Warnecke et al., 2007; Wheeler et al., 2007; Zhou et al., 2007). The 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 the symbiotic microbiota including hydrogen metabolism, carbon dioxide-reductive acetogenesis, and ni-

trogen fixation (Warnecke et al., 2007). Overall, the development of metagenomics, metatranscripomics, and metaproteomics during the past decades has been focused on the better understanding of microbial diversity and function in the eco-environment, and has been driven by the increasing demands for biocatalysts and biomolecules for applications like biorefinery (Schmeisser et al., 2007). We hereby review the application of these 'omics' platforms to study the 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.

IV. 1 Metagenome Sequencing and Next Generation Sequencing

There are two principal metagenomic strategies for metagenomics, the sequence-based metagenomics approaches and functional metagenomics (Figure 1). Sequence-based metagenomics involves the 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 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). About 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, fifteen fosmids were selected for further sequencing and training of the dataset. The data has led to a comprehensive coverage and quantification of the microbial composition in the 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 aphids species can be conferred by the 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 the metagenome analysis (Marusina, 2006). When next generation sequencing is used, the approach can be the shot-gun sequencing of metagenomic DNA directly. Up until now, four major next generation sequencing platforms are available.

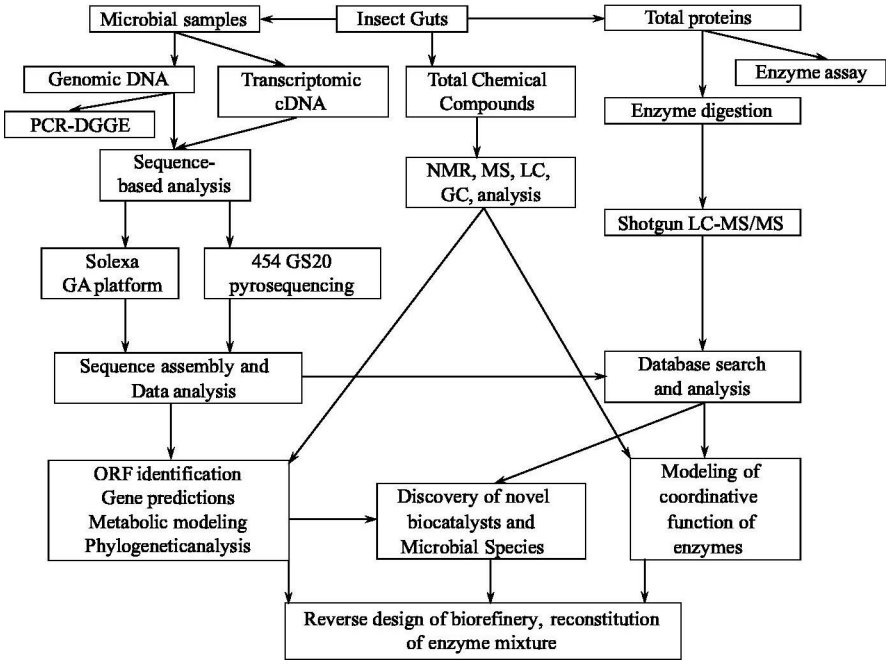

Figure 1 'Omics' analysis of insect gut as a natural biocatalyst system

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 the *de novo* sequencing (Shendure and 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 sequence 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, whilst 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 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 the better assembly of the sequence, and the deeper coverage of the genome (Ansorge, 2009; Shendure and Ji, 2008). Despite the limitations of the next generation sequencing techniques, they have been broadly used for the metagenome sequencing of environmental microbial communities from different niche including soil (Blaha et al., 2007; Tringe et al., 2005; Voget et al., 2003), the human gastrointestinal (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 (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 (Shi et al., unpublished). The analysis has led to the discovery of numerous novel biocatalysts.

IV. 2 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 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 and Streit, 2005). For 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 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 adults *Anopheles gambiae* females (Shen and JacobsLorena, 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 the carboxypetidase activity (Bown et al., 1998). Girard and Jouanin isolated a cDNA encoding chitinase of *Pheadon cochleariae* from a larval gut library (Girard and Jouanin, 1999). For the bioenergy research, novel xylanases with distinct domains have been discovered using metagenomic libraries of microbiota in several insects belong 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 for 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 his colleagues at the University of Wisconsin constructed a metagenomic library consisting of DNA extracted directly from the 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 green fluorescent protein (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 monooxygenase 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 cDNA library allows us to identify novel enzymes and genes for a particular application, however, the analysis is limited by the available probe 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 community and annotate the metatranscriptome to discover the novel genes.

IV. 3 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 high percentage of rRNA signal. Different strategies have been developed to remover 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 the microflora associated with honey bee *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 the host health and showed how the metatranscriptome can be applied to study insect gut systems. The advantage for metatranscriptome sequencing is that it can reflect the dynamics and function of the insect gut symbionts better.

IV. 4 Metaprotoeomics Techniques for Insect Gut Symbiont Studies

Another way to explore the systems dynamics is to study the metaproteomics of insect gut symbionts. Like any genome sequencing project, the 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 the microbial community like 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 and Weir, 1999; Wilmes and Bond, 2004). The proteomics platform can be gener-

ally classified as gel-free or gel-based system (Kan et al., 2005). The traditional approach is to analyze the protein sample with two-dimensional polyacrylamide gel electrophoresis (2D-gel) at first and then further cut the spot for mass spectrometry (MS)-based protein identification. The mass spectrometry techniques that can be used for protein identification includes both MALDI (Matrix-assisted laser desorption ionization) and ESI (Electrospray ionization). MALDI is often coupled with TOF (Time of flight) mass analyzer, whilst 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 the tandem MS (MS/MS) analysis, where individual peptide 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 of fragment pattern 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 the low sensitivity, low coverage of proteome, and the difficulties in quantification. For all these reasons, the gel-based proteomics has been gradually replaced with the gel-free proteomics, which mainly relies on the LC-MS/MS (liquid chromatography-mass spectrometry/mass spectrometry) platform. The most popular approach for gel-free proteomics is the MudPIT (Multidimensional Protein Identification Technology)-based shot-gun proteomics (Delahunty and Yates, 2007; Lohrig and 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 liquid chromatography. 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 and Yates, 2007). MudPIT can also be used as a label-free platform, where peptide quantification can be based on the total ion counts and number of peptides (Delahunty and Yates,

2007). Despite the broad application of the proteomics techniques in various studies, the use of proteomics in the analysis of insect gut symbiotic microbiota is still very limited. In the aforementioned the 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 the three-dimensional LC-MS/MS analysis for protein expression analysis. The fragment ion patterns from metaproteomics were searched against a sequence database derived from the metagenome sequencing for protein identification. The study has revealed that the expression of glycosyl hydrolases are regulated at the protein level, and the enzymes in the metagenome were not expressed at the same time and same level (Warnecke et al., 2007). The further study of the metaproteome in the natural biocatalyst systems like termite gut will allow us to understand how enzymes coordinate to degrade plant cell wall. Metaproteomics analysis will be based on the metagenome sequencing data and will help to further the understanding of insect gut symbiotic microbes to the proteome level.

V 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 the 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 the insect gut symbionts have experienced dramatic changes during the past two decades. The initial study 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 and 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 the molecular biotechniques independent of microbial culturing. Methods like DGGE, SSCP, RFLP, and FISH allowed us to explore the complexity of natural microbial communities better. These techniques provided some speculations of the microbial community composition, dynam-

ics, and function. However, the traditional molecular techniques still cannot provide a comprehensive view of the composition and dynamics of insect symbiotic microbial community. 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 in 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 the information will lead to a systems level understanding of insect gut as the system for biomass deconstruction, nutrient biosynthesis and such. Despite the significant progresses, several aspects of research need to be emphasized to better exploit the 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 capacity of degrading lignocellulosic biomass (Sun and Zhou 2009). The cellullolytic enzyme activity in grasshopper gut is actually comparable to that of the termite gut (Shi et al., 2009a). The comparative analyses of different insect gut systems will allow us to identify the common and unique features for degrading different lignocellulosic biomass in various insect gut systems. The study 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, the sequence assembly for metagenome is much more challenging than single species, in particular for the complex system. The more microbe species in a community, the more complexity and overall genome size will be for insect gut symbiotic microbiota. Illumina genome analyzer has the most potential of increasing sequence coverage due to the higher sequencing throughput and lower per base cost. However, the short sequence read length together with the large overall genome size from this technology makes it extremely challenging to assemble the metagenome sequences. The recent development of several assemblers for short sequences like SSAKE, VEVELT, ABySS, and Euler have provided the solutions for the assembly of short sequence reads genome sequencing (Scheibye-Alsing et al., 2009). However, the conditions used for the single genome assembly were 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 softwares 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 a particular important role in the lignin degradation (Tartar et al., 2009). In previous studies, the metabolism of monoaromatic model compounds by termites and their gut microflora was 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 the 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; Shi et al., unpublished data). Metaproteomics will provide a powerful solution toward the observation of how biocatalysts from the host and the 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 biorefinery 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 the 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 the genome, proteome, metabolome and transcriptome analysis (Rachlin et al., 2006; Vieites et al., 2009). The accumulation of different 'omics' data regarding the insect gut systems will allow us to investigate how the different components and biocatalysts work together to fulfill various functions including biomass degradation.

Overall, we are at a golden age to address the 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 the unprecedented understanding of insect gut symbiotic microbiota and its interaction with the hosts. The systems level integration of the tremendous information will enable the in-depth understanding of natural biocatalyst systems like insect guts toward providing novel solutions for the next generation biorefinery.

Reference

Adams I P, Glover, R H, Monger W A, et al. 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, et al. 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 et al. 2005. Community genomics in microbial ecology and evolution. Nature Reviews Microbiology, 3: 489-498.

Allen H K, Moe L A, Rodbumrer J, et al. 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 et al. 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, et al. 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: 6371-6.

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

Backhed F, Ley R E, Sonnenburg J L, et al. 2005 Host-Bacterial Mutualism in the Human Intestine. Science, 307: 1915-1920.

Bauer S, Tholen A, Overmann J, et al. 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, Schonhuber W, et al. 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, et al. 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, Weir M P. 1999. Proteomics: quantitative and physical mapping of cellular proteins. Trends In Biotechnology, 17: 121-7.

Blaha C A G, Santos I L V D, Pereira M S, et al. 2007. Metagenomic evaluation of indigenous microbial diversity in tropical mangrove and semi-arid soils from petroliferous basin. Journal of Biotechnology. 131: S239-S239.

Bown D P, Wilkinson H S, 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. 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, et al. 2003. Metagenomic analyses of an uncultured viral community from human feces. Journal of Bacteriology, 185: 6220-6223.

BrennanY, Callen W N, Christoffersen L, et al. 2004. Unusual microbial xylanases from insect guts. Applied and Environmental Microbiology, 70: 3609-3617.

Breznak J A. 2004. Invertebrates-Insects. (eds. A. T. Bull), Microbial diversity and bioprospecting. ASM Press, WashingtonD.C:

Brinkmann N, Martens R, and Tebbe 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, et al. 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 ,et al .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 , Miambi E , Breznak JA .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 .1998 .Termite guts : the world's smallest bioreactors . Trends in Biotechnology ,16 : 16-21 .

Brune ,A . 2003 .Symbionts aiding digestion . (eds . V . H . Rush & R . T . Cardé) Encyclopedia of insects . Academic Press , Inc . , New York :N .Y .

Casu R E , Eisemann C H , Vuocolo T ,et al .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 ,et al .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 et al .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 . (eds . Canberra) , Australian Biological Resources Study .

Chaves S , Neto M , 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 ,et al .1977 .Gene mapping by fluorescent in situ hybridization . Cell Biology International Report ,1 :255-62 .

Chougule N P , Giri A P , Sainani M N ,et al .2005 .Gene expression patterns of *Helicoverpa armigera* gut proteases . Insect Biochemistry and Molecular Biology , 35 :355-367 .

Colebatch G , Cooper P ,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 ,et al . 2005 . Metagenomic gene discovery : past , present and future . Trends in Biotechnology ,23 :321-329 .

Cox-Foster D L , Conlan S , Holmes E C ,et al . 2007 . A metagenomic survey of microbes in honey bee colony collapse disorder . Science ,318 :283-287 .

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

Culley A I , Lang A S ,Suttle C A .2006 .Metagenomic analysis of coastal RNA virus communities . Science ,312 :1795-1798 .

da Mota F F , Gomes E A , Paiva E ,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 ,Chang K P Endosymbionts of insects . *In* : a . J G H N R Krieg , (Ed .) , Bergey's Manual of Systematic Bacteriology , 811-833 .

de Vries E J , Breeuwer J A J , Jacobs G ,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 ,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 ,Yates J R .2007 . MudPIT : multidimensional protein identification technology . Biotechniques ,43 : 563 , 565 , 567 .

Demaneche S, David M M, Navarro E, et al. 2009. Evaluation of functional gene enrichment in a soil metagenomic clone library. Journal of Microbiological Methods, 76: 105-7.

Dethlefsen L, Eckburg P B, Bik E M, et al. 2006. Assembly of the human intestinal microbiota. Trends in Ecology and Evolution, 21: 517-523.

Dillon R J, Webster G, Weightman A J, et al. 2008. Composition of Acridid gut bacterial communities as revealed by 16S rRNA gene analysis. Journal of Invertebrate Pathology, 97: 265-272.

Dillon R J, Dillon V M. 2004. The gut bacteria of insects: Nonpathogenic interactions. Annual Review of Entomology, 49: 71-92.

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

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.

Donovan S E, Purdy K J, Kane M D, et al. 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.

Douglas A E. 1994. Symbiotic interactions. Oxford University Press, Oxford, UK.

Egert M, Stingl U, Bruun L D, et al. 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, et al. 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, et al. 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: 4247-57.

Falkowski P G, Fenchel T, Delong E F. 2008. The microbial engines that drive Earth's biogeochemical cycles. Science, 320: 1034-1039.

Fall S, Hamelin J, Ndiaye F, et al. 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.

Finegold S M, Molitoris D, Song Y, et al. 2002. Gastrointestinal Microflora Studies in Late-Onset Autism. Clinical Infectious Diseases, 35: S6-S16.

Fisher M, Miller D, Brewster C, et al. 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.

Flint H J, Bayer E A, Rincon M T, et al. 2008. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. Nature Reviews Microbiology, 6: 121-131.

Friedrich M W, Schmitt-Wagner D, Lueders T, et al. 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, Sekiguchi Y. 1998. Specific detection of intracellular symbiotic bacteria of aphids by oligonucleotide-probed in situ hybridization. Applied Entomology and Zoology, 33: 461-472.

Gil R, Latorre A, Moya A. 2004. Bacterial endosymbionts of insects: insights from comparative genomics. Environmental Microbiology, 6: 1109-1122.

Gill S R, Pop M, DeBoy R T, et al. 2006. Metagenomic analysis of the human distal gut microbiome. Science, 312: 1355-1359.

Giovannoni S J, Britschgi T B, Moyer C L, et al. 1990. Genetic Diversity in Sargasso Sea Bacterioplankton. Nature, 345: 60-63.

Girard C, 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, Whitaker R J. 2008. Microbial Biogeography: from taxonomy to Traits. Science, 320: 1039-1043.

Guan C H, Ju J J, Borlee B R, et al. 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, et al. 2009. Metagenomics as a new technological tool to gain scientific knowledge. World Journal of Microbiology and Biotechnology, 25: 945-954.

Handelsman J, Liles M, Mann D, et al. 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, et al. 1998. Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. Chemistry & Biology, 5: 245-249.

Handelsman J. 2004. Metagenomics: application of genomics to uncultured microorganisms. Microbiology and Molecular Biology Reviews, 68: 669-85.

Harada H, Ishikawa H. 1993. Gut Microbe of Aphid Closely-Related to Its Intracellular Symbiont. Biosystems, 31: 185-191.

Harmon J P, Moran N A, 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, et al. 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, 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, et al. 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, et al. 1988. DNA Probe Method for the Detection of Specific Microorganisms in the Soil Bacterial Community. Applied and Environmental Microbiology, 54: 703-711.

Holben W E, Harris D. 1995. DNA-Based Monitoring of Total Bacterial Community Structure in Environmental-Samples. Molecular Ecology, 4: 627-631.

Hongoh Y, Ekpornprasit L, Inoue T, et al. 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, et al. 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, et al. 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, Odonnell A G. 1991A Dispersion and Differential Centrifugation Technique for Representatively Sampling Microorganisms from Soil. Soil Biology & Biochemistry, 23: 217-225.

Hosokawa T, Kikuchi Y, Nikoh N, et al. 2006. Strict host-symbiont cospeciation and reductive genome evolution in

insect gut bacteria . Plos Biology ,4 :1841-1851 .

Hurst M R H ,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 ,et al . 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 ,et al .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 ,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 ,et al . 2005 . Metaproteomic analysis of Chesapeake Bay microbial communities . Saline Systems ,1 : 7 .

Kane M D ,Pierce N E . 1994aDiversity within diversity : molecular approaches to studying microbial interactions with insects . EXS ,69 :509-24 .

Kane M D ,Pierce N E .1994b .Diversity within diversity : molecular approaches to studying microbial interactions with insects . (eds .Schierwater B S B ,Wagner G ,Desalle R), Molecular Methods in Ecology and Evolution . Berlin , Birkhauser Verlag .

Kaufman M G , Walker E D , Odelson D A ,et al .2000 .Microbial Community Ecology & Insect Nutrition . American Entomologist ,46 :173-185 .

Kauppinen J , Mantyjarvi R , 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 ,Zengler K .2004 .Tapping into microbial diversity . Nature Reviews Microbiology ,2 : 141-150 .

Khachane A N , Timmis K N , 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 ,et al .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 , 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 .

Ley R E , Peterson D A , and Gordon J I . 2006 .Ecological and Evolutionary Forces Shaping Microbial Diversity in the Human Intestine . Cell ,124 :837-848 .

Li L , Frohlich J , Pfeiffer P ,et al .2003 .Termite gut symbiotic archaezoa are becoming living metabolic fossils . Eukaryotic Cell ,2 :1091-1098 .

Lilburn T G , Schmidt T M ,Breznak J A .1999 .Phylogenetic diversity of termite gut spirochaetes . Environmental Microbiology ,1 :331-345 .

Lindh J M , Borg-Karlson A K ,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 ,et al .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 ,Wolters D . 2009 . Multidimensional protein identification technology . Methods in Molecular Biology ,564 : 143-53 .

Lorenz P ,Eck J .2005 .Metagenomics and industrial applications . Nature Reviews Microbiology ,3 :510-516 .

Lynd L R, Laser M S, Brandsby D, et al. 2008. How biotech can transform biofuels. Nature Biotechnology, 26: 169-172.

MacDonald T T, Monteleone G. 2005. Immunity, Inflammation, and Allergy in the Gut. Science, 307: 1920-1925.

Mackenzie L M, Muigai A T, Osir E O, et al. 2007. Bacterial diversity in the intestinal tract of the fungus-cultivating termite *Macrotermes michaelseni* (Sjostedt). African Journal of Biotechnology, 6: 658-667.

Manefield M, Whiteley A S, Griffiths R I, et al. 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, 243: 133-41

Margulies M, Egholm M, Altman W E, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 7057: 376-80.

Maron P A, Ranjard L, Mougel C, et al. 2007. Metaproteomics: A new approach for studying functional microbial ecology. Microbial Ecology, 53: 486-493.

Marusina K. 2006. The next generation of DNA sequencing. Genetic Engineering News, 26: 28-31.

McIntosh D, Ji B, Forward B S, et al. 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, et al. 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, et al. 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, et al. 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, 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, Telang A. 1998. Bacteriocyte-associated symbionts of insects - A variety of insect groups harbor ancient prokaryotic endosymbionts. Bioscience, 48: 295-304.

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. 2006. Symbiosis. Current Biology 16, R866-R871.

Moran N A. 2007. Symbiosis as an adaptive process and source of phenotypic complexity. Proceedings of the National Academy of Sciences, 104: 8627-8633.

Moran NA, McCutcheon JP, Nakabachi A. 2008. Genomics and evolution of heritable bacterial symbionts. Annual Review of Genetics, 42: 165-190.

More M I, Herrick J B, Silva M C, et al. 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 International Journal of General and Molecular Microbiology, 73: 127-141.

Muyzer G, Dewaal E C, Uitterlinden A G. 1993. Profiling of complex microbial-populations by denaturing gradient gel-electrophoresis analysis of polymerase chain reaction-Amplified Genes-Coding for 16s Ribosomal-Rna. Applied and Environmental Microbiology, 59: 695-700.

Nakashima K, Watanabe H, Saitoh H, et al. 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. (eds. Zengler K). Accessing uncultivated microorganisms., Washington ASM Press,, DC:153-170

Nicholson J K, Connelly J, Lindon J C, et al. 2002. Metabonomics: a platform for studying drug toxicity and gene function. Nature Reviews Drug Discovery, 1:153-61.

Noda S, Ohkuma M, Usami R, et al. 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, Moran N A. 2001. Genes lost and genes found: Evolution of bacterial pathogenesis and symbiosis. Science, 292:1096-1098.

Ogram A, Sayler G S, 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, Kudo T. 1996. Phylogenetic diversity of the intestinal bacterial community in the termite *Reticulitermes speratus*. Applied and Environmental Microbiology, 62:461-468.

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

Oliver J D. 2000. Problems in detecting dormant (VBNC) cells, and the role of DNA elements in this response. Tracking Genetically-Engineered Microorganisms1-15.

Orozco Medina C, Lopez Cortes A, 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, 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, et al. 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, et al. 1991. Rapid method for processing soil samples for polymerase chain reaction amplification of specific gene sequences. Applied and Environmental Microbiology, 57: 2283-6.

Pittman G W, Brumbley S M, Allsopp P G, et al. 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, et al. 2008b. "Endomicrobia" and other bacteria associated with the hindgut of *Dermolepida albohirtum* larvae. Applied and Environmental Microbiology, 74:762-767.

Pootanakit K, Angsuthanasombat C, 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, et al. 2006. Biological context networks: a mosaic view of the interactome. Moecular Systems Biology, 2: 66.

Raymond B, Lijek R S, Griffiths R I, et al. 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, et al. 2003. Application of 16S rDNA-DGGE to examine the microbial ecology associated with a social wasp *Vespula germanica*. Insect Molecular Biology, 12:85-91.

Riesenfeld C S, Schloss P D, Handelsman J. 2004. Metagenomics: Genomic analysis of microbial communities. Annual

Review of Genetics ,38 :525-552 .

Robe P , Nalin R , Capellano C ,et al . 2003 . Extraction of DNA from soil . European Journal of Soil Biology ,39 :183-190 .

Roose Amsaleg G S , C L , E . Harry , M . 2001 .Extraction and purification of microbial DNA from soil and sediment samples . Applied Soil Ecology ,18 :14 .

Roussel E G , Bonavita M A , Querellou J ,et al .2008 .Extending the Sub-Sea-Floor Biosphere . Science ,320 :1046 .

Santo Domingo J W , Kaufman M G , Klug M J ,et al .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 , 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 ,et al .2009 .Sequence assembly . Computational Biology and Chemistry , 33 :121-36 .

Schloss P D , Delalibera I , Handelsman J ,et al .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 ,Streit W R .2007 . Metagenomics , biotechnology with non-culturable microbes . Applied Microbiology and Biotechnology ,75 :955-62 .

Schmitt Wagner D , Friedrich M W , Wagner B ,et al .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 ,et al .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 , Kaise K ,et al .2005 .A proteomic fingerprint of dissolved organic carbon and of soil particles . Oecologia ,142 :335-343 .

Schwieger F ,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 C ,JacobsLorena M .1997 .Characterization of a novel gut-specific chitinase gene from the human malaria vector *Anopheles gambiae* . Journal of Biological Chemistry ,272 :28 895-28 900 .

Shendure J ,Ji H .2008 .Next-generation DNA sequencing . Nature Biotechnology ,26 :1135-1145 .

Shi W B , Ding SY ,Yuan S Y .2009a .Comparison of insect gut cellulase and xylanase activity across different insect species with distinct food sources . Bioenergy Research . (In review)

Shi W B , Uzuner U , Jesudhasan P R , et al .2009b .Diversity of gut bacteria from different insect species revealed by denaturing gradient gel electrophoresis (DGGE) . Applied Environmental Microbiology . (In review)

Shinzato N , Muramatsu M , Matsui T ,et al . 2005 . Molecular phylogenetic diversity of the bacterial community in the gut of the termite *Coptotermes formosanus* . Bioscience Biotechnology and Biochemistry ,69 :1145-1155 .

Shinzato N , Muramatsu M , Matsui T ,et al .2007 .Phylogenetic analysis of the gut bacterial microflora of the fungus-growing termite *Odontotermes formosanus* . Bioscience Biotechnology and Biochemistry ,71 :906-915 .

Singleton I , Merrington G , Colvan S ,et al .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 , Oros Sichler M , Milling A ,et al .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 .

Smalla K .2004 .Culture-Independent microbiology . (eds . A . T . Bull) , Microbial diversity and bioprospecting . Washington D .C :ASM Press

Spiteller D , Dettner K ,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: 267-270.

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

Stingl U, 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, et al. 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 Zhou X G. 2009. Utilization of lignocellulose-feeding insects for viable biofuels: an emerging and promising area of entomologcal science. (eds. L Kang, T X Liu), Recent Advances in Entomological Research., Beijing Higher Education Press: 251-291

Tamas I, Klasson L, Canback B, et al. 2002. 50 million years of genomic stasis in endosymbiotic bacteria. Science, 296: 2376-2379.

Tamas I, Wernegreen J J, Nystedt B, et al. 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, et al. 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, 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.

Todaka N, Moriya S, Saita K, et al. 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.

Tokuda G, 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, et al. 2005. Comparative metagenomics of microbial communities. Science, 308: 554-557.

Turnbaugh P J, Hamady M, et al. 2008. A core gut microbiome in obese and lean twins. Nature, 457: 480-484.

Turroni F, Ribbera A, Foroni E, et al. 2008. Human gut microbiota and bifidobacteria: from composition to functionality. Antonie Van Leeuwenhoek, 94: 35-50.

Tyson G W, Chapman J, Hugenholtz P, et al. 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature, 428: 37-43.

Van Elsas J D, Smalla K, Tebbe C C. 2000. Extraction and analysis of microbial community nucleic acids from environmental matrices. (eds. J. D. v. J.K. Jansson, and M.J. Bailey), Tracking Genetically-Engineered Microorganisms. pp. 29-51. Eurekah, Austin, Texas.

van Ham R, Kamerbeek J, Palacios C, et al. 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, et al. 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, et al. 2004. Environmental genome shotgun sequencing of the Sargasso Sea.

Science ,304 :66-74 .

Vieites J M , Guazzaroni M E , Beloqui A ,et al .2009 .Metagenomics approaches in systems microbiology . FEMS Microbiol Reviews ,33 :236-55 .

Voget S , Leggewie C , Uesbeck A , Raasch C ,et al .2003 .Prospecting for novel biocatalysts in a soil metagenome . Applied and Environmental Microbiology ,69 :6235-6242 .

Warnecke F , Luginbuhl P , Ivanova N ,et al . 2007 . Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite . Nature ,450 :560-565 .

Warnecke F ,Hugenholtz P . 2007 .Building on basic metagenomics with complementary technologies . Genome Biology ,8 :231 .

Webster G , Newberry C J , Fry J C ,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 ,et al .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 ,et al . 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 ,et al . 2005 . Intracellular screen to identify metagenomic clones that induce or inhibit a quorum-sensing biosensor . Applied and Environmental Microbiology ,71 :6335-44 .

Wilmes P ,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 :911-20 .

Yandell D W . 1991 . SSCP (single strand conformation polymorphism) analysis of PCR fragments . Japanese Journal of Cancer Research ,82 :1468-9 .

Yoon H S , Lee C S , Lee S Y ,et al .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 ,et la .2008 . Plant systems biology comes of age . Trends in Plant Science ,13 :165-71 .

Zhang H ,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 ,et al .2007 . Correlation of cellulase gene expression and cellulolytic activity throughout the gut of the termite *Reticulitermes flavipes* . Gene ,395(1-2): 29-39 .

作者简介

袁戎华博士,2008 年 7 月任美国德克萨斯 A&M 大学植物病理与微生物学院担任助理教授。主要从事生物信息学与系统生物学方向的科学研究。于 2007 年获美国田纳西大学植物、昆虫与土壤专业博士学位。2001 年在亚利桑那大学获得植物科学专业硕士学位,1997 年在复旦大学获得生物学学士学位,同时还辅修了国际经济专业。目前,主要从事与生物能源技术相关领域的研究,包括纤维素酶的开发与修饰,植物、昆虫与微生物的相互作用等多个研究

领域。特别是开发了新型的蛋白质相互作用分析平台,用以研究植物宿主在病原体识别过程中的相互作用,以及蛋白相互作用。另外,在应用技术研究方面,开发了为新一代测序技术、系统生物学建模,以及质谱分析应用的新数据分析软件。现兼任田纳西大学诺克斯维尔分校植物科学部担任基因组学科学家以及 UTIA 基因组学中心的主任。

Dr. Joshua S. Yuan joined the Department of Plant Pathology and Microbiology at Texas A&M University as an Assistant Professor, located at College Station effective July 2008. He is directing the Laboratory of Bioenergy, Systems Biology and Bioinformatics in Borlaug Center, also known as Institute for Plant Genomics and Biotechnology. Dr Yuan received his Ph.D. in Plants, Insects, and Soils with Focus on Functional Genomics (Major) and Statistics (Minor) from the University of Tennessee (2007). He also holds an M.S. in Plant Sciences from the University of Arizona (2001) and a B.S. in Biology (Major) and International Economics (Minor) from Fudan University (1997). Most recently, Dr. Yuan was a Genomics Scientist in the Department of Plant Sciences, and Director of UTIA Genomics Hub, University of Tennessee, Knoxville. Before that, Dr. Yuan was the microarray core manager at Gallo Center in University of California San Francisco. The Yuan lab aims to develop active programs from applied, basic, and technical perspectives. For the applied research, we are focusing on bioenergy to provide novel solutions for reverse design of natural biomass utilization systems to improve biorefinery procedures. His lab also developed novel strategies for structure dynamics guided biocatalyst improvement. For the fundamental research, we are employing systems biology approaches to study the plant, insect, and microbe interaction. From the technology development perspective, we are developing novel packages for next generation sequencing data analysis, systems biology modeling and mass spectra analysis.

孙建中,1958 年生于江苏南通,2002 年获美国路易斯安那州立大学博士学位。在美国学习工作 15 年,其中,2004～2009 年曾任美国密西西比州立大学昆虫系/海岸研究与推广中心助理教授/副教授、博士生导师;2009 年 5 月任美国华盛顿州立大学兼职教授;自 2009 年 7 月任江苏大学特聘教授、博导、江苏大学生物质能源研究所所长。为美国科学学会、工业微生物学会、昆虫学会会员。曾主持美国农业部、能源部、交通部等有关白蚁生物学,以及白蚁生物质能源利用研究等多个项目;受邀担任美国农业部美国 2008～2013 年第三个五年科技发展计划起草与评议专家、美国农业部重大课题评审专家、美国国家植物生物技术评审委员会生物质能源和环境研究项目评审专家(CPBR)、中国国家自然科学基金评审委员会重点项目评审专家;受邀担任 10 多种国际昆虫类、农业类、生物质能源类 SCI 专业杂志审稿人、编委、特邀主编;作为发起人和主持人,2008 年成功主持美国昆虫学会年会首届"昆虫与生物质能源应用"国际性专题学术研讨会。2011 年成功申请召开国家香山科学会议"高效降解生物质的自然生物系统资源利用与仿生"。2009 被列入江苏省"创新创业人才

千人计划”和“六大人才高峰计划”,目前的主要研究领域为白蚁木质纤维素的高效生物转化机制以及催化资源的工程化开发利用,最终目标是构建白蚁的高效仿生生物反应器系统实现生物质的经济与高效生物转化。

Jianzhong Sun

Ph.D. in Entomology from Louisiana State University, USA.

Distinguished Professor, Jiangsu University, P. R. China.

Adjunct professor, Dept. of biological systems engineering, Washington State University, USA.

Director of Biofuels Institute, Jiangsu University, China

Guest editor of Insect Science (published by Wiley-Blackwell) and also serve as a referee for 7 international journals associated with the entomology, agriculture, plant, and bioenergy, etc.

Panel members for grant review, including USA and China, e.g. 1)."Bioenergy and Environment" of the Consortium for Plant Biotechnology Research, Inc.("CPBR"), USA. 2). USDA National program 104 for veterinary, medical and urban entomology, USA. 3). National scientific foundation of China, etc.

Research interests

Dr. Sun' research interests are mainly focused on the termite cellulolytic systems and other insect systems as well with the function to efficiently degrade lignocellulosic biomass. With the studies on these fantastic systems in utilization of a variety of recalcitrant biomasses, he currently leads an enthusiastic research team with a variety of disciplines to mimic the termite cellulolytic systems to realize an artificial bioreactor that would breakdown biomass efficiently and economically .Further contact information: jzsun1002@ hotmail.com or jzsun1002@ ujs.edu.cn , address: Biofuels Institute, College of the environment, Jiangsu University, 301 Xuefu Rd., Zhenjiang 212013, Jiangsu, P. R. China, Tel: 86-0511-88796122.