Metagenomic Approaches for Insect Symbionts

6

Mani Chellappan and M. T. Ranjith

Contents

M. Chellappan (✉) · M. T. Ranjith
AINPVPM, Department of Agricultural Entomology, College of Agriculture, Vellanikkara, Kerala Agricultural University, Thrissur, Kerala, India
e-mail: mani.chellappan@kau.in

Omkar (ed.), *Microbial Approaches for Insect Pest Management*,
https://doi.org/10.1007/978-981-16-3595-3_6

Abstract

Insects, the most successful groups in animal kingdom, harbor diverse groups of microbes, such as bacteria, archaea, fungi, protozoa, and viruses, which profoundly influence their survival and adaptations over a wide range of ecological niches. These microbes are associated with their host insects permanently or transiently and such associations may be beneficial or harmful to the host insect under various instances. Attempts were made earlier to characterize insect microbiome by isolation and cultivation techniques and polymerase chain reaction (PCR)-based cloning methods that resulted in identification of a few groups of microbes. The metagenomic approaches under the next-generation sequencing platforms provide unparallel opportunities to understand the composition of the microbiome and their functional role in the biology of the insects, thus expanding our understanding from a single microbial species to the whole community. These approaches provide an ample opportunity to understand the components of the microbiome that can potentially and collectively affect the behavior and physiological traits of insects through genetic and metabolic interactions. For instance, endosymbionts (*i.e.*, microbes that live inside host cells or tissues) depend on the insect hosts for obtaining nutrients, provide fitness advantages to their insect hosts in terms of the breakdown of plant cell wall components, viz, cellulose, lignocelluloses, and xylan, supplying essential amino acids and vitamins to host insects, thereby upgrading the nutrient status of their diet, detoxification of lethal insecticide molecules, plant defensive compounds such as phenolics, and production of anti-microbial peptides against insect pathogens. However, in some instances, the microbes may also be pathogenic to the insect hosts by producing insecticidal toxins, which reduce viability and cause morbidity.

The culture-independent metagenomic approach allows us to characterize a variety of genes that microbes possess or are expressing, which signifies 'what they are doing' within the host. It also enables us to compare the performance of

insect with changes in their microbiome composition. These approaches have a wide range of applications apart from the study of insect's microbial ecology. The microbiota associated with wood-feeding beetles can be exploited as source of novel enzymes in industrial bioprocesses. The information on microbial genes and enzymes involved in cellulose hydrolysis, vitamin production, and nitrogen fixation can be useful in improving the reliability and efficiency of industrial processes. Furthermore, insect–microbe relationships could be manipulated to improve pest control, by decreasing pest's fitness or by increasing the efficacy of pest management programs.

Keywords

Metagenome · Next Generation Sequencing · Microbial community · 16S rRNA gene

Learning Objectives

1. Insects harbour diverse group of microbes belonging to various taxa, which profoundly influence their survivability and adaptations over a wide range of ecological niches.
2. Over the past two decades, the insect microbiome analyses were carried out primarily by the classical approach involving the isolation cum cultivation techniques and polymerase chain reaction (PCR)-based cloning methods that resulted in identification of only a few groups of microbes.
3. However, the recent advances in metagenomic strategies and sequencing techniques revolutionized the study of insect microbiome and provided unparalleled opportunities to understand the composition and functional diversity of insect microbiome.
4. The insect metagenome analysis under next-generation sequencing (NGS) platforms offers valuable information that could be helpful in formulating novel approaches in pest management by manipulating the insect–microbe relationships.

6.1 Introduction

Approximately 4–6 $\times$ 10^{30} microorganisms are present on earth (Sleator et al. 2008). Out of the total, nearly 99% are not amenable for culture plate, but play an important role in a variety of environment, namely soil, water, atmosphere, plant and animal systems. Metagenomics (also called as ecogenomics or environmental genomics or community genomics) is the scientific study of DNA sequences collected directly from an environment to know the diversity and ecology of microorganisms of that

specific environment. According to Chen and Pachter (2005), it is the application of modern genomic techniques for the study of microbial communities from an environment directly without actually culturing them (on earth remain uncultured). The study may help in our understanding on the microbial diversity in a specified environment, interaction between the communities and higher animals, and the biology, as a whole. Studies on the uncultured microorganisms will not only give in-depth details about their ecology, it also helps in the identification of novel enzymes, signal mimics, smart molecules and new generation antibiotics (Rajagopal 2009; Krishnan et al. 2014).

Insects, which represent more than half of all the biodiversity in the world, are one of the most diverse and successful organisms in the history of life on earth. The remarkable success is due to the abilities of insects to colonize highly diverse niches and the metamorphosis in its biology. A poikilothermic form, insects cannot regulate their body temperature, but with varied adaptations insects can survive both hot as well as sub-zero conditions (Finn et al. 2015).

Insect's digestive system harbours numerous microorganisms, which dictate the growth, development, adaptation and general fitness of the host. Alimentary canal of insects contains approximately 10 times more microorganisms than the total body cells of the insect (Rajagopal 2009). Microbes get into the digestive system of the insect through the food and reside inside as commensals or parasites or symbionts. The gut microbiota influences all aspects of insect physiology, ecology and evolution (the beneficial microbes help the insect in digestion and metabolism like cellulose and xylan hydrolysis, vitamin production, nitrogen fixation, insecticide resistance, antibiotic resistance, signal molecules like quorum sensors, etc.). The gut microbiota of insects also involves in food digestion, pesticide detoxification, growth and development of the organism, pathogen resistance, intra-specific communication and general physiology (Engel and Moran 2013; Douglas 2015; Jing et al. 2020). The contributions of these gut microorganisms in relation to insect functions are highly relevant in the field of public health and veterinary medicine, agriculture and ecology.

6.2 History and Milestones in the Metagenomic Research

Microorganisms occur in almost all habitats in nature, even in extreme environments, namely polar regions, desert, hot geysers, deep sea and inhospitable rocks. They play crucial roles in biology, palaeontology, soil health etc. The study of microorganisms is based on morphological features, growth and selection of some biochemical profiles in vogue for the past 300 years since the invention of microscope by Antonie Philips van Leeuwenhoek in 1676 (Schierbeek 1959; Roszak et al. 1984). Over the years, microbiologists realized that bulk of microorganisms (99%) cannot be cultured by routine culture media. The proposal to use ribosomal RNA genes as molecular markers for biological classification (Woese and Fox 1977) and automated sequencing method invented by Sangers et al. (1977a, b), in fact, revolutionized the study and classification of microorganisms in the late 1970s.

Stahl et al. (1984) demonstrated the direct analysis of 5S and 16S rRNA genes to describe the microbial diversity without culturing the microorganisms per se. This led to the subsequent isolation and cloning of DNA from environmental samples. Begon et al. (1986) proposed the microbial community concept as the set of microorganisms coexisting in the same space and time. During this period, the microbiologists conclusively learnt that the number of observed microorganisms in a microscope did not correspond with number of microorganisms obtained in culture plates (Staley and Konopka 1985). Several advances have been made in the ensuing decade, like polymerase chain reaction (PCR), rRNA genes cloning and sequencing, fluorescent in situ hybridization (FISH), denaturing gradient gel electrophoresis (DGGE and TGGE), restriction-fragment length polymorphism and terminal restriction-fragment length polymorphism (T-RFLP). Handelsman and co-workers in the late 1990s defined the study and characterization of uncultivable microorganism as metagenomics; it is the theoretical collection of all genomes from members in a microbial community from a specific environment (Handelsman et al. 1998). Metagenomics laid the foundation of the '-omics' techniques and revolutionized research in microbial ecology (Handelsman 2004).

Venter et al. (2004a, b) carried out the first ever large-scale metagenomic project by sequencing samples from Sargasso Sea close to Bermuda, which yielded a whopping 1.6 billion base pairs of DNA and led to the discovery of 1.2 million new genes. Following this, many studies on the microbial community structure in varied environment had been carried out to uncover several unknown facts as microbes continue to play a crucial part in those habitats.

6.3 Insect Microbiome and its Functional Role

Insect gut microbiome, the collective genome of the native microbiota of the gut has multiple relationships with their insect hosts, which range from the obligate mutualism to pathogenic (Dillon and Dillon 2004). The gut microbes play vital roles in the digestion of food ingested, produce essential vitamins, make the host to survive on suboptimal food sources, resisting hostile pathogen infection, aid in the detoxification of xenobiotics compounds, growth and development of overall physiology of the host insect (Jing et al. 2020). Acquisition, colonization and transmission of microbiome determine the success of the insect in an environment (Gupta and Nair 2020). In a mutual relationship, insects present distinctive habit for a variety of microorganisms to colonize and the gut microbes in turn provide numerous benefits to their insect hosts (Douglas 2015). In rare conditions, established symbionts can become opportunistic pathogen also, if conditions become so. A number of factors, namely digestive enzymes, pH, redox potential of the gut, type of food the host has ingested and the secondary plant compounds present in the food, dictate the microbial density and diversity in the insect gut (Dillon and Dillon 2004).

6.3.1 Nutritional Symbioses of Gut Microbes

Symbiosis is essential for the survival of insects in extreme environmental conditions (Gupta and Nair 2020). The insect host can be primarily (obligate) or secondarily (facultative) dependant on the microbial symbionts to get nutrition and protection from their natural enemies, respectively (Moran et al. 2008). Major role of gut microbes is to provide the essential nutrients to the host (Jing et al. 2020). For instance, spirochetes through acetogenesis and nitrogen fixation provide required carbon, nitrogen and energy requirements of termite (Breznak 2004). The symbionts also enable the host to overcome barriers, like plant allelochemicals (Dowd and Shen 1990), nutritionally poor diets and recalcitrant food resources (Slaytor 1992).

6.3.2 Protection against Natural Enemies of the Insect Host

Another important beneficial function of gut microbiota in insect is to provide a buffering action to help prevent the proliferation of pathogens (Kodama and Nakasuji 1971; Charnley et al. 1985; Dillon et al. 2002). Resident gut microbiota protect their insect hosts against invaders by multiple mechanisms including restricting nutrients or space, production of toxins and activation of insect immune system functions that are more deleterious to the invader than the resident (Douglas 2015).

6.3.3 Gut Microbes in Detoxification of Xenobiotics

Insect resistance or tolerance to xenobiotics is mostly mediated by the insect genome rather than the gut microbes (Dillon and Dillon 2004; Douglas 2015). However, a compelling evidence of gut microbiota, *Burkholderia*-mediated fenitrothion resistance emerged in *Riptortus pedestris* (Kikuchi et al. 2012).

6.3.4 Gut Microbes in Insect Communication and Mating

Inter- and intra-communication of insect may be mediated by the microorganisms associated with insects (Ezenwa et al. 2012; Gupta and Nair 2020). Gut microbe activities result in the production of some compounds, which may act as kairomones or pheromones. Aggregation pheromone in grasshopper, *Schistocerca gregaria,* is produced by the gut microbe *Pantoea agglomerans* (Dillon et al. 2002). In *Drosophila melanogaster*, mating preference is dictated by the gut microbiota where the flies mate preferentially with individuals harbouring similar microbiota (Sharon et al. 2010, 2011).

6.3.5 Trophic Interactions

The insect gut microbiota is involved in the behavioural aspects of the interactions between insects, their natural enemies and the host (Campbell 1990). In course of evolution, insects have evolved many strategies to feed on plants mediated by mutualistic symbionts (Frago et al. 2012). Insect symbionts have been reported to benefit their hosts; the best-known example is ambrosia beetles and their mutualistic fungi of bark, which make wood digestible for their host (Paine et al. 1997).

6.3.6 Interaction of Gut Microbiota in Productive Insects

Gut bacteria promote populations of beneficial insects by improving the fitness of productive insects, pollinators and biocontrol agents. In irradiated sterile male flies of the Mediterranean fruit fly, *Cerratitis capitata,* mating competence can be improved by feeding diet enriched in *Klebsiella oxytoca* (Lance et al. 2000; Ami et al. 2010) as irradiation causes shift in the microbial community and results in fitness decrease. Feeding on fortified diet significantly increased the sexual competitiveness of irradiated males, enhanced their survival and inhibited sexual receptivity of female flies (Gavriel et al. 2011). Pollinators, like bumble bees, are prone to the attack of parasitoid *Crithidia bombi* and depend upon the gut microbiota (Koch and Schmid-Hempel 2012).

6.4 Insect Microbiome Analysis: From Genomics to Metagenomics

Over the past two decades, the insect microbiome analysis was carried out primarily by the classical approach involving isolation of microorganisms from the various insect physiological systems, culturing them on solid or liquid growth medium containing appropriate sources of carbon, energy and electron acceptor and phenotypic characterization of isolates. This approach solely depended on the physiological conditions under which the organism isolated and sometimes the optimal conditions provided in the laboratory might impose selection pressure, thereby inhibiting the growth of a large number of microorganisms (Staley and Konopka 1985). The media used to isolate insect gut microorganisms were frequently the same as those employed in medical studies. However, some bacteria that were found to be numerically dominating in these media may be physiologically insignificant (Dillon and Dillon 2004). Thus, the focus was on developing the media for culturing of microorganisms, which satisfy the environmental factors, such as pH and available nutrients encountered in the insect physiological system. Further, both the simple morphological and physiological traits in most of the microbes provide only a few identification clues (Pace et al. 1986) and have revealed a large discrepancy between the relatively few culturable microorganisms and the significant diversity present in insect gut (Pace 1997; Head et al. 1998). It was also recognized

that approximately 99% of microbes in the environment cannot be cultured (Amann et al. 1995) and due to limitations of culture methods**,** it was envisioned that most of the microbes associated with the insect gut were still to be identified (Stokes et al. 2001).

From late 1970s onwards, remarkable works were carried out to exploit the ribosomal RNA genes as molecular markers for classifying the life system (Woese and Fox 1977) and this approach in association with the Sanger automated sequencing (Sanger et al. 1977a, b) method revolutionized the study and classification of microorganisms. However, with the advancement in molecular techniques, three traditional molecular approaches, namely 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) have been employed to investigate the insect gut microbial communities (Stokes et al. 2001), whereas the recent advances in sequencing techniques and the metagenomic strategies replaced the above techniques and revolutionized both gene discovery and biodiversity analysis of the insect gut symbiotic microbiota (Fig. 6.1).

6.4.1 Traditional Molecular Approaches in Microbiome Analysis

Gene-Specific PCR

This technique employs gene-specific primers to specifically amplify the target genes, such as conserved *16S rRNA* gene or a gene of specific functional interest from insect gut symbionts. Kane and Pierce (1994) were among the first to conceive the idea of using PCR-based ribosomal DNA sequencing to explore gut microbial communities of termites. Further, McKillip et al. (1997) analysed the composition of the microbiome in the midgut of leaf roller, *Pandemis pyrusana* Kearfott, using both PCR and culturing techniques. Later on, Lilburn et al. (1999) sequenced 98 clones of near-full-length *16S rDNA* of *Spirochaetes* in the gut of termite species, *Reticulitermes flavipes* and observed a substantial phylogenetic diversity in the termite gut. Schmitt Wagner et al. (2003) carried out phylogenetic analysis of *16S rRNA* genes recovered from the hindgut of soil-feeding termites and revealed an enormous diversity of bacteria in the different gut compartments, whereas Ohkuma and Kudo (1996) did the PCR targeting of *16S rRNA* in gut of termite species, *Reticulitermes speratus* and found that most of the gut microbial *16S rRNA*s amplified were unknown. Most of the earlier studies targeting *16S rRNA* gene analyses revealed a significant number of unknown bacterial species at the time.

Apart from *16S rRNA* gene analysis, gene-specific PCR has also been widely applied to identify genes from microbial communities, which are involved in various metabolic pathways. Gene targeting method was followed to clone a number of cellulases belonging to glycosyl hydrolase family 45 from the flagellates *Koruga bonita* and *Deltotrichonympha nana*, which are associated with termite gut (Li et al. 2003). Further, Inoue et al. (2005) identified a cellulase gene from lower termite hindgut using PCR with gene-specific primers and in situ hybridization.

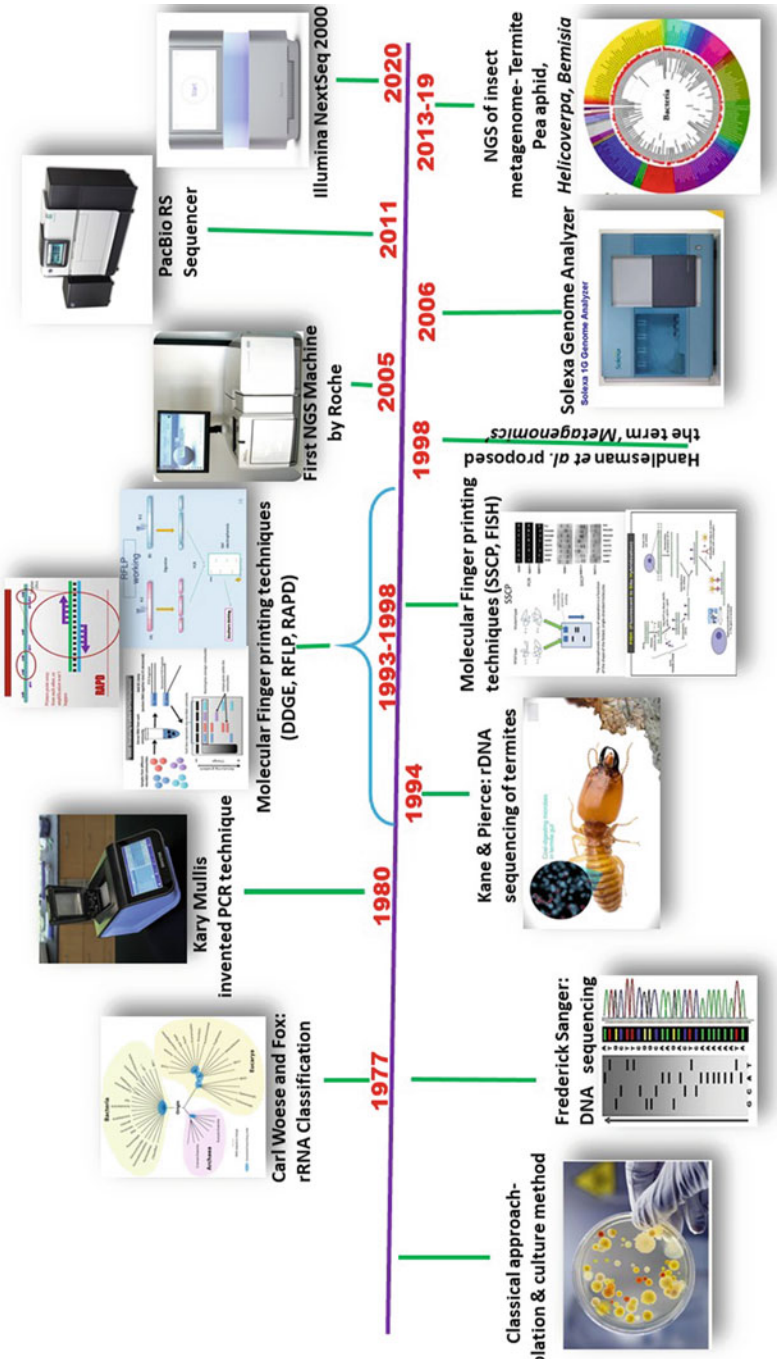

Fig. 6.1 Insect metagenomics—timeline and milestones

In addition to gene-targeting PCR of DNA samples, reverse transcriptase PCR (RT-PCR) from RNA has also been employed to clone genes from environmental samples (Manefield et al. 2002). Casu et al. (1996) identified a major excretory/secretory protease from *Lucilia cuprina* larvae by combining both RT-PCR and immune blotting technique. Later on, Noda et al. (1999) amplified a nitrogen fixing gene from microbial RNA in the gut of the termite, *Neotermes koshunensis* by RT-PCR method. Nakashima et al. (2002) carried out RT-PCR experiments and revealed that five GHF9 EG (Glycosyl Hydrolase Family 9 Endoglucanase) homologs were expressed in the salivary glands and the midgut of termites. Further, the RT-PCR technique was also employed for identifying the genes from gut bacterial communities of *Helicoverpa armigera* and *Manduca sexta,* respectively (Chougule et al. 2005; Brinkmann et al. 2008).

Though gene-specific PCR was proven to be effective for gene discovery and microbial diversity analysis, two major limitations have restricted the application of this technique. The gene-targeting PCR techniques depend on existing sequence information to design primers for PCR amplification and normally only partial sequence of the genes could be cloned; the cloning of full-length genes would have to involve further PCR-based chromosome walking, which greatly limited the application of this technique (Cowan et al. 2005).

Molecular Fingerprinting Techniques

Apart from the sequence, library-based gene targeting PCR and some other PCR-based techniques have also been widely used to analyse microbial diversity in various environmental samples. The molecular fingerprinting techniques used for microbiome analysis include denaturing or temperature gradient gel electrophoresis (DGGE or TGGE), restriction fragment length polymorphisms, single strand conformation polymorphism and random amplified polymorphic DNA (Muyzer et al. 1993; Lee et al. 1996; Liu et al. 1997; Kauppinen et al. 1999). These techniques are used to analyse the sequence of *16S rRNA* gene from different microbial species, where both molecular fingerprints and phylogenetic affiliation of microbial species can be generated. These above-mentioned molecular fingerprinting techniques have been found useful in exploring the microbial diversity associated with insect gut also (Smalla 2004).

Among the molecular fingerprinting techniques, DGGE is the most commonly used method to study insect gut microbial diversity that provides a much more comprehensive understanding of insect symbionts (da Mota et al. 2005). Reeson et al. (2003) analysed the microbial communities associated with wasp larva, *Vespula germanica* based on DGGE profiling and found that the wasp larvae are not solely dependent on one particular type of mutualist. Further, analysis of gut bacterial communities in Mediterranean fruit fly using both culture-dependent and culture-independent approaches, such as DGGE, revealed that the family Enterobacteriaceae is the most dominant species in the gut of fruit fly (Behar et al. 2005). The DGGE method is also used to analyse gut microorganisms in wood feeding termites (Hayashi et al. 2007), soil-feeding termites and their mounds (Fall et al. 2007), hindguts of scarab beetle larvae (Pittman et al. 2008; Vasanthakumar

et al. 2006), gut of grasshoppers and locusts (Dillon et al. 2008), and diamond back moth (Raymond et al. 2008).

Restriction fragment length polymorphism (RFLP) analysis differentiates the homologous DNA sequences based on the distinct DNA fragment patterns generating from the sequence specificity toward restriction enzymes (Esumi et al. 1982). Harada and Ishikawa (1993) used RFLP to analyse *16S rRNA* gene from the group of prokaryote microbes in the gut of the pea aphid and the result suggested that gut microbes have a close relationship with aphid intracellular symbionts. However, due to the technical limitations and low resolution of traditional RFLP technique, terminal restriction fragment length polymorphism (T-RFLP) has been employed to study microbial diversity in insect gut (Shinzato et al. 2005). T-RFLP separates 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 as in the case of bacterial *16S rRNA* genes analysis of the midguts of European cockchafer (*Melolontha melolontha*) larvae (Egert et al. 2005), soil-feeding termites (Kohler et al. 2008) and fungus-growing termites (Shinzato et al. 2007).

Another traditional molecular fingerprinting technique is random amplified polymorphic DNA (RAPD) and the analysis is based on amplification of genomic DNA using random primers. RAPD-PCR was carried out to compare the gut microbial composition between different generations of Western flower thrips, *Frankliniella occidentalis,* and the results revealed that some bacteria in the thrips could be passed from generation to generation for up to 50 generations (de Vries et al. 2001a, b). However, application of RAPD is very limited in gut microbiota analysis due to technical complexity and low reproducibility of the technique.

Single-strand conformation polymorphism (SSCP) is another technique that uses electrophoresis to separate single-strand DNA to differentiate the homologous sequences (Yandell 1991). Mohr and Tebbe (2006) used SSCP to study the diversity and phylogenetic relationship of bacteria in the guts of three bee species at the same oilseed rape field, whereas Brinkmann et al. (2008) used combination of PCR-SSCP, RT-PCR SSCP and stable isotope probing (SIP) to study the diversity of metabolically active bacteria in the larval gut of *Manduca sexta.*

Fluorescent in Situ Hybridization

Fluorescent in situ hybridization (FISH) is one of the most common techniques 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 microbiota studies often involves fluorescently labelled probes targeting *16S rRNA* genes with sequences specific for a bacterial species or genus (Turroni et al. 2008). FISH has been used to detect, visualize and characterize the intracellular symbiotic bacteria of insects, such as aphids (Fukatsu et al. 1998), crickets (Domingo et al. 1998) and termites (Berchtold et al. 1999). The approach has been shown to be particularly useful in studying uncultivated microbes to observe the dynamics of microbiota (Santo Domingo et al. 1998). However, the analysis of complex bacterial communities from environmental samples by FISH with rRNA-targeted probes

often encounters several technical problems and thus the detailed composition of the microbiota cannot be revealed. In addition, bacteria lives in less nutrient-rich environments with low ribosome content, could affect the sensitivity of detection (Smalla 2004).

To complement to FISH, DAPI (4,6-diamidino-2- phenylindole) and GFP (green fluorescent protein) can been used to visualize microbial communities inhabiting insect gut. DAPI staining of bacterial cells highlights the significant differences in the number of bacterial cells among different insect species, when reared under the similar conditions (Cazemier et al. 1997). Similarly, GFP can be used to track target microbial species in the host. Hurst and Jackson (2002) used GFP to show that the colonization of bacterium, *Serratia entomophila* in the gut of the host, *Costelytra zealandica,* is not confined to a specific site in the gut.

6.4.2 Metagenomics

Though the various traditional molecular techniques have greatly advanced our understanding of insect gut microbial communities, 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 insect gut. In order to understand the biological processes involved in biomass degradation, a detailed understanding on the biocatalysts, pathways and compositions of insect gut symbionts is required. However, the high throughput 'metagenomic' approaches allow us to understand the complex properties of the microbiota, their dynamics and function in the natural system. Various metagenomic approaches answer fundamental questions, such as which organisms are present? (Taxonomic diversity), and what roles they play? (Functional metagenomics) (Vieites et al. 2008).

The term 'metagenomics' was coined in 1998 (Handelsman et al. 1998). It helps us to investigate complex microbial communities sampled directly from the environment, without culturing or isolating a single organism. The so-called 'metagenomics' often involves sequence-based, compositional and/or functional analyses of the combined microbial genomes contained within an environmental sample, such as the insect gut (Handelsman et al. 1998). The amplification of specific targeted genes, such as (V1toV9) of *16S rRNA, 18S rRNA,* ribosomal *ITS, NifH,* among others, by PCR before sequencing permit diversity analysis (Morgan and Huttenhower 2012). The diversity, composition and dynamics of a microbial community largely define its effectiveness, specificity and reactivity for a certain function related to life, biogeochemical cycles and environmental mitigation (Allen and Banfield 2005; Falkowski et al. 2008). In the past two decades, significant works have been carried out to explore the components of microbial communities from different niches at the molecular, organismic and ecological levels to reveal novel enzymes, functional pathways and requisite organisms for various applications (Green et al. 2008; Roussel et al. 2008).

Metagenome sequencing has also become important approaches for exploring biomass degrading mechanisms in wood-feeding insects. Several works have been carried out to explore the diversity of microbiota inhabiting the mid- and hindgut of higher (Warnecke et al. 2007) and lower termites (Todaka et al. 2007). However, some studies revealed that symbiotic bacteria and protozoa in the hindgut of the termite play an important role in the hydrolysis of cellulose and hemicellulose (Nakashima et al. 2002; Tokuda and Watanabe 2007; Warnecke et al. 2007; Zhou et al. 2007). Recently, numerous studies have been carried out using metagenomic approach under next-generation sequencing platform to explore the microbial communities associated with insects and their role in its survival and host–insect relationship. In addition to metagenomic approach, metatrascriptomics (refers to sequencing analysis of mRNA from a microbial population) and metaproteomics (refers to the quantification and identification of all the proteins in a microbial community) approaches have also been employed for better understanding of microbial diversity and function in the environment.

6.5 Types and Approaches in Metagenomics

One of the most significant developments in the field of microbial ecology in the past decade has been the advent of metagenomics and it is the explicit method of direct analysis of genomes present in an environmental sample. The field initially started with the cloning of environmental DNA, followed by functional expression screening (Handelsman et al. 1998) and was then quickly complemented by direct random shotgun sequencing of DNA from various environmental samples (Tyson et al. 2004; Venter et al. 2004a, b). Metagenomics provides an insight into the composition of functional genes present in microbial communities and gives a much broader description than usual phylogenetic surveys, which are based only on the diversity analysis of one gene, i.e., *16S rRNA* gene. It provides the valuable genetic information on potentially novel biocatalysts or enzymes involved in various metabolic pathways, genomic linkages between function and phylogeny of uncultured organisms, and evolutionary relationship of community function and structure. The types and approaches in metagenome analysis followed in various environmental fields are discussed hereunder.

6.5.1 Types of Metagenome Analysis

Two types of metagenomic analysis commonly used to unravel the microbial identity and their composition for high throughput sequencing data are: (i) amplicon-based analysis, which includes 16S ribosomal RNA for bacteria, internal transcribed spacer (ITS) and 18S region for fungi and eukaryotes, respectively, and (ii) whole metagenomic shotgun sequencing.

Amplicon-Based Analysis

16S sequencing is a widely used technique that relies on the variable regions (V1-V9) of the bacterial *16S rRNA* gene to make community-wide taxonomic assignments (Chakravorty et al. 2007). It is also used for microbial diversity analysis and for various environmental samples, such as soil (Chong et al. 2012), human gut (Dethlefsen et al. 2008) and various insect gut specimens (Hirsch et al. 2012a, b; Malathi et al. 2018). Some degree of divergence is allowed during the sequence similarity assessment stage of the analysis; typically, nearly identical sequences (>97%) are clustered into Operational Taxonomical Units (OTU) (Morgan and Huttenhower 2012). The limitation of this method is that if any two organisms have the same *16S rRNA* gene sequence, they may be classified as the same species in a 16S analysis, even if they are from different species. Because 16S analysis is based on the *16S rRNA* gene and OTUs are designated as taxa, it is difficult to discriminate between strains and, in some circumstances, closely related species. For example, 16S analysis cannot distinguish *Escherichia coli* O157:H7 from *E. coli* K-12 (Weinstock, 2012) but it can separate *Shigella flexneri* from *E. coli* (Hilton et al., 2016). Similarly, the *18S rRNA* is mainly used for taxonomic studies of fungi, while the *ITS* region is widely adopted for analysing fungal diversity in environmental samples (Bromberg et al. 2015).

Shotgun Metagenome Analysis

Shotgun metagenomic analysis has the ability to identify the majority of the organisms (culturable and unculturable bacteria) in the environmental sample. It helps to create a community biodiversity profile, which can be further utilized for functional composition analysis of organism lineages (i.e., genera or taxa) (Tringe et al. 2005). Before initiating a whole metagenomic study, an understanding of the potential microbial diversity and the relative abundance of species in the environmental sample is very important. Chen et al. (2018) carried out comparative shotgun metagenome analysis of silkworm, *Bombyx mori,* and the sequence datasets not only provide first insights into all bacterial genes in silkworm guts, but also help to generate hypotheses for subsequent analysis of functional traits of gut microbiota. A higher sequencing depth is required to detect a rare taxa from the given environmental sample (Sharpton 2014). This makes shotgun metagenomic sequencing much more expensive than 16S sequencing (Quail et al. 2012).

6.5.2 Approaches in Metagenome Analysis

There are two principal approaches in metagenome analysis: (i) the sequence-based metagenomics, and (ii) functional metagenomics. Sequence-based metagenomics involves metagenome sequencing and downstream data analysis, whereas functional metagenomics involves screening of DNA or cDNA library for gene discovery.

Sequence-Based Metagenomics

Sequence-based analysis of metagenomic DNA from insect gut has been well explored during the past decade to mine out the associated microbial communities.

However, metagenome analysis was first carried out with the conventional Sanger sequencing techniques, which are mainly used toward the *16S rRNA* library or metagenomic DNA library preparations (Smalla 2004). Warnecke et al. (2007) prepared the metagenomic DNA library of termite hindgut symbiotic microbiota with Sanger sequencing techniques, where approximately 71 million pairs of sequence data were generated and assembled, but they are highly fragmented in nature. In order to have a better understanding, 15 fosmids were selected for further sequencing and analysis through shotgun method. The data have led to a comprehensive coverage and quantification of the microbial composition in termite gut symbionts.

The advances in next-generation sequencing technology have offered the potential to revolutionize metagenome analysis (Marusina 2006). When next-generation sequencing is used, the approach can be the direct shotgun sequencing of metagenomic DNA. 454 sequencing technology is the first available next-generation sequencing technique and the platform is based on 'pyrosequencing' and emulsion PCR amplification (Margulies et al. 2005). The sequence read length for 454 sequencing can be up to 400 bases and the throughput is relatively lower at 400 million bases per run. The advantage of the 454 sequencing is the read length, which makes it easier for the sequence assembly in de novo sequencing (Shendure and Ji 2008; Yuan et al. 2008).

Illumina MiSeq, formerly known as Solexa, is based on the concept of 'sequencing by synthesis' (SBS) (Mardis 2008; Adams et al. 2009). With the latest development of the technology, Illumina genome analyser can generate pairwise end sequencing of 100 base pairs and 40 gigabase sequences per run. The two NGS platforms are ABSOLiD and Helocus, both of which have similar sequencing throughput and less sequence read-length (Mardis 2008). Thus, 454 and Illumina have been the major approaches for metagenome sequencing, where 454 offer the longer read length, while the strength of Illumina is the sequence throughput (Stangier 2009). Recently, next-generation sequencing-based metagenome analysis was carried out to explore microbial communities associated with major insect pest of global importance (Hirsch et al. 2012a, b; Scully et al. 2013; Ranjith et al. 2016; Jones et al. 2019; Harish et al. 2019).

Functional Metagenomics

Functional metagenomics involves screening for target genes in a library constructed with metagenomic DNA or RNA (Allen et al. 2009). Generally, metagenomic DNA can be stored stably as a DNA library for further investigation. Similarly, RNA can be reverse transcribed to build a cDNA library. The information available 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).

In order to construct metagenomic DNA library, the basic steps include the extraction of metagenomic DNA, the generation of suitably sized DNA fragments, and the cloning of these fragments into an appropriate vector (Cowan et al. 2005). For the construction of metagenomic 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 technique using the probes of target genes or homologous genes (Demaneche et al. 2009). The approach has been widely used to search for various genes from insect guts. Shen and Jacobs-Lorena (1997) were the first to clone the chitinase gene from a cDNA library through screening and showed that it got expressed exclusively in the midgut of *Anopheles gambiae* adult females using Northern Blot techniques.

One of the major limitations of the traditional screening strategy is the need for specific probes to a certain gene. The sensitivity and reproducibility often also depend 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.

One of the recent developments in the functional metagenomics is the use of biosensor technology for gene discovery from the insect symbionts. Guan et al. (2007) constructed a metagenomic DNA library of midgut microbiota of gypsy moth, and analysed it using an intracellular screen named as METREX. The biosensor detects compounds that induce the expression of GFP from a bacterial quorum promoter by fluorescence microscopy or fluorescence-activated cell sorting (Williamson et al. 2005). Further, they identified an active metagenomic clone encoding a monooxygenase homologue that mediates a pathway of indole oxidation. Further, the metagenomic analysis of whole gut microbiota in four subspecies of termites revealed that they have shared conserved functional and carbohydrate-active enzyme profile and specialized in cellulose and chitin degradation (Grieco et al. 2019).

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

6.6 Steps Involved in Metagenomic Studies

Metagenomics is the study of collective genomes and genes from the members of a microbiota residing at a particular environment. This collection is obtained through sequencing of DNA extracted from an environmental sample followed by annotating the sequence data in silico, thereby increasing the understanding of the dynamics of the microbial community understudy. The various steps involved in insect metagenomic studies (Fig. 6.2), and tools and techniques used in metagenome sequence data analysis are discussed hereunder.

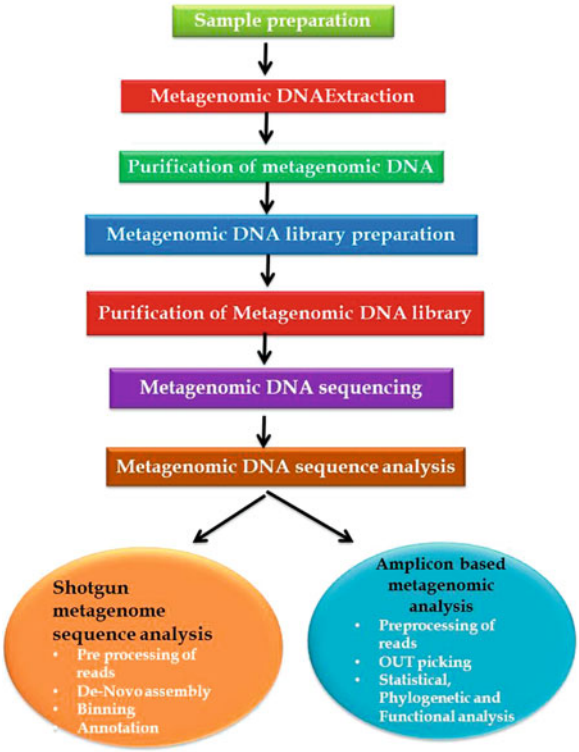

Fig. 6.2 Schematic diagram of steps involved in metagenomic studies

6.6.1 Sample Preparation

The sample preparation step for metagenomic analysis is crucial and must be carefully designed, with immediate analysis or freezing of samples for late analysis. Proper care must be taken to avoid multiple freeze-thaw cycles, which can alter the profile of the microbial community under investigation (Quince et al. 2017). The insect specimens collected from field should be kept overnight for starving and further immobilize them by treating with chloroform (100%). In order to make the surface of insect free from other extraneous microbial communities, which may interfere during the downstream analysis of gut microbiota, it should be surface sterilized with an antibacterial agent, such as streptomycin (0.05%) for approximately 1 h. Further, the antibacterial agent should be removed by surface washing with sufficient quantity of sterile water. As per the objective of study, the gut regions of the insects should be dissected out with utmost care under aseptic conditions.

6.6.2 Metagenomic DNA Extraction

Most of the insect gut metagenomic DNA extraction procedure has been adopted from soil DNA isolation methods (Zhou et al. 1996) with slight modifications. In metagenomic DNA isolation, two major strategies have been employed viz., cell recovery method and the direct lysis method (Roose-Amsaleg et al. 2001). In the cell recovery method, prior to cell lysis, the intact microorganisms associated with insect gut will be isolated by following either frequent homogenization or differential centrifugation or by gradient centrifugation in media, such as percoll or sucrose (Hopkins et al. 1991; Robe et al. 2003). However, in the direct lysis method, either SDS or CTAB containing buffer is used for extracting the metagenomic DNA. Some commercially available kits can also be used for the extraction of metagenomic DNA from uncultured organisms. However, the extraction protocol must be standardized under laboratory conditions, since most of these kits are not designed specifically for insect metagenomic DNA isolation.

6.6.3 Purification of Metagenomic DNA

If the total content of microorganisms from samples may not be efficiently extracted, it may lead to loss of DNA diversity (Josefsen et al. 2015). Since the extracted metagenomic DNA is prone to degradation by nucleases from the external environment, its integrity needs to be protected by inhibiting those enzymes with denaturing agents, which are commonly available in commercial kits. It is also necessary to remove the metal ions to avoid interference with DNA purification steps based on ion exchange. Silica-based columns are also used to bind DNA under high pH and salt concentrations, which helps to remove metal ion interferents (Bag et al. 2016). The DNA interference from dead microbial cells may be eliminated by treatment with propidium monoazide (PMA) or ethidium monoazide (EMA) before DNA extraction. These are DNA intercalating agents that pass only through ruptured membranes and after exposure of the treated cells to ultraviolet light; these agents prevent PCR amplification of the DNA of dead cells (Mayo et al. 2014).

6.6.4 Metagenomic DNA Library Preparation

After metagenomic DNA extraction and purification, the DNA fragmentation and insertion of adapters into the end regions of fragments will be carried out according to various protocols, depending on the sequencing platform (Van Djick et al. 2014). The DNA fragmentation can be performed with physical methods (i.e., ultrasonication), chemical reagents and enzymes with or without transposase activity (Head et al. 2014). Enzymes with transposase activity are highly advantageous because they perform both fragmentation and insertion of labelled or unlabelled sequencing adapters simultaneously, depending on the protocol of choice.

The sequencing adapters will be inserted in the DNA fragments and are specific to each sequencing platform (Van Djick et al. 2014). The adapters are ligated to a support or solid surface to enable spatial separation of fragments. Each fragment will serve as a template for the synthesis of new fragments in the amplification phase and different samples can be sequenced simultaneously during the process (Metzker 2010). The use of DNA indexes allows the processing of a pool of samples and correlates a given fragment with its original sample. The Illumina® platform has a unique indexing process that combines both the adapter and the indexes (barcodes) instead of adding the indexes to the ends of the each mould molecule, as performed for other sequencing platforms (Meyer and Kircher 2010).

For the preparation of libraries, two different approaches can be adopted: (i) paired-end, and (ii) mate-pair. Libraries with short-sized inserts are called paired-end libraries, whereas libraries with long-sized inserts are called mate-pair libraries. Both libraries support the sequencing data to discriminate the physical distance between two reads aligned in the reference genome. According to Van Nieuwerburgh et al. (2012), the success of de novo assembly from short reads depends on the determination of physical distance of the fragment, which is very important to specify the order and orientation of a *contig* in the genome. Thus, the preparation of a paired-end library is highly recommended to complete the regions of the genome containing small gaps, because the short-sized fragments can easily fill empty spaces and provide confirmation for the closing of a draft genome.

6.6.5 Purification of Metagenomic DNA Library

The generated libraries need to be purified before sequencing by selecting appropriately sized fragments and removing free adapters, dimers of adapters and other possible artifacts. This step can be performed with magnetic beads or agarose gel. If dimers of adapters are not removed, they can form clusters in the flow cell and lead to the generation of unwanted sequencing data (Head et al. 2014).

6.6.6 Metagenomic DNA Sequencing

The first step in metagenomic DNA sequencing is to choose a sequencing platform of a particular manufacturer, with due attention given to the set of data generated from the platform in each run (output). Among the companies that market sequencing platforms, Illumina® currently stands out for offering a variety of highly compatible platforms (Goodwin et al. 2016). In addition, Illumina® platforms provide the highest high-throughput per run and the lowest cost per sequenced base among all companies (Van Djick et al. 2014).

Illumina® platform uses the Sequencing By Synthesis (SBS) technique coupled with bridge amplification process in the flow cell (Shokralla et al. 2012). SBS sequencing uses the enzymes, such as DNA polymerase or DNA ligase, for the massive parallel amplification of template DNA. During the operation of SBS

platforms, DNA polymerase adds labelled dNTPs on real time uninterruptedly, which are easily distinguishable from nucleotides not incorporated into the template DNA with the aids of an optical reader (Fuller et al. 2009).

Either single-end (SE) or paired-End (PE) sequencing can be opted and this profoundly influences the downstream analysis. SE sequencing refers to sequencing from a single end of the library fragment, whereas PE sequencing refers to sequencing from both ends of the fragment in a two-way elongation process (Van Djick et al. 2014). Paired-end sequencing is the most common approach and it is cost-effective, because it generates two reads for the same fragment per run.

6.6.7 Metagenomic Sequence Data Analysis

Various pipelines are used for downstream analysis in different metagenomic methods and the requisite bioinformatics' tools (Table 6.1).

Shortgun Metagenome-Sequence Analysis

Pre-Processing of Sequence Reads

The raw reads generated from the next-generation sequencing platform are subjected to adapter trimming, quality filtration and de-replication. If the metagenomic sample is isolated from a host organism, then host contamination is typically removed by aligning to the reference genome of the host organism, using Bowtie2 or other short-read mapper (Oulas et al. 2015).

De Novo Assembly

Assembly is computationally expensive and it requires sophisticated algorithms based on de Bruijn graphs. Tools that are specifically designed for metagenomic applications are mainly built on de Bruijn graph algorithms. A few common metagenomics assembly tools include CLC workbench, Meta-Ray, MetaVelvet-SL, MetaVelvet, Meta-IDBASOAP and metaSPAdes (Nurk et al. 2017; Luo et al. 2012). If an appropriate reference metagenome is available in the database, a reference-based assembly may be performed (Nagarajan et al. 2010).

Binning

Binning is the process of clustering the reads or contigs into a highly similar groups, and assigning the groups to specific taxa, such as species, subspecies or genera. Two types of algorithms are available: (a) composition-based binning, and (b) similarity-based binning. Certain binning tools make use of hybrid approaches, which run both kinds of algorithms. In composition-based binning, the groups occur in a supervised or semi-supervised manner, where the DNA fragments are with similar composition, whereas in similarity-based binning, it aligns the DNA fragments to database or reference sequences (Leung et al. 2011).

Table 6.1 Bioinformatic tools used in metagenomic analysis

Sl. No.	Name	Application
1	**QIAGEN CLC Main workbench** digitalinsights.qiagen.com/products-overview/discovery-insightsportfolio/analysis-andvisualization/qiagen-clc-mainworkbench/	For gene expression analysis, primer design, molecular cloning, phylogenetic analyses and sequence data management
2	**MetaVelvet** metavelvet.dna.bio.keio.ac.jp	For de novo metagenomic assembly
3	**MaxBin2 online** https://kbase.us/applist/apps/kb_maxbin/run_maxbin2/release?gclid=EAIaIQobChMI8tY2IqH7AIVxm4qCh2B6Qr6EAAYASAAEgIxc_D_BwE	For Metagenome Assembly Binning
4	**Pfam** https://pfam.xfam.org/	The Pfam database is a large collection of protein families, each represented by *multiple sequence alignments* and *hidden Markov models (HMMs)*
5	**TIGRFAMs** http://tigrfams.jcvi.org/cgi-bin/index.cgi	For protein sequence classification, and associated information designed to support automated annotation of (mostly prokaryotic) proteins.
6	**WebMGA** http://weizhonglilab.org/webMGA/server/tigrfam/	Web service for metagenomic analysis
7	**Ribosomal database project (RDP)** http://rdp.cme.msu.edu/	For quality-controlled, aligned and annotated bacterial and archaeal 16S rRNA sequences, and fungal 28S rRNA sequences
8	**Phylosift** https://phylosift.wordpress.com/	Phylogeny-driven metagenomic classification and comparison
9	**metAMOS** https://github.com/marbl/metAMOS	Metagenomic assembly and classification pipeline
10	**MG-RAST** https://www.mg-rast.org/	Open source, open submission web application server that suggests automatic phylogenetic and functional analysis of metagenomes.
11	**MEGAN6** https://uni-tuebingen.de/fakultaeten/mathematisch-naturwissenschaftliche-	Comprehensive toolbox for interactively analysing microbiome data

(continued)

Table 6.1 (continued)

Sl. No.	Name	Application
	fakultaet/fachbereiche/informatik/lehrstuehle/algorithms-in-bioinformatics/software/megan6/	
12	**Qiime** http://qiime.org/	For performing microbiome analysis From raw DNA sequencing data
13	**UCLUST** https://www.drive5.com/usearch/manual/uclust_algo.html	It is an algorithm designed to cluster nucleotide or amino-acid sequences into clusters based on sequence similarity
14	**AmphoraNet** https://pitgroup.org/amphoranet/	For metagenomic and genomic phylotyping
15	**Silva** https://www.arb-silva.de/	A comprehensive on-line resource for quality checked and aligned ribosomal RNA sequence data
16	**PICRUSt** http://picrust.github.io/picrust/	To predict metagenome functional content from marker gene (e.g., 16S rRNA) surveys and full genomes.
17	**PyNAST** https://biocore.github.io/pynast/	For adding new 16 s rRNA sequences to existing 16 s rRNA alignments.
18	**SRA (sequence read archive)** https://www.ncbi.nlm.nih.gov/sra	It is the largest repository of high throughput sequencing data (including metagenome sequencing data)
19	**Greengenes** https://greengenes.secondgenome.com/	It is a full-length 16S rRNA gene database that provides a curated taxonomy based on de novo tree inference
20	**Krona tools** https://hpc.nih.gov/apps/kronatools.html	It allows hierarchical data to be explored with zooming, multi-layered pie charts

Annotation

Annotation is the prediction of CDS (coding DNA sequences) of the genome, followed by its functional assignment based on similarity searches of query sequences against databases containing a known functional and/or taxonomic information. The taxonomic information can be displayed using Krona, which plays hierarchical data as an interactive multi-layered pie-chart (Ondov et al. 2011). The predicted genes are annotated to identify homologous genes using Gene ontology terms, KEGG pathways, protein families using Pfam or TIGRfams, clusters of orthologous genes (COGs/KOGs) or orthologous families and functional motifs using Inter Pro. Some tools, such as Kaiju, assign taxonomy status using a reference database, and also integrate the Krona tool for visualization of taxonomic composition, whereas COGNIZER can be used for functional annotation, which applies a new approach of search strategy that helps in reducing the computational requirements (Gosh et al. 2018).

Amplicon-Based Metagenomic Analysis

Pre-Processing of Reads for Amplicon Analysis

During this process, the raw files generated from the next-generation sequencing platform will be subjected to de-multiplexing, adapter trimming, and quality filtration (Plummer et al. 2015) and the detection of PCR chimera and its removal will be carried out using UCHIME algorithm (Sinclair et al. 2015).

OTU Picking and Taxonomic Assignment

OTU picking groups are the similar sequences by clustering or a similarity-based method. OUT picking in the most popular tool QIIME is performed using the UCLUST programme. The UCLUST program uses the algorithm USEARCH to assign the sequences to clusters (Edgar 2010). Each OUT represents a cluster of sequences with similarity greater than a threshold, typically 97–98%, which is then assigned to a corresponding taxonomic group. There are various OUT picking strategies: (1) De novo, wherein the reads are clustered without reference to known sequences; (2) Closed-reference, where the reads are clustered based on the alignment to a reference database; or (3) Open reference method, where clusters read against a reference database and also clusters unaligned reads using a de novo approach. All these methods are incorporated in the tool, QIIME (Oulas et al. 2015).

Statistical Analysis

The taxonomic tree in Newick format can be obtained from QIIME tool and it can be visualized using any tree display tool, such as FigTree. The alpha diversity measures the variability within a single population, which measures the richness, dominance and evenness. Rarefaction analysis is used to assess the coverage of the microbial community contained in the sample and the resultant rarefaction curves plot the sample size versus the estimated number of genera (Jaenicke et al. 2011).

Beta diversity measures the diversity across many samples or populations, which is calculated using various matrices, such as weighted and unweighted UniFrac and

PCoA (Principal Coordinate Analysis). It includes the absolute or relative overlap between the samples for estimating the taxa shared among them. The calculation of both the alpha and beta diversity is well supported by QIIME tool.

Phylogenetic Analysis

The phylogenetic analysis helps in identifying the species and its lineages at taxonomic levels. The various tools used for analysing the phylogenetic relationship in metagenomes are AmphoraNet, TIPP (taxonomic identification and phylogenetic profiling) and Phylosift (Gosh et al. 2018).

Functional Analysis

In order to predict the functional composition of microbial communities from the 16S profile, the tool PICRUSt can be used. It employs an extended ancestral-state reconstruction algorithm, which predicts the gene families and further combines the gene families to estimate the composite metagenome. The annotation of the predicted gene family counts can be obtained from orthologous groups of gene families, KOGs, COGs, NOGs, or Pfam families (Langille et al. 2013).

6.7 Metagenome Analysis of Insect Pests: An Overview

Nowadays, the insect-associated microbial communities are attracting wide attention mainly because of their ecological and economic importance. Microorganisms have been investigated for their profound influence on their host partner by directly mediating interactions with other species or indirectly by impacting the host genetic diversity. Moreover, microorganisms can help insects to counteract the defence mechanisms offered by the host plants, provide protection against natural enemies, influence the reproductive ability and help to survive on nutritionally marginal diets (Ferrari and Vavre 2011). Recently, the study of host-microorganism interactions has attracted a wide attention with the introduction of metagenomic techniques. A wide range of research described the insect-associated microbial community using metagenomic tools and the glimpse on metagenome analysis of insect pests of global importance are discussed hereunder.

6.7.1 Termites

Termites pose serious threat to a wide range of agricultural crops, structures, especially wooden materials and prove themselves a major insect-pest to humankind. The gut of termite is a rich reservoir of microbes, belongs to Bacteria, Archea and Euckarya and the higher termites are capable of digesting the lignocellulose in various stages of humification with the help of an array of symbiotic prokaryotic microbiota housed in their compartmented intestinal tract. The metagenomic profiling of hindgut pouches of wood (*Amitermes wheeleri*) and dung (*Nasutitermes corniger*) feeding termites based on *16S rRNA* pyro-sequencing revealed that

Firmicutes and *Spirochaetes* were the most abundant phyla in *A. wheeleri* in contrast to *N.corniger* where *Spirochaetes* and *Fibrobacteres* dominated. Further, functional analysis revealed that the microbiota associated with *A. wheeleri* involved in hemicellulose breakdown and fixed-nitrogen utilization, whereas, those associated with *N. corniger* possess glycoside hydrolases attacking celluloses and nitrogen fixation genes (He et al. 2013).

Later on, metagenomic profiling of highly compartmented hindgut of six wood or soil feeding termite reveals that P1 compartment of the most of termite species is dominated by Firmicutes, whereas P4 is generally more diverse when compared to other compartments and displayed an increasing abundance of Bacteroidetes (Rossmassler et al. 2015). Metagenomic analysis of whole gut microbiota in seven species of termites (Termitidae) with different feeding habits from four locations at Brazil reveals that in termite species feeding on litter, the bacteria belong to the phylum Firmicutes are abundant, whereas in humus feeding termite species, the bacteria belonging to the phylum Proteobacteria are abundant. The gut microbiota of all four examined subfamilies of termites shared a conserved functional carbohydrate-active enzyme profile specialized for cellulose and chitin degradation (Grieco et al. 2019).

6.7.2 Pea Aphid, *Acyrthosiphon pisum*

The pea aphid, *Acyrthosiphon pisum*, a pest of legume crops represents a well-studied case of symbiotic associations. The 454 pyro-sequencing of pea aphid resulted in a range of 2838–16,637 sequence reads with a median of 4199 reads per sample. In total, *Buchnera* sequences comprised an average of 88.4% of the sequence reads followed by *Serratia symbiotica* with an average sequence read abundance of 4.3%. The X-type, Rickettsia, *H. defensa* and *R. insecticola* were next in read abundance, ranging from 1.2 to 1.7% of sequence reads, on average, across all samples (Russell et al. 2013).

The diversity analysis of bacterial communities associated with nine biotypes of the pea aphid complex using pyro-sequencing of *16S rRNA* genes reveals that *Spiroplasma* was the most dominant taxon in number of sequences (48%) followed by *Rickettsia* (25%) and *Buchnera* (21%) (Gauthier et al. 2015). Cariou et al. (2018) compared both *16S rDNA* amplicon sequencing and hybridization capture for pea aphid microbiota diversity analysis and found that both the methods provide description of 8 bacterial taxa, namely *Buchnera aphidicola*, *Hamiltonella defensa, Rickettsiella viridis, Rickettsia* sp., *Regiella insecticola, Fukatsuia*, *Serratia symbiotica* and *Spiroplasma* sp. and considered as qualitatively and quantitatively robust on such a sample with low microbial complexity.

6.7.3 Boll Worm, *Helicoverpa Armigera*

Helicoverpa armigera (Hübner), commonly known as American bollworm or gram caterpillar or tomato fruit borer, is a polyphagous insect pest known to infest many economically important crops throughout the world. The dreaded nature of this pest is attributed to number of factors, among which the gut microbiota also play a major role to thrive in various crop ecosystem. T-FRLP analysis of the gut bacterial community associated with *H. armigera* from tomato, chickpea and cotton crops at different locations showed that among the 12 bacterial phylotypes detected, *Enterococcus faecalis* and *Enterobacter* sp. were the major phylotypes found in all the larvae regardless of the crop or location of samples collected including artificial diets (Priya et al. 2012).

Further, Ranjith et al. (2016) analysed the composition and diversity of gut bacterial communities associated with *H. armigera* based on Illumina Next-Generation Sequencing (NGS) of 16S ribosomal RNA. The NGS dataset consisted of 864,813 high-quality paired end sequences with mean length of 150 base pairs. A highly diverse groups of bacteria were present in the sample with an approximate of 2303 operational taxonomic units (OTUs). A total of 17 bacterial phyla, 34 classes, 84 orders, 173 families, 334 genera, and 707 species were deduced from the sequence analysis. *Actinobacteria* was the most dominant taxon, followed by *Proteobacteria* and *Firmicutes*. Dar et al. (2018) identified cellulose degrading bacteria *Klebsiella* sp. MD21 from the gut of cotton bollworm, *Helicoverpa armigera* based on *16S rRNA* gene sequencing and demonstrated that *H. armigera* can be used as source of cellulolytic bacteria, which can be utilized in both biorefinery and pulp industries.

6.7.4 Whitefly, *Bemisia tabaci*

Silverleaf whitefly, *Bemisia tabaci* (Gennadius), is one of the polyphagous sucking insect pests, infesting more than 900 species of plants and serve as a vector for spreading more than 200 viral diseases. Harish et al. (2019) studied the composition of bacterial communities associated with whitefly infesting cassava from two different zones (zone P: plains; zone H: high ranges) of Kerala, India, using the next-generation sequencing of *16S rDNA*. Sequence analysis revealed a marked difference in the relative abundance of gut inhabiting bacteria present in the populations. In the P population, the taxonomic status of bacteria identified were 16 phyla, 27 classes, 56 orders, 91 families, 236 genera and 409 species, whereas in H population, it was earmarked as 16, 31, 60, 88, 225 and 355, respectively. The most dominant bacterium present in P population was *Arsenophonus* sp. (Enterobacteriaceae), which aids in virus transmission, whereas in the H population, *Bacillus* sp. was found relatively abundant. This study pinpoints the association between whitefly biotypes and secondary symbionts and the role of bacteria in modifying the host characteristics, such as transmission of various virus groups, expanding the host range, imparting the insecticide resistance and speciation.

The comparative analysis of endosymbionts present in 21 globally collected species in the *B. tabaci* complex, and two samples of *B. afer* using PacBio sequencing of full-length bacterial *16S rRNA* gene amplicons revealed the new putative bacteria and one among them was *Halomonas,* first confirmed to be present in MED *B. Tabaci* (Indiragandhi et al. 2010). Similarly, new secondary endosymbiotic strains of *Rickettsia* and *Arsenophonus* were also found associated with the whitefly samples collected from different locations (Wang et al. 2019).

Shah et al. (2020) characterized bacterial communities present in wild adult *B. tabaci* infesting cotton plants in eight major cotton growing districts of southern Punjab, Pakistan based on 16S rDNA next-generation sequencing and identified 50 known and 7 unknown genera of bacteria belonging to 10 phyla, 20 classes, 30 orders and 40 families. Proteobacteria was the most abundant phylum followed by Bacteriodetes, Firmicutes and Actinobacteria.

6.7.5 Diamond Back Moth, *Plutella xylostella (L.)*

The diamond back moth, *Plutella xylostella* (L.) (Lepidoptera: Plutellidae), is one of the most destructive insect pests infesting the cruciferous vegetables, such as cabbage, broccoli and cauliflower across the globe. The first report of high-throughput DNA sequencing of the entire microbiota of DBM reveals that more than 97% of the bacteria were from three orders, namely Enterobacteriales, Vibrionales and Lactobacillales. Both chlorpyriphos and fipronil resistant lines used in the study had more Lactobacillales and the much scarcer taxa Pseudomonadales and Xanthomonadales with fewer Enterobacteriales compared with the susceptible strain and this is consistent with the hypothesis that Lactobacillales or other scarcer taxa play a role in conferring DBM insecticide resistance (Xia et al. 2013).

Metagenomic analysis of diamond back moth reveals that the phylum, Proteobacteria was the dominant taxon in the *P. xylostella* gut microbiota, followed by Firmicutes. Functional metagenome analysis reveals the role of gut bacteria in metabolic activities associated with glycans, carbohydrates, amino acids, vitamins, xenobiotics and terpenoids, which are linked to digestion, nutrition and detoxification. The most enriched functions within these activities were carbohydrate metabolism and amino acid metabolism (nutrition), followed by xenobiotic degradation and terpenoid metabolism (detoxification of plant defensive compounds) (Xia et al. 2017).

6.8 Application of Metagenomics in Insect Pest Management

Metagenomics has wide range of application from clinical to environmental samples, from food safety to industrial waste and also in identifying the pathogens, which can infest various hosts including humans and animals. Metagenome analysis provides the information on both the diversity and function of microbiota associated

with insect pest. These inherent gut microbiota play a crucial role in its insect survival by upgrading the nutrient status of diet, aids in digestion of recalcitrant food, protection from parasites, pathogens, lethal insecticidal molecules and development and maintenance of host immune system (Gill et al. 2004; Wernegreen 2002). The metagenomics offers new technologies and conceptual approaches to the entomologists by facilitating the study of impacts of microbes on insect function and to adopt various pest control strategies based on manipulation of microbial partners.

The application of insect metagenomic studies in formulating various novel approaches in pest management and few are discussed hereunder.

6.8.1 Improve Biosurveillance Programme

Metagenomics techniques can be used to improve bio-surveillance programmes, as a tool to detect the arrival, origin, invasion pathways and adaptation traits of invasive species (Roe et al. 2018) in different ecosystems. It can be employed in the monitoring of critical areas viz., port of entry where massive trapping is the most common practice followed to identify the arrival of invasive insect species (Rassati et al. 2018; Poland and Rassati 2018; Rassati et al. 2015). This regular mass trapping is time-consuming and laborious process, often requires extensive taxonomic knowledge of different systematic groups. However, the metagenome analysis simplifies the process by analysing the entire genetic pool of single traps, and detecting not only the arrival of an invasive insect species, but also likely plant pathogens (Malacrinò et al. 2017; Roe et al. 2018).

6.8.2 Suppression of Vector Competence of Insects

On the basis of comparative metagenome analysis, Hajeri et al. (2014) developed a novel method of RNAi mediated control for the Asian citrus psyllid, *Diaphorina citri*, a vector for multiple citrus diseases, by modifying the genome of the citrus tristeza virus into a stable vector producing dsRNA. The progeny of *D. citri* individuals feeding on plants infected with the modified virus showed increased mortality.

6.8.3 Manipulation of Host Range of Insect Pests

Microbial symbionts play a major role in determining the host range of phytophagous insects, but persuasive evidence is rare (Hansen and Moran 2014). However, exceptional phenomenon has been noticed in two plataspid stinkbugs, *Megacopta punctatissima* and *M. cribraria*. In their native range in Japan, *M. punctatissima* is an agricultural pest, especially of soybean crops, but *M. cribraria* performs very poorly on soybean. Metagenome profiling reveals that the capacity of *M. punctatissima* to utilize soybean is mediated by the bacterial symbiont *Ishikawaella* localized to the

distal portion of the insect gut. When the newly hatched nymphs of *M. punctatissima* were administered with *Ishikawaella* symbiont of the other species, *M. cribaria* leads to poor consumption with high mortality on soybean (Hosokawa et al. 2007). The *Megacopta* association has great potential for manipulation to suppress the infestation of soybean crops because heterologous associations can be generated very easily by feeding neonate nymphs on symbionts from a different insect species and the acquired partner is then transmitted vertically with high fidelity (Hosokawa et al. 2005).

6.8.4 Heterologous Symbionts those Are Insecticidal

Many insect-microbial associations are co-evolved, with the implication that certain microorganisms that are benign in their native insect host may be deleterious when introduced to a different insect. Incompatibility can occur naturally on hybridization between two related insect species with maternally inherited symbionts. This phenomenon has been demonstrated very clearly through metagenomic profiling of two species of jewel wasps, *Nasonia vitripennis* and *N. girauldi*, with genetic evidence that interspecific crosses yield incompatibilities between a maternally inherited 'factor' and the nuclear genome of the hybrid (Breeuwer and Werren 1995). Identification of the microbial symbionts as the 'factor' comes from the finding that antibiotic treatment protects against hybrid lethality and that lethality is revived by adding back specific gut bacteria, *Providencia* sp. and *Proteus mirabilis*, derived from each of the two parental jewel wasp species (Brucker and Bordenstein 2013). The high populations of *P. mirabilis* in hybrid insects are indicative of immunological dysfunction even though the underlying mechanisms are not understood fully (Chandler and Turelli 2014).

For application to control insect pests, there are two key requirements: (1) the pest and non-pest species hybridize under field conditions, yielding viable progeny; and (2) an association can be constructed between the non-pest species and microorganism(s) that are benign in the non-pest species but lethal to hybrids between the non-pest and pest species, as well as posing no risk to other species or the wider environment. Under these conditions, implementation would have many parallels to the sterile insect technique, but with the mass release of the non-pest species bearing the microorganisms instead of sterile conspecific male insects.

6.8.5 Paratransgenesis and Induced Lethality in Insect Pests

The most developed application of genetically modified microorganisms in insect pest control is paratransgenesis, which can be defined as the alteration of insect traits by genetic manipulation of associated microorganisms (Beard et al. 1998). The potential of this technology in insect pest control has been appreciated for more than twenty years (Beard et al. 1993), especially in relation to mosquito vectors of human disease agents. The key requirements for paratransgenesis are that the

microbial partner is culturable under ideal condition and amenable to genetic manipulation as well as readily transmitted among insects to facilitate the transfer of the desired trait (Beard et al. 2002). The possibility of using this technique to manage the agriculturally important insect pests is to be thoroughly explored with metagenome analysis.

6.8.6 Genetically Modified Microorganisms as Insecticides

Genetic technologies can be applied to modify microorganisms to express traits that are virulent to the insect. The use of microorganisms for delivery of dsRNA relates to the promise of RNA-interference (RNAi) to target insect pests by suppressing the expression of essential insect genes. *In planta* RNAi is now used widely in research on herbivorous insects, with an insecticidal RNAi against the Western corn rootworm *Diabrotica virgifera virgifera* reported to be close to commercial release in transgenic corn. In addition, the environmental release of dsRNA against insect pests found associated with soils, water and other natural habitats are being promoted by advanced encapsulation technologies (Scott et al. 2013).

6.8.7 Elimination of Vertically Transmitted Obligate Microbial Partner

The goal of targeting microbial partners is to control insect pests by eliminating the microorganisms required for sustained insect growth, reproduction and survival. Unlike the use of heterologous or genetically modified microorganisms, which involve the administration of microorganisms to insects, this strategy involves the use of specific symbiocides, i.e., effectors that perturb the resident microbial partners and their interactions with the insect.

The insect systems ideally suited to this strategy involve bacteria that are localized to specialized insect cells known as bacteriocytes. Because the bacterial partners are obligately vertically transmitted and unknown apart from their insect hosts (Buchner 1965; Douglas 1989) a treated insect cannot regain the association horizontally from other insects or the environment.

A strong motivation to develop methods that target the bacteriocyte symbioses comes from the expectation of specific molecular targets linked to the coevolutionary interactions between the participating insect and microbial lineages (Douglas 2015). This can be explicitly studied based on metagenome analysis of pest taxa potentially amenable to this strategy include sap feeding hemipterans (aphids, whiteflies, planthoppers, leafhoppers etc.), and various xylophagous and stored product coleopteran pests (many curculionids and chrysomelids, the anobids and bostrychids).

Grape plants transformed with constructs coding the anti-microbial peptide, cecropin B, with either melittin or elastase, reduced the *Xylella* abundance and disease symptoms in the plants (Dandekar et al. 2012; Li et al. 2015). The

antimicrobials circulating in the xylem sap of these plants are presumably ingested by the xylem-feeding insects, including the leafhoppers that vector *Xylella,* but their activity against the obligate bacterial symbionts in the leafhoppers (Wu et al. 2006) needs to be investigated. The feasibility of selective symbiocides is supported further by the relative ease with which orally delivered antibiotics and antibodies can cross the gut wall to the hemocoel and internal organs of insects (Bonning and Chougule 2014; Jeffers and Roe 2008).

6.8.8 Elimination of Horizontally Transmitted Obligate Microbial Partner

This strategy combines the use of genetically modified microorganisms as a delivery vehicle to target obligate microbial partners. Gut metagenome analysis reveals that the lower termites are absolutely dependent on cellulose degrading trichomonad and hypermastigote protists in their hindgut, providing an opportunity to control the pest by targeting the protist symbionts. Various antimicrobial peptides, including melittin, cecropin, and the synthetic product Hecate have been demonstrated to lyse these protists, but their application has been constrained by challenges in their delivery to the hindgut. This limitation has been overcome by using microorganisms as the delivery vehicle to target the obligate symbiont of the insect pest instead of the insect.

The commercially available yeast, *Kluyveromyces lactis,* engineered to express melittin was used as microbial delivery vehicle and administrated to the termite *Coptotermes formosanus* was found effective in eliminating the protists without detectable direct damage to the insect gut (Husseneder et al. 2016). It was also demonstrated that a bacterial isolate from *C. formosanus*, *Trabulsiella odontotermitis*, is genetically transformable and transmitted efficiently among termites (Tikhe et al. 2016) and offers a route to use a natural symbiont as the delivery vehicle for the toxic peptides.

6.9 Future Perspective

The study of insect metagenome analysis yields valuable information on role of different groups of microbes in insect physiology, pest management, evolutionary relationships and the tritrophic interactions existing in the nature. It also gives an insight into the various microbe derived novel biocatalysts, which can be used for various 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 both host insect and symbiont derived enzymes plays an important role in biomass processing and degradation. Thus, study of insect gut symbiotic microbiota at the systems level will enable us to design the next-

generation biorefinery for various levels of industrial applications. Similarly, the insect microbiome analysis provides the role of different microbiota in insect survival and development in a particular environment.

The insect metagenome analysis has experienced dramatic changes during the past two decades. The initial studies of insect gut microbes were based on culture-dependent platforms, which provided a very limited information on the diversity and functions. The culture-dependent analysis was quickly replaced and complemented by the advancement in molecular techniques, which is a key partner in culture independent methods. The molecular fingerprinting techniques, like DGGE, SSCP, RFLP and FISH, allow us to better explore the complexity of natural microbial communities present in an ecosystem. However, the development of metagenomic approaches and the advancements in next-generation sequencing techniques allow us to explore the metagenomes from insect gut symbiotic microbiota to an unprecedented depth and comprehensiveness.

In addition to metagenomics, metatranscriptomic and metaproteomic profilings are also providing important information regarding the function of insect hosts and symbionts from different perspectives. The integration of information will lead to a systems-level understanding of insect gut as the system for biomass deconstruction, nutrient biosynthesis and to formulate various novel approaches in pest management. Despite significant progresses, several aspects of research need to be emphasized to better exploit insect gut systems for various biotechnology applications.

Though the metagenomic approach provided a thorough knowledge on the microbial census in the insect gut, identification of novel genes and the development of potential biotechnological applications is a great challenge due to the presence of both diverse microbial communities and the variability existing in their genomes. Most of the bioinformatics programmes are designed for collecting and depositing of the metagenomic sequence composition, and their respective data management. However, more sophisticated bioinformatics tools are yet to be developed to analyse the *hitherto* unexplored microbial genes of insect gut metagenomics. Though the new high throughput next sequencing technologies enable us for identifying a novel candidate gene, the assay for protein function exemplify one of the most important and inimitable tools for identifying their target genes. Thus, the development of high throughput functional screening methods will also be necessary to assess the functional role in particular system.

Most of the insect metagenomic studies are focusing merely on exploring the gut microbial composition and their functional diversity. However, metagenome analysis also provides valuable information to formulate various pest management strategies based on manipulation of insect-associated microorganisms. The status of the various strategies varies from generalized concepts and experimental proof-of-principle under defined laboratory conditions to products suitable for field application and ongoing field trials in multiple countries. Furthermore, knowledge of the molecular basis of most strategies on microbial manipulations offers the opportunity for modification of the product in response to a novel insect pest and resistance evolution in insect pests.

6.10 Conclusions

Insect-associated microbial communities are attracting increasing interest nowadays, mainly because of their ecological and economic importance. They play essential roles in the growth, development, pathogenesis and environmental adaptation of host insects. At present, we are capable of exploring the microbial communities associated with insects, their composition, diversity and interaction with their hosts. In particular, the modern molecular techniques, metagenomics revolutionized the field with enormous data to enable unprecedented understanding of insect gut symbiotic microbiota and their interactions with hosts. The metagenome approaches together with the recent advancements in next-generation sequencing provide enormous sequencing information, allowing in-depth microbial diversity analysis and modelling of pathways for biological processes, such as biomass degradation. In addition, insect gut metagenome analysis data also provide conceptual approaches to the plant protection specialists to formulate various novel pest management strategies based on manipulation of insect-associated microorganisms. Certainly, metagenomics in combination with metaproteomic and metatranscriptomic approaches and modern bioinformatics tools enable us to retrieve pivotal information that can effectively be used in combating ravages of insect pests.

6.11 Points to Be Remember

- Insects, the most successful groups in animal kingdom, harbour diverse groups of microbes viz., bacteria, archaea, fungi, protozoa and viruses.
- These microbes are associated with their host insects permanently or transiently and such associations may be beneficial or harmful to the host insects under various instances.
- The initial studies of insect microbiome were based on culture-dependent platforms, which provided very limited information for the diversity and functions of insect gut symbiotic microbiota.
- This classical approach was quickly replaced and complemented by traditional molecular approaches, like gene specific PCR, molecular fingerprinting techniques (DGGE or TGGE, RFLP, SSCP and RAPD) and FISH allowed us to better explore the complexity of natural microbial communities.
- The recently developed metagenome sequencing techniques, in particular, the advancements in next-generation sequencing techniques allow us to explore the metagenomes from insect gut symbiotic microbiota to an unprecedented depth and comprehensiveness.
- Two types of metagenomic analysis are commonly used to unravel the identity and composition of microbes for a high throughput sequencing data are amplicon-based analysis and shotgun metagenome analysis and the two principal approaches in metagenome analysis are sequence-based metagenomics and functional metagenomics.

- The major steps involved in metagenomic studies include sample preparation, extraction and purification of metagenomic DNA, metagenomic DNA library preparation, sequencing under various next-generation platforms, sequence data analysis and interpretation.
- Recently, a number of researches have been carried out to explore metagenome of insect pest of global importance, viz. termite, pea aphid, cotton bollworm, silver leaf whitefly, diamond back moth, etc. reveals the crucial role played by the microbes in insect nutrition, protection from parasites, pathogens, lethal insecticidal molecules and development and maintenance of immune system.
- Insect metagenomic research will aid in formulation of various novel pest management approaches viz, improved bio-surveillance programme, suppression of vector competence of insects, manipulation of insect host range, use of heterologous symbionts, paratransgenesis and induce lethality in insects, genetically modified microorganisms as insecticides and elimination of both vertically and horizontally transmitted obligate microbial partners.

References

Adams IP, Glover RH, Monger WA, Mumford R, Jackeviciene E, Navalinskiene M, Samuitiene M, Boonham N (2009) Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. Mol Plant Pathol 10(4):537–545

Allen EE, Banfield JF (2005) Community genomics in microbial ecology and evolution. Nat Rev Microbiol 3(6):489–498

Allen HK, Moe LA, Rodbumrer J, Gaarder A, Handelsman J (2009) Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. ISME J 3(2):243–251

Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Rev 59(1):143–169

Aminov RI, Walker AW, Duncan SH, Harmsen HJ, Welling GW, Flint HJ (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*. Appl Environ Microbiol 72(9):6371–6376

Ami E, Yuval B, Jurkevitch E (2010) Manipulation of the microbiota of mass-reared Mediterranean fruit flies Ceratitis capitata (Diptera: Tephritidae) improves sterile male sexual performance. ISME J 4:28–37. https://doi.org/10.1038/ismej.2009.82

Bag S, Saha B, Mehta O, Anbumani D et al (2016) An improved method for high qualitymetagenomics DNA extraction from human and environmental samples. Sci Rep 6. https://doi.org/10.1038/srep26775

Beard CB, O'Neill SL, Tesh RB, Richards FF, Aksoy S (1993) Modification of arthropod vector competence via symbiotic bacteria. Parasitol Today 9(5):179–183

Beard CB, Durvasula RV, Richards FF (1998) Bacterial symbiosis in arthropods and the control of disease transmission. Emerg Infect Dis 4(4):581–591

Beard CB, Cordon-Rosales C, Durvasula RV (2002) Bacterial symbionts of the triatominae and their potential use in control of Chagas disease transmission. Annu Rev Entomol 47:123–141

Begon M, Harper JL, Townsend CR (1986) Ecology: individuals, populations and communities. Blackwell Scientific Publications, Oxford

Behar A, Yuval B, Jurkevitch E (2005) Enterobacteriamediated nitrogen fixation in natural populations of the fruit fly Ceratitis capitata. Mol Ecol 14(9):2637–2643

Berchtold M, Chatzinotas A, Schonhuber W, Brune A, Amann R, Hahn D, Konig H (1999) Differential enumeration and in situ localization of microorganisms in the hindgut of the

lower termite *Mastotermes darwiniensis* by hybridization with rRNA-targeted probes. Arch Microbiol 172:407–416
Bonning BC, Chougule NP (2014) Delivery of intrahemocoelic peptides for insect pest management. Trends Biotechnol 32(2):91–98
Breeuwer JAJ, Werren JH (1995) Hybrid breakdown between two haplo diploid species: the role of nuclear and cytoplasmic genes. Evolution 49(4):705–717
Breznak JA (2004) Invertebrates-insects. In: Bull AT (ed) Microbial diversity and bioprospecting. ASM Press, Washington, D.C., pp 191–203
Brinkmann N, Martens R, Tebbc CC (2008) Origin and diversity of metabolically active gut bacteria from laboratorybred larvae of *Manduca sexta* (Sphingidae, Lepidoptera, Insecta). Appl Environ Microbiol 74(23):7189–7196
Bromberg JS, Fricke WF, Brinkman CC, Simon T, Mongodin EF (2015) Microbiota: implications for immunity and transplantation. Nat Rev Nephrol 11(6):342–353
Brucker RM, Bordenstein SR (2013) The hologenomic basis of speciation: gut bacteria cause hybrid lethality in the genus Nasonia. Science 341(6146):667–669
Buchner P (1965) Endosymbioses of animals with plant microoorganisms. John Wiley and Sons, Chichester, UK
Campbell BC (1990) On the role of microbial symbiotes in herbivorous insects. In: Bernays EA (ed) Insect-plant interactions. CRC Press p, Boca Raton, FL, pp 1–44
Cariou M, Ribière C, Morlière S et al (2018) Comparing 16S rDNA amplicon sequencing and hybridization capture for pea aphid microbiota diversity analysis. BMC Res Notes 11:461. https://doi.org/10.1186/s13104-018-3559-3
Casu RE, Eisemann CH, Vuocolo T, Tellam RL (1996) The major excretory/secretory protease from *Lucilia cuprina* larvae is also a gut digestive protease. Int J Parasitol 26(6):623–628
Cazemier AE, Hackstein JHP, Op den Camp HJM, Rosenberg J, van der Drift C (1997) Bacteria in the intestinal tract of different species of arthropods. Microb Ecol 33:189–197
Chakravorty S, Helb D, Burday M, Connell N, Alland D (2007) A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. J Microbiol Methods 69(2):330–339
Chandler JA, Turelli M (2014) Comment on "the hologenomic basis of speciation: gut bacteria cause hybrid lethality in the genus Nasonia". Science 345(6200):1011
Charnley AK, Hunt J, Dillon RJ (1985) The germ-free culture of desert locusts, Schistocerca gregaria. J Insect Physiol 31(6):477–485
Chaves S, Neto M, Tenreiro R (2009) Insect-symbiont systems: from complex relationships to biotechnological applications. Biotechnol J 4(12):1753–1765
Chen K, Pachter L (2005) Bioinformatics for whole-genome shotgun sequencing of microbial communities. PLoS Comput Biol 1(2):e24. 106–112
Chen B, Yu T, Xie S, Du K, Liang X, Lan Y, Sun C, Lu X, Shao Y (2018) Comparative shotgun metagenomic data of the silkworm Bombyx mori gut microbiome. Sci Data 5:180285. https://doi.org/10.1038/sdata.2018.285
Cheung SW, Tishler PV, Atkins L, Sengupta SK, Modest EJ, Forget BG (1977) Gene mapping by fluorescent in situ hybridization. Cell Biol Int Rep 1(3):255–262
Chong CW, Pearce DA, Convey P, Yew WC, Tan IKP (2012) Patterns in the distribution of soil bacterial 16S rRNA gene sequences from different regions of Antarctica. Geoderma 181:45–55
Chougule NP, Giri AP, Sainani MN, Gupta VS (2005) Gene expression patterns of *Helicoverpa armigera* gut proteases. Insect Biochem Mol Biol 35(4):355–367
Cowan D, Meyer Q, Stafford W, Muyanga S, Cameron R, Wittwer P (2005) Metagenomic gene discovery: past, present and future. Trends Biotechnol 23(6):321–329
da Mota FF, Gomes EA, Paiva E, Seldin L (2005) Assessment of the diversity of *Paenibacillus* species in environmental samples by a novel rpoB-based PCR-DGGE method. FEMS Microbiol Ecol 53(2):317–328

Dandekar AM, Gouran H, Ibanez AM, Uratsu SL, Aguero CB, McFarland S, Borhani Y, Feldstein PA, Bruening G, Nascimento R et al (2012) An engineered innate immune defense protects grapevines from Pierce disease. Proc Natl Acad Sci U S A 109(10):3721–3725

Dar MA, Shaikh AA, Pawar KD, Pandit RS (2018) Exploring the gut of *Helicoverpa armigera* for cellulose degrading bacteria and evaluation of a potential strain for lignocellulosic biomass deconstruction. Process Biochem 73:142–153

de Vries EJ, Breeuwer JAJ, Jacobs G, Mollema C (2001a) The association of western flower thrips, *Frankliniellaoccidentalis*, with a near *Erwinia* species gut bacteria: transient or permanent? J Invertebr Pathol 77(2):120–128

de Vries EJ, Jacobs G, Breeuwer JAJ (2001b) Growth and transmission of gut bacteria in the western flower thrips, Frankliniella occidentalis. J Invertebr Pathol 77(2):129–137

Demaneche S, David MM, Navarro E, Simonet P, Vogel TM (2009) Evaluation of functional gene enrichment in a soil metagenomic clone library. J Microbiol Methods 76(1):105–107

Dethlefsen L, Huse S, Sogin ML, Relman DA (2008) The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. PLoS Biol 6(11):e280

Dillon RJ, Dillon VM (2004) The gut bacteria of insects: non-pathogenic interactions. Annu Rev Entomol 49:71–92

Dillon RJ, Vennard CT, Charnley AK (2002) A note: gut bacteria produce components of a locust cohesion pheromone. J Appl Microbiol 92:759–763

Dillon RJ, Webster G, Weightman AJ, Dillon VM, Blanford S, Charnley AK (2008) Composition of Acridid gut bacterial communities as revealed by 16S rRNA gene analysis. J Invertebr Pathol 97(3):265–272

Domingo JWS, Kaufman MG, Klug MJ, Tiedje JM (1998) Characterization of the cricket hindgut microbiota with fluorescently labeled rRNA-targeted oligonucleotide probes. Appl Environ Microbiol 64(2):752–755

Douglas AE (1989) Mycetocyte symbiosis in insects. Biol Rev 64(4):409–434

Douglas AE (2015) Multiorganismal insects: diversity and function of resident microorganisms. Annu Rev Entomol 60:17–34

Dowd PF, Shen SK (1990) The contribution of symbiotic yeast to toxin resistance of the cigarette beetle (Lasioderma serricorne). Entomologia Experimentalis et Applicata 56:241–248

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26 (19):2460–2461

Egert M, Stingl U, Bruun LD, Pommerenke B, Brune A, Friedrich MW (2005) Structure and topology of microbial communities in the major gut compartments of *Melolonthamelolontha* larvae (Coleoptera: Scarabaeidae). Appl Environ Microbiol 71(8):4556–4566

Engel P, Moran NA (2013) The gut microbiota of insects – diversity in structure and function. FEMS Microbiol Rev 37:699–735

Esumi H, Takahashi Y, Sato S, Sugimura T (1982) Restriction fragment length polymorphism of the rat albumin gene in Sprague-Dawley rats and its application in genetic study of analbuminemia. Nucleic Acids Res 10(14):4247–4257

Ezenwa VO, Gerardo NM, Inouye DW, Medina M, Xavier JB (2012) Microbiology: animal behaviorand the microbiome. Science 338:198–199

Falkowski PG, Fenchel T, Delong EF (2008) The microbial engines that drive Earth's biogeochemical cycles. Science 320(5879):1034–1039

Fall S, Hamelin J, Ndiaye F, Assigbetse K, Aragno M, Chotte JL, Brauman A (2007) Differences between bacterial communities in the gut of a soil-feeding termite. Appl Environ Microbiol 73 (16):5199–5208

Ferrari J, Vavre F (2011) Bacterial symbionts in insects or the story of communities affecting communities. Philos Trans R Soc Lond Ser B Biol Sci 366(1569):1389–1400. https://doi.org/10.1098/rstb.2010.0226

Finn RN, Chauvigné F, Stavang JA, Belles X, Cerda J (2015) Insect glycerol transporters evolved by functional co-option and gene replacement. Nat Commun 6:7814

Frago E, Dicke M, Godfray HC (2012) Insect symbionts as hidden players in insect-plant interactions. Trends Ecol Evol 27:705–711

Fukatsu T, Watanable K, Sekiguchi Y (1998) Specific detection of intracellular symbiotic bacteria of aphids by oligonucleotide-probed in situ hybridization. Appl Entomol Zool 33(3):461–472

Fuller CW, Middendorf LR, Benner SA, Church GM, Harris T, Huang X, Jovanovich SB, Nelson JR, Schloss JA, Schwartz DC, Vezenov DV (2009) The challenges of sequencing by synthesis. Nat Biotechnol 27(11):1013–1023. https://doi.org/10.1038/nbt.1585

Gauthier JP, Outreman Y, Mieuzet L, Simon JC (2015) Bacterial communities associated with host-adapted populations of pea aphids revealed by deep sequencing of 16S ribosomal DNA. PLoS ONE 10(3):e0120664. https://doi.org/10.1371/journal.pone.0120664

Gavriel S, Jurkevitch E, Gazit Y, Yuval B (2011) Bacterially enriched diet improves sexual performance of sterile male Mediterranean fruit flies. J Appl Entomol 135:564–573

Gill R, Latorre A, Moya A (2004) Bacterial endosymbionts of insects: insights fromcomparative genomics. Environ Microbiol 6(11):1109–1122

Goodwin S, McPherson JD, McCombie R (2016) Coming of age: tenyears of next-generation sequencing technologies. Nat Rev Genet 17:333–351. https://doi.org/10.1038/nrg.2016.49

Gosh A, Mehta A, Khan AM (2018) Metagenomic analysis and its application. Encyclopedia Bioinform Comput Biol:1–9. https://doi.org/10.1016/B978-0-12-809633-8.20178-7

Green JL, Bohannan BJ, Whitaker RJ (2008) Microbial biogeography: from taxonomy to traits. Science 320:1039–1043

Grieco MB, Lopes FAC, Oliveira LS (2019) Metagenomic analysis of the whole gut microbiota in Brazilian Termitidae termites *Cornitermes cumulans, Cyrilliotermes strictinasus,Syntermes dirus, Nasutitermes jaraguae, Nasutitermes aquilinus,Grigiotermes bequaerti*, and *Orthognathotermes mirim*. Curr Microbiol 76(6):687–697

Guan CH, Ju JJ, Borlee BR, Williamson LL, Shen B, Raffa KF, Handelsman J (2007) Signal mimics derived from a metagenomic analysis of the gypsy moth gut microbiota. Appl Environ Microbiol 73(11):3669–3676

Gupta A, Nair S (2020) Dynamics of insect-microbiome interaction influence host and microbial symbiont. Front Microbiol 11:1357

Hajeri S, Killiny N, El-Mohtar C, Dawson WO, Gowda S (2014) Citrus tristeza virus-based RNAi in citrus plants induces gene silencing in *Diaphorina citri*, a phloem-sap sucking insect vector of citrus greening disease (Huanglongbing). J Biotechnol 176:42–49

Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem Biol 5 (10):45–49

Handelsman J (2004) Metagenomics: application of genomics to uncultured microorganisms. Microbiol Mol Biol Rev MMBR 68(4):669–685. https://doi.org/10.1128/MMBR.68.4.669-685.2004

Hansen AK, Moran NA (2014) The impact of microbial symbionts on host plant utilization by herbivorous insects. Mol Ecol 23:1473–1496

Harada H, Ishikawa H (1993) Gut microbe of aphid closely-related to its intracellular symbiont. Biosystems 31(2-3):185–191

Harish ER, ManiChellappan K, Makeshkumar T, Mathew D, Ranjith MT, Girija D (2019) Next-generation sequencing reveals endosymbiont variability in cassava whitefly, Bemisia tabaci, across the agro-ecological zones of Kerala, India. Genome 62(9):571–584

Hayashi A, Aoyagi H, Yoshimura T, Tanaka H (2007) Development of novel method for screening microorganisms using symbiotic association between insect (Coptotermes formosanus Shiraki) and intestinal microorganisms. J Biosci Bioeng 103(4):358–367

He S, Ivanova N, Kirton E, Allgaier M, Bergin C, Scheffrahn RH et al (2013) Comparative metagenomic and Metatranscriptomic analysis of hindgut paunch microbiota in wood- and dung-feeding higher termites. PLoS One 8(4):e61126

Head IM, Saunders JR, Pickup RW (1998) Microbial evolution, diversity, and ecology:Adecade of ribosomal RNA analysis of uncultivated microorganisms. Microb Ecol 35:1–21

Head SR, Komori HK, LaMere SA, Whisenant T, Van Nieuwerburgh F, Salomon DR, Ordoukhanian P (2014) Library construction for nextgeneration sequencing: overviews and challenges. Biotech 56:61–64., n 66, 68, passim. https://doi.org/10.2144/000114133

Hilton SK, Castro-Nallar E, Perez-Losada M et al (2016) Metataxonomic and metagenomic approaches vs. culture-based techniques for clinical pathology. Front Microbiol 7:484. https://doi.org/10.3389/fmicb.2016.00484

Hirsch J, Strohmeier S, Pfannkuchen M, Reineke A (2012a) Assessment of bacterial endosymbiont diversity in Otiorhynchus spp. (Coleoptera: Curculionidae) larvae using a multitag 454 pyrosequencing approach. BMC Microbiol 12(Suppl 1):S6

Hirsch J, Strohmeier S, Pfannkuchen M et al (2012b) Assessment of bacterial endosymbiont diversity in *Otiorhynchus*spp. (Coleoptera: Curculionidae) larvae using a multitag 454 pyrosequencing approach. BMC Microbiol 12:S6. https://doi.org/10.1186/1471-2180-12-S1-S6

Hopkins DW, Macnaughton SJ, O'Donnell AG (1991) A dispersion and differential centrifugation technique for representatively sampling microorganisms from soil. Soil Biol Biochem 23 (3):217–225

Hosokawa T, Kikuchi Y, Meng XY, Fukatsu T (2005) The making of symbiont capsule in the plataspid stinkbug Megacoptan punctatissima. FEMS Microbiol Ecol 54(3):471–477

Hosokawa T, Kikuchi Y, Shimada M, Fukatsu T (2007) Obligate symbiont involved in pest status of host insect. Proc Biol Sci 274(1621):1979–1984

Hurst MRH, Jackson TA (2002) Use of the green fluorescent protein to monitor the fate of *Serratia entomophila* causing amber disease in the New Zealand grass grub, *Costelytrazealandica*. J Microbiol Methods 50(1):1–8

Husseneder C, Donaldson JR, Foil LD (2016) Genetically engineered yeast expressing a lytic peptide from bee venom (Melittin) kills symbiotic protozoa in the gut of Formosan subterranean termites. PLoS One 11(3):e0151675

Indiragandhi P, Yoon C, Yang JO et al (2010) Microbial communities in the developmental stages of B and Q biotypes of sweetpotato whitefly, Bemisia tabaci (hemiptera: Aleyrodidae). J Korean Soc Appl Biol Chem 53:605–617. https://doi.org/10.3839/jksabc.2010.093

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

Jaenicke S, Ander C, Bekel T (2011) Comparative and joint analysis of two metagenomic data sets from a biogas fermenter obtained by 454-pyrosequencing. PLoS One 6(1):e14519

Jeffers LA, Roe RM (2008) The movement of proteins across the insect and tick digestive system. J Insect Physiol 54(2):319–332

Jing TZ, Qi FH, Wang ZY (2020) Most dominant roles of insect gut bacteria: digestion, detoxification, or essential nutrient provision? Microbiome 8:38

Jones AG, Mason CJ, Felton GW, Hoover K (2019) Host plant and population source drive diversity of microbial gut communities in two polyphagous insects. Sci Rep 9:2792. https://doi.org/10.1038/s41598-019-39163-9

Josefsen MH, Andersen SC, Christensen J, Hoorfar J (2015) Microbial food safety: potential of DNA extraction methods for use in diagnostic metagenomics. J Microbiol Methods 114:30–34. https://doi.org/10.1016/j.mimet.2015.04.016

Kane MD, Pierce NE (1994) Diversity within diversity: molecular approaches to studying microbial interactions withinsects. In: Schierwater B, Streit B, Wagner GP, DeSalle R (eds) Molecular methods in ecology and evolution. Birkhauser Verlag, Berlin, Germany, pp 509–524

Kauppinen J, Mantyjarvi R, Katil ML (1999) Mycobacterium malmoense-specific nested PCR based on a conserved sequence detected in random amplified polymorphic DNA fingerprints. J Clin Microbiol 37(5):1454–1458

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

Koch H, Schmid-Hempel P (2012) Gut microbiota instead of host genotype drives the specificity in the interaction of a natural host-parasite system. Ecol Lett 15:1095–1103
Kodama R, Nakasuji Y (1971) Further studies on the pathogenic mechanism of bacterial diseases in gnotobiotic silkworm larvae. IFO Res Commun 5:1–9
Kohler T, Stingl U, Meuser K, Brune A (2008) Novel lineages of Planctomycetes densely colonize the alkaline gut of soil-feeding termites (*Cubitermes* spp.). Environ Microbiol 10(5):1260–1270
Krishnan M, Bharathiraja C, Pandiarajan J, Prasanna VA, Rajendhran J, Gunasekaran P (2014) Insect gut microbiome: an unexploited reserve for biotechnological application. Asian Pacific J Trop Biomed 4(1):516–522
Lance DR, McInnis DO, Rendon P, Jackson CG (2000) Courtship among sterile and wild Ceratitis capitata (Diptera: Tephritidae) in field cages in Hawaii and Guatemala. Ann Entomol Soc Am 93:1179–1185
Langille MG, Zaneveld J, Caporaso JG (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol 31:814–821
Lee DH, Zo YG, Kin SJ (1996) Nonradioactive method to study genetic profiles of natural bacterial communities by PCR-single-strand-conformation polymorphism. Appl Environ Microbiol 62 (9):3112–3120
Leung HC, Yiu SM, Yang B (2011) A robust and accurate binning algorithm for metagenomic sequences with arbitrary species abundance ratio. Bioinformatics 27(11):1489–1495
Li L, Frohlich J, Pfeiffer P, Konig H (2003) Termite gut symbiotic archaezoa are becoming living metabolic fossils. Eukaryot Cell 2(5):1091–1098
Li ZT, Hopkins DL, Gray DJ (2015) Overexpression of antimicrobial lytic peptides protects grapevine from Pierce's disease under greenhouse but not field conditions. Transgenic Res 24 (5):821–836
Lilburn TG, Schmidt TM, Breznak JA (1999) Phylogenetic diversity of termite gut spirochaetes. Environ Microbiol 1(4):331–345
Liu WT, Marsh TL, Cheng H, Forney LJ (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl Environ Microbiol 63(11):4516–4522
Luo R, Liu B, Xie Y (2012) SOAP denovo2: an empirically improved memory-efficient short-readdenovoassembler. Gigascience 1(1):18
Malacrinò A, Rassati D, Schena L, Mehzabin R, Battisti A, Palmeri V (2017) Fungal communities associated with bark and ambrosia beetles trapped at international harbours. Fungal Ecol 28:44–52. https://doi.org/10.1016/j.funeco.2017.04.007
Malathi VM, More RP, Anandham R, Gracy GR, Mohan M, Venketesan T, Samaddar S, Jalali SK, Tongmin (2018) Gut bacterial diversity of insecticide-susceptible and –resistant nymphs of the Brown Planthopper Nilaparvata lugens Stål (Hemiptera: Delphacidae) and elucidation of their putative functional roles. J Microbiol Biotechnol 28(6):976–986
Manefield M, Whiteley AS, Griffiths RI, Bailey MJ (2002) RNA stable isotope probing, a novel means of linking microbial community function to phylogeny. Appl Environ Microbiol 68 (11):5367–5373
Mardis ER (2008) The impact of next-generation sequencing technology on genetics. Trends Genet 24(3):133–141
Margulies M, Egholm M, Altman WE et al (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437(7057):376–380
Marusina K (2006) The next generation of DNA sequencing. Genet Eng News 26:28–31
Mayo B, Rachid CTCC, Alegría A, Leite AMO, Peixoto RS, Delgado S (2014) Impact of next generation sequencing techniques in food microbiology. Curr Genomics 15:293–309. https://doi.org/10.2174/1389202915666140616233211
McKillip JL, Small CL, Brown JL, Brunner JF, Spence KD (1997) Sporogenous midgut bacteria of the leafroller, *Pandemis pyrusana* (Lepidoptera: Tortricidae). Environ Entomol 26 (6):1475–1481

Metzker ML (2010) Sequencing technologies- the next generation. Nat Rev Genet 11:31–46. https://doi.org/10.1038/nrg2626
Meyer M, Kircher M (2010) Illumina sequencing library preparation forn highlymultiplexed target capture and sequencing. Cold Spring Harb Protoc. https://doi.org/10.1101/pdb.prot5448
Mohr KI, Tebbe CC (2006) Diversity and phylotype consistency of bacteria in the guts of three bee species (*Apoidea*) at an oilseed rape field. Environ Microbiol 8(2):258–272
Moran NA, McCutcheon JP, Nakabachi A (2008) Genomics and evolution of heritable bacterial symbionts. Annu Rev Genet 42:165–190
Morgan XC, Huttenhower C (2012) Human microbiome analysis. PLoS Comput Biol 8(12): e1002808
Muyzer G, de Waal EC, Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel-electrophoresis analysis of polymerase chain reactionamplified genes coding for 16s rRNA. Appl Environ Microbiol 59(3):695–700
Nagarajan N, Cook C, DiBonaventura M (2010) Finishing genomes with limited resources: lessons from an ensemble of microbial genomes. BMC Genomics 11(1):242
Nakashima K, Watanabe H, Saitoh H, Tokuda G, Azuma JI (2002) Dual cellulose-digesting system of the wood-feeding termite, Coptotermes formosanus Shiraki. Insect Biochem Mol Biol 32 (7):777–784
Noda S, Ohkuma M, Usami R, Horikoshi K, Kudo T (1999) Culture-independent characterization of a gene responsible for nitrogen fixation in the symbiotic microbial community in the gut of the termite *Neotermes koshunensis*. Appl Environ Microbiol 65(11):4935–4942
Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017) metaSPAdes:A new versatile metagenomica ssembler. Genome Res 27(5):824–834
Ohkuma M, Kudo T (1996) Phylogenetic diversity of theintestinal bacterial community in the termite *Reticulitermessperatus*. Appl Environ Microbiol 62(2):461–468
Ondov BD, Bergman NH, Phillippy AM (2011) Interactive metagenomic visualization in a web browser. BMC Bioinform 12:385
Oulas A, Pavloudi C, Polymenakou P (2015) Metagenomics:tools and insights for analysing next-generation sequencing data derived from biodiversity studies. Bioinform Biol Insights 9(9):75
Pace NR (1997) A molecular view of microbial diversity and the biosphere. Science 276 (5313):734–740
Pace NR, Stahl DA, Lane DJ, Olsen GJ (1986) The analysis of natural microbial populations by ribosomal RNA sequences. Adv Microb Ecol 9:1–55
Paine TD, Raffa KF, Harrington TC (1997) Interactions among Scolytid bark beetles, their associated fungi, and live host conifers. Annu Rev Entomol 42:179–206
Pittman GW, Brumbley SM, Allsopp PG, O'Neill SL (2008) "Endomicrobia" and other bacteria associated with the hindgut of *Dermolepida albohirtum* larvae. Appl Environ Microbiol 74 (3):762–767
Plummer E, Twin J, Bulach DM, Garland SM, Tabrizi SN (2015) A comparison of three bioinformatics pipelines for the analysis of preterm gut microbiota using 16S rRNA gene sequencing data. J Proteomics Bioinformatics 8(12):283
Poland TM, Rassati D (2018) Improved biosecurity surveillance of non-native forest insects: a review of current methods. J Pest Sci:1–13. https://doi.org/10.1007/s10340-018-1004-y
Priya NG, Ojha H, Kajla MK, Raj A, Rajagopal R (2012) Host plant induced variationin gut bacteria of *Helicoverpa armigera*. PLoS One 7:e30768
Quail MA, Smith M, Coupland P et al (2012) A tale of three next generation sequencing platforms: comparison of IonTorrent, Pacific biosciences and Illumina MiSeq sequencers. BMC Genomics 13:341. https://doi.org/10.1186/1471-2164-13-341
Quince C, Walker AW, Simpson JT, Loman NJ, Segata N (2017) Shotgun metagenomics, from sampling to analysis. Nat Biotechnol 35:833–844. https://doi.org/10.1038/nbt.3935
Rajagopal R (2009) Beneficial interactions between insects and gut bacteria. Indian J Microbiol 49: 114–119

Ranjith MT, Chellappan M, Harish ER, Girija D, Nazeem PA (2016) Bacterial communities associated with the gut of tomato fruit borer, *Helicoverpa armigera* (Hubner) (Lepidoptera: Noctuidae) based on Illumina next-generation sequencing. J Asia Pac Entomol 19(2):333–340

Rassati D, Faccoli M, Petrucco Toffolo E, Battisti A, Marini L (2015) Improving the early detection of alien wood-boring beetles in ports and surrounding forests. J Appl Ecol 52:50–58. https://doi.org/10.1111/1365-2664.123

Rassati D, Faccoli M, Marini L, Haack RA, Battisti A, Petrucco Toffolo E (2018) Exploring the role of wood waste landfills in early detection of non-native wood-boring beetles. J Pest Sci 88:563–572. https://doi.org/10.1007/s10340-014-0639-6

Raymond B, Lijek RS, Griffiths RI, Bonsall MB (2008) Ecological consequences of ingestion of *Bacilluscereus* on Bacillus thuringiensis infections and on the gut flora of a lepidopteran host. J Invertebr Pathol 99(1):103–111

Reeson AF, Jankovic T, Kasper ML, Rogers S, Austin AD (2003) Application of 16S rDNA-DGGE to examine the microbial ecology associated with a social wasp *Vespulagermanica*. Insect Mol Biol 12(1):85–91

Robe P, Nalin R, Capellano C, Vogel TM, Simonet P (2003) Extraction of DNA from soil. Eur J Soil Biol 39:183–190

Roe AD, Torson AS, Bilodeau G, Bilodeau P, Blackburn GS, Cui M, Cusson M, Doucet D, Griess VC, Lafond V et al (2018) Biosurveillance of forest insects: part I—integration and application of genomic tools to the surveillance of non-native forest insects. J Pest Sci:1–20. https://doi.org/10.1007/s10340-018-1027-4

Roose-Amsaleg CL, Garnier-Sillam E, Harry M (2001) Extraction and purification of microbial DNA from soil and sediment samples. Appl Soil Ecol 18(1):47–60

Rossmassler K, Dietrich C, Thompson C et al (2015) Metagenomic analysis of the microbiota in the highly compartmented hindguts of six wood- or soil-feeding higher termites. Microbiome 3:56. https://doi.org/10.1186/s40168-015-0118-1

Roszak DB, Grimes DJ, Colwell RR (1984) Viable but nonrecoverable stage of *salmonella enteritidis* in aquatic systems. Can J Microbiol 30:334–338

Roussel EG, Bonavita MAC, Querellou J, Cragg BA, Webster G, Prieur D, Parkes RJ (2008) Extending the sub-sea-floor biosphere. Science 320(5879):1046–1046

Russell JA, Weldon S, Smith AH, Kim KL, Hu Y, Łukasik P, Oliver KM (2013) Uncovering symbiont-driven genetic diversity across north American pea aphids. Mol Ecol 22 (7):2045–2059

Sanger F, Nicklen S, Coulson AR (1977a) DNAsequencing with chain- terminating inhibitors. Proc Natl Acad Sci U S A 74(12):5463–5467. https://doi.org/10.1073/pnas.74.12.5463

Sanger F, Nicklen S, Coulson AR (1977b) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci U S A 74:5463–5467

Santo Domingo JW, Kaufman MG, Klug MJ, Holben WE, Harris D, Tiedje JM (1998) Influence of diet on the structure and function of the bacterial hindgut community of crickets. Mol Ecol 7 (6):761–767

Schierbeek A (1959) Measuring the invisible world: the life and works of Antoni van Leeuwenhoek. Abelard-Schuman, London. Available online at: https://www.questia.com/library/73684/measuring-the-invisible-world-thelife-and-works (Accessed Sep 12, 2020)

Schmitt Wagner D, Friedrich MW, Wagner, B.& Brune, A. (2003) Phylogenetic diversity, abundance, and axial distribution of bacteria in the intestinal tract of two soil-feeding termites (*Cubitermes* spp.). Appl Environ Microbiol 69(10):6007–6017

Scott JG, Michel K, Bartholomay LC, Siegfried BD, Hunter WB, Smagghe G, Zhu KY, Douglas AE (2013) Towards the elements of successful insect RNAi. J Insect Physiol 59(12):1212–1221

Scully ED, Geib SM, Hoover K, Tien M, Tringe SG et al (2013) Metagenomic profiling reveals lignocellulose degrading system in a microbial community associated with a wood-feeding beetle. PLoS One 8(9):e73827. https://doi.org/10.1371/journal.pone.0073827

Shah SHJ, Malik AH, Zhang B, Bao Y, Qazi J (2020) Metagenomic analysis of relative abundance and diversity of bacterial microbiota in Bemisia tabaci infesting cotton crop in Pakistan. Infect Genet Evol 84:104381

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

Sharon G, Segal D, Zilber-Rosenberg I, Rosenberg E (2011) Symbiotic bacteria are responsible for diet-induced mating preference in Drosophila melanogaster, providing support for the hologenome concept of evolution. Gut Microbes 2:190–192

Sharpton TJ (2014) An introduction to the analysis of shotgun metagenomic data. Front Plant Sci 5: 209. https://doi.org/10.3389/fpls.2014.00209

Shen Z, Jacobs-Lorena M (1997) Characterization of a novel gut-specific chitinase gene from the human malaria vector Anopheles gambiae. J Biol Chem 272:28895–28900

Shendure J, Ji H (2008) Next-generationDNAsequencing. Nat Biotechnol 26:1135–1145

Shinzato N, Muramatsu M, Matsui T, Watanabe Y (2005) Molecular phylogenetic diversity of the bacterial community in the gut of the termite Coptotermes formosanus. Biosci Biotechnol Biochem 69(6):1145–1155

Shinzato N, Muramatsu M, Matsui T, Watanabe Y (2007) Phylogenetic analysis of the gut bacterialmicroflora of the fungus-growing termite *Odontotermes formosanus*. Biosci Biotechnol Biochem 71:906–915

Shokralla S, Spall JL, Gibson JF, Hajibabaei M (2012) Nextgenerationsequencing technologies for environmental DNA research. Mol Ecol 21(8):1794–1805. https://doi.org/10.1111/j.1365-294X.2012.05538.x

Sinclair L, Osman OA, Bertilsson S, Eiler A (2015) Microbial community composition and diversity via 16S rRNA gene amplicons: evaluating the illumina platform. PLoS One 10(2): e0116955

Slaytor M (1992) Cellulose digestion in termites and cockroaches: what role do symbionts play? Comp Biochem Physiol B: Comp Biochem 103(4):775–784

Sleator RD, Shortall C, Hill C (2008) Under the microscope metagenomics. Lett Appl Microbiol 47: 361–366

Smalla K (2004) Culture-independent microbiology. In: Bull AT (ed) Microbial diversity and bioprospecting. ASM Press, Washington, D.C., pp 88–99

Stahl DA, Lane GJ, Pace NR (1984) Analysis of hydrothermal vent-associated symbionts by ribosomal RNA sequences. Science 224(4647):409–411

Staley JT, Konopka A (1985) Measurement of *in situ* activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. Annu Rev Microbiol 39:321–346

Stangier KA (2009) Next-generation sequencing: a short comparison. LaboratoriumsMedizin-J Lab Med 33:153–157

Steele HL, Streit WR (2005) Metagenomics: advances in ecology and biotechnology. FEMS Microbiol Lett 247(2):105–111

Stokes HW, Holmes AJ, Nield BS, Holley MP, Nevalaine KMH, Mabbutt BC, Gillings MR (2001) Gene cassette PCR: sequence-independent recovery of entire genes from environmental DNA. Appl Environ Microbiol 67(11):5240–5246

Tikhe CV, Martin TM, Howells A, Delatte J, Husseneder C (2016) Assessment of genetically engineered Trabulsiella odontotermitis as a 'Trojan Horse' for paratransgenesis in termites. BMC Microbiol 16:202. https://doi.org/10.1186/s12866-016-0822-4

Todaka MS, Saita K, Hondo T, Kiuchi I, Takasu H, Ohkuma M, Piero C, Hayashizaki Y, Kudo T (2007) Environmental cDNA analysis of the genes involved in lignocellulose digestion in the symbiotic protist community of *Reticulitermes speratus*. FEMS Microbiol Ecol 59(3):592–599

Tokuda G, Watanabe H (2007) Hidden cellulases in termites: revision of an old hypothesis. Biol Lett 3(3):336–339

Tringe SG, VonMering C, Kobayashi A et al (2005) Comparative metagenomics of microbial communities. Science 308(5721):554–557

Turroni F, Ribbera A, Foroni E, van Sinderen D, Ventura M (2008) Human gut microbiota and bifidobacteria: from composition to functionality. Antonie Van Leeuwenhoek 94(1):35–50

Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM, Solovyev VV, Rubin EM, Rokhsar DS, Banfield JF (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature 428(6978):37–43

Van Djick EL, Auger H, Jaszczyszyn Y, Thermes C (2014) Ten years of next-generation sequencing technology. Trends Genet 30:418–426. https://doi.org/10.1016/j.tig.2014.07.001

Van Nieuwerburgh F, Thompson RC, Ledesma J, Deforce D, Gaasterland T, Ordoukhanian P, Head SR (2012) Illumina mate-paired DNA sequencing-library preparation using Cre-lox recombination. Nucleic Acids Res 40(3). https://doi.org/10.1093/nar/gkr1000

Vasanthakumar A, Delalibera I, Handelsman J, Klepzig KD, Schloss PD, Raffa KF (2006) Characterization of gut-associated bacteria in larvae and adults of the southern pine beetle, Dendroctonus frontalis Zimmermann. Environ Entomol 35(6):1710–1717

Venter JC, Remington K, Heidelberg JF et al (2004a) Environmental genome shotgun sequencing of the Sargasso Sea. Science 304(5667):66–74

Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D et al (2004b) Environmental genome shotgun sequencing of the Sargasso Sea. Science 304:66–74

Vieites JM, Guazzaroni ME, Beloqui A, Golyshin PN, Ferrer M (2008) Metagenomics approaches in systems microbiology. FEMS Microbiol Rev 33(1):236–255

Wang H, Lei T, Xia W et al (2019) Insight into the microbial world of Bemisia tabaci cryptic species complex and its relationships with its host. Sci Rep 9:6568. https://doi.org/10.1038/s41598-019-42793-8

Warnecke F, Luginbuhl P, Ivanova N et al (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. Nature 450(7169):560–565

Wernegreen JJ (2002) Genome evolution in bacterial endosymbionts of insects. Nat Rev Genetics 3 (11):850–861

Williamson LL, Borlee BR, Schloss PD, Guan CH, Allen HK, Handelsman J (2005) Intracellular screen to identify metagenomic clones that induce or inhibit a quorumsensing biosensor. Appl Environ Microbiol 71(10):6335–6344

Weinstock GM (2012) Genomic approaches to study human microbiota. Nature 489:250–256

Woese CR, Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci U S A 74(11):5088–5090. https://doi.org/10.1073/pnas.74.11.5088

Wu D, Daugherty SC, Van Aken SE, Pai GH, Watkins KL, Khouri H, Tallon LJ, Zaborsky JM, Dunbar HE, Tran PL et al (2006) Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters. PLoS Biol 4(6):e188

Xia X, Zheng D, Zhong H, Qin B, Gurr GM, Vasseur L et al (2013) DNA sequencing reveals the midgut microbiota of diamondback moth, *Plutella xylostella* (L.) and a possible relationship with insecticide resistance. PLoS ONE 8(7): e68852. https://doi.org/10.1371/journal.pone.0068852

Xia X, Gurr GM, Vasseur L et al (2017) Metagenomic sequencing of diamondback moth gut microbiome unveils key Holobiont adaptations for herbivory. Front Microbiol 8:663. https://doi.org/10.3389/fmicb.2017.00663

Yandell DW (1991) SSCP (single strand conformation polymorphism) analysis of PCR fragments. Jpn J Cancer Res 82:1468–1469

Yuan JS, Galbraith DW, Dai SY, Griffin P, Stewart CN (2008) Plant systems biology comes of age. Trends Plant Sci 13(4):165–171

Zhou J, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. Appl Environ Microbiol 62(2):316–322

Zhou XG, Smith JA, Oi FM, Koehler PG, Bennett GW, Scharf ME (2007) Correlation of cellulase gene expression and cellulolytic activity throughout the gut of the termite Reticulitermes flavipes. Gene 395(1-2):29–39