COST Action FA1105. Vegetable diseases diagnostic tools and control methods under greenhouse organic farming. Practical training.

CEBAS-CSIC

TRAINING SCHOOL

Vegetable diseases diagnostic tools and control methods under greenhouse organic farming. Practical training.

Rapid methods for plant pathogen diagnosis based on molecular approaches

Josefa Blaya, Jose Antonio Pascual, Margarita Ros*

Grupo de Enzimología y Biorremediación de Suelos y Residuos Orgánicos
Campus Universitario de Espinardo, Espinardo, 30100 Murcia (Spain)
***margaros@cebas.csic.es**

1. Techniques to study soil-borne plant pathogens and the microbial communities in composts

Today, a vast number of techniques are used to improve plant disease management. We are going to focus in the techniques which regard the detection and quantification of soil-borne plant pathogens as well as the diversity of their population.

1.1. Detection and quantification methods for soil-borne plant pathogens

Plant pathogenic fungi and oomycetes cause serious and economically important plant diseases around the world including horticultural, ornamental and fruit tree species. Rapid detection and accurate identification of plant pathogens is one of the most important strategies for controlling plant diseases as is required for taking appropriate disease management measures. An early detection of the pathogen, even before of the onset of the symptoms, is of special interest in seeds, nursery plants and propagative plant material to avoid the introduction and further spreading of new pathogens in growing areas (Lievens *et al.*, 2008; Capote *et al.*, 2012).

Different approaches, from classical methods, traditionally used in microbiology, to the molecular methods, which have gained more importance in the last decades, are available for diagnosis and detection of soil-borne fungal and oomycete pathogens in environmental samples.

1.1.1. Classical methods

Traditional techniques are often based on identification of disease symptoms, isolation and culturing of environmental organisms, and subsequent laboratory identification by morphology and biochemical tests (Atkins & Clark, 2004).

Identification found on observation of disease symptoms has several disadvantages since different pathogens which affect the same crops may cause similar or identical symptoms (wilt, blight, root rot) leading to mistaken diagnosis or even sometimes, specific strains are needed to distinguish among similar structures. In this sense, **microscopic techniques** have been widely used for the identification of soil-borne pathogens based on morphological characteristics.

The serial **dilution plate technique** in selective media not only allows the identification of the pathogen, but also the quantification of the colony forming units presented per gram of soil or culture substrate. The use of selective media is based on the ability of certain microorganisms to use specific substrates or their ability to grow under the presence of certain antibiotics and fungicides. For instance, the culture media Komada (Komada, 1976) has been used as a *Fusarium*-selective medium and reported to distinguish between *F. oxysporum*, *F. solani*, *F. moniliforme*, and *F. roseum* by the color of the colonies. In the case of detection of *Phytophthora* spp., different culture media are available, such as V8 juice agar, pea agar or corn meal agar amended with antibiotics against bacterial growth (penicillin,

polymixin B, rifampicin, pentachloronitrobenzene (PCNB), ampicillin, vancomycin) or with antifungal effect (nystatin and pimaricin) (Tsao & Ocana, 1969).

Sometimes isolation of phytopathogens from soils remains difficult. Baiting techniques may be used for isolation of *Phytophthora* and *Pythium* species directly from infected soils (Tsao & Ocana, 1969). Once the pure-culture is ready, the identification is based primarily on the shape of the sporangia, mycelium and in the case of sexual fungi, the morphological features of the sexual structures. Other criteria widely used to distinguish species are cardinal growth temperature, growth rate, morphological (growth) characteristics in culture, and mating behavior. However, a reliable and accurate identification within species level is often complicated due to intraspecific variation and overlapping characters, even for specialists (Ippolito *et al.*, 2002; Darine *et al.*, 2007). These techniques are time and labour-consuming, preclude the handling of large number of samples and require extensive knowledge of classical taxonomy. Besides, not all the microorganisms are considered to be cultured. It is thought that only the 1 % is able to grow under laboratory conditions (Hugenholtz, 2002).

Chemical and **physiological methods** have also been used to quantify the fungal biomass in soil. The chemical methods are based on the measure of different fungal compounds such as chitin and ergosterol (Djajakirana *et al.*, 1996; Ekblad *et al.*, 1997). The physiological ones rely on respiration rate measures after the addition of different substrates, where the

CO_2 produced or the O_2 consumed is determined (Weaver *et al.*, 1994). **Immunological techniques** have been employed for detection, differentitation and quantification of fungal pathogens rapidly and are based on recognition, by antibodies, of specific antigens either present on the surface of the pathogens or secreted by them (McCartney *et al.*, 2003). Most of them use the enzyme-linked immunosorbent assay (ELISA) (Clark & Adams, 1977), although new formats are available. However, the problems associated with the production of specific antibodies that may be effective (difficulty and costs) limits the use of these techniques (Narayanasamy, 2011).

1.1.2. Molecular tools

In the last decades polymerase chain reaction (PCR) has emerged as a powerful tool for the identification and study of phytopathogenic fungi and oomycetes, contributing to solve some problems related to the detection, control and containment of plant pathogens (Schena *et al.*, 2004). PCR-based detection methods are characterized for its sensitivity, selectivity, robustness, rapidity and ease of methodology (Justé *et al.*, 2008). Moreover these techniques can overcome some of the shortcoming of the traditional ones, since generally it is not necessary the culturing step of the pathogen from the infected material, reducing the diagnosis time from weeks to hours and allowing the detection and identification of non-culturable pathogens (Capote et al; 2012; Schena *et al.*, 2013). The PCR consists in the repetitive amplification (depending on the number of cycles) of a specific fragment of

the genome after the extraction of DNA from the target microorganism. Ideally, DNA extraction protocol should enable to obtain a good quality DNA with a low concentration of substances inhibiting PCR reactions (Schena *et al.*, 2013). Different methodologies have been development for obtaining the nucleic acids (DNA) from plant material, fungi, oomycetes or soil/compost (Doyle & Doyle, 1987; van Burick *et al*., 1998; Robe *et al*., 2003). Most of them are based on a first step, consisting on the physical disruption (grinding samples with liquid nitrogen or bead beater) of microbial cell walls to release nucleic acids, followed by separation of nucleic acids from the soil or tissue particles and their extraction with one or more organic solvents (mainly phenol and chloroform), then concentrated by alcohol/salt precipitation according to standard procedures (Sambrook & Russell, 2011).

The presence of natural compounds such as polysaccharides, tannins and phenolics compounds in plant tissue samples and fulvic and humic compounds in soil/compost samples, may affect PCR efficiency (Wilson *et al.*, 1997). For this reason, these compounds need to be removed during DNA extraction or dilute the samples to reduce further inhibition during the PCR reaction. Nowadays, the nucleic acid extraction protocols are almost all kit-based (commercial), so that a high variety of commercial kits either general or specifically designed for plant material, fungi/oomycetes or soil are available in the market. Examples are: Ultraclean kit (MOBIO), FastDNA Spin kit for soil (MP Biomedicals), DNeasy Plant Kit (Qiagen Sciences, USA), Nucleon Phytopure (Amersham Biosciences Europe GmbH,

Germany) (Capote *et al.*, 2012). Major advantages of commercial kits are their simplicity and rapidity together with the absence of harmful chemical compounds.

The identification of appropriate target DNA regions and the subsequent development of specific PCR primers to target organisms is a crucial step in PCR assay development. A good target gene should be sufficiently variable to enable the differentiation of closely related species but, at the same time, should not contain intraspecific variation that would jeopardize the detection of all strains (Schena *et al.*, 2013). Molecular studies with fungi and oomycetes have largely concentrated on the ribosomal RNA gene cluster. This gene cluster consists of multiple copies (up to 200 copies per haploid genome) arranged in tandem repeats comprising three ribosomal RNA subunit genes (18S small subunit, 5.8S and the 28S large subunit) separated by internally transcribed spacers (ITS1 and ITS2), and intergenic spacers (IGS), the spacer between the small and large subunit (Bridge & Spooner, 2001). The nuclear-encoded ribosomal RNA genes (rDNA) with conserved as well as variable sequences can be used to develop broad specificity PCR primers (Bruns *et al.*, 1990; White *et al.*, 1990). The ITS regions of ribosomal RNA genes are the most commonly regions used for designing PCR diagnostic assays because molecular methods based on this target are potentially very sensitive (Tooley *et al.*, 2006; Schena *et al.*, 2013). After PCR amplification, molecular identification of plant pathogens is

accomplished by direct sequencing and BLAST searching in GenBank or other databases (White *et al.,* 1990).

Real-time PCR (qPCR)

Not only detection but also quantification of pathogens is becoming more and more significant. Accurate quantification of DNA can be performed using real-time PCR (qPCR) (Heid *et al.,* 1996). Elimination of the required post-amplification processing steps significantly reducing the time, assay labor and risk of carryover contaminations are some of de advantages of qPCR (Schena *et al.,* 2004). Moreover, this technique allows the accurate quantification of the target pathogen, by interpolating the quantity measured to a standard curve with known amounts of target copies (Garrido *et al.,* 2009). In addition, qPCR seems to be more sensitive than conventional PCR which is essential to detect soil-borne fungi and oomycetes that can be present at very low levels (Schena *et al.,* 2004).

The qPCR allows the monitoring of the reaction during the amplification process by the use of a fluorescent signal that increases proportionally to the number of amplicons generated. Different methodologies can be used to measure the PCR product obtained after each cycle of the PCR with the use of inespecific or specific fluorescens dyes. The **inespecific methods** are based on the use of binding dyes, such as SYBR Green (Morrison *et al.,* 1998) which is a fluorescence intercalating dye with a high affinity for double-stranded DNA. Although the no need of probe reduces costs, special attention must be paid in the formation of non-specific

amplicons and dimmers, since this dye does not discriminate between the different double-stranded DNA. On the other hand, the **specific methods** are based on the use of oligonucleotides probes labelled with a donor fluorophore (reporter) covalently attached to the 5' end and an acceptor dye (quencher) attached to the 3' end. The fluourophore does not emit fluorescence in the presence of the quencher, which dissipates the energy by proximal quenching or by fluorescent resonance energy transfer (FRET). Once the primers and probe specifically hybridise to the DNA, the 5'-3' exonuclease activity of the *Taq* DNA polymerase cleaves the probe causing the liberation of the fluorophore, which start emitting fluourescence. The fluorescence detected in the qPCR thermal cycler is directly proportional to the fluorophore released and the amount of DNA template present in the PCR. Among labelled probed, TaqMan system is one of the most used (Livak *et al.*, 1995). The TaqMan probes were designed to increase the specificity of the reaction because detection and accurate quantification require high complementarity with the target sequence. Probes may include fluorophores such as FAM (6-carboxyfluorescein), TET (tetrachloro-6-carboxyfluorescein), JOE (2,7-dimethoxi-4,5-dichloro-6-carboxyfluorescein), ROX (6-Carboxyl-X-Rhodamine), Cy5 (indo-dicarbocyanine) and quenchers such as TAMRA (6-carboxytetramethylrhodamine), BHQ (black hole quencher) and MGB (minor grove binder) (Schena *et al.*, 2004). For instance, the MGB probes that include a MGB group at the 3' end raising the *Tm* (melting temperature) of the hybrid allows the use of shorter and more specific probes. The high

specificity of MGB probes make them very suitable for specific detection of fungal species based on SNPs (Massart *et al.,* 2005).

Reverse transcription PCR (RT-PCR)

One of the major drawbacks of PCR-based approaches is the inability to discern between DNA obtained from active and non-active fungal material (conidia, dead mycelia, etc.) which result in the overestimation of populations (Bridge & Spooner, 2001; Lievens *et al.,* 2006). The use of reverse transcription of RNA combined with qPCR (qRT-PCR) may solve this problem and could provide a more meaningful assessment than current DNA approaches (Atkins *et al.,* 2003; Beaulieu *et al.,* 2011).

Digital PCR (dPCR)

Digital PCR (dPCR) is an emerging technique which offers an unique approach to qPCR for measuring nucleic acids that may be particularly suited for low-level detection (Vogelstein & Kinzler, 1999; Dube *et al.,* 2008). The dPCR is based on the principle that an absolute count of amplified targets can be achieved working on the premise that every molecule of target is successfully amplified, and therefore, this new technique should, in theory, provide the most accurate method of molecular quantification (Sanders *et al.,* 2011).

The dPCR involves performing PCR with real-time or end-point data collection in a large number of separate reaction chambers, also termed partitions. Single molecules are isolated by dilution and individually

amplified by PCR; each product is then analyzed separately. This is achieved by partitioning a sample prior to PCR amplification such that at least some of the partitions contain no copies of the target sequence(s) of interest. Results are obtained by counting the number of partitions containing target sequence detected by fluorescence (regarded as positive) and the number of partitions in which there is no fluorescence (regarded as negative). Poisson statistical analysis of the numbers of positive and negative partitions yields absolute quantitation of the target sequence. This compensates of the fact that more than one copy of template may be present in some partitions (Dube *et al.*, 2008; Huggett *et al.*, 2013) (Fig. 1).

Today, dPCR instruments achieve partitioning either on chips or throught water-in-oil emulsions or droplets. The dPCR based on chips is performed in small-volumen, solid partitions that allow either real-time or end-point analysis of the individual reactions. Compared to the droplet instruments, the number or partitions that are available, typically on the order of a few thousand, is fewer.

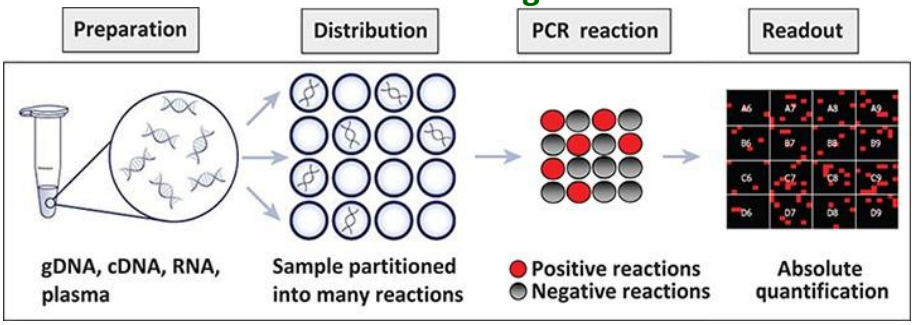

Figure 1. Absolute quantification of DNA using dPCR.

Source: www.lifetechnologies.com

Emulsion (or droplet) dPCR occurs in partitions made up of water-in-oil emulsion droplets. One of the advantages of these instruments is the great number of partitions achieved and lower running costs that most chip-based instruments. The higher number of partitions leads to an increased dynamic range, allowing analysis of a great range of sample concentrations for any given precision. However, these instruments require more technical complexity and the need for post-PCR sample manipulation (Huggett *et al.*, 2013).

The most outstanding advantage of dPCR over qPCR is to enable the absolute concentrations of DNA concentrations without external calibrators. The dPCR not only allows absolute quantification of target genes without any standards but also is considered less susceptible to PCR inhibitors present in the DNA extracts than qPCR (Hoshino & Inagaki, 2012).

The dPCR has already been used in clinical diagnosis (Schulz *et al.*, 2014; Wang *et al.*, 2014), especifically for rare variant measurement, which

has made of this technique a potential tool in several scenarios (e.g. the diagnosis and staging of cancer) (Vogelstein & Kinzler, 1999). However, its applicability in plant pathology for the detection of pathogens in environmental samples (soil, composts or plants) remains uncertain.

Oligonucleotide array

Nucleic acid arrays offer the possibility to analyze a specifically selected group of microorganisms, concerning their presence or absence in a particular environmental sample, in a single experiment (Bodrossi *et al.*, 2003). In DNA arrays, DNA extracted from samples is amplified and subsequently fluorescently labeled and hybridized to the array. This technology is the most suitable technique to detect several target organisms simultaneously. This technology has been already successfully applied in diagnostics of human and animal (Lievens & Thomma, 2005) and in plant pathology, it has been successfully applied to identify DNA from pure cultures of oomycetes, nematoded and bacterial pathogens (Lévesque *et al.*, 1998; Uehara *et al.*, 1999; Fessehaie *et al.*, 2003) and from multiple tomato wilt pathogens (Lievens *et al.*, 2003, 2005).

1.2. Techniques to assess the population genetics of soil-borne plant pathogens

In spite of the last advances in molecular biology and the new available technology, the effective management of diseases caused by plant soil-borne pathogens is still a challenge for plant pathologists. One of the

main reasons for this limited success is the lack of knowledge about the variability of the genetic structure of pathogen populations (Martin & English, 1997). Knowledge of the genetics of populations of plant pathogens may provide information about the evolutionary potential of pathogens to overcome management strategies (Southwood *et al.,* 2012). The genetic structure of a population reflects its evolutionary history and its potential to evolve. In this sense, knowledge of the spatial distribution of genotypes within populations can proved information about their dispersal potential within fields (McDonald, 1997).

Multiples tools and techniques have been used to the study of population genetics including DNA fingerprints (RAPDs, RFLPs and AFLPs) and multilocus genotyping.

1.2.1. Fingerprinting techniques

Fingerprinting techniques have been extensively used to study the phylogenetic structure of plant pathogens populations and to differentiate strains of the same species with different host range, virulence, compatibility group or mating type (Capote *et al.,* 2013). DNA fingerprinting techniques include r**andom amplified polymorphic DNA** (RAPDs) (Williams *et al.,* 1990), **restriction fragment length polymorphism** (RFLPs) (Botstein *et al.,* 1980) and **amplified fragment length polymorphism** (AFLPs) (Vos *et al.,* 1995). The PCR-based approach of RAPD-PCR has been successfully used to identify soil-borne pathogens such as the fungal pathotypes of *P. nicotianae* that cause tobacco black shank (Zhang *et al.,* 2001) and several

forma speciales and races of *F. oxysporum* (Lievens *et al.*, 2008). On the other hand, RFLPs in nuclear and mitochondrial genomes have been used in many studies of plant-pathogenic fungi. However, because of this technique is based on DNA-DNA hybridization is more reproducible but technically more difficult than RAPDs (McDonald, 1997). AFLPs are relatively costly and a rather complicate technical procedure (Lievens *et al.*, 2008).

Although most of these techniques have proved valuable within a particular study, results obtained with such fingerprinting tools are not always easily reproducible in different laboratories (Cooke & Lees 2004).

1.2.2. Multilocus genotypes

Genetic diversity and evolution of populations of a certain pathogen can be studied using multi-gene sequence phylogenies (Southwood *et al.*, 2012). Nucleotide sequence data offers the possibility of reconstructing patterns of descent among genotypes within a species or among populations of one or more species (Goss *et al.*, 2009). The comparison of the Single-Nucleotide Polymorphisms (SNPs) among the different isolates is ideal for evolutionary studies because they are widespread in the genome as well as easy to screen (Abott *et al.*, 2010). Closely related pathogens showing different host ranges or pathogenicity often differ in only a single to a few base pairs in target genes.

Different suitable markers have been proposed. In spite of the fact that ITS regions (Cooke *et al.*, 2000) are the most commonly sequenced for fungi

and oomycetes for phylogenetic studies, these regions are usually conserved within species. In this sense, the Intergenic Spacer (IGS) DNA region has emerged as a suitable alternative to the ITS region when closely related taxa or even different strains of the same species need to be differentiated. The IGS region evolves faster and consequently, more sequence polymorphisms are present (Martin & Tooley, 2003). The wide utilization of the IGS region as target for developing specific molecular markers is primarily limited by difficulties in amplifying long fragments (approximately 2-4 kbp in fungi and 3-5 kbp in oomycetes) and the lack of effective universal primers (Capote *et al.*, 2012). Another alternative to differentiate closely related species is the use of mitochondrial genes such as the mitochondrial encoded cytochrome oxidase I (cox I) and II (cox II) and their intergenic region (mt-IGS) (Martin & Tooley, 2003; Nguyen & Seifert, 2008; Seifert *et al.*, 2007). It usually has a higher rate of evolution than nuclear DNA (White *et al.*, 1990). In the absence of sexual combination the use of mitochondrial markers, which are maternally inherited, would be useful to study clonally reproducing populations of the pathogen (Mamella *et al.*, 2011; Martin & Coffey, 2012). A general disadvantage of mitochondrial DNA is the very high AT/GC ratio which is generally more difficult to amplify and requires a higher concentration of $MgCl_2$ compared to genomic DNA.

In recent years, multi-locus approaches have been used to study the genetic diversity and evolution of different *formae speciales* of *F. oxysporum* (O'Donnell *et al.*, 1998, 2004; Southwood *et al.*, 2012) as well as different

species of *Phytophthora*, namely *P. infestans* (Cárdenas *et al.,* 2011), *P. ramorum* (Goss *et al.,* 2009), *P. capsici* (Hurtado-Gonzales *et al.,* 2008) and *P. nicotianae* (Martin *et al.,* 2014).

1.2.3. Other molecular markers

Microsatellites (SSRs) are units of 1 to 4 bp long which are repeated 10 to 100 or more times in the genome of eukaryotic (this event is rare in prokaryotes). These repeats have a tendency to change in number when DNA is replicated due to DNA polymerase slippage. Size differences in the repeat length can be visualized by radiolabel or fluorescent molecules incorporated into the PCR products during amplification or by sequencing. SSRs have been recognized as one of the most powerful alternative molecular approaches to study intraspecific variability among populations (Cooke *et al.,* 2007) and have been shown to be useful for differentiating diverse *formae specialis* of *F. oxysporum* (Leyva-Madrigal *et al.,* 2014). SSRs have been widely utilized for those species whose genome has been partially or completely sequenced since it is necessary to know the DNA sequence of the SSR flanking regions to design specific primers (Mamella *et al.,* 2012).

1.3. Methods to characterize compost microbial communities

It is well-known that microbes play key roles in the suppressive effect of composts and the appearance of some reflects the successful suppressiveness achieved against different soil-borne pathogens (Noble &

Coventry, 2005). Numerous methods have been developed to more fully characterize microbial communities in soils, and specifically in composts. There are two kinds of approaches, the based on culture-based methods and the independent-culture ones.

1.3.1. Cultivation-dependent methods

Traditionally, cultivation-based methods have been used to study the microbial diversity in environmental samples. A wide array of culture media has been designed so as to maximize the recovery of diverse microbial groups (Figure 2). For instance, Suárez-Estrella *et al.* (2007; 2013a,b) used Sodium Caseinate Agar, nutrient agar and Rose-Bengal Chloramphenicol agar plates for the isolation of actinobacteria, bacteria and fungi respectively from different composts to test their potential antagonistic activity against several plant pathogens.

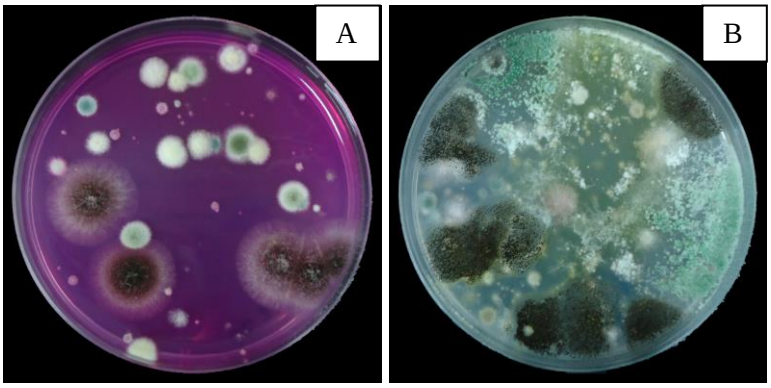

Figure 2. Different culture media used to characterize compost microbial community.A) Rose-Bengal Chlorampenicol agar plates; B) Nutrient agar plates.

Besides, a Biolog-based method, for directly analyzing the potential activity of soil microbial communities, denotes community-level physiological profiling (CLPP) (Garbeva *et al.*, 2004). Because of the inability of these methods to detect non-culturable species, which are able to be present or even been predominant in certain composts, other methods, mainly molecular ones, have become popular.

1.3.2. Cultivation-independent methods

A broad range of cultivation-independent techniques has been applied to the study of microbial communities presented in composts. These methods have been classified into two major categories: partial community analysis and whole community analysis approaches.

1.3.2.1. Partial community approaches

Partial community analyses are based on the use of PCR-based methods to amplify total DNA/RNA extracted from an environmental sample. The 16S rDNA genes and ITS regions from bacterial and fungal communities respectively have been widely used for the study of microbial communities (Mehta et al., 2014). These approaches include genetic fingerprinting techniques, fluorescence in situ hybridisation, clone libraries, DNA microarrays, qPCR and DNA/RNA stable isotope probing.

Microbial community fingerprinting techniques

The techniques that have been developed to fingerprint compost microbial communities include denaturing or temperature gradient gel electrophoresis (DGGE/TGGE), amplified rDNA restriction analysis (ARDRA), terminal restriction fragment length polymorphism (T-RFLP), single-strand conformational polymorphism (SSCP), ribosomal intergenic spacer length polymorphism (RISA) and phospholipid fatty acids analysis (PLFAs). **DGGE** is based on the separation of the PCR producs depending on their nucleotide sequences on a polyacrylamide gel containing a linear gradient of DNA denaturant (Muyzer *et al.,* 1993). In **TTGE** a temperature gradient is applied instead of the chemical denaturant. The ability to excise, reamplify and sequence particular bands in the patterns allows the identification of the microbial types or genes that underly these bands (van Elsas & Boersma, 2011). DGGE is the most widely used among the methods to study microbial communities in environmental samples, and in particular in composts (Ishii *et al.,* 2000; Garbeva *et al.,* 2004; Danon *et al.,* 2008). **ARDRA** is based on the PCR amplification of genes coding for rRNA, mostly 16S rRNA and 18S rRNA genes, and the consequent digestion with tetracutter restriction endonuclease. Restricted fragments are resolved on agarose or polyacrylamide gels to produce a specific community pattern (Vaneechoutte *et al.,* 1992). ARDRA technique has been applied to the analysis of the microbial community in compost (Uchiyama *et al.,* 2002). **RISA** approach involves PCR amplification of a portion of the intergenic region present between the small (16S) and large (23S) ribosomal subunits in

bacteria and the transcribed spacers and the 5.8S rRNA genes (ITS1-5.8S-ITS2) in fungi (Jensen *et al.,* 1993; Ranjard *et al.,* 2001). An automated RISA (ARISA) involves the use of a fluorescence-tagged oligonucleotide primer for PCR amplification and the subsequent electrophoresis in an automated system (Ranjard *et al.,* 2001). This approach has been used successfully to assess the structure bacterial and fungal communities in compost (Schloss *et al.,* 2003; Hansgate *et al.,* 2005). Another technique that studies the DNA sequence variations present in PCR-amplified bacterial 16S and fungal 18S rDNA genes is **T-RFLP** (Liu *et al.,* 1997). This approach is based on the detection of PCR products previously labeled with the use of one 5' fluorescently labeled primer during the PCR reaction. The resulting PCR products are digested with restriction enzymes, and terminal restriction fragments (T-RFs) are separated on an automated DNA sequencer (Marsh *et al.,* 2000). Tiquia (2005) analyzed the microbial community structure and diversity in manure composts at different stages of composting using this technique. In **SSCP**, PCR products are denatured followed by electrophoretic separation of single-strand DNA fragments on a nondenaturing polyacrylamide gel. Single-stranded DNAs fold into secondary structures according to their nucleotide sequences and their physicochemical environment leading to measurable differences in mobility in the gel (Shwieger & Tebbe, 1998). SSCP has been used for genetic profiling of microbial communities including those involved during the composting process (Lee *et al.,* 1996; Peters *et al.,* 2000). **PLFA** analysis is based on the extraction of PLFA biomarkers, which are identify by gas chromatography

with flame ionization detection (GC-FID) and confirmed by mass spectroscopy (MS), if necessary. Among the profiling techniques, PLFA analysis is characterized for giving quantitative information about community structure (Ebersberger *et al.*, 2004) and has proved to be a useful tool for monitoring microbial community dynamics (Klamer & Baath, 1998). It has been used to characterize microbes in diverse systems such as marine sediments, soils, plant rhizospheres and composts (Herrmann & Shann, 1997; Bastida *et al.*, 2008) being sensitive to changes in composition and during the composting process (Amir *et al.*, 2010; Kindo *et al.*, 2012).

Fluorescence *in situ* hybridisation (FISH)

Fluorescence *in situ* hybridisation (FISH) is a technique used to evaluate the phylogenetic identity, morphology, number, and spatial arrangements of microorganisms in different environmental samples. This technique is based on the design of fluorescently labeled 16S rRNA-oligonucleotide probes specific for the organism of interest (Hugenholtz, 2002). FISH has been used to study the influence of composts in bacterial root colonization (Iverson & Maier, 2009) as well as the spatial and temporal distribution of *Bacillus* and *Clostridium histolyticum* during composting of swine manure (Yi *et al.*, 2011).

Clone libraries

Clone libraries are useful to identify and characterize the dominant bacterial and fungal types in composts and may reveal the identity of

uncultured and yet-unknown composting microorganisms (Garbeva *et al.*, 2004; Franke-Whittle *et al.*, 2009; Partanen et *al*., 2010; Blaya *et al*., 2013). Briefly, PCR-generated amplicons are ligated into a suitable vector plasmid. Later, the resulting constructs are transformed into *Escherichia coli*. After the growth of positive colonies, cloned amplicons can be isolated by plasmid extraction, and after sequencing them, the results are analyze comparing with databases such as GenBank, Ribosomal Database Project or Greengenes (van Elsas & Boerma, 2011). The clone libraries have a high resolution considering that they have to be quite large to accurately describe the microbial diversity within a compost sample (Garbeva *et al.,* 2004).

Oligonucleotide array

This technology has been successfully used in the study of microbial communities in diverse ecosystems. Some examples are the GEOChip (He *et al.,* 2007), a microarray for investigating biogeochemical, ecological and environmental processes; the PhyloChip (DeSantis *et al.,* 2007; Schatz *et al.,* 2010), designed to detect and quantify abundance of bacterial and archaeal taxa using signature probes targeting all known 16S rRNA gene sequences, the ANAEROCHIP (Franke-Whittle *et al.,* 2009), which study sludge methanogenic communities, the COMPOCHIP (Franke-Whittle *et al*., 2009b), an array which gives information about the characteristic microbiota of composts; and other arrays specifically designed to detect landfill methanotroph communities (Bodrossi *et al.,* 2003; Stralis-Pavese *et al.,* 2004), municipal wastewater microorganisms (Lee *et al.,* 2006), or disease

suppressive microorganisms (Lievens *et al.*, 2007). The COMPOCHIP microarray has been used, together with other molecular techniques (DGGE or clone libraries), to characterize the bacterial communities involved of different composts (Danon *et al.*, 2008; Fernández-Gómez *et al.*, 2012).

Real Time PCR (qPCR)

The qPCR is currently widely applied to soil/compost-extracted DNA, allowing the quantitative detection of target genes such as 16S rRNA genes or of functional genes like *amoA*, *nifH*, *pmoA*, or *dsrA* (van Elsas & Boerma, 2011). The qPCR may be well employed to assess to what extent, local conditions affect gene and gene expression levels. As PCR based on soil/compost DNA extracts, qPCR may be subjected to some biases.

Stable isotopes probing (SIP)

Stable isotopes probing (SIP) offers a powerful technique for identifying microorganisms that are actively involved in specific metabolic processes (Radajewski *et al.*, 2000). Different methods have been proposed such as the incorporation of 5-bromo-2'-deoxyuridine (BrdU) into DNA followed by fingerprinting of the active communities (Artursson *et al.*, 2005), or the incorporation of ^{13}C (stable isotope labeling) into cellular biomarkers (DNA, RNA, PFLAs) followed by separation and fingerprinting (Boschker *et al.*, 1998; Radajewski *et al*., 2000; Manefield *et al.*, 2002). Recently, Peng *et al.* (2013) used $^{13}C4$-pyrene to study the microorganisms responsible for the degradation of pyrene during composting. The coupling of molecular

biological methods with stable-isotope abundance in biomarkers has provided a cultivation-independent means of linking the identity of microorganisms with their function in the environment (Radajewski *et al.*, 2000).

1.3.2.2. Whole community analysis approaches

These techniques attempt to analyze all the genetic information present in total DNA extracted from an environmental sample, in comparison to PCR-based molecular approaches that target only a single or few genes. These approaches include DNA-DNA hybridisation, Guanine-plus-citosine (G+C) content fractionation, whole genome sequencing and metagenomics.

DNA-DNA hybridisation

Whole-genome DNA–DNA hybridization (DDH) offers true genome-wide comparison between organisms. This technique measures the degree of genetic similarity between pools of DNA sequences. A value of 70% DDH was proposed as a recommended standard for bacterial species delineation (Goris *et al.*, 2007). Typically, bacterial species having 70% or greater genomic DNA similarities usually have >97% 16S rRNA gene sequence identity. Although DDH techniques have been originally developed for pure culture comparisons, they have been modified for its use in whole microbial community analysis (Rastogi & Sani, 2011).

Guanine-plus-cytosine (G+C) content fractionation

Guanine-plus-citosine (G+C) content technique is based on the fact that different prokaryotic groups differ in their guanine-plus-cytosine (G + C) content of DNA, and phylogenetically related bacterial groups only vary by 3–5% in their G + C content (Nüsslein & Tiedje, 1999). The total community DNA is physically separated by density-gradient centrifugation, into highly purified fractions, each representing a different G + C content that can be analyzed by additional molecular techniques (e.g. DGGE or ARDRA) to better assess total community diversity (Rastogi & Sani, 2011). G + C fractionation has been widely applied in investigation of soil microbial communities to evaluate for instance, the effect of compost application on soil microbial communities (Rastogi & Sani, 2011).

Whole-microbe-genome sequencing

Exploring microbial systems through whole-genome analysis is a comprehensive and integrated approach to understand microbial ecology and function. Whole microbial genomes are sequenced using a shotgun cloning method that involves (1) extraction of DNA from pure cultures, (2) random fragmentation of obtained genomic DNA into small fragments of ~2 kb, (3) ligation and cloning of DNA fragments into plasmid vectors, and (4) bidirectional sequencing of DNA fragments. Once the sequences are obtained, they are aligned and assembled into finished sequences using specialized computer programs such as MEGAN (*ME*ta*G*enome*AN*alyzer) (Huson *et al.,* 2007). The classical shotgun sequencing was based on the

Sanger sequencing method. This method is based on the use of dideoxynucleotides (ddNTPs) in addition to the normal nucleotides (dNTPs). The extension of a newly synthesized DNA strand terminates every time the corresponding ddNTP is incorporated. Moreover, one of the nucleotides or the sequencing primer should be radioactively or fluorescently labeled so that the final product can be detected. The large-scale sequencing technologies, which are discussed in the following section, are useful for whole-genome sequencing as well as metagenomics and metatranscriptomics.

Metagenomics (next-generation sequencing methods)

Metagenomics is the direct investigation of collective microbial genomes contained within an environmental sample (Riesenfeld *et al.*, 2004). The field initially started with cloning environmental DNA, being the Sanger sequencing method one of the most widely used during the last decades. It is still in use but in a small scale. Recently, few novel and powerful sequencing techniques (Next-generation sequencing methods) namely 454-based/pyrosequencing and Illumina/Solexa's Genome Analyzer sequencing, have been developed. They consist of multiparallel sequencing by synthesis, in which the pyrophosphate that is released is detected in an enzymatic cascade ending in luciferase and followed by the detection of the emitted light. Meanwhile pyrosequencing allows the production of hundreds of thousands to millions of 450-bp read in a run, Solexa platform offers a higher magnitude of reads but lower lengths (van Elsas & Boerma, 2011). In the last

years, new systems namely Ion Torrent PGM, PacBio RS and MiSeq, have revolutionized the genome studies of collective microbial communties. These next generation plataforms are faster and less expensive than traditional Sanger sequencing (Metzker, 2010).

Some of the limitationss of these techniques are the high costs required for the analysis of the inmense amount of data obtained which may be limited by human capability and the ability of databases to deal with errors. On the other hand, the ultra-high throughput and lack of biases of these methods will allow the discovery of many new sequences, which is very important in the case of compost. Compost communities are extremely diverse and may contain a large microbiota represented by an enormous number of low-abundance unique taxa, which may be involved in compost suppressiveness. This fact highlights the importance of large-scale sequencing techniques in investigating the highly diverse compost microbial communities (Lauber *et al.*, 2009; Neher *et al.*, 2013). Metagenomic investigations have been conducted in an extent variety of environments including composts (Dougherty *et al.*, 2012; DeGannes *et al.*, 2013a,b; Yeh *et al.*, 2013) and have provided access to phylogenetic and functional diversity of uncultured microorganisms (Handelsman, 2004).

1.3.2.3. Postgenomic approaches

The inability of DNA-based molecular techniques to provide information of the gene expression (functionality) as it occurs under *in situ* conditions (Wilmes & Bond 2006), has prompted the development of

postgenomic approaches such as metatranscriptomics, metabolomics and metaproteomics. These new approaches reveal the link between genetic potential and functionality in microbial communities (Rastogi & Sani, 2011).

Metatranscriptomics

Metatrascriptomic (or environmental transcriptomic) involves random sequencing of microbial mRNA allowing monitoring of microbial gene expression profiles in natural environments at a particular time and place (Moran, 2009). This approach, compared to metaproteomic, has a high-resolved view of instanteous regulatory responses (Moran, 2009). The major challenge in metatranscriptomics is the fact that prokaryotic microbial mRNA transcripts are not polyA tailed, so obtaining complementary DNA (cDNA) is not easy. This results in coextraction of more abundant rRNA molecules in the total RNA pool, which can lead to overwhelming background sequences in a largescale sequencing analysis (Rastogi & Sani, 2011).

Metabolomics

Metabolomic allows monitoring of low molecular weight metabolites produced by the extant microbial communities in certain environment conditions. The set of metabolites synthesized by an organism constitute its metabolome (Oliver *et al.,* 1998). Metabolites are the end products of cellular regulatory processes, and their levels can be regarded as the ultimate response of biological systems to genetic or environmental changes (Fiehn,

2002). Therefore, this approach has a high potential to elucidate changes in the levels of metabolites due to its sensibility.

Methodologically, the study of metabolome involves firstly the extraction of the metabolites and secondly, their detection. Different methods of metabolites extraction may be used. Frequently, polar organic solvents like methanol, methanol-water mixtures or ethanol are added to the sample, followed by an additional step of using non-polar solvents such as chloroform (Fienh, 2002). The metabolites detection is performed using analytical techniques such as Nuclear Magnetic Resonance (NMR) or mass spectrometry (MS). MS can be performed without previous separation of metabolites or after separation through gas chromatography (GC), High Performance Liquid Chromatography (HPLC) and Ultra Performance Liquid Chromatography-mass spectrometry (UPLC).

Metaproteomics

Metaproteomics reveal information on proteins expressed by environmental microbial communities at a given point in time (Wilmes & Bond, 2006). Protein biomarkers are more reliable, compared to other cell molecules (eg. lipids or nucleic acids), and provide a clearer picture of metabolic functions than functional genes or even the corresponding mRNA transcripts of microbial communities (Wilmes & Bond, 2006; Beendorf *et al*., 2007). Metaproteome analysis involves the extraction of total proteins from an environmental sample, separation by 1-D or 2-D electrophoresis and followed by digestion of protein spots to identify the proteins by MS or

chromatography. Once the proteins are identified, it could be possible to link them with their metabolic functions (eg. involved with suppressive activities) and at the same time those to individual microbial species.

Despite all progress made in this area, one of the pitfalls of metaproteomics is the lack of available sequences in the dabatases which hinder the idenfitication of isolated proteins (Bastida *et al*., 2009). Moreover, metaproteomics of environmental samples is a challenge regarding resolution and yield of proteins (Keiblinger *et al*., 2012). In this respect, the preparation of protein extracts is of paramount importance and its application to soil or compost requires considerably improved protocols of protein extraction and sample preparation (Benndorf *et al*., 2007). Until now, several authors have reported and reviewed different methods to extract proteins from environmental microbial communities including soil and water which could be used for the study of compost microbial communities as well (Benndorf *et al*., 2007; Keiblinger *et al*., 2012; Bastida *et al*., 2014)