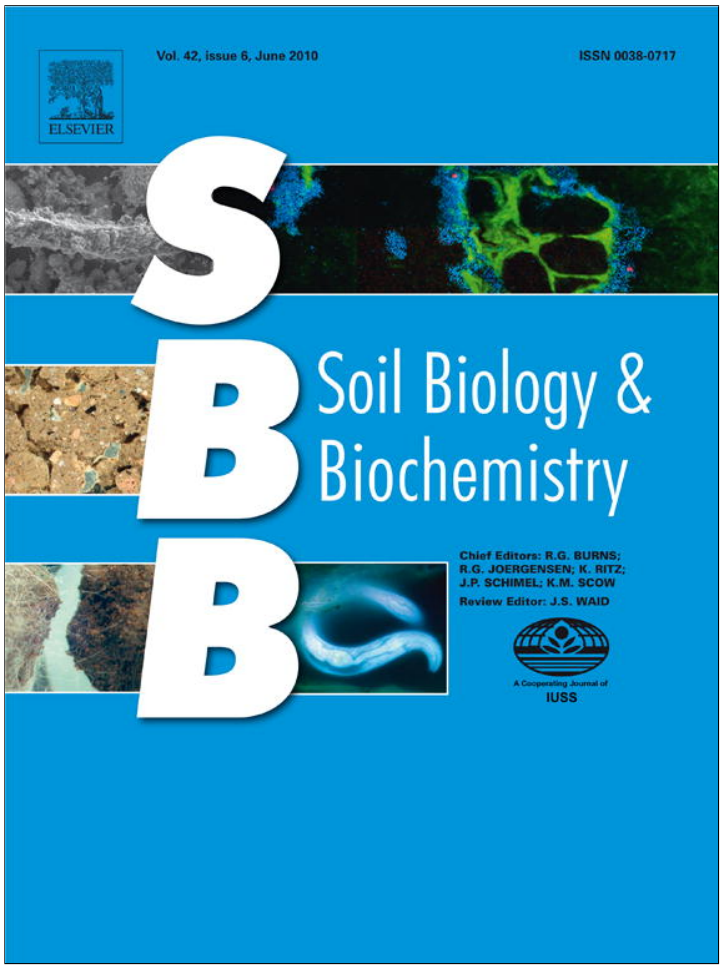

Soil Biology & Biochemistry 42 (2010) 878–887

Contents lists available at ScienceDirect

Soil Biology & Biochemistry

journal homepage: www.elsevier.com/locate/soilbio

Review

Culture-independent molecular techniques for soil microbial ecology

Penny R. Hirsch [a,b,*], Tim H. Mauchline [a], Ian M. Clark [a,b]

[a] Plant Pathology & Microbiology Department, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK
[b] Cross Institute Programme for Sustainable Soil Function, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK

ARTICLE INFO

Article history:
Received 19 August 2009
Received in revised form
23 December 2009
Accepted 19 February 2010
Available online 17 March 2010

Keywords:
Microbial communities
Molecular diagnostics
Genetic diversity
DNA
RNA
Microarrays
High-throughput sequencing
Metagenomics

ABSTRACT

The advent of nucleic acid-based molecular methods, in particular the polymerase chain reaction (PCR), has revolutionised the study of soil microbial ecology, previously constrained by an inability to culture the majority of cells detected by direct microscopic observation. Extraction of DNA and RNA directly from cells in soil circumvents the requirement to grow microorganisms in laboratory culture, avoiding problems associated with the differential growth rates of the estimated 1% that can be grown routinely. However, not all cells that contain DNA are capable of growth, and in some conditions such as air-dried soil, DNA can be extracted from non-viable microorganisms after 140 years of storage. To investigate the active microbial community, RNA can also be isolated directly from soil. Analysis of ribosomal RNA (rRNA) indicates the dominant active population in any particular set of conditions and the large, constantly increasing electronic database of gene sequences for the small subunit of rRNA (16S for prokaryotes, 18S for eukaryotes) provides identification of many soil bacteria, archaea and fungi with varying degrees of certainty to the genus, species or sub-species level. More precise information on which functional genes are active can be obtained from messenger RNA (mRNA). Newer methods including high-throughput (massively parallel) sequencing and microarrays offer further advances. We describe a range of molecular techniques used to investigate soil microbial communities, discuss how they relate to other methods for investigating bacterial and fungal activity, and explore their drawbacks and limitations.

1. Introduction

Together, the vast biodiversity of living microorganisms in soil can be considered collectively as the "soil microbial biomass", a concept which has proved useful for modelling nutrient and energy cycling (Stockdale and Brookes, 2006). The microbially-mediated activities in soil include properties perceived by humans as beneficial (e.g. carbon and nitrogen cycling; symbiotic interactions with plants) or deleterious (nitrous oxide and methane emissions; plant diseases), but nonetheless important. To better manage soil and minimise negative environmental impacts, there is a need for more detailed and predictive understanding of the microbial communities responsible for these activities and how they may respond to environmental stress and climate change. However, all investigations of soil microorganisms are hampered by the heterogeneity of the soil composition, the vast numbers (c. 10^9 individual cells) and diversity (>10^6 distinct taxa) of microorganisms present in each gram of soil (Curtis and Sloan, 2005; Gans et al., 2005), and the paucity of knowledge concerning the majority of the microbiota. Only a minority of soil bacteria, usually estimated at around 1% of the total number of cells observed by direct counting can be cultured routinely in laboratory media (Skinner et al., 1952). In areas of relatively high nutrient availability such as the rhizosphere, this can rise to >10% (Sørensen, 1997) but despite recent improvements in the design of culture medium for fastidious soil bacteria (Davis et al., 2005), the majority of bacterial cells present in soil are unable to grow in culture. Some of these will belong to groups for which optimal culture conditions have not yet been defined, including dependence on other microbes for growth, but others are probably moribund. Even for the culturable bacteria, there is no guarantee that activity measured in the laboratory is relevant to that which occurs under the range of conditions in soil. For soil fungi, even less is known about their diversity and activity. The mycelial growth of fungi, together with the production of many resting spores, can impede any interpretation of data on their abundance and activity.

Direct observation of microorganisms *in situ* in soil is very time consuming. A summary of the main methods follows but detailed

* Corresponding author at: Plant Pathology & Microbiology Department, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK. Tel.: +44 1582 763133; fax: +44 1582 760981.
E-mail address: penny.hirsch@bbsrc.ac.uk (P.R. Hirsch).

0038-0717/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.soilbio.2010.02.019

discussion is beyond the scope of this review. Stains can discriminate live and dead cells (on the basis of membrane integrity) and Gram positive or negative bacteria (Abaye et al., 2005). Also, it is possible to monitor microorganisms containing bioluminescent and other reporter genes in soil (Jansson, 2003). Techniques for labelling specific cells with RNA or DNA probes in soil using fluorescent *in situ* hybridization (FISH) have been applied to soil with limited success (Bouvier and del Giorgio, 2003).

To avoid the need for laboratory culture, bacteria and fungi can be lysed directly in soil to extract biochemical markers. Phospholipids derived from cell membranes, characterized by different acyl chains, have proved useful as broad indicators of shifts in bacterial community structure in soils subjected to various stresses (Zelles et al., 1992). They are mostly extracted from viable microbial cells, degrading rapidly upon cell death. The membranes of bacteria and eukaryotic cells contain phospholipid fatty acids (PLFAs), whereas archaeal cell membranes contain phospholipid etherlipids (PLELs), a feature unique to this group of prokaryotes (Gattinger et al., 2002). Similarly ergosterol, a component of fungal cell membranes, can provide estimates of viable fungal biomass in soil, and there are several bacterial cell wall and membrane specific biomarkers (Zelles and Alef, 1995). However, none of these methods targets groups with specific, key functions in soil, in contrast to the genetic material which is the basis for the properties of an organism. Considered as an entity, the genetic material isolated from soil is sometimes described as the soil metagenome and methods to study it as "metagenomics" although this term is also used to describe the isolation and cloning of relatively large intact DNA (10–100 kb) fragments from microbial communities, that might be predicted to contain several genes and operons (Rondon et al., 2000).

2. Sampling and extraction of nucleic acids from soil

An overview of the processes leading to preparation of labelled or unlabelled DNA and RNA is shown in Fig. 1; RNA can be converted to DNA where necessary for subsequent cloning or for PCR analyses (Fig. 2). More details are given below.

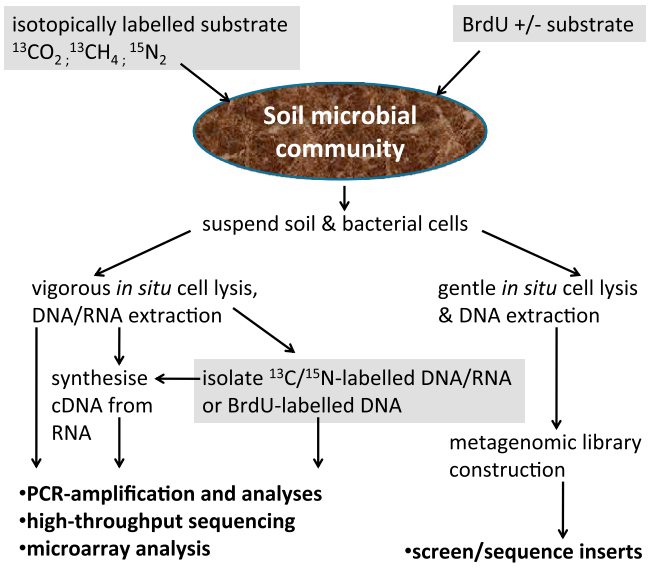

Fig. 1. Flow chart showing DNA and RNA extraction from soil. The soil microbial community will include bacteria, archaea, fungi, protists and microfauna.

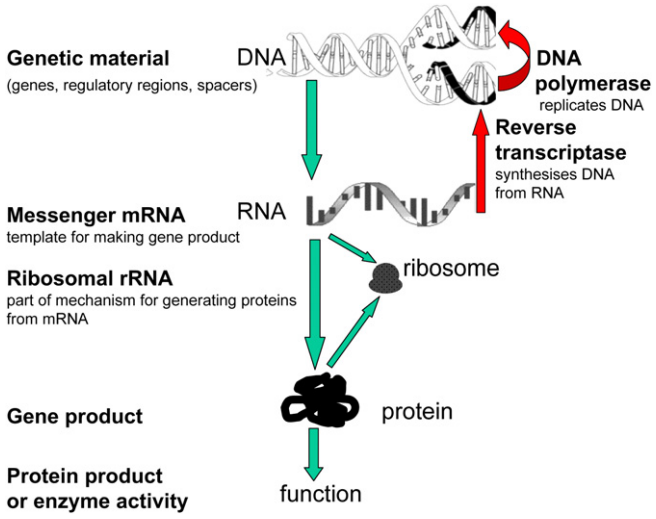

Fig. 2. The relationship between DNA, forms of RNA, protein and functions. The analysis of DNA and RNA is described as genomics; transcriptomics refers to studies of gene expression using rRNA and mRNA; proteomics and metabolomics refer to the analysis of proteins and the products of enzyme pathways, respectively.

2.1. Sampling

There are well-established sampling protocols to provide a representative sample of any site although sample size, the number of replicate samples, whether sampling is randomized or at regular intervals, and whether or not there is bulking and mixing of sub-samples depends on soil properties, target biota and the analytical methods to be used (Schleuß and Müller, 2001; Remenant et al., 2009). However, questions remain concerning the spatial scaling of microbial biodiversity in soil. In most studies, estimates of microbial diversity (species or taxa richness) increase with the area sampled (Curtis and Sloan, 2004; Green and Bohannan, 2006). There is also a lack of consensus over the definition of "species" in prokaryotes as classification based on 16S rRNA gene (16S ribotyping) sequences eludes differential functional attributes encoded by genomes although it is generally accepted as an appropriate taxonomic identifier (Rosselló-Mora and Amann, 2001; Prosser et al., 2007). However, the generation of new high-throughput methods that greatly increase the number of samples that can be compared is beginning to address these longstanding questions.

2.2. Extraction and purification

The extraction of nucleic acids from soil is complicated by the presence of humic substances, organic matter and clay particles that are able to bind to nucleic acids and inhibit their purification (Moran et al., 1993). Moreover, humic and fulvic acids are also capable of inhibiting post-extraction downstream enzymatic analyses (Tebbe and Vahjen, 1993), thus an important step is the removal of organic contaminants that co-extract with nucleic acids (Cullen and Hirsch, 1998). An additional consideration for RNA extractions is the ubiquitous RNases present in soil which must be deactivated rapidly to prevent degradation of the sample. Usually, RNA extractions involve an additional step to inactivate RNase before extraction of nucleic acids (Ogram et al., 1995; Mendum et al., 1998; Saleh-Lakha et al., 2005).

Despite these difficulties, numerous techniques have been developed for DNA and RNA extraction from soils, many of which are now available commercially. Protocols can be divided into indirect and direct nucleic-acid extractions. The first step of an

indirect extraction is to extract the microbial cells from the soil matrix before cell lysis. This type of extraction has been criticised for showing bias, for example, methane oxidising and ammonia oxidising bacteria have been shown to be more difficult to dislodge from soil particles compared to the majority of other soil bacteria, and actinomycete spores may be under-represented (Priemé et al., 1996; Aakra et al., 2000; Courtois et al., 2001). On the other hand, indirect extraction enabled isolation of high molecular weight DNA for metagenomic library construction (Berry et al., 2003) and revealed a different community structure, in comparison to direct extraction based on *in situ* lysis of microbes in soil (Bertrand et al., 2005).

An important step for either approach is the method of microbial cell lysis, and this choice depends on the type of biological structure being primarily extracted from (spores or vegetative material), the size of DNA fragments required for downstream applications (metagenomic library construction requires larger fragments than, for example, 16S ribotyping) and the properties of the soil. Soils with high organic matter can pose particular problems due to their high humic acid content. Similarly, soils with a high clay content present more problems when extracting nucleic acids than sandy soils due to their higher organic matter content and their hygroscopic propensity to bind water. Aggressive methods such as bead-beating of soil samples provide effective disruption of cell walls and membranes *in situ* (Cullen and Hirsch, 1998) and increasing vigour is needed as the soil humic acid and clay content increases (Saleh-Lakha et al., 2005) notwithstanding a consequential decrease in the average DNA fragment size. Other methods of cell lysis include sonication, grinding–freezing–thawing, solubilisation of cell walls and membranes by detergents or degradation of these structures by boiling and/or enzymatic means (Robe et al., 2003). It is important to recognize that biological structures differ in their susceptibility to aggressive disruption methods such as bead-beating, and so the protocol adopted for a given sample may depend on the target community. For example, extraction of DNA from spores may require vigorous disruption but this will result in highly fragmented DNA from easily-lysed cells. A protocol optimised to extract genomic DNA from the majority microbial community will be biased against both tougher and more fragile propagules: in order to maximize the diversity represented by genomic DNA extracted from soil, a variety of methods may be necessary. Once cells have been lysed, a number of methods may be used to extract and purify the nucleic acids, several using commercial kits (Roose-Amsaleg et al., 2001; Robe et al., 2003; Thakuria et al., 2009).

2.3. Whole genome amplification

When the size of a sample is limited (for example in the case of microsamples from roots, soil aggregates or meagre archived material) the amount of DNA can be increased post-extraction using "Whole Genome Amplification" (WGA) using one of several possible methods that rely on enzymic replication of DNA with short random primers (Hawkins et al., 2002). Now an established technique in forensic science and human medicine (in particular for embryo pre-implantation diagnostics), it is likely that WGA methods will become increasingly important in environmental microbial ecology in the coming years. Concerns arise from the potential for bias during amplification leading to misleading results and there is not yet a general consensus in the literature on the best method for applications in soil. Nevertheless, the potential of one method, "Multiple Displacement Amplification" (MDA) using φ29 DNA polymerase has been demonstrated with soil-extracted DNA, enabling subsequent PCR amplification, microarray hybridization and metagenomic cloning, albeit with a certain degree of bias (Gonzalez et al., 2005; Abulencia et al., 2006; Chen et al., 2008).

3. *In situ* labelling to link functions with identity

3.1. Isotopic labelling

Identifying the microbial communities responsible for particular activities is often circumstantial, correlating increases in the abundance of particular groups or genes with the activity in question. More tangible evidence has been obtained when isotopically-labelled substrates are traced to a particular group (Dumont and Murrell, 2005). Addition to soil of radiolabelled substrates or nucleotides such as ^{3}H-labelled thymidine enables microautoradiography (MAR) of cells that have incorporated the label. Combined with FISH this can identify active cells *in situ* but application to soil faces the same practical constraints as other direct imaging methods (Wagner et al., 2006; Rogers et al., 2007).

Incorporation of label into PLFAs or nucleic acids that can be extracted from soil is more encouraging, allowing identification of specialised groups. The best known example of this "stable isotope probing" (SIP) is the association of ^{13}C with PLFAs or DNA from methane oxidizers when soil is exposed to $^{13}CH_4$ (Bull et al., 2000; Radajewski et al., 2000). Both DNA and rRNA of actively-growing cells incorporate ^{13}C from the chosen substrates. Following extraction of nucleic acids from soil, the ^{13}C-enriched fraction can be separated from the unlabelled fraction using density gradient centrifugation. Methods for SIP, reviewed by Kreuzer-Martin (2007), work most effectively for investigations of distinctive subgroups that consume specific substrates such as methane but can also provide more general information on the actively-growing members of communities in soil. For example, a range of ^{13}C substrates have been used to probe different groups in soil using both DNA and RNA-SIP (Manefield et al., 2002; Lueders et al., 2004; Whiteley et al., 2006).

3.2. BrdU labelling

Actively-growing cells can also be labelled in soil with the thymidine nucleotide analogue, 5-bromo-2′-deoxyuridine (BrdU), in conjunction with the addition of any, or no, exogenous substrate (Borneman, 1999; Yin et al., 2000). DNA extracted from cells that incorporated BrdU can be isolated by immunocapture and then compared (using any methods appropriate for community DNA analyses) to the unlabelled DNA from the less active majority. We have used BrdU to compare the total and actively-growing communities in soils from different sites incubated at 4 or 25 °C (Fig. 3). Results show that profiles from total DNA do not group according to the source of the soil or the temperature of incubation, in contrast to the DNA from actively-growing bacteria that incorporated BrdU, which group according to both criteria. Other promising applications of BrdU include the identification of bacteria associated with arbuscular mycorrhizal fungal hyphae (Artursson and Jansson, 2003) and of active microbial community members in soils suppressive to plant pathogens (Hjort et al., 2007).

4. Nucleic-acid targets

4.1. DNA or RNA?

DNA extracted from soil represents the total metagenome, including components that are no longer viable, whereas RNA is synthesised only by actively-growing cells and it degrades relatively rapidly once produced: it arises from, and can thus identify, the functioning members of soil microbial communities. In

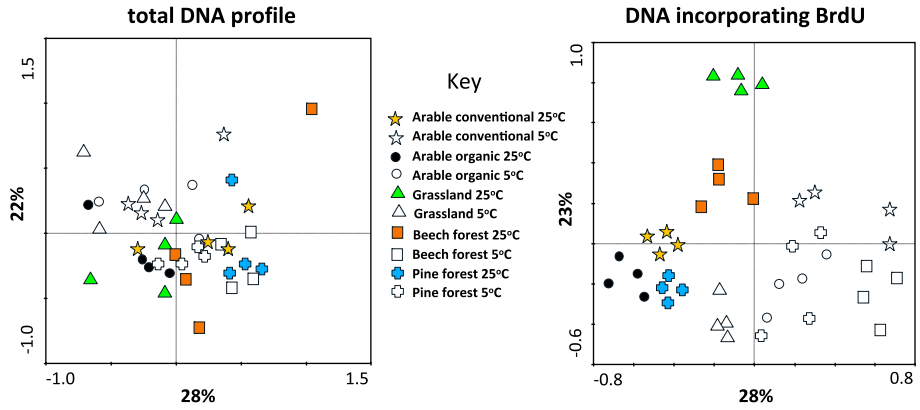

Fig. 3. Canonical variate analysis of 16S rRNA gene DGGE profiles comparing total bacterial communities and the active components identified by incorporation of BrdU. The active components group according to land management and the temperature at which soil samples were incubated with BrdU; in contrast the total community is not resolved between treatments.

prokaryotes, messenger RNA (mRNA) is usually very short-lived and indicates which genes are active at the time of extraction but ribosomal RNA (rRNA) is more stable as it possesses secondary structure and is associated with ribosomal proteins so in theory, it could survive for months in moribund or dead cells in soil. However, in cells that are, or have recently been active, there are many thousands of molecules of rRNA. Thus, analysis of rRNA abundance and diversity has been used to indicate the dominant active population in soil, despite recognition that the number of ribosomes varies between groups (Janssen, 2006). More precise information relevant to particular functions can be obtained from mRNA but it presents more technical difficulties during extraction from soil (see above); nevertheless it opens exciting possibilities for future investigation (Fleming et al., 1998; McGrath et al., 2008). Both mRNA and rRNA can be converted to DNA using the enzyme reverse transcriptase (RT), but can also be hybridized to microarrays directly if sufficient material is obtained. The relationship between DNA, the forms of RNA, protein and functions is shown in Fig. 2.

4.2. Target genes for analysis of microbial communities

A large (and constantly increasing) electronic database of small subunit rRNA gene sequences (16S in prokaryotes, 18S in eukaryotes) provides identification of many soil bacteria, archaea and fungi to the genus, species and less frequently to the sub-species level, although many remain unassigned to known phyla and are classified as "uncultured" (Table 1). However, if the relative abundance of particular groups of bacteria is to be estimated from the numbers of different 16S rRNA genes in a sample, it is important to consider that rRNA gene copy number varies in different bacterial species (typically 10 copies in bacilli, 7 in enterobacteria, 4 in pseudomonads, 1 in nitrifiers and 1 in the majority of archaea that have been sequenced), as listed on the ribosomal rRNA operon copy number database *rrn*DB at http://ribosome.mmg.msu.edu/rrndb/search.php (Lee et al., 2009). There is insufficient information available on genome organisation of soil fungi to know if this problem affects estimation of fungal population structure based on rRNA genes.

Although small subunit (16S or 18S) rRNA is the molecule best characterized for microbial molecular systematics, other protein-encoding genes are now being evaluated for this purpose. Because rRNA genes evolve slowly, they provide a framework for assigning sequences to genera and species, appropriate for investigation of microbial community diversity, but cannot always resolve species and provide insufficient discrimination at the sub-species level (Rosselló-Mora and Amann, 2001). For intra-species variation, protein-encoding genes that have higher levels of sequence variation will permit differentiation of closely related individuals. An added advantage is that most of these genes occur with only

Table 1
Soil prokaryote and eukaryote rRNA gene sequences listed in the NCBI nucleotide sequences database on 20 may 2009.

Search term "16S soil"	
uncultured bacterium	194944
uncultured archaeon (archaea)	15589
uncultured soil bacterium	8482
uncultured Acidobacteria bacterium	6066
uncultured actinobacterium	4720
uncultured beta proteobacterium	3455
uncultured Firmicutes bacterium	2589
uncultured alpha proteobacterium	2581
uncultured proteobacterium	2319
uncultured gamma proteobacterium	1738
unidentified bacterium	1687
uncultured Bacteroidetes bacterium	1465
uncultured planctomycete	1355
uncultured delta proteobacterium	1016
uncultured crenarchaeote (archaea)	913
uncultured Verrucomicrobia bacterium	811
uncultured Acidobacteriales bacterium	670
uncultured Gemmatimonas sp.	601
uncultured Gemmatimonadetes bacterium	571
uncultured forest soil bacterium	537
All other taxa	43385
Search term "18S soil"	
uncultured fungus	6564
uncultured soil fungus	5976
uncultured eukaryote	2342
unidentified nematode Sourhope farm	*2113*
uncultured Ascomycota	844
uncultured Glomus	591
uncultured basidiomycete	479
uncultured Eimeriidae (protist)	*384*
uncultured Agaricomycetes	263
Phaseoleae environmental sample (plant)	*242*
uncultured Boletaceae	226
uncultured Pezizomycotina	179
uncultured Tomentella	171
uncultured nematode	*165*
uncultured cercozoan (protist)	*149*
Acrobeloides sp. *Sourhope farm (nematode)*	*132*
Oxytrichidae environmental sample (protist)	*130*
uncultured glomeromycete	124
uncultured Agaricomycetidae	122
uncultured Boletales	115
All other taxa	8584

The general terms "bacterium", "archaeon", "fungus" include the named bacterial, archaeal and fungal groups; 18S results in italics are not fungi.

a single copy per geome, and therefore give a more reliable indication of the relative abundance of different microbial goups than the 16S rRNA genes. The DNA sequence can be translated into the respective amino acid sequence to provide a more conserved identifier (as up to six different codons are used for the same amino acid), enabling comparisons between more distantly related organisms (Watanabe et al., 2001). Sufficient DNA sequence data on some protein-encoding genes is now available in the publically-available databases EMBL (http://www.ebi.ac.uk/embl/) and NCBI (http://www.ncbi.nlm.nih.gov/Genbank/) to provide alternative means of generating phylogenies, including *gyrB* (and its paralogue *parE*) which encodes DNA gyrase subunit B (a bacterial type II topoisomerase), *rpoD* (RNA polymerase σ70) and *fliC* (encodes a structural flagellin gene). The most extensive dataset is for *gyrB* (1.4×10^4 nucleotide sequence accessions in the NCBI database) in contrast to more than 1.4×10^6 for 16S rRNA genes, illustrating the continuing importance of the latter.

5. Analysis of the soil metagenome

The vast numbers and great diversity of soil microorganisms, together with the heterogeneity of the soil environment, pose a major practical constraint and results from the limited number of samples that it is possible to analyse for any particular study require careful interpretation. All the methods for investigation of microbial community diversity and activity contain inherent biases and it is necessary to understand the underlying mechanisms to appreciate the strengths and weaknesses of each approach. Table 2 provides a summary of molecular methods based on nucleic acids, discussed in more detail below along with some examples of the application of these methods in soil microbial ecology by our group.

5.1. PCR and quantitative PCR

PCR exploits the semi-conservative replication of DNA to enable exponential amplification of the target sequence and can produce $>10^9$ copies after 30 cycles of DNA synthesis. The first applications of PCR to identify particular groups of organisms in soil were only useful to determine if DNA from that organism was present or absent in the sample under investigation and did not provide information on abundance. Nevertheless, this enabled monitoring of soil-borne plant pathogens (Volossiouk et al., 1995) and genetically modified bacteria following field release (Cullen et al., 1998). Subsequently, PCR methods which were quantitative were developed. Initially, amount of product was compared to a synthetic fragment amplified by the same primers and added to the reaction at a known concentration and this has enabled, for example, measurement of the population dynamics in the field of non-culturable ammonia oxidising bacteria following fertilizer additions (Mendum et al., 1999) and of nematophagous fungi used for biological control of nematodes (Mauchline et al., 2002; Atkins et al., 2003). There are now several different commercial systems that offer quantitative real-time PCR (qPCR), using fluorescently-labelled probes and dyes to measure the progression of DNA amplification during the reaction. It is also possible to apply qPCR to measure RNA if it is first converted to DNA. This makes it possible to assess many different samples simultaneously and has improved the efficiency with which abundance of specific soil microbial populations can be measured (Sharma et al., 2007). We have used real-time qPCR with primers specific for the 16S rRNA gene of the genus *Pseudomonas*, to investigate the presence of this group in fresh or archived (air-dried, milled, stored for >100 years) soils from Rothamsted long-term experiments (Clark and Hirsch, 2008). No colonies were obtained from the archived soils on an agar medium selective for *Pseudomonas,* but results (Fig. 4) show how comparison of the amplification rate with standard reactions enabled estimation of the number of *Pseudomonas* genomes g^{-1} soil, demonstrating survival of genomic DNA, although up to 100-fold more was present in fresh soil.

5.2. Assessing the diversity of PCR products

The diversity of microbial communities can be investigated by separating out PCR products derived from universal primers. The rRNA genes have been used extensively for this purpose: primers that bind to regions that are conserved in the majority of eukaryotic or prokaryotic genes, but which flank a variable region, will generate a product (amplicon) with an internal sequence that will vary according to the source from which it was amplified. The same principle applies to any other genes where sufficient sequence information is available for designing primers. Whether derived from rRNA or functional genes, amplicons can be separated by cloning and identified by subsequent DNA sequence analysis, or hybridization to specific probes, but the number that it is possible to analyse has been limited by practical considerations and only the

Table 2
Comparison of molecular methods.

		Presence	Diversity	Abundance	Activity	Resolution
DNA	PCR	+ gels	+ DGGE; TRFLP; ARISA	+ qPCR	−	Sensitive detection: to isolate level with specific primers
	Microarrays (functional and phylogenetic genes)	+	+	(+)	−	To functional group, genus or species - depends on microarray
	High-throughput direct sequencing	+	+	(+)	-	To genus or species if abundant; depends on sequencing effort
RNA (mRNA; rRNA)	Reverse transcriptase (RT) PCR	+	+ DGGE; TRFLP	+ RT-qPCR	(+)	To genus/species with 16S rRNA; accentuates the most abundant
	Microarrays (functional & phylogenetic genes)	+	+	(+)	(+)	16S rRNA similar to DNA; mRNA to functional group
	High-throughput direct sequencing	+	+	(+)	(+)	To genus or species if abundant; depends on effort
SIP and BrdU	Label actively-growing cells	+	+	(+)	+	Sensitive detection to genus, species or functional group
Protein	Immunological assays may detect certain proteins (technically difficult in soil)					In some cases, sensitive to pathotype or species
Cell/system	Functional assays	(+)	(+)	+	+	To functional group only

+ suitable method; (+) indirect or semi-quantitative information; − inappropriate method.

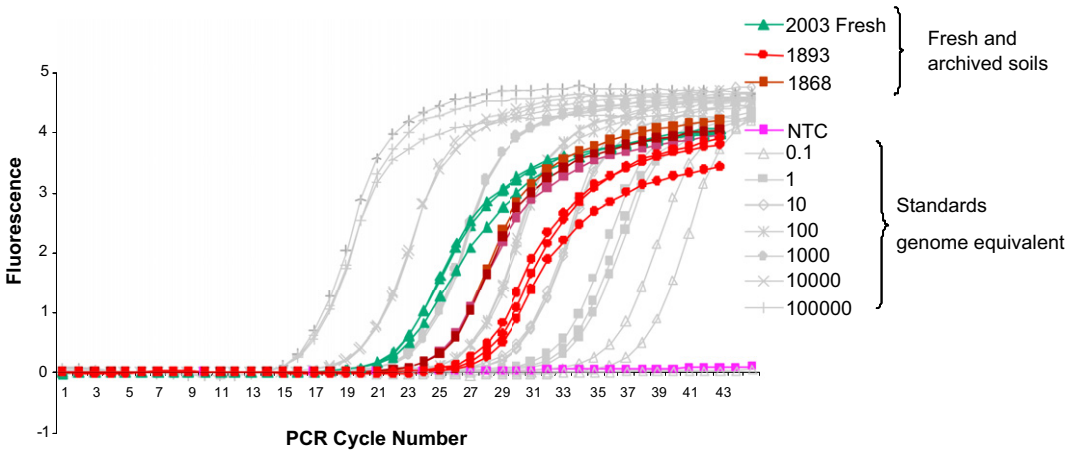

Fig. 4. Survival of *Pseudomonas* DNA in archived and fresh soils using quantitative real-time PCR. The rate of amplification of the target sequence in each sample is detected by fluorescence, and compared to standards with a known amount of *Pseudomonas* DNA. The results show here that *Pseudomonas* genomes are more abundant in the 2003 sample and least abundant in the 1893 sample.

most abundant types could be identified. However, new methods for massively parallel high-throughput sequencing (see Section 5.4), used to target particular amplicons, can now provide much greater coverage of diversity and better estimates of the relative abundance of types. Amplicons can also be separated by denaturing gradient gel electrophoresis (DGGE), where migration in the gel depends on sequence composition (Muyzer et al., 1993). The DGGE-separated bands, most containing many copies of a single amplicon (although some will be composed of multiple amplicons which co-migrate) can be excised and sequenced. Cloning of the excised DNA prior to sequencing of a number of clones will reveal if more than one amplicon is present in a single band. Alternatively one or both primers can be fluorescently-labelled and the PCR products digested with restriction enzymes to reveal terminal restriction fragment polymorphisms (TRFLP) which are then separated by capillary or gel electrophoresis (Avaniss-Aghajani et al., 1994; Liu et al., 1997). Several different fluorescently-labelled PCR targets can be run in the same reaction (multiplexing), for example to compare bacterial, fungal and archaeal communities (Macdonald et al., 2008). The inclusion of internal size markers in each run provides a more reproducible fingerprint pattern than DGGE and the labelled fragment (TRF) size can be compared to a database from known phyla to identify the source organism although a problem arises from the chance occurrence of identical TRFs arising from unrelated phyla. A similar method that provides more discrimination between groups is automated ribosomal intergenic spacer analysis (ARISA). This compares the length of the variable spacer regions separating the ribosomal genes (Fischer and Triplett, 1999; Danovaro et al., 2006). A less commonly used method to differentiate between same-sized PCR products is single strand conformation polymorphism (SSCP) which relies on differential electrophoretic mobility of single-stranded DNA products (Schweiger and Tebbe, 1998). These methods provide an estimate of the diversity within a community, based on the number of amplicons of each type (each representing an "operational taxonomic unit" or OTU, assumed to be equivalent to a bacterial species or genus) that can be observed. A recent comparison of several soils found that DGGE and SSCP gave similar results and that TRFLP detected fewer types but gave more reproducible results so was more suitable for comparing large numbers of samples (Smalla et al., 2007). The shortcomings of these methods arise from the relatively small number of types that can be distinguished because only a limited number of amplicon types (bands or peaks, depending on the system) can be detected, and because of the potential for co-migration by different amplicons. Detection depends not only on the abundance and species diversity (often referred to as "species richness") but also on the relative abundance compared to the total population (referred to as "evenness"). If only a low fraction of the present populations is detected, a change in the number of DGGE bands or TRFLP peaks could reflect a change in rank-abundance of populations (i.e. in the number of populations above the threshold of detection) rather than a change in richness or diversity. The methods and problems in interpretation have been reviewed recently (Forney et al., 2004; Nakatsu, 2007). We illustrate the problems of DGGE interpretation with Fig. 5, where the

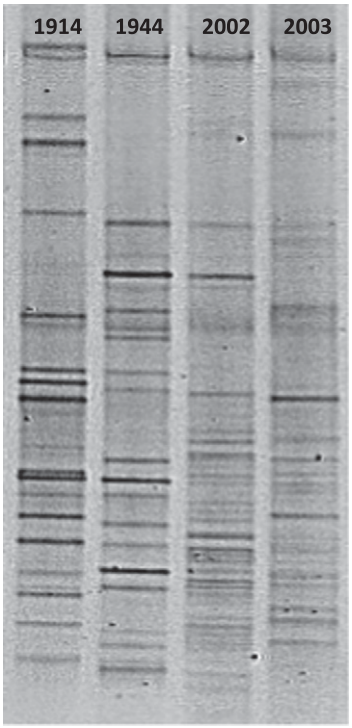

Fig. 5. DGGE profiles of 16S rRNA genes in archived soils. The recently archived samples appear to have more bands of similar relative intensity compared to the older samples where fewer, more intense bands indicate differential survival of DNA from some bacterial groups.

16S rRNA gene diversity might, on first glance, appear greater in the old archived soils from long-term experiments at Rothamsted. There are many, "blurred" bands in the recently archived 2003 and 2002 samples, which are difficult to count but on close inspection are more abundant compared to those stored since 1944 and 1914.

Despite their limitations, these methods have provided a new dimension to the study of soil microbial ecology, in particular for assessing community diversity based on the 16S or 18S rRNA genes. The methods can also be applied to investigate the structure of specific communities if genes diagnostic for that community are known and there is sufficient information to design primers that recognize conserved sequences that flank variable regions. Examples include functional genes *amoA* – ammonia monooxygenase (Bernhard et al., 2005; Nicol et al., 2008; Principi et al., 2009); *nirK*, *nirS* – nitrite reductases (Sharma et al., 2006); *nifH* – nitrogenase (Eilmus et al., 2007; Ogilvie et al., 2008) and genus-specific 16S rRNA gene sequences (Clark and Hirsch, 2008). Although there are programmes to design degenerate primers (i.e. primers that tolerate some sequence variation) for genes other than 16S rRNA, using the NCBI or EMBL databases, there are limits to the degree of primer degeneracy that can be tolerated, in particular at the 3′ ends (Kwok et al., 1994). Also, there is a lack of suitable software to compare these to the databases to determine their true specificity. Consequentially, some degenerate primer sets are not fully specific for their target gene, but amplify other targets that may or may not be related; only sequence analysis of amplicons can confirm this.

5.3. The potential of microarrays in soil

Over the last few years, microarrays have been developed, based on bacterial sequences deposited in public databases including GenBank at the NCBI. Microarray applications for soil microbial ecology have been reviewed recently (Sessitsch et al., 2003). These include several phylogenetic microarrays based on 16S rRNA gene sequence databases such as the PhyloChip developed by Gary Andersen and his team at the Lawrence Berkeley National Lab, USA, a high-density array with 500,000 probes in total, identifying ~9000 species/taxa. These microarrays are regularly updated as more information becomes available from the rapidly expanding DNA databases. They are reported to reveal greater soil bacterial phylum diversity and many more individual taxa, compared to more conventional cloning/sequencing methods (DeSantis et al., 2007). In addition to identifying the diversity of communities, such arrays can give some indication of the relative abundance of different taxa in different treatments, for example in the rhizosphere compared to bulk soil (DeAngelis et al., 2009). If rRNA is used rather than genomic DNA, it can indicate the active components of the community. However, it cannot inform on the diversity of any as yet unknown microbial groups not represented in the array, nor functional genes responsible for important activities in soil. In contrast to phylochips, the GeoChip, developed by Jizhong Zhou at the University of Oklahoma, USA, is based on functional gene sequences. GeoChip 2, which theoretically can detect >10,000 genes in >150 functional groups, has been used, for example, to demonstrate a general trend of increasing number and diversity of genes involved in organic C decomposition in soils with increasing C content (Zhang et al., 2007). The probes identify the functional genes from various different taxa, and thus provide some information on overall diversity. An updated version, GeoChip 3, should detect 47,000 genes in 292 gene families; this is likely to increase as more sequence data becomes available. If soil-extracted RNA is used rather than DNA, it can give an indication of which genes are active immediately prior to sampling. However, microarrays depend on prior knowledge of DNA sequences and so they cannot identify as yet unknown groups that might be essential to particular functions. Similarly, there is little information on the taxonomic diversity and functional gene sequences of soil fungi and therefore microarrays for the fungal communities in soil are some way off.

5.4. High-throughput sequencing

The conventional cloning and Sanger sequencing methods are time consuming and limit the number of samples that can be processed. Methods are now being developed to automate the diagnostic process ("lab-on-a-chip") which provide the possibility of investigating microbial functional diversity in many thousands of samples. Currently, massively parallel high-throughput pyrosequencing methods can process hundreds of thousands of sequences simultaneously. An example of this approach is the technology "454 sequencing" (http://www.454.com/enabling-technology/the-system.asp), where, in 2009, the current "GS FLX Titanium Series" system claims to sequence one million fragments with an average read length of 400 b per read and aims to read 800 b per read in the near future. Other commercial high-throughput sequencing systems provide a greater number of shorter sequences - company websites provide specifications for the latest models. For example, at the time of writing, the Illumina Inc. Genome Analyser$_{IIx}$ claims to sequence 20–25 Gb per 9–10 day run with reads of up to 75 b and the Applied Biosystems Inc. SOLiD™ 3 system claims 10–15 Gb per 6–7 day run with reads of up to 50 b. This technology is changing rapidly (Mardis, 2008), new methods such as nanopore sequencing are being developed (Stoddart et al., 2009) and various internet discussion forums compare the benefits and drawbacks of each system, a task beyond the scope of this review.

DNA extracted from soil can be subjected to PCR amplification prior to sequencing using selective primers that are specific for particular genes, and this will change the nature of the DNA that is to be sequenced. For example, universal 16S rRNA gene primers pre-selected bacterial and archaeal genes prior to sequencing to establish the relative number of bacterial and archaeal 16S rRNA genes (the mean from four soils was ~140,000 bacterial; ~9000 archaeal sequences) as well as to establish the overall taxonomic diversity of these organisms in soil (Roesch et al., 2007). A limitation of this approach is that it assumes that all prokaryotes possess 16S rRNA gene sequences that are homologous to the primers used for the PCR amplification step. This is not necessarily the case as there are some examples of bacteria and archaea that differ to the consensus primer sequence for this gene, which are based on known, mostly culturable organisms. Nevertheless, as the technology improves, and longer fragments can be sequenced in each run, it will become even more useful as a tool.

Current methods for high-throughput sequencing cannot detect less abundant but ecologically essential groups without pre-selection, thus negating the benefit of direct and unbiased sampling. However, with the new/improved methods being developed, sequencing will become increasingly efficient and important and in the future is likely to be limited only by the bioinformatic analysis of sequence data (Pop and Salzberg, 2008).

5.5. Bioinformatic analysis

The metagenomic methods described above produce very large data sets, presenting considerable challenges for analysis. The human genome project was concerned with a single species; the soil metagenome includes millions, many of which are not currently described. The initial computational approach depends on the sequencing "platform" (i.e. method); individual software

packages have been developed for each system for the initial analysis, but to identify the likely genes and taxonomic groups represented by the sequence data and the relative abundance with which they occur, programmes need to be linked together to create automated "pipelines". When single genomes are sequenced using short reads, overlapping stretches are identified to allow "assembly" of fragments into longer sections; this is made simpler when the genome sequence of a related organism is available as a "scaffold". Longer sequences can be identified by comparison to the NCBI DNA and protein databases (after translation to the corresponding amino acids) using the BLAST (Basic Local Alignment Search Tool) family of programmes (Altschul et al., 1990). The optimal alignment of multiple sequences prior is critically important for subsequent analyses. Until recently, the most widely used alignment programme has been CLUSTALW (Thompson et al., 1994) but several improved alternatives have been developed including T-COFFEE and PROBCONS (Edgar and Batzoglou, 2006). The closest relatives identified and phylogenetic trees can then be inferred using the PHYLogeny Inference Package PHYLIP (Felsenstein, 1981). The process of identifying genes and thus assigning biological information to sequences is termed "genome annotation". However, short sequences are difficult to assign with any certainty and there is no appropriate software that can assemble de novo genomes from short reads (i.e. <50 b), with significant inaccuracies predicted with longer reads of 100 b (Womack et al., 2008). Thus at present, reconstruction of microbial genomes from highly diverse mixed populations (for which there may or may not be related reference genomes in existing databases to act as scaffolds) is not possible using high-throughput, short read-length technologies. Rather, a combination of approaches using short, medium and long read-length methods would be required. However, over time, as more reference microbial genomes are added to the databases, and as read lengths increase, these problems may diminish.

With the ever increasing size of metagenomic projects and publically-available databases the requirement for high-performance computing and automated software has become a limiting factor. To overcome this, several groups are setting up freely available pipelines for the analysis of metagenomic sequence data. Initially set up for genome annotation of single microbes, some now allow the automated annotation of sequence fragments from environmental metagenomic projects, their subsequent phylogenetic classification and an initial metabolic pathway reconstruction. One example is mg-RAST, the metagenomics RAST server (Meyer et al., 2008) which has been used to identify discriminatory metabolic profiles of nine different biomes (Dinsdale et al., 2008). These were sampled from marine, freshwater, animal and subterranean environments but did not include soil! At the time of writing mg-RAST utilises 47 archaeal, 725 bacterial and 29 eukaryotic genomes held at the National Microbial Pathogen Data Resource. Sequence data is also analysed using BLASTX (in translation), ribosomal databases (GREENGENES, RDP-II and the European 16S rRNA database) and the chloroplast, mitochondria and mobile elements database ACLAIME. A different approach is taken by the freely available stand alone program MEGAN, which provides visual outputs and can rapidly assess the biodiversity of metagenomic samples to allow comparative studies of different data sets that can include some functional assessments and metadata (Huson et al., 2007).

An important requirement for submitting metagenome data to mg-RAST and to any other database should be the inclusion of appropriate metadata MIGS (Minimum Information about a Genome Sequence). This should include soil and climatic data, land management and any biological process measurements and will enable comparison of different metagenomic projects and to identify the relative impacts of these parameters.

6. Future prospects

The ultimate goal of the soil microbial ecologist is to achieve sufficient understanding of the interactions within microorganism communities and with their environment to be able to predict impacts of change and manage soil functions. Soil metagenomic data will reveal more of the range of potential activities present in microbial communities as methods improve. High-throughput studies based on RNA to indicate gene activity, "Transcriptomics", are being developed in parallel with the high-throughput sequencing and microarray methods described above, requiring a preliminary step to convert RNA to DNA (Urich et al., 2008). However, the other "omics", proteomics and metabolomics, valuable in describing interactions within organisms or in simple interactions, have had limited success in soil up to now. The constraints are mainly technical: proteins have many different affinities with different mineral and organic constituents of soil; the diverse molecular components of cells defined as the "metabolome" pose even greater problems. Nevertheless, there are already techniques for extracting biochemical markers from soil, as described earlier and the coupling of genomic and proteomic approaches has been successful in examining less complex microbial communities such as those found in biofilms in acid mine drainage systems (Ram et al., 2005). Promising techniques include protein-SIP (Jehmlich et al., 2008) and isolation of specific fractions of the soil proteome (Schulze et al., 2005). The use of post-genomic techniques for complex communities are currently in their infancy, but likely to become more and more prominent, not only when reliable universal extraction methods become available but also when highly sophisticated computer software is available to process the wealth of data that these techniques will undoubtedly produce.

7. Summary

The vast numbers and great diversity of soil microorganisms, together with the heterogeneity of the soil environment, pose major problems for analysing microbial population diversity and structure and linking them to functional processes. Results from the limited number of samples that it is possible to analyse for any particular study require careful interpretation, however, new methods will greatly increase the number of samples that can be analysed in the future. All methods for the investigation of microbial community diversity and activity contain inherent biases and it is necessary to understand the underlying mechanisms in order to be aware of the drawbacks and limitations, and to appreciate the strengths and weaknesses of each approach. Nevertheless, these methods are starting to dissect the soil microbial biomass and the soil metagenome, and will, in the future, enable a greatly improved understanding of microbial community dynamics and interactions relevant to soil functions.

Acknowledgements

Rothamsted Research receives grant aided support from the Biotechnology and Biological Sciences Research Council. We thank the Lawes Agricultural Trust for funding the Rothamsted Sample Archive and gratefully acknowledge collaboration with Richard Cookson (University of Western Australia) for the BrdU experiment.

References

Aakra, A., Hesselsoe, M., Bakken, L.R., 2000. Surface attachment of ammonia oxidizing bacteria in soil. Microbial Ecology 39, 222–235.

Abaye, D.A., Lawlor, K., Hirsch, P.R., Brookes, P.C., 2005. Changes in the microbial community of an arable soil caused by long-term metal contamination. European Journal of Soil Science 56, 93–102.

Abulencia, C.B., Wyborski, D.L., Garcia, J.A., Podar, M., Chen, W., Chang, S.H., Chang, H.W., Watson, D., Brodie, E.L., Hazen, T.C., Keller, M., 2006. Environmental whole-genome amplification to access microbial populations in contaminated sediments. Applied and Environmental Microbiology 72, 3291–3301.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403–410.

Artursson, V., Jansson, J.K., 2003. Use of bromodeoxyuridine immunocapture to identify active bacteria associated with arbuscular mycorrhizal hyphae. Applied and Environmental Microbiology 69, 6208–6215.

Atkins, S.D., Clark, I.M., Sosnowska, D., Hirsch, P.R., Kerry, B.R., 2003. Detection and quantification of *Plectosphaerella cucumerina*, a potential biological control agent of potato cyst nematodes, by using conventional PCR, real-time PCR, selective media, and baiting. Applied and Environmental Microbiology 69, 4788–4793.

Avaniss-Aghajani, E., Jones, K., Chapman, D., Brunk, C., 1994. A molecular technique for identification of bacteria using small subunit ribosomal RNA sequences. BioTechniques 17, 144–149.

Bernhard, A.E., Donn, T., Giblin, A.E., Stahl, D.A., 2005. Loss of diversity of ammonia-oxidizing bacteria correlates with increasing salinity in an estuary system. Environmental Microbiology 7, 1289–1297.

Berry, A.E., Chiocchini, C., Selby, T., Sosio, M., Wellington, E.M.H., 2003. Isolation of high molecular weight DNA from soil for cloning into BAC vectors. FEMS Microbiology Letters 223, 15–20.

Bertrand, H., Poly, F., Van, V.T., Lombard, N., Nalin, R., Vogel, T.M., Simonet, P., 2005. High molecular weight DNA recovery from soils prerequisite for biotechnological metagenomic library construction. Journal of Microbiological Methods 62, 1–11.

Borneman, J., 1999. Culture-independent identification of microorganisms that respond to specified stimuli. Applied and Environmental Microbiology 65, 3398–3400.

Bouvier, T., del Giorgio, P.A., 2003. Factors influencing the detection of bacterial cells using fluorescence in situ hybridization (FISH): a quantitative review of published reports. FEMS Microbiology Ecology 44, 3–15.

Bull, I.D., Parekh, N.R., Hall, G.H., Ineson, P., Evershed, R.P., 2000. Detection and classification of atmospheric methane oxidizing bacteria in soil. Nature 405, 175–178.

Chen, Y., Dumont, M.G., Neufeld, J.D., Bodrossy, L., Stralis-Pavese, N., McNamara, N. P., Ostle, N., Briones, M.J., Murrell, J.C., 2008. Revealing the uncultivated majority: combining DNA stable-isotope probing, multiple displacement amplification and metagenomic analyses of uncultivated *Methylocystis* in acidic peatlands. Environmental Microbiology 10, 2609–2622.

Clark, I.M., Hirsch, P.R., 2008. The impact of soil organic matter on the diversity and long-term survival of viable bacteria and bacterial DNA in archived soils from the Rothamsted Broadbalk experiment. Soil Biology and Biochemistry 40, 1090–1102.

Courtois, S., Frostegård, A., Göransson, P., Depret, G., Jeannin, P., Simonet, P., 2001. Quantification of bacterial subgroups in soil: comparison of DNA extracted directly from soil or from cells previously released by density gradient centrifugation. Environmental Microbiology 3, 431–439.

Cullen, D.W., Hirsch, P.R., 1998. Simple and rapid method for direct extraction of microbial DNA from soil for PCR. Soil Biology and Biochemistry 30, 983–993.

Cullen, D.W., Nicholson, P.S., Mendum, T.A., Hirsch, P.R., 1998. Monitoring genetically-modified rhizobia in field soils using the polymerase chain reaction. Journal of Applied Microbiology 84, 1025–1034.

Curtis, T.P., Sloan, W.T., 2004. Prokaryotic diversity and its limits: microbial community structure in nature and implications for microbial ecology. Current Opinion in Microbiology 7, 221–226.

Curtis, T.P., Sloan, W.T., 2005. Exploring microbial diversity – a vast below. Science 309, 1331–1333.

Danovaro, R., Luna, G.M., Dell'Anno, A., Pietrangeli, B., 2006. Comparison of two fingerprinting techniques, terminal restriction fragment length polymorphism and automated ribosomal intergenic spacer analysis, for determination of bacterial diversity in aquatic environments. Applied and Environmental Microbiology 72, 5982–5989.

Davis, K.E.R., Joseph, S.J., Janssen, P.H., 2005. Effects of growth medium, inoculum size and incubation time on culturability and isolation of soil bacteria. Applied and Environmental Microbiology 71, 826–834.

DeAngelis, K.M., Brodie, E.L., DeSantis, T.Z., Andersen, G.L., Lindow, S.E., Firestone, M.K., 2009. Selective progressive response of soil microbial community to wild oat roots. ISME Journal 3, 168–178.

DeSantis, T.Z., Brodie, E.L., Moberg, J.P., Zubieta, I.X., Piceno, Y.M., Andersen, G.L., 2007. High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. Microbial Ecology 53, 371–383.

Dinsdale, E.A., Edwards, R.A., Hall, D., Angly, F., Breitbart, M., Brulc, J.M., Furlan, M., Desnues, C., Haynes, M., Li, L., McDaniel, L., Moran, M.A., Nelson, K.E., Nilsson, C., Olson, R., Paul, J., Brito, B.R., Ruanm, Y., Swan, B.K., Stevens, R., Valentine, D.L., Thurber, R.V., Wegley, L., White, B.A., Rohwer, F., 2008. Functional metagenomic profiling of nine biomes. Nature 452, 629–632.

Dumont, M.G., Murrell, J.C., 2005. Stable isotope probing – linking microbial identity to function. Nature Reviews Microbiology 3, 499–504.

Edgar, R.C., Batzoglou, S., 2006. Multiple sequence alignment. Current Opinion in Structural Biolopgy 16, 368–373.

Eilmus, S., Rösch, C., Bothe, H., 2007. Prokaryotic life in a potash-polluted marsh with emphasis on N-metabolizing microorganisms. Environmental Pollution 146, 478–479.

Felsenstein, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. Journal of Molecular Evolution 17, 368–376.

Fischer, M.M., Triplett, E., 1999. Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. Applied and Environmental Microbiology 65, 4630–4636.

Fleming, T., Yao, W.H., Sayler, G.S., 1998. Optimization of differential display of prokaryotic mRNA: application to pure culture and soil microcosms. Applied and Environmental Microbiology 69, 3698–3706.

Forney, L.J., Zhou, X., Brown, C.J., 2004. Molecular microbial ecology: land of the one-eyed king. Current Opinion in Microbiology 7, 210–220.

Gans, J., Wolinsky, M., Dunbar, J., 2005. Computational improvements reveal great bacterial diversity and high metal toxicity in soil. Science 309, 1387–1390.

Gattinger, A., Ruser, R., Schloter, M., Munch, J.C., 2002. Microbial community structure varies in different soil zones in a potato field. Journal of Plant Nutrition and Soil Science 165, 421–428.

Gonzalez, J.M., Portillo, M.C., Saiz-Jimenez, C., 2005. Multiple displacement amplification as a pre-polymerase chain reaction (pre-PCR) to process difficult to amplify samples and low copy number sequences from natural environments. Environmental Microbiology 7, 1024–1028.

Green, J., Bohannan, B.J.M., 2006. Spatial scaling of microbial biodiversity. Trends in Ecology and Evolution 21, 501–507.

Hawkins, T.L., Detter, J.C., Richardson, P.M., 2002. Whole genome amplification – applications and advances. Current Opinion in Biotechnology 13, 65–67.

Hjort, K., Lembke, A., Speksnijder, A., Smalla, K., Jansson, J.K., 2007. Community structure of actively growing bacterial populations in plant pathogen suppressive soil. Microbial Ecology 53, 399–413.

Huson, D.H., Auch, A.F., Qi, J., Schuster, S.C., 2007. MEGAN analysis of metagenomic data. Genome Research 17, 377–386.

Janssen, P.H., 2006. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. Applied and Environmental Microbiology 72, 1719–1728.

Jansson, J.K., 2003. Marker and reporter genes: illuminating tools for environmental microbiologists. Current Opinion in Microbiology 6, 310–316.

Jehmlich, N., Schmidt, F., von Bergen, M., Richnow, H.H., Vogt, C., 2008. Protein-based stable isotope probing (Protein-SIP) reveals active species within anoxic mixed cultures. ISME Journal 2, 1122–1133.

Kreuzer-Martin, H.W., 2007. Stable isotope probing: linking functional activity to specific members of microbial communities. Soil Science Society of America Journal 71, 611–619.

Kwok, S., Chang, S.-Y., Sninsky, J.J., Wang, A., 1994. A guide to the design and use of mismatched and degenerate primers. Genome Research 3, S39–S47.

Lee, Z.M., Bussema 3rd, C., Schmidt, T.M., 2009. *rrn*DB: documenting the number of rRNA and tRNA genes in bacteria and archaea. Nucleic Acids Research 37, D489–D493.

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

Lueders, T., Manefield, M., Friedrich, M.W., 2004. Enhanced sensitivity of DNA- and rRNA-based stable isotope probing by fractionation and quantitative analysis of isopycnic centrifugation gradients. Environmental Microbiology 6, 73–78.

Macdonald, C.A., Campbell, C., Bacon, J., Singh, B.K., 2008. Multiple profiling of soil microbial communities identifies potential genetic markers of metal-enriched sewage sludge. FEMS Microbiology Ecology 65, 555–564.

Manefield, M., Whiteley, A.S., Ostle, N., Ineson, P., Bailey, M.J., 2002. Technical considerations for RNA-based stable isotope probing: an approach to associating microbial diversity with microbial community function. Rapid Communications in Mass Spectrometry 16, 2179–2183.

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

Mauchline, T.H., Kerry, B.R., Hirsch, P.R., 2002. Quantification in soil and the rhizosphere of the nematophagous fungus *Verticillium chlamydosporium* by competitive PCR and comparison with selective plating. Applied and Environmental Microbiology 68, 1846–1853.

McGrath, K.C., Thomas-Hall, S.R., Cheng, C.T., Leo, L., Alexa, A., Schmidt, S., Schenk, P.M., 2008. Isolation and analysis of mRNA from environmental microbial communities. Journal of Microbiological Methods 75, 172–176.

Mendum, T.A., Sockett, R.E., Hirsch, P.R., 1998. The detection of Gram-negative bacterial mRNA from soil by RT-PCR. FEMS Microbiology Letters 164, 369–373.

Mendum, T.A., Sockett, R.E., Hirsch, P.R., 1999. Use of molecular and isotopic techniques to monitor the response of autotrophic ammonia-oxidizing populations of the subdivision of the class Proteobacteria in arable soils to nitrogen fertilizer. Applied and Environmental Microbiology 65, 4155–4162.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., Edwards, R.A., 2008. The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics 9, 386.

Moran, M.A., Torsvik, V.L., Torsvik, T., Hodson, R.E., 1993. Direct extraction and purification of rRNA for ecological studies. Applied and Environmental Microbiology 59, 915–918.

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

Nakatsu, C.H., 2007. Soil microbial community analysis using denaturing gradient gel electrophoresis. Soil Science Society of America Journal 71, 562–571.

Nicol, G.W., Leininger, S., Schleper, C., Prosser, J.I., 2008. The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea and bacteria. Environmental Microbiology 10, 2966–2978.

Ogilvie, L.A., Hirsch, P.R., Johnston, A.W., 2008. Bacterial diversity of the broadbalk 'classical' winter wheat experiment in relation to long-term fertilizer inputs. Microbial Ecology 56, 525–537.

Ogram, A., Sun, W., Brockman, F.J., Fredrickson, J.K., 1995. Isolation and characterisation of RNA from low-biomass deep-surface sediments. Applied and Environmental Microbiology 61, 763–768.

Pop, M., Salzberg, S.L., 2008. Bioinformatics challenges of new sequencing technology. Trends in Genetics 24, 142–149.

Priemé, A., Sitaula, J.I.B., Klemedtsson, A.K., Bakken, L.R., 1996. Extraction of methane-oxidizing bacteria from soil particles. FEMS Microbiology Ecology 21, 59–68.

Principi, P., Villa, F., Giussani, B., Zanardini, E., Cappitelli, F., Sorlini, C., 2009. The effect of copper on the structure of the ammonia-oxidizing microbial community in an activated sludge wastewater treatment plant. Microbial Ecology 57, 215–220.

Prosser, J.I., Bohannan, B.J.M., Curtis, T.P., Ellis, R.J., Firestone, M.K., Freckleton, R.P., Green, J.L., Green, L.E., Killham, K., Lennon, J.E., Osborn, A.M., Solan, M., van der Gast, C.J., Young, J.P.W., 2007. The role of ecological theory in microbial ecology. Nature Reviews Microbiology 5, 384–392.

Radajewski, S., Ineson, P., Parekh, N.R., Murrell, J.C., 2000. Stable-isotope probing as a tool in microbial ecology. Nature 403, 646–649.

Ram, R.J., Verberkmoes, N.C., Thelen, M.P., Tyson, G.W., Baker, B.J., Blake 2nd, R.C., Shah, M., Hettich, R.L., Banfield, J.F., 2005. Community proteomics of a natural microbial biofilm. Science 308, 1915–1920.

Remenant, B., Grundmann, G.L., Jocteur-Monrozier, L., 2009. From the micro-scale to the habitat: assessment of soil bacterial community structure as shown by soil structure directed sampling. Soil Biology and Biochemistry 41, 29–36.

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

Roesch, L.F.W., Fulthorpe, R.R., Riva, A., Casella, G., Hadwin, A.K.M., Kent, A.D., Daroub, S., Camargo, F.A.O., Farmerie, W.G., Triplett, E.W., 2007. Pyrosequencing of 16S rRNA reveals contrasts in microbial diversity of four soils. The ISME Journal 1, 283–290.

Rogers, S.W., Moorman, T.B., Ong, S.K., 2007. Fluorescent in situ hybridization and micro-autoradiography applied to ecophysiology in soil. Soil Science Society of America 71, 620–631.

Rondon, M.R., August, P.R., Bettermann, A.D., Brady, S.F., Grossman, T.H., Liles, M.R., Loiacono, K.A., Lynch, B.A., MacNeil, I.A., Minor, C., Tiong, C.L., Gilman, M., Osburne, M.S., Clardy, J., Handelsman, J., Goodman, R.M., 2000. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Applied and Environmental Microbiology 66, 2541–2547.

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

Rosselló-Mora, R., Amann, R., 2001. The species concept for prokaryotes. FEMS Microbiology Reviews 25, 39–67.

Saleh-Lakha, S., Miller, M., Campbell, R.G., Schneider, K., Elahimanesh, P., Hart, M.M., Trevors, J.T., 2005. Microbial gene expression in soil: methods, applications and challenges. Journal of Microbiological Methods 63, 1–19.

Schleuß, U., Müller, F., 2001. Requirements for soil sampling in the context of ecosystem research. Science of the Total Environment 264, 193–197.

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

Schweiger, F., Tebbe, C.C., 1998. A new approach to utilise PCR-single-strand-conformation polymorphism for 16S rRNA gene-based microbial community analysis. Applied and Environmental Microbiology 64, 4870–4876.

Sessitsch, A., Hackl, E., Wenzl, P., Kilian, A., Kostic, T., Stralis-Pavese, N., Sandjong, B.T., Bodrossy, L., 2003. Diagnostic microbial microarrays in soil ecology. New Phytologist 171, 719–7836.

Sharma, S., Szele, Z., Schilling, R., Munch, J.C., Schloter, M., 2006. Influence of freeze-thaw stress on the structure and function of microbial communities and denitrifying populations in soil. Applied and Environmental Microbiology 72, 2148–2154.

Sharma, S., Radl, V., Hai, B., Kloos, K., Fuka, M.M., Engel, M., Schauss, K., Schloter, M., 2007. Quantification of functional genes from prokaryotes in soil by PCR. Journal of Microbiological Methods 68, 445–452.

Skinner, A., Jones, P.C.T., Mollison, J.E., 1952. A comparison of a direct and a plate-counting technique for the quantitative estimation of soil microorganisms. Journal of General Microbiology 6, 261–271.

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

Sørensen, J., 1997. The Rhizosphere as a Habitat for Soil Microorganisms. In: van Elsas, J.D., et al. (Eds.), Modern Soil Microbiology. Marcel Dekker, Inc, New York, pp. 21–45.

Stockdale, E.A., Brookes, P.C., 2006. Detection and quantification of the soil microbial biomass – impacts on the management of agricultural soils. Journal of Agricultural Sciences 144, 285–302.

Stoddart, D., Heron, A.J., Mikhailova, E., Maglia, G., Bayley, H., 2009. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. Proceedings of the National Academy of Sciences 106, 7702–7707.

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

Thakuria, D., Schmidt, O., Mac Siúrtáin, M., Egan, D., Doohan, F.M., 2009. Importance of DNA quality in comparative soil microbial community structure analyses. Soil Biology and Biochemistry 40, 1390–1403.

Thompson, J.D., Higgins, D.G., Gibson, D.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673–4680.

Urich, T., Lanzén, A., Qi, J., Huson, D.H., Schleper, C., Schuster, S.C., 2008. Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. PloS One 3, e2527.

Volossiouk, T., Robb, E.J., Nazar, R.N., 1995. Direct DNA extraction for PCR-mediated assays of soil organisms. Applied and Environmental Microbiology 61, 3972–3976.

Wagner, M., Nielsen, P.H., Loy, A., Nielsen, J.L., Daims, H., 2006. Linking microbial community structure with function: fluorescence in situ hybridization-micro-autoradiography and isotope arrays. Current Opinion in Biotechnology 17, 83–91.

Watanabe, K., Nelson, J., Harayama, S., Kasai, H., 2001. ICB database: the *gyrB* database for identification and classification of bacteria. Nucleic Acids Research 29, 344–345.

Whiteley, A.S., Manefield, M., Lueders, T., 2006. Unlocking the 'microbial black box' using RNA-based stable isotope probing technologies. Current Opinion in Biotechnology 17, 67–71.

Womack, K.E., Bhavsar, J., Ravel, J., 2008. Metagenomics: read length matters. Applied and Environmental Microbiology 74, 1453–1463.

Yin, B., Crowley, D., Sparovek, G., De Melo, W.J., Borneman, J., 2000. Bacterial functional redundancy along a soil reclamation gradient. Applied and Environmental Microbiology 66, 4361–4365.

Zelles, L., Alef, K., 1995. Biomarkers. In: Alef, K., Nannipieri, P. (Eds.), Methods in Applied Soil Microbiology and Biochemistry. Academic Press, London, pp. 422–439.

Zelles, L., Bai, Q., Beck, T., Beese, F., 1992. Signature fatty acids in phospholipids and lipopolysaccharides as indicators of microbial biomass and community structure. Soil Biology and Biochemistry 24, 317–323.

Zhang, Y., Zhang, X., Liu, X., Xiao, Y., Qu, L., Wu, L., Zhou, J., 2007. Microarray-based analysis of changes in diversity of microbial genes involved in organic carbon decomposition following land use/cover changes. FEMS Microbiology Letters 266, 144–151.