Resolving functional diversity in relation to microbial community structure in soil: exploiting genomics and stable isotope probing

Elizabeth MH Wellington*, Andrew Berry and Martin Krsek

The microbial ecology of soil still presents a challenge to microbiologists attempting to establish the ways in which bacteria and fungi actively metabolise substrates, link into food webs and recycle plant and animal remains and provide essential nutrients for plants. Extraction and *in situ* analysis of rRNA has enabled identification of active taxa, and detection of mRNA has provided an insight into the expression of key functional genes in soil. Recent advances in genomic analysis and stable isotope probing are the first steps in resolving the linkage between structure and function in microbial communities.

Addresses
Department of Biological Sciences, University of Warwick, Coventry, CV4 7AL, UK
*Corresponding author
e-mail: ewellington@bio.warwick.ac.uk

Current Opinion in Microbiology 2003, **6**:295–301

This review comes from a themed issue on
Techniques
Edited by Jo Handelsman and Kornelia Smalla

1369-5274/03/$ – see front matter

DOI 10.1016/S1369-5274(03)00066-3

Abbreviations

chi	chitinase
FISH	fluorescence *in situ* hybridisation
HGT	horizontal gene transfer
PLFA	phospholipid fatty acid
RT-PCR	reverse transcription polymerase chain reaction
SIP	stable isotope probing

Introduction

Understanding the ways in which microbial communities function in soil is still a major objective for microbial ecologists. Progress has been made by the application of molecular techniques, which have allowed detailed studies of bacterial and, to a lesser extent, fungal diversity in response to varying environmental factors. The objective has been to target a suitable phylogenetic marker such as 16S rRNA and resolve diversity, which in some cases can be related to functional groups such as nitrifying bacteria [1••], but our lack of knowledge about uncultured groups and their functions means this approach has only limited application. In addition, some important activities in soil are adaptive and will be subject to horizontal gene transfer (HGT) and selection, for example antagonism and antibiotic production [2] and degradation of xenobiotics [3]. Soil is a highly complex environment that acts as a reservoir for microorganisms, with their activity varying over space and time. The complexity of soil biofilms has been reviewed [4] and provides a daunting prospect for elucidation of structure–function relationships. However, recent developments in analysis of RNA are beginning to identify where and when functional groups are active and functional genes expressed. Molecular analysis combined with substrate labelling techniques such as stable isotope probing (SIP) can now explore the identity of specific functional groups. In parallel, some functional data are being retrieved from the uncultured communities inhabiting soil by the construction of environmental gene libraries; a method frequently referred to as metagenomics.

Active versus dormant communities

Much of our knowledge about microbial diversity in soil is based on the extraction and analysis of DNA. This approach has enabled comparative analysis of changes in community structure in relation to environmental factors such as vegetation [5]. For functional gene analysis, it provides information on diversity and potential activity [6•], which is useful for studying impacts of soil treatments on functional potential. Total community DNA can be extracted from cells in various states of activity and many may be starved or dormant structures such as spores. It is unlikely that dead cells will contribute significantly to community DNA as they show a reduced residence time in soil [7] but some may be protected from degradation by adsorption to clays, as is the case with small amounts of free DNA [8]. The rapid developments in reverse transcription-PCR (RT-PCR) and improvement in RT enzymes has provided opportunities for evaluation of active communities inferred from analysis of rRNA and mRNA. The number of ribosomes in a cell is known to correlate with growth rate, but in slow growing bacteria there are exceptions to this such as streptomycete spores the ribosome number of which is approximately 8×10^3 compared with 22×10^3 for actively growing mycelium [9]. Felske *et al.* [10] were the first to use RT-PCR to demonstrate that uncultured soil bacteria were active in soil. Subsequently, many diversity studies have used both DNA and RNA for community structure analysis [1••].

The detection of mRNA is a definitive indicator of activity but it may not always directly equate with phenotypic expression of the targeted gene because of the possibility of post-transcriptional modification of proteins. Attempts have been made to estimate specific activities on the basis of *in situ* derived catabolic mRNA levels [11–13]. This approach can work well if the level of regulation for the functional gene is understood and

detection of mRNA can be correlated specifically with measurement of the activity. The analysis of soil treatments on activities of nitrifying bacteria was assessed by targeting *nifH* and relating mRNA expression to bulk N-fixing activity [14]. Where degenerate primers have been designed there is a possibility that structure–function linkage in relation to activity can be revealed. This would work well with the chitinase (*chi*) gene primers used by Metcalfe *et al.* [6•] for studies with total community DNA, as mRNA clone libraries would reveal active components of the *chi* gene community. The phylogeny of *chi* genes was found to be taxon-specific so the host background for detected *chi* gene sequences could be inferred. This relationship may be subject to disruption by HGT. The diversity of microniches available for colonisation in soil means that there will be hot spots for microbial activity as illustrated for HGT in the rhizosphere [15] so activity analysis at the single-cell level *in situ* would be a feasible way to study spatial relationships of active and inactive groups. The combination of microautoradiography and fluorescence *in situ* hybridisation (FISH) [16] has achieved this but applications so far have been in aquatic environments [17]. Soil presents considerable difficulties for FISH analysis because of background fluorescence and low metabolic activity of soil communities (refer to the review by Wagner *et al.* in this issue).

The recent combination of baiting key functional groups with isotopically labelled substrates with recovery of labelled DNA (SIP) has made a significant breakthrough in linking function to microbial identity.

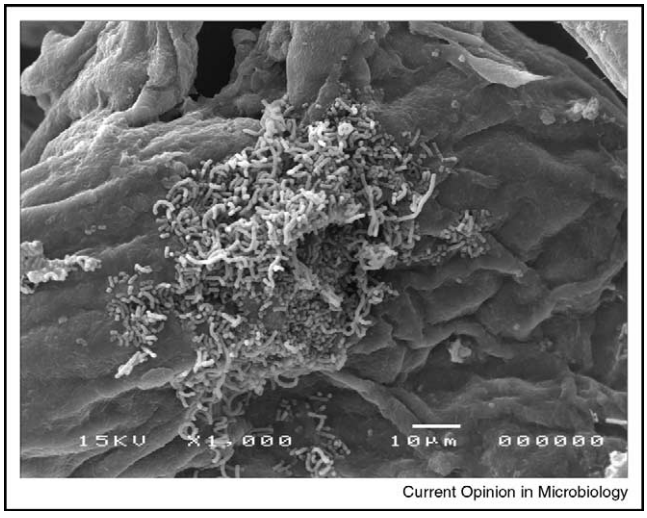
Functional genes in soil

Expansion in our knowledge of key soil processes has provided some target genes for diversity analysis. The nitrogen cycle provides a good example where functional gene microarrays have been developed to monitor dynamics of nitrification and denitrification [18•]. The arrays rely on hybridisation with oligonucleotides encompassing different variants of the genes involved in nitrogen fixation (*nifH*), the first stage of nitrification involving ammonia oxidation (*amoA*), and denitrification (*nirK* and *nirS*). Nitrification is a key microbial-driven process in soil and the oxidation of ammonia to nitrite is attributed to ammonia oxidisers followed by oxidation to nitrate by nitrite oxidisers. Molecular studies of nitrification have concentrated mainly on ammonia oxidisers, and 16S rRNA PCR primers were developed for the detection of known members of this functional group [19]. The diversity within the α-subunit of the ammonia monooxygenase (*amo*A) has been used to evaluate nitrifying potential in grassland [20]. Phylogenetic comparison of 16S rRNA genes and *amo*A genes indicated similar evolutionary relationships, implying that ammonia oxidation is a conserved trait not subject to HGT, and sequence diversity could indicate the origin of the gene thus providing both structural and functional information. Denitrification occurs in four stages via nitrite to NO, N_2O and N_2. Each step of this process is catalysed by different reductases and, in contrast to nitrification, which is achieved by a few well-studied genera of bacteria, a large number of bacteria can reduce nitrate to nitrite in the absence of oxygen. The reduction of NO to N_2O and N_2 is attributed to a more restricted range of groups. Most denitrifiers are common heterotrophic soil genera, such as *Pseudomonas*, *Bacillus* and *Alcaligenes*, which possess NO_3^-, NO_2^- and N_2O reductases, but some lack NO_3^- reductase, and others (principally *Pseudomonas*) lack N_2O reductase. Clearly, it will be difficult to derive structure–function relationships from such a plethora of species, and it is likely that the genes are subject to HGT.

The widespread possession of denitrification genes in many distantly related species of Bacteria (and Archaea), has meant that 16S rRNA-based methods cannot be used to study the functional diversity of the process. Thus, only genes encoding denitrification enzymes can provide functional markers for ecological studies. Complete denitrification requires the sequential action of four enzymes, often biochemically distinct enzymes catalyse the same reaction and it is common for bacteria to have only a part of the pathway. This complexity makes it difficult to use a single gene marker to denote denitrification ability. In a recent study, Gregory *et al.* [21] used two degenerate primer pairs for the membrane-associated nitrate reductase (*narG*) together with a 16S rRNA approach for characterisation of a nitrate-respiring bacterial community. Both 16S rRNA and *narG* phylogenies were only partly congruent, suggesting HGT of the *narG* gene. Results of 16S rRNA-based studies revealed unexpected species richness within a community. The *narG* sequences amplified directly from environmental DNA gave a very different picture to the culture-based results, some grouping with Gram-negative groups of denitrifiers and others creating a new deep-branching clade not affiliated with any known *narG* sequences. The host background for these genes is most likely to be resolved by metagenomics studies.

Chitinolytic ability is ubiquitous among soil microorganisms (and also numerous higher organisms) and chitin is prevalent as an important source of nitrogen in most environments. Many bacteria possess multiple chitinase genes and fungal chitinases may have dual roles in nutrition and hyphal cell wall growth. To enrich for such a diverse and complex functional group, baiting has been used to facilitate retrieval of functional genes. Insoluble substrates such as chitin are amenable to this approach and litter bags have been used to bury and subsequently retrieve substrate in field conditions (Figure 1). Analysis of the substrate by extraction of total community DNA and RNA will provide functional and structural data on major colonists. Degenerate primers were designed to study *chi* gene diversity [22]. Primers specific for family 18

Figure 1



A scanning electron micrograph showing colonised chitin recovered after two months buried in soil and used as bait for microbial chitinase producers (encased in litter bags of 20 mm nylon mesh).

subgroup A bacterial chitinases were used to compare diversity of chitinolytic populations in grassland receiving different treatments [6•]. Clone libraries recovered from total community DNA extracted from chitin baits buried at the grassland site were dominated by *Arthrobacter*-like *chi* sequences in plots receiving sludge, whereas libraries from the untreated plots contained streptomycete-like and *Stenotrophomonas maltophilia*-like chitinase sequences. No arthrobacter isolates were recovered from the baits but they were represented in 16S rRNA libraries from community DNA. It will not be possible to definitively confirm the identity of the dominant functional groups until the recovery of linked phylogenetic markers with the *chi* sequences is achieved; metagenomics holds promise to realise this objective.

Activity measurements of functional genes in the environment can be achieved by bulk assays of the substrate that can be compared with data derived from gene expression via detection of mRNA [23]. Nogales *et al.* [24•] used degenerate primers for five denitrification genes (*narG*, *napA*, *nirS*, *nirK* and *nosZ*) to detect denitrification activity in two river sediment samples. Interestingly, all primer pairs yielded PCR products of the correct size when using community DNA extracted from both samples, confirming the presence of relevant genes in both environments, but the RT-PCR product was obtained only for *nirS* and *nosZ* genes. This could be attributed to low levels of expression despite high abundance of the undetected genes. Extensive sequence diversity for *nirZ* sequences was observed, indicating its presence in diverse bacterial groups. It is likely that detection limits for mRNA are not as low as for DNA because of instability, degradation and the additional RT step.

A major advance in linking functional activity to community structure came with the development of SIP, which relies on the labelling of DNA with ^{13}C, resulting in the separation of heavier labelled DNA during density gradient centrifugation [25]. Labelled DNA can then be analysed for functional and taxonomic marker genes. Radajewski *et al.* pioneered this approach for the study of methylotrophs in soil. *Methylobacterium extorquens* growing on $^{13}CH_3OH$ produced DNA with an increased buoyant density compared with the control growing on unlabelled methanol. Labelled substrate was fed as bait to microbial communities in forest soil microcosms and after gradient centrifugation and PCR with universal 16S rRNA primers and primers for *mxaF* (encodes the active-site subunit of methanol dehydrogenase) sequence libraries were made from the labelled DNA. Two groups were detected in the labelled DNA fraction, one belonging to the α-*Proteobacteria* and the other resembling members of the *Acidobacterium* division. The former was expected, but none of the cultivated members of the latter have been reported to utilise methanol. The DNA-SIP approach can derive activity data from 16S rDNA analysis because to become labelled the bacteria must be actively utilising the substrate. The analysis of functional genes including *mxaF*, *mmoX* and *pmoA* (active site subunits of the soluble and particulate methane monooxygenase) can provide more data on the community structure and diversity in addition to confirming community data derived from 16S rRNA analysis as illustrated by Morris *et al.* [26••].

Labelled CO_2 has also been used to identify active autotrophic ammonia oxidisers in aquatic environments [27•] where specific 16S rRNA primers were used to confirm the activity of a functional group.

SIP was used with extracted labelled RNA by Manefield *et al.* [28••] to identify bacteria responsible for phenol degradation in an anaerobic industrial bioreactor. The use of RNA was proposed as a more responsive biomarker as its turnover is much higher than that of DNA in active cells. Other biomarkers such as phospholipid fatty acid (PLFA) can also be used but the resolution for diversity analysis is less powerful than for sequence analysis. Pelz *et al.* [29] evaluated the combination of isotope (^{13}C labelling of PLFA) and molecular techniques (16S rRNA gene probing) to link toluene degradation to specific microbial groups in denitrifying aquifer microcosms. Results of PLFA analysis agreed with 16S rRNA probing, designating *Azoarcus* spp as a major toluene degrader. The possibility of using mRNA for SIP has yet to be explored.

The SIP approach is not without limitations as there is the necessity of heavy labelling of RNA/DNA for gradient

centrifugation. This can be achieved only using virtually 100% labelled substrate, which can be expensive and often not available for complex insoluble substrates. The second precondition is that the labelled substrate is used almost exclusively for growth, otherwise the label would become too diluted to detect. Heavy labelling of biomolecules also requires long incubation times, which increases the danger of cross-feeding and false results. From this point of view, rRNA labelling, because of its faster turnover and natural abundance, might offer better possibilities for tracking active functional groups in soil.

One common feature of all the above-mentioned approaches is dependence on quality and quantity of data in sequence databases that serve as a source of information for primer/probe design or are directly used for determining the affiliation of sequences detected. Fungal biodiversity is poorly represented in sequence databases and this will limit progress in the study of their ecology.

Detection of function in relation to structure

The contribution thus far, and the future potential of genetic libraries of environmentally isolated DNA as tools to study the ecological relationships, activities and phylogenetic structures of simple or complex microbial communities from oceanic, freshwater, sediment, soils or other environments is in itself an interesting consideration that has been recently reviewed [30]. In essence, DNA isolated from an environment is archived in the form of bacteriophage lambda, cosmid, fosmid or bacterial artificial chromosome (BAC) library. A major advantage of this strategy is that genomic fragments, from both cultivable and uncultivables, are represented in what has been termed a 'metagenomic' library [31]. Therefore, metagenomic libraries, especially BAC libraries, which can maintain inserts of around 100 kb, are a crucial technical link, facilitating the focus of powerful genomics techniques into the realm of the unculturables, which are thought to comprise around 99.9% of all species [32]. The salient and demanding question in this context centres on how data from metagenomic libraries can engage questions relevant to what is often called 'the structure–function relationship' of microbial communities. An immediate consideration is whether it will be feasible to construct metagenomic libraries that are large enough to adequately represent all members of microbial communities, especially in soil environments that, when at their most intricate, can be composed of many thousands of species. Thus, there is a danger that low-abundance species, which are nevertheless involved in important processes, will be overlooked [32]. The metagenomic approach is sequence and informatics intensive, but retains the advantage that it is more independent of *a priori* sequence data of genes key to the 'process' or 'functions' of interest for probe/primer design than PCR-based approaches; and so is not immediately biased by limitations of nucleic acid databases, derived primarily from the culturable world. As a result, it is more likely to reveal unexpected 'genomic based' discoveries of 'novel functionality' or processes. It should be stressed here that metagenomic libraries are not the only method by which whole genes can be retrieved from the environment [33].

In a landmark study, Rondon *et al.* [31] successfully constructed two large BAC libraries from total community DNA isolated from soil. One library contained 24 576 clones with an average insert size of 45 kb, which equates to over 1 Gb of DNA. BAC clones representing 16S rDNA genes affiliated to the proteobacteria, *Cytophagales*, low G+C Gram-positives and *Acidobacteria* were characterised. The *Acidobacterium* are of particular interest as they occur frequently in 16S rDNA-based soil diversity studies, but there are only three cultivated representatives of this phylum. In an analogous study, Béjà *et al.* [34], used BAC libraries to characterise the microbial community from a marine microbial assemblage. They found that the phylogenetic structure of the community as assessed by identifying 16S rRNA gene containing BAC clones did not differ significantly from that observed in parallel 16S rRNA gene libraries. Interestingly, in both cases techniques were developed that effectively overcame the problem of detecting the host 16S rDNA genes when BAC clones were screened for 16S rDNA genes. The crucial point in this context is that both studies have clearly demonstrated the ability of metagenomic BAC libraries as a PCR-bias-free route to access the phylogenetic structure of a community. A natural progression was to sequence the large inserts of BAC clones containing 16S rDNA genes to reveal open reading frames, and perhaps operons that could be ascribed function by sequence comparison, thus providing a direct link between 16S rDNA-based phylogeny and functional genes. Although this is a very sequence and bioinformatics intensive approach it will provide crucial sequence data for microarray design [35]. A striking example of such a linkage was reported by Béjà *et al.* [36]. A 130 kb fragment from the Monterey Bay surface water BAC library included both a 16S rRNA gene, affiliated with the as-yet uncultured SAR86 group of marine γ-*Proteobacteria*, and a gene encoding a light-driven proton-pump similar to bacteriorhodopsins. The gene was expressed in *Escherichia coli* and its functionality demonstrated. This suggests that rhodopsins may have been transferred horizontally from the halophilic archaea to the Eubacteria. Furthermore, it suggests that the SAR86 group, shown by FISH to be abundant in the ocean, has a photoautotrophic or photoheterotrophic lifestyle. This has far-reaching implications for global carbon cycling. This is an excellent example of how 16S rDNA-containing BAC clones provide gene sequence data that has shed light on an important structure–function relationship; but it should also be emphasised that it is fortuitous that the functional 'gene of interest' was located in proximity to the 16S rDNA gene.

Unsurprisingly, no analogous example has yet been found in BAC libraries constructed from soil DNA, although a large 16S rDNA-containing insert has been sequenced revealing many functional genes [31]. The potential remains to identify and characterise genes key to important processes in terrestrial environments. However, only a fraction of functional genes and operons will be located within 100 kb of a 16S rRNA gene and this is not a systematic approach to discovery of phylogenetic structure or functional genes. An observed recovery rate of 3–4% is consistent with theoretical estimations [34]. Perhaps the use of other carefully chosen taxonomic marker genes (possibly *gyrB* or *rpoD*[4]) would increase the possibility of linking a phylogenetic marker to functional genes and operons on a 100 kb insert.

It will not be possible to predict function on the basis of sequence comparison in all cases; even in the genome sequences of well-characterised organisms, a high proportion (30–40%) of the ORFs cannot be ascribed function. Secondly, conclusions about gene function should be treated with care, but the presence of the gene on a BAC clone offers the possibility of expression studies to confirm or deny the predicted function. In a subsequent study, a Monterey Bay surface water BAC library was screened for evidence of bacteriochlorophyll-containing aerobic anoxygenic phototrophic bacteria. A number of BAC clones containing large operons that included genes encoding the photosynthetic reaction centre, carotenoid and bacteriochlorophyll biosynthesis were sequenced. This analysis showed that the most prevalent phototrophs were not those predicted from characterised and cultivated strains [37]. Phenotypical predictions made solely on the basis of gene sequence information can be highly suspect; expression studies or suitable *in situ* data are necessary to attempt to demonstrate this link. The analysis of 16S rDNA containing BAC clones has provided examples to support the assertion that knowledge of microbial ecosystems is seriously restricted by cultivation problems. Although it has already provided striking discoveries of ecological relevance, it is not a systematic strategy to link phylogeny to function and is reliant on the prediction of gene function.

A weakness of the metagenomic library approach as mentioned earlier is that it may quite easily overlook parts of the community that are in low abundance but nevertheless perform essential processes. Thus any enrichment strategy to focus the library on genomes encoding defined processes would be beneficial. There is a natural potential to combine a metagenomic approach with SIP to extend this general concept and link metabolic activity to structure. SIP has been effective in fractionating the genomes of members of the community active in a given biochemical pathway that have utilised and incorporated a ^{13}C substrate into their DNA. 'Heavy DNA' can subsequently be used for metagenomic library construction. This possibility is being explored by Radajewski *et al.* [38]. Analogous enrichment strategies are conceivable where appropriately labelled substrates are available. Similarly, libraries enriched for decomposer activity could be envisaged, although in this case the genomes would not be isotopically fractionated. For example, addition of a complex substrate such as chitin to soil in a spatially defined way to achieve a 'colonisation zone' from which total community DNA could be extracted and used to make metagenomic libraries is feasible [6•].

Functional analysis without cultivation

Whilst the function of cloned genes can be inferred by sequence database comparisons, it is often more satisfactory to demonstrate the activity of the encoded protein. Differences in codon usage and regulatory signals can hinder gene expression in heterologous hosts. Despite this, there have been many reports of both ecology-focused studies and those driven by natural product discovery that utilise expression. An important example in an ecological context is that of Cottrell *et al.* [39]. They used a fluorescent substrate (4-methylumbelliferyl beta-D-*N*,*N*′-diacetylchitobioside) to screen two metagenomic libraries constructed from DNA isolated from two marine sources and discovered and characterised 11 chitinases. Screening by expression has also been effective in BAC libraries. Rondon *et al.* [31], successfully detected DNase, lipase and amylase activities by relatively simple screens on solid media. Later, screening of the same 24 546 clone library led to the characterisation of two closely related antibiotics, Turbomycin A and B; novel triaryl cations with broad spectrum activity [40•]. The production of Turbomycin A and B in *E. coli* was not a result of the cloning of a complete biosynthetic pathway but rather the cloning of a single enzyme gene. The activity of this gene, which showed extensive similarity to 4-hydroxyphenolpyruvate deoxygenase, interacts with the host's tyrosine degradation pathway, resulting in the synthesis of Turbomycin. Brady *et al.* [41] reported the cloning of a four-gene operon encoding the production of the broad-spectrum antibiotic violacein and also deoxyviolacein. The clone was identified by the fact that its expression imparted a blue coloration to the *E. coli* colony. The cluster was shown to be similar to a gene sequence from *Chromobacterium violaceum*.

As is perhaps expected, much of soil metagenomics is driven by natural product discovery and interest is focused on the *Actinobacteria*, which are historically a very rich source of antimicrobial, antitumour and immunosuppressant compounds [42]. Sosio *et al.* [43] have developed BAC shuttle vectors, termed ESACS (*E. coli–Streptomyces* artificial chromosomes) that can replicate autonomously in *E. coli* and integrate into the *Streptomyces* genome at a phage attachment site. This allows cloning, screening and sequencing to be performed conveniently in *E. coli*, while

protein expression can be studied at single copy in a Gram-positive background. Courtois *et al*. have screened a 5000 clone cosmid library to reveal novel polyketide synthases and the production of novel fatty acid dienes [44•]. Wang *et al*. [45] have discovered a novel class of compounds called the terragines by screening large fragment libraries in *Streptomyces lividans*. At present, much of the emphasis in metagenomics research in terrestrial systems is placed on natural product discovery. However, as metagenomics technology develops further and becomes a greater provider of sequence data, thus facilitating microarray design, it may perhaps find a more important and central role in aiding our understanding of complex interactions and processes in terrestrial microbial communities.

Conclusions

There now exists the potential to interrogate total community DNA extracted from soil, which in the context of this review could be referred to as the soil metagenome, to relate the functional capabilities of the microbial community to the biodiversity defined within it. This is important as it allows assessment of community resilience to perturbations such as pollutant inputs and global climate change. There are conflicting theories on the importance of species biodiversity in relation to resilience as it has been proposed that functional diversity is more important than species diversity [46]. It appears from this review that we are discovering some functional redundancy in key microbially driven soil processes but it remains to be seen if this is due to our lack of information on the precise conditions of the soil microniche.

Acknowledgements

We acknowledge financial support from the European Community Grant QLRT-2000-01783 and the Natural Environment Research Council NERC: NER/B/S/2000/00730.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Freitag TE, Prosser JI: **Community structure of ammonia-**
•• **oxidizing bacteria within anoxic marine sediments**. *Appl Environ Microbiol* 2003, **69**:1359-1371.
Natural microbial communities of ammonia-oxidising bacteria were characterised by analysis of 16S rRNA genes amplified using group-specific primers by PCR- and RT-PCR amplification of 16S rDNA and RNA, respectively. Phylogenetic analysis of sequences suggests dominance of *Nitrosospira* cluster-1-like organisms and of strains belonging to a novel cluster.

2. Egan S, Wiener P, Kallifidas D, Wellington EMH: **Analysis of streptomycin biosynthetic gene clusters in streptomycetes isolated from soil**. *Appl Environ Microbiol* 1998, **64**:5061-5063.

3. van der Meer JR, de Vos WM, Harayama S, Zehnder AJ: **Molecular mechanisms of genetic adaptation to xenobiotic compounds**. *Microbiol Rev* 1992, **56**:677-694.

4. Burns RG, Stach JEM: **Microbial ecology of soil biofilms: substrate bioavailability, bioremediation and complexity**. *Developments in Soil Science* 2002, **28B**:17-42.

5. Nüsslein K, Tiedje JM: **Characterization of the dominant and rare members of a young Hawaiian soil bacterial community with small-subunit ribosomal DNA amplified from DNA fractionated on the basis of its guanine and cytosine composition**. *Appl Environ Microbiol* 1998, **64**:1283-1289.

6. Metcalfe AC, Krsek M, Gooday GD, Prosser JI, Wellington EMH:
• **A molecular analysis of a bacterial chitinolytic community in an upland pasture**. *Appl Environ Microbiol* 2002, **68**:5042-5050.
Litter bags filled with chitin were used to enrich for chitinolytic bacterial community structure, which was studied using 16S rRNA genes and chitinase encoding genes. Sequence data revealed a major contribution of unexpected genera to the chitin degradation. Sludge application increased chitinolytic activity and concomitantly decreased diversity of chitinolytic communities.

7. Turpin PE, Maycroft KA, Wellington EMH: **Viable but nonculturable *Salmonellas* in soil**. *J Appl Bacteriol* 1993, **74**:421-427.

8. Lorenz MG, Wackernagel W: **Bacterial gene transfer by natural genetic transformation in the environment**. *Microbiol Rev* 1994, **58**:563-602.

9. Quiros LM, Parra F, Hardisson C: **Structural functional analysis of ribosomal subunits from vegetative *Mycelium* and spores of *Streptomyces antibioticus***. *J Gen Microbiol* 1989, **135**:1661.

10. Felske A, Rheims H, Wolterink A, Stackebrandt E, Akkermans AA: **Ribosome analysis reveals prominent activity of an uncultured member of the class *Actinobacteria* in grassland soils**. *Microbiology* 1997, **143**:2983-2989.

11. Fleming JT, Sanseverino J, Sayler GS: **Quantitative relationship between naphthalene catabolic gene frequency and expression in predicting PAH degradation in soils at town gas manufacturing sites**. *Environ Sci Technol* 1993, **27**:1068-1074.

12. Ogunseitan OA, Olsen BH: **Effect of 2-hydroxybenzoate on the rate of naphthalene mineralization in soil**. *Appl Microbiol Biotechnol* 1993, **38**:799-807.

13. Sayler GS, Fleming JT, Nivens DE: **Gene expression monitoring in soils by mRNA analysis and gene lux fusions**. *Curr Opin Biotechnol* 2001, **12**:455-460.

14. Burgmann H, Widmer F, Sigler WV, Zeyer J: **mRNA extraction and reverse transcription-PCR protocol for detection of *nif*H gene expression by *Azotobacter vinelandii* in soil**. *Appl Environ Microbiol* 2003, **69**:1928-1935.

15. van Elsas JD, Fry J, Hirsch P, Molin S: **Ecology of plasmid transfer and spread**. In *The Horizontal Gene Pool*. Edited by Thomas CM. Amsterdam: Harwood Academic Publishers; 2000:175-199.

16. Lee N, Nielsen PH, Andreasen KH, Juretschko S, Nielsen JL, Schleifer KH, Wagner M: **Combination of fluorescent *in situ* hybridization and microautoradiography - a new tool for structure-function analyses in microbial ecology**. *Appl Environ Microbiol* 1999, **65**:1289-1297.

17. Gray ND, Howarth R, Pickup RW, Jones JG, Head IM: **Use of combined microautoradiography and fluorescence *in situ* hybridization to determine carbon metabolism in mixed natural communities of uncultured bacteria from the genus *Achromatium***. *Appl Environ Microbiol* 2000, **66**:4518-4522.

18. Taroncher-Oldenburg G, Griner EM, Francis CA, Ward BB:
• **Oligonucleotide microarray for the study of functional gene diversity in the nitrogen cycle in the environment**. *Appl Environ Microbiol* 2003, **69**:1159-1171.
This paper describes the development and optimisation of a DNA microarray method for the detection and quantification of functional genes in the environment and reports on their application to the study of the denitrification gene *nir*S.

19. McCaig AE, Embley TM, Prosser JI: **Molecular analysis of enrichment cultures of marine ammonia oxidisers**. *FEMS Microbiol Lett* 1994, **120**:363-368.

20. Webster G, Embley TM, Prosser JI: **Grassland management regimens reduce small-scale heterogeneity and species diversity of β-proteobacterial ammonia oxidizer populations**. *Appl Environ Microbiol* 2002, **68**:20-30.

21. Gregory LG, Bond PL, Richardson DJ, Spiro S: **Characterization of a nitrate-respiring bacterial community using the nitrate**

reductase gene (*nar*G) as a functional marker. *Microbiology* 2003, **149**:229-237.

22. Williamson N, Brian P, Wellington EMH: **Molecular detection of bacterial and streptomycete chitinases in the environment**. *Antonie van Leeuwenhoek* 2000, **78**:315-321.

23. Burgmann H, Widmer F, Sigler WV, Zeyer J: **mRNA extraction and reverse transcription-PCR protocol for detection of *nif*H gene expression by *Azotobacter vinelandii* in soil**. *Appl Environ Microbiol* 2003, **69**:1928-1935.

24. • Nogales B, Timmis KN, Nedwell DB, Osborn AM: **Detection and diversity of expressed denitrification genes in estuarine sediments after reverse transcription-PCR amplification from mRNA**. *Appl Environ Microbiol* 2002, **68**:5017-5025.
Five denitrification genes were analysed by RT-PCR of mRNA from river sediments. Active expression of *nirS* and *nosZ* was confirmed. High diversity of *nirS* was observed and distribution of sequence variants was specific to sampling sites.

25. Radajewski S, Ineson P, Parekh NR, Murrell JC: **Stable-isotope probing as a tool in microbial ecology**. *Nature* 2000, **403**:646-649.

26. •• Morris SA, Radajewski S, Willison TW, Murrell JC: **Identification of the functionally active methanotroph population in a peat soil microcosm by stable-isotope probing**. *Appl Environ Microbiol* 2002, **68**:1446-1453.
A detailed work on the active population of low-affinity methanotrophs in soil microcosms incubated with $^{13}CH_4$. Using PCR amplification of the labelled RNA/DNA, sequences related to extant type I and II methanotrophs and to the β-subclass *Proteobacteria* were identified, suggesting that novel groups of bacteria may also be involved in the CH_4 cycling in this soil.

27. • Whitby CB, Hall G, Pickup R, Saunders JR, Ineson P, Parekh NR, McCarthy A: **C-13 incorporation into DNA as a means of identifying the active components of ammonia-oxidizer populations**. *Lett Appl Microbiol* 2001, **32**:398-401.
Using group-specific 16S rRNA primers, this paper shows the identification of different active autotrophic ammonia-oxidising bacteria in freshwater enrichment cultures.

28. •• Manefield M, Whiteley AS, Griffiths RI, Bailey MJ: **RNA stable isotope probing, a novel means of linking microbial community function to phylogeny**. *Appl Environ Microbiol* 2002, **68**:5367-5373.
This paper deals with the RNA-SIP approach combined with profiling techniques of RT-PCR products recovered from the density gradient centrifugation fractions and shows how active phenol-degrading bacteria can be identified in an industrial bioreactor.

29. Pelz O, Chatzinotas A, Andersen N, Bernasconi SM, Hesse C, Abraham WR, Zeyer J: **Use of isotopic and molecular techniques to link toluene degradation in denitrifying aquifer microcosms to specific microbial populations**. *Arch Microbiol* 2001, **175**:270-281.

30. DeLong E: **Microbial population genomics and ecology**. *Curr Opin Microbiol* 2002, **5**:520-524.

31. Rondon MR, August PR, Bettermann AD, Brady SF, Grossman TH, Liles MR, Loiacono KA, Lynch BA, MacNeil IA, Minor C *et al.*: **Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms**. *Appl Environ Microbiol* 2000, **66**:2541-2547.

32. Torsvik V, Ovreas L: **Microbial diversity and function in soil: from genes to ecosystems**. *Curr Opin Microbiol* 2002, **5**:240-245.

33. Holmes AJ, Gillings MR, Neild BS, Mabbutt BC, Nevalainen KM, Stokes HW: **The cassette metagenome is a basic resource for bacterial genome evolution**. *Environ Microbiol* 2003, **5**:383-394.

34. Béjà O, Suzuki MT, Koonin EV, Aravind L, Hadd A, Nguyen LP, Villacorta R, Amjadi M, Garrigues C, Jovanovich SB *et al.*: **Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage**. *Environ Microbiol* 2000, **2**:516-529.

35. Zhou J: **Microarrays for bacterial detection and microbial community analysis**. *Curr Opin Microbiol* 2003, in press.

36. Béjà O, Spudich EN, Spudich JL, Leclerc M, DeLong EF: **Proteorhodopsin phototrophy in the ocean**. *Nature* 2001, **411**:786-789.

37. Béjà O, Suzuki MT, Heidelberg JF, Nelson WC, Preston CM, Hamada T, Eisen JA, Fraser CM, DeLong EF: **Unsuspected diversity among marine aerobic anoxygenic phototrophs**. *Nature* 2002, **415**:630-633.

38. Radajewski S, McDonald IR, Murrell JC: **Stable isotope probing of nucleic acids: a window to the function of uncultured organisms**. *Curr Opin Biotechnol*, in press.

39. Cotterall MT, Moore JA, Kirchman DL: **Chitinases from uncultured marine microorganisms**. *Appl Environ Microbiol* 1999, **65**:2553-2557.

40. • Gillespie DE, Brady SF, Bettermann AD, Cianciotto NP, Liles MR, Rondon MR, Clardy J, Goodman RM, Handelsman J: **Isolation of antibiotics turbomycin A and B from a metagenomic library of soil microbial DNA**. *Appl Environ Microbiol* 2002, **68**:4301-4306.
A description of the screening of a large BAC library that led to the discovery of two closely related triaryl cation antibiotics with a broad spectrum of activity. The capacity of the host, *E. coli*, to produce these antibiotics was stimulated by the cloning of a single operon. The interaction of the gene product with the tyrosine degradation pathway of the host results in its production.

41. Brady SF, Choa CJ, Handelsman J, Clardy J: **Cloning and heterologous expression of a natural product biosynthetic gene cluster from eDNA**. *Org Lett* 2001, **13**:1981-1984.

42. Donadio S, Monciardini P, Alduina R, Mazza P, Chiocchini C, Cavaletti L, Sosio M, Puglia AM: **Microbial technologies for the discovery of novel bioactive metabolites**. *J Biotechnol* 2002, **99**:187-198.

43. Sosio M, Giusino F, Cappellano C, Bossi E, Puglia AM, Donadio S: **Artificial chromosomes for antibiotic-producing actinomycetes**. *Nat Biotechnol* 2000, **18**:343-345.

44. • Courtois S, Cappellano CM, Ball M, Francou FX, Normand P, Helynck G, Martinez A, Kolvek SJ, Hopke J, Osburne MS *et al.*: **Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products**. *Appl Environ Microbiol* 2003, **69**:49-55.
Total community DNA from soil was cloned into a cosmid vector designed so that it could be transformed into *S. lividans*. PCR screening revealed eight novel PKS I genes.

45. Wang GYS, Graziani E, Waters B, Pan W, Li X, McDermont J, Meurer G, Saxena G, Anderson RJ, Davies J: **Novel natural products from soil DNA libraries in a streptomycete host**. *Org Lett* 2000, **2**:2401-2404.

46. Tilman D, Naeem S, Knops J, Reich P, Siemann E, Wedin D, Ritchie M, Lawton J, Wardle DA, Zackrisson O, Hornberg G, Gallet C: **Biodiversity and ecosystem properties**. *Science* 1997, **278**:1866-1867.