Ecosystem processes and interactions in a morass of diversity

James I. Prosser
Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen, UK

Abstract

High diversity in natural communities is indicated by both traditional, cultivation-based methods and molecular techniques, but the latter have significantly increased richness estimates. The increased ease and reduced cost associated with molecular analysis of microbial communities have fuelled interest in the links between richness, community composition and ecosystem function, and raise questions about our ability to understand mechanisms controlling interactions in highly complex communities. High-throughput sequencing is increasing the depth of sequencing but the relevance of such studies to important ecological questions is often unclear. This article discusses, and challenges, some of the often implicit assumptions made in community studies. It suggests greater focus on ecological questions, more critical analysis of accepted concepts and consideration of the fundamental mechanisms controlling microbial processes and interactions in situ. These considerations indicate that many questions do not require deeper sequence analysis and increased phylogenetic resolution but, rather, require analysis at smaller spatial scale, determination of phenotypic diversity and temporal, rather than snapshot, studies. Increasing realisation of the high richness of microbial communities, and potentially high physiological diversity, also require new conceptual approaches.

Introduction

Axenic microbial cultures are rare in natural environments. The vast majority of microorganisms, outside laboratories, spend their lives interacting with other microorganisms, often with plants and animals and always with their environment. Biological interactions are obviously essential for many of the ecosystem processes performed by microorganisms. These range from biogeochemical cycling processes, in which the product of one organism is the substrate for another, to more specific interactions, traditionally exemplified by lichens, rumen communities and symbiotic associations between plants and nitrogen fixers or mycorrhizas.

Interactions are often described and explained in terms of interactions between two species or functional groups. For example, rhizobia in root nodules fix atmospheric nitrogen, while the plant supplies rhizobia with organic carbon. Even complex interactions that comprise biogeochemical cycles are often considered in terms of interactions between two functional groups. For example, methanogens produce methane that is used by methanotrophs. This view of microbial interactions in biogeochemical cycles is qualified by realisation that all cells contain many elements, cycles of all elements are interlinked and each organism contributes to all biogeochemical cycles. It is also, traditionally, moderated by realisation that diversity exists within microbial functional groups and communities. Cultivated representatives of functional groups exhibit physiological diversity, and this may provide the basis for distributions of different representatives of functional groups and their contributions to ecosystem function.

‘New, ridiculous diversity’

Traditional cultivation-based techniques demonstrated high bacterial richness: large numbers of morphologically distinct colonies grow on solid medium inoculated with a soil suspension or seawater and further richness is uncovered by the use of different growth media and incubation conditions. Laboratory characterisation of these organisms reveals their high taxonomic and physiological diversity. Estimates of richness have increased as
molecular techniques for community analysis have developed. For example, cultivation-independent, 16S rRNA gene-based techniques led to the estimates of $10^{3-6}$ 'distinct genomes' (species?) g$^{-1}$ soil or L$^{-1}$ seawater (Curtis et al., 2002; Gans et al., 2005). The ability to sequence genomes of individual cells extracted from environmental samples (Rodrique et al., 2009) enables the application of techniques, such as multimlocus sequence typing (MLST) and analysis, that increase phylogenetic resolution. These will doubtless significantly increase richness estimates. It is also possible to envisage characterisation of communities by the analysis of complete genome sequences of individual cells rather than the sequences of single or several genes. We may therefore soon have the technical ability, if not the need, to distinguish $10^8-10^9$ species or operational taxonomic units (OTUs) or phylotypes g$^{-1}$ soil. If this genome diversity is linked to physiological diversity, we may find that, like us, each individual microbial cell is different and behaves differently.

The consequences of these developments can be exemplified by the 'simple' commensal interaction between an autotrophic nitrite oxidiser growing on nitrite produced by an autotrophic ammonia oxidiser. Laboratory studies of a relatively small number of bacterial ammonia oxidisers reveal differences in potentially important ecophysiological characteristics, including maximum specific growth rate, yield on, and affinity for ammonia and oxygen, carbon yield, pH and temperature growth optimum and range, ammonia and nitrite inhibition, ureolytic activity and salinity (Prosser, 1989; Koops & Pommerening-Röser, 2001). (Additional physiological diversity in heterotrophs is reflected in the large number of organic carbon compounds used in their phenotype-based taxonomy and identification.)

Molecular analysis of single genes (not MLST or genome sequences) suggests that we must consider many hundreds of different bacterial ammonia oxidiser phylotypes and their physiological diversity. We also have to consider the role of both archael and bacterial ammonia oxidisers (Prosser & Nicol, 2008), and nitrite oxidisers presumably exhibit similar diversity (Daims et al., 2001; Freitag et al., 2005). This diversity may (or may not) relate to combinations of environmental characteristics to which each is perfectly suited, determining their environmental distribution and activity.

We can never be certain that we have measured all of the phylogenetic or physiological diversity in a functional group, such as ammonia or nitrite oxidisers, or in 'total' bacterial or archael communities. This unavoidable inadequacy, coupled with the dramatic decrease in DNA sequencing costs, can lead to deeper and deeper sequence-based characterisation of communities, merely because it is possible. (Unfortunately, there is no such drive to achieve more in-depth analysis of phenotypic diversity, which is equally important but technically more difficult.) In some senses, the diversity now being measured is ridiculously high, in terms of its relevance to ecological questions and our ability to deal with it technically and, more importantly, conceptually. Are there important ecological questions that require characterisation of every 16S rRNA gene sequence in a sample? Will the major deficiencies in our understanding of microbial interactions be remedied by considering interactions between thousands of phylogenetically and phenotypically distinct organisms? The remainder of this article will consider some of the questions we ask, the approaches adopted and the relevance of high levels of richness to these questions.

**The ‘big’ questions**

There are many important ecological questions, and views of their relative importance will vary, but two fundamental questions are highlighted by the ability to characterise previously uncultivated organisms and molecular identification and phylogenetic analysis of natural microbial communities.

1. What determines diversity? What are the mechanisms determining the number and relative abundances of different phylotypes?

2. So what? Does it matter which and how many phylotypes are present. To what extent do they (all?), and their interactions, influence ecosystem functions, such as nutrient cycling?

These two questions are linked, as activities and interactions will depend on environmental characteristics, which may determine richness, but richness will also depend on immigration, generation of genetic and phenotypic diversity, rates of 'speciation' and removal through emigration and death. Ultimately, we are interested in prediction and control. Do we have sufficient understanding to predict ecosystem function quantitatively on the basis of which and how many phylotypes are present and can we manage ecosystems by changing these?

**Terminology**

To avoid the vexed and unresolvable questions of whether and how we can or need to define bacterial and archael species, and eukaryotic microbial species, I will use the term phylotype. This is generally what we attempt to measure when characterising communities using molecular techniques and classifying cultivated organisms. It has greater relevance than OTU, which by definition has...
operational value, whereas phylotype describes a group that may have evolutionary significance.

It is also necessary to consider the term ‘diversity’ and to distinguish between richness, evenness, community composition and community structure. Richness (the total number of phylotypes) and evenness (the distribution of individuals between the different phylotypes) are actually combined in the most commonly used measure of diversity, the Shannon–Weaver diversity index, and opposing changes in richness and evenness can result in the same value for this, and other indices. Thus, for example, a reduction in evenness, through growth of, and dominance by a particular component within the community, results in reduced values for diversity indices, even though species richness is unchanged. The term ‘diversity’ is therefore ambiguous, but is often used to mean richness, without explicitly stating this. It is feasible to assess dominance and evenness of microbial communities, even using molecular fingerprinting techniques, although this is rarely attempted. It is not feasible to measure richness with these techniques (Bent et al., 2007), as numbers of T-RFLP peaks or DGGE bands will not detect or distinguish the total number of phylotypes present in most environmental samples at the resolution at which they are normally studied. Rarefaction analysis can provide estimates of richness but depend strongly on assumptions of the nature of phylotype abundance distribution curves (Gans et al., 2005). Their use with high-throughput sequencing data is now beginning to provide more meaningful estimates of richness, as depth of coverage increases, but is still limited by ignorance of these distributions.

Coverage of richness increases as phylogenetic resolution decreases; for example, it requires much less sequencing effort to obtain 50% coverage of phyla than of genera. At the resolution usually employed (97% or 99% 16S rRNA gene (fragment) sequence diversity), we have only a vague idea of richness, but in reality this matters little. Most environmental studies characterise communities in samples (e.g. 1 g soil, 1 L seawater) that contain hundreds of thousands of phylotypes. The number of phylotypes present is usually not important, either \textit{a priori} (when devising the study) or \textit{a posteriori} (when analysing the data), and attention usually focuses on community composition.

Community composition and community structure

Community composition and community structure are commonly used when describing microbial communities. Community composition comprises the phylotypes present and their relative abundances and is what is actually estimated. The term ‘community structure’, additionally, implies that the different phylotypes are ordered, through interactions, into a network specifically selected in a particular environment, potentially providing optimal ‘efficiency’. That is, the term suggests that the whole is different to the sum of the individual parts and also implies that if one member of this network is removed, or ceases to function, then the ecosystem will work less ‘efficiently’, be less robust or cease to function. This will obviously be the case if at least one representative of a particular group performing an essential ecosystem function is absent or becomes extinct. The nitrogen cycle will obviously be affected if denitrifiers or ammonia oxidisers are completely absent. It is less obvious when considering extinction of one of the myriad of phylotypes that can perform these functions.

Ecosystem function

Microbial ecosystem function typically relates to biogeochemical cycling, but many others exist, for example, maintenance of soil structure and quality and pathogen control (Ritz et al., 2009). These are important functions, but are frequently ignored in microbial diversity-ecosystem function studies because they are difficult to associate with a particular phylogenetic or functional gene. Similarly, physiological characteristics targeted in molecular studies are usually restricted to those associated with specific functional genes, rather than potentially more important characteristics, for example, survival, dormancy (Lennon & Jones, 2011), growth rate, substrate affinity or susceptibility to predation.

Evolution and ecology

Niche specialisation and differentiation

The majority of microbial community composition studies are based implicitly, but rarely explicitly, on links between evolution and ecology, and between phylogenetic and phenotypic diversity, embraced in the concept of niche differentiation. They commonly involve analysis of community composition in environments with different characteristics or under different conditions. For example, bacterial community composition might be measured in soils following different fertiliser treatments, in different regions or before and after pollution.

The rationale for such studies is that:
(1) Recombination and mutation generate new phylotypes with different phenotypes.
(2) Phylotypes whose phenotypes are better adapted to the environment will be selected.
(3) Selection increases the relative abundance of better-adapted phylotypes and, potentially, leads to the development of distinctive phylogenetic groups possessing the new phenotype, that is, we have niche specialisation and differentiation.

(4) Dispersal of members of the new phylotype leads to its colonisation of, and selection in other, similar environments.

(5) Sites with similar environmental characteristics will be dominated by phylotypes whose phenotypes are adapted to those characteristics.

(6) Analysis of 16S rRNA genes provides information on phylogeny, enabling both detection and identification of adapted phylotypes and giving information on their evolution.

(7) 16S rRNA gene analysis of communities will identify links between phylogeny, phenotype and environmental characteristics.

(8) Analysis of functional genes provides similar tracking of evolution and adaptation, but within particular functional groups.

For various reasons, this experimental approach has limited value.

(1) A tenet of 16S rRNA gene analysis is that phylogeny provides only limited information on function. Philippot et al. (2010) discuss this with respect to ecological coherence and taxonomic rank. Sometimes there are patterns, often there are not and, even more often, we do not know.

(2) There are many examples of phenotypes (functions) of environmental significance that are distributed among many phylogenetic groups. Denitrification is represented within approximately 50% of phylogenetic groups, ammonia oxidisers are found within the Beta- and Gammaproteobacteria and the thaumarchaea and nitrite oxidisers occur within the Alpha-, Beta- and Deltaproteobacteria, Nitrospirae and Nitrospina.

(3) Links between phylogeny and function are seriously disrupted by lateral gene transfer, over both evolutionary and ecological time scales.

(4) Plasmids often determine important, selective environmental characteristics, notably antibiotic resistance and heavy metal resistance, but many more.

(5) 16S rRNA gene-based phylogeny discriminates at the level of traditionally defined genera, but is much less useful for the discrimination of species or at higher taxonomic resolution. Significant phenotypic variation occurs within genera; this is how species are traditionally distinguished.

(6) Finer scale resolution is possible using other molecular approaches and indicates fine scale physiological diversity. For example, MLST analysis of cultivated organisms demonstrates that phylotypes differing in very few, or just a single gene, can have critically different characteristics, for example, the presence or absence of virulence. There are good reasons to expect that similar differences will exist in organisms preforming important ecosystem functions.

(7) Functional gene analysis restricts phylogenetic analysis to particular functional groups, but gives no more information than 16S rRNA genes on phenotypic diversity, unless adaptation results from changes in the gene product itself.

(8) Analysis of nucleic acid-based phylogeny will not detect epigenetic effects, in which gene expression or phenotype is determined by mechanisms that are not based on DNA sequence. Epigenetics is rarely considered by microbial ecologists, but is increasingly considered to be important in plant and in animal ecology (Bossdorf et al., 2008).

(9) It is difficult to detect effects of biological interactions (e.g. predation) by analysing sequences.

This list is not exhaustive but demonstrates that we know very little of the extent to which 16S rRNA gene or functional gene diversity tell us about phenotypic diversity. Thus, even if community composition is driven by adaptation, phenotypic diversity, differences in environmental characteristics and selection, it will be very difficult to detect links between phylotype diversity, phenotypic diversity and ecosystem function by the analysis of 16S rRNA or functional genes.

There is also considerable, apparent, functional redundancy within microbial communities. Certainly, there are many hundreds of phylotypes at the traditional species level that can perform basic ecosystem processes, for example, metabolism of simple organic carbon compounds, nitrification and denitrification. However, even if we find two strains that appear to have identical characteristics and, by implication, identical ecosystem function, we cannot eliminate the possibility that they will respond differently to a new, previously untried set of environmental characteristics. It is therefore never possible to rule out or disprove functional redundancy within phylotypes. This ‘inability to disprove’ makes meaningless the frequently cited concept that ‘everything is everywhere but the environment selects’ (Bass Becking, 1934). This may be useful as a way of thinking of about niche differentiation but, as a hypothesis, it cannot be refuted and therefore lacks value. It is not possible to measure everything (undiscovered phylotypes may be present) and it is not possible to say with confidence that any phylotype is absent, only that is below the detection limit. In addition, it is usually quoted with no attempt to define ‘everything’ or ‘everywhere’. If ‘everywhere’ refers to 1 mL of lake water, it is not possible for such a volume to contain everything, if ‘everything’ refers to every known 16S rRNA gene-defined phylotype.
Neutral theory

Niche differentiation combined with interactions, notably competition, is intuitively attractive as a mechanism determining community composition, even if it is difficult to assess and demonstrate. Neutral theory (Bell, 2000; Hubbell, 2001) presents an alternative mechanism in which all phylotypes in a community are assumed to follow the same rules, regardless of phenotype, and their relative abundances are determined by random immigration, birth, death and speciation. If speciation is considered to be negligible, relationships between numbers of phylotypes and area (or volume) can be predicted from knowledge of the size of the community and two parameters. The first, $m$, is the probability of replacing a dying organism with one entering randomly from the source community in the surrounding environment, rather than reproduction within the local community. The second, $\theta$, describes the log-series distribution that is assumed to exist for phylotype abundance; $\theta$ increases with total richness in the source community. Neutral theory therefore combines both phylotype abundance distributions and phylotype area distributions.

Woodcock et al. (2007) tested neutral theory predictions of the relationship between phylotype abundance in bacterial communities (characterised by 16S rRNA-DGGE analysis) and the volume (rather than area) of tree holes (the water-filled holes remaining after tree death). The two parameters, $m$ and $\theta$, were estimated from data for the smallest tree hole and then used to predict phylotype abundance in other tree holes. Figure 1 shows the quality of fit of predicted and experimental phylotype abundance distributions over a wide volume range. Relationships between phylotype abundance and tree-hole volume determined empirically were almost identical to those calculated by neutral theory using the estimated values of $m$ and $\theta$ (Fig. 2a and b) and there was a good fit between observed and predicted phylotype richness (Fig. 2c). Most of the variability in richness could therefore be explained solely on the basis of two parameters, without consideration of environmental differences (similar ecosystems in close proximity were characterised), links between phylogenetic and phenotypic diversity, niche differentiation or interactions. Predictions were based on the assumption...
that phylotype abundance is described by a log-series distribution, which is not possible to determine using DGGE (as discussed earlier), but neutral theory offers the most parsimonious explanation for phylotype abundance relationships with tree-hole volume.

Ofituru et al. (2010) extended this approach using T-RFLP time series data on ‘total’ bacterial communities and bacterial ammonia oxidiser communities in a wastewater treatment system. This gave more information on phylotype abundance distributions, and the basic neutral model was extended to allow variation in growth rates of different phylotypes, potentially through differences in phenotype associated with particular environmental factors. Variability arising from a purely neutral model and additional variability because of environmental factors could therefore be estimated. The proportion of variability indicated by the purely neutral model was 0.2 and 0.23 for the bacteria and ammonia oxidisers, respectively, and increased to respective values of 0.28 and 0.37 with inclusion of major environmental factors. Substantial variation remained unexplained, suggested reasons being effects of unmeasured environmental factors, more complex effects of environmental factors, measurement error and lack of relationship between T-RFLP-defined phylotypes and ecotypes (i.e. between phylogeny and function).

Neutral theory and niche theory are not necessarily mutually exclusive but the former is usually ignored. It is sometimes criticised for its simplicity, but simplicity is arguably a major aim of models, hypotheses and theories. At the other extreme, consideration of the many possible phenotypic characteristics that might be influenced by many environmental factors in different ways in many thousands of phylotypes is unlikely to be technically or conceptually tractable, and unlikely to generate insight, understanding or predictive power.

**Temporal heterogeneity and scale**

The degree to which communities reflect environmental characteristics is also influenced by differences in temporal scales of community change and environmental change. Some factors, for example, ocean salinity or subsurface temperature, are stable over relatively long time scales. Others change regularly over short time scales, for example, diurnal cycles of light and temperature in surface waters or upper soil layers. Other events are much less predictable and more variable, for example, changes in concentrations of many nutrients, invasion, predation and rainfall. Niche differentiation and selection through competitive interactions assume relatively constant conditions, or at least environmental conditions that prevail for sufficiently long to enable adaptation, selection and detectable community change. Most attempts to link phylogeny and function involve ‘snapshot’ sampling and correlation of community composition and environmental characteristics. Temporal investigations are of greater value, but we know little of the rates of invasion, adaptation and selection in natural environments and, importantly, even less about the rates of decrease in abundance of phylotypes that are ‘de-selected’ following environmental change. In this respect, survival, dormancy and resuscitation rate may be more important in determining community composition than growth parameters such as maximum specific growth rate and substrate affinity or more easily measured physiological characteristics associated with functional genes.

**The impact of individual phylotypes**

It is therefore dangerous to assume that phylogeny is linked to function and that links, even if they exist, can be determined using experimental approaches currently or imminently available. Nevertheless, let us consider a hypothetical stable community containing phylotypes whose phenotypic characteristics are perfectly adapted to their immediate, stable environment. How might we quantify the impact of a particular phylotype on ecosystem function, in an attempt to assess the importance of richness in complex natural ecosystems?

One approach is to apply the concept of metabolic control theory (MCT; Kacser et al., 1995). This can be illustrated (Fig. 3) by imagining conversion of a substrate, Xo, through several intermediate substrates, S1–S3, to a product S4, catalysed by enzymes E1–E4. MCT quantifies the influence of each enzyme on flux through the pathway as a flux control coefficient, defined as the proportional increase in flux resulting from a proportional increase in enzyme concentration. Enzymes at low...
concentration might be expected to have greater control than those at high concentration. All flux control coefficients will sum to 1, and the average flux control coefficient for the pathway in Fig. 3, which contains four enzymes, will be 0.25. If E₂ has a flux control coefficient of 0.97, and therefore significantly greater control (equivalent to the traditional, qualitative concept of a rate-limiting step), the remaining enzymes will have an average control coefficient of 0.01, and negligible control. Doubling the amount of E₂ will significantly affect flux, while changing amounts of E₁, E₃ or E₄ will have little effect.

If we imagine a typical metabolic map of a cell, and assume 100 enzymes, the average flux control coefficient for each enzyme will be 0.01. If we apply this concept to an equivalent map of all biogeochemical cycling processes occurring within an ecosystem, the average control of a particular functional group will equal the inverse of the number of functional groups contributing to ecosystem function. So, for example, production of N₂ will depend on denitrifiers, but also on other groups active in the nitrogen and carbon cycles that provide them with their major substrates. All will therefore have flux control coefficients for N₂ production. The impact of denitrifier biomass, per se, on N₂ flux will therefore depend on flux control coefficients for functional groups directly involved in nitrate reduction, but also those involved in supply of carbon and other essential nutrients. The average impact of an individual process or functional group on an ecosystem function will therefore be very low.

The situation is much more complex in microbially mediated ecosystems than metabolic pathways because each step in a nutrient cycle is mediated by many 10s or 100s or even 1000s of phylotypes. The influence and importance of a particular phylotype involved in a function will therefore be, on average, very, very low. One phylotype may be important, that is, may have a high flux control coefficient, but, if so, the average contribution of other phylotypes will be negligible. This approach has been tested for simple microbial interactions in a ‘stable’ system (Allison et al., 1993) but is more difficult to apply to non-equilibrium, growing systems (Röling, 2007), and flux control coefficients will also be subject to environmental change. Nevertheless, it predicts that most phylotypes will have negligible impact on ecosystem function and if we double the biomass of the average phylotype, it will have no measurable impact. Analysis of the effects of richness or community composition on ecosystem function may therefore not require increasing expenditure on the analysis of phylotypes that are present at low relative abundance and greater effort should be expended on finding ways to determine which phylotypes control function. This has particular implications for our ability to manipulate and manage microbial communities.

Does richness matter for ecosystem function and interactions?

**Network analysis**

Interactions within an ecosystem have been represented as networks linking pairs of phylotypes based on the strength of associations between each pair, quantified by the correlation of phylotype relative abundances in time series data (Fuhrman & Steele, 2008; Steele et al., 2011) and following environmental change (He et al., 2012). Fuhrman & Steele (2008) used network analysis to investigate interactions within marine bacterial communities. A threshold value (equivalent to a P-value associated with correlation analysis) was chosen to distinguish phylotypes that were interacting ‘significantly’ with each other and also with environmental variables that had greatest influence. Analysis is illustrated by networks of SAR11 OTUs, and biotic and abiotic environmental variables (Fig. 4). An interesting feature is that a single phylotype interacts with only a limited number of others, a maximum of 11 in this example. One interacted with only one other phylotype and two showed no significant interactions, although they presumably interact with phylotypes below the detection limit or not targeted. The analysis highlights the importance of quantifying the strength of interactions to assess which are significant and shows that the majority of interactions will be insignificant and immeasurable. He et al. (2012) describe a similar approach, Random-Matrix Theory, using microarray data on soil communities subjected to elevated carbon dioxide.

Basic network analysis is empirical and suffers the disadvantages of any association- or correlation-based method. The complexity of the network will depend on the threshold value used to assess significance of associations and phylotypes and their links with abiotic factors will be restricted to those that are measured, which may not include the major drivers. It does have the potential, however, to uncover unsuspected interactions between phylotypes that might then be the subject of more rigorous and critical investigation. More importantly, exploration of network characteristics may reveal mechanistic control of system properties. For example, complexity (average numbers of significant associations per phylotype) may be linked to stability; connectance (the proportion of potential associations observed) may be linked to behaviour; focal phylotypes (those associated directly with a high proportion of other phylotypes) may be critical for ecosystem functioning (analogous to keystone species); and other aspects of networks (nestedness, compartmentalisation) may reflect different behaviours and interactions. Network analysis therefore provides the potential
for investigation of fundamental ecological concepts using microbial communities.

**Richness reduction**

While there is evidence that many microbial communities are sensitive to environmental change and perturbation, the importance of community composition, and particularly richness, for ecosystem function is less clear (Allison & Martiny, 2008). This is partly because of technical difficulties in distinguishing effects of environmental conditions and community composition on function. Richness of complex natural communities can be reduced by removal of a portion of the community, through partial killing or dilution, followed by incubation to enable reestablishment of the original biomass concentration. In one example (Griffiths et al., 2000) progressive reduction in richness by chloroform fumigation did not affect broad-scale function (e.g. decomposition) but reduced rates of more specific functions (nitrification, denitrification, methane oxidation). Resilience of all processes, however, decreased with decreasing richness. Girvan et al. (2005) reduced diversity by addition of benzene or copper sulphate to soil. Benzene had a greater effect on microbial community composition than copper sulphate, and there was evidence for a link between diversity and resistance to perturbation. Richness reduction again did not affect broad-scale function (wheat shoot mineralisation) but reduced narrow-niche function (mineralisation of 2,4-dichlorophenol), with greater resilience in the more diverse soil.

Fumigation methods are potentially selective and require uniform penetration of killing agents throughout soil. These limitations may be reduced by inoculation of sterilised soil with dilutions of a soil suspension obtained from the same soil. Using this approach, Griffiths et al. (2001) and (Wertz et al., 2006) found that richness did not influence a wide range of broad and specific soil functions, unless dilution led to complete extinction or biomass recovery was incomplete (Wertz et al., 2006). Richness also had no effect on resistance and resilience of soil denitrifier and nitrite oxidiser communities (Wertz et al., 2007).

Both of these approaches have disadvantages and reestablishment of communities may select for faster growing organisms. Nevertheless, the results, particularly those from dilution-based experiments, suggest that the species present at high relative abundance are capable of performing major ecosystem functions, even under transient conditions, as long as they are present and at sufficient biomass concentration. It suggests, again, that greater evidence for the importance of less abundant species on ecosystem function is required before significant resources are expended on their characterisation.

**Community construction**

Richness–ecosystem function relationships can also be investigated by the construction of communities with
defined composition. For example, Langenheder et al. (2010) isolated six bacteria from a soil sample that could grow on glucose, xylose or galactose as the sole carbon and energy source. These were used to construct communities with all levels of richness (1–6) in all possible combinations. Each isolate was present at the same biomass concentration, and the system was designed to maintain evenness during incubation on all possible combinations of the three carbon sources. This enabled investigation of the effect of richness, species composition and environmental (substrate) complexity on broad ecosystem function (respiratory activity). Richness–function relationships were initially linear but then became saturating at species and substrate richness values of 4 (Fig. 5) and 2, respectively. Species richness saturation was presumably due to functional redundancy or negative interactions among competitors. Effects of species and resource richness were independent, suggesting complex, species–substrate-specific interactions that, unfortunately, prevent widely applicable prediction of behaviour.

These strains were isolated from the same soil and potentially competed for the same organic carbon substrates, but results indicate that high natural richness will not influence the rate of organic C degradation. Changes in relative abundance or extinction of phylotypes may therefore be less important than physiological flexibility within existing communities, which may maintain ecosystem function without the requirement for community change. Using a similar approach, Wittebolle et al. (2009) constructed microcosms with communities containing 18 denitrifier strains, but with differences in initial evenness. Denitrification activity during incubation and growth of the communities increased with evenness, particularly under stress conditions. This study highlights the different aspects of diversity that must be considered and potential differences between growing communities, in which relative abundances will change, and ‘stable’ communities.

**Spatial heterogeneity and microenvironments**

Microbial cells in natural environments are rarely distributed uniformly. This is most evident in soil and sediments, but is also true for planktonic marine communities, where particulate material (marine snow) provides foci for microbial growth and activity. Spatial distribution in soil is illustrated in Fig. 6 (Nunan et al., 2003). In topsoil (Fig. 6b), cells appear to form communities of different size containing (presumably) interacting cells. In subsoil, cells are rare (Fig. 6c) and ‘uniform’ colonisation is only approached when soil is supplemented with high levels of organic carbon (Fig. 6a). Only a fraction of the soil surface is colonised (Grundmann, 2004), and the number of cells in an area of 0.282 mm² ranged from approximately one in subsoil to six to eight in cropped soil (Nunan et al., 2003). If cells are homogeneously distributed, therefore, distances between cells will be great in relation to cell size. Heterogeneous distribution means that interacting cells and microcolonies will be separated even further and these calculations apply to all cells that can be detected microscopically. Intercellular distances between members of particular subgroups or
cell concentrations (10^7–9 cells g⁻¹ soil) and high numbers of phylotypes. There is very little information on community composition of soil microcolonies, but presumably some will contain monocultures and others mixed cultures of interacting phylotypes.

Colonisation of 1 g soil by unicellular microbes may be seen as analogous to our colonisation of the earth. Each system contains several billion individuals and each contains communities of varying size. Interactions within both communities can be envisaged, but interactions between spatially separated organisms will obviously be restricted and dependent on a wide range of factors. The probability of two cells within a gram of soil meeting may be similar to that of two humans meeting, but it will certainly be very low and the scope for significant interaction will be similarly low.

If microbial interactions are mediated by soluble compounds (substrates, products, antibiotics, signalling compounds), the probability of these compounds reaching target organisms at effective concentrations must be considered. In soil, this will depend on intercellular distance, diffusion, production rate, path length, tortuosity, pore structure, moisture content, bulk flow and compound degradation. As a consequence, the majority of such interactions may be restricted to the 10–100 µm scale, with organisms interacting within soil pores or microcolonies. If so, interactions are likely to involve only a limited number of phylotypes and the observations described earlier (‘saturation’ of effects on ecosystem function at low richness) may apply at the scale over which interactions are taking place.

These considerations have enormous implications, conceptually and experimentally. The latter are obvious, but usually ignored. If interactions operate at the 10–100 µm scale, then it will be difficult to gain understanding using techniques that destroy spatial structure (i.e. lysing all cells and solubilising nucleic acids) and analysing 1 g soil samples. (It would be equally difficult to understand human interactions, or interactions between nations, by homogenising the total human population and analysing a small subsample.) Network analysis may give hints of interactions but more rapid progress requires study at the spatial scale at which interactions operate. This may be more difficult technically, but much more relevant and potentially much more rewarding.

Signalling compounds are thought to mediate many microbial social interactions (see reviews by Keller & Surette, 2006; West et al., 2007). Such compounds control production of exoenzymes, biosurfactants, antibiotics and exopolysaccharides that are presumed to be of ecological significance. Discovery of their widespread occurrence (e.g. acyl-homoserine lactone and oligopeptide quorum sensing molecules in Gram-negative and

---

**Fig. 6.** Bacterial colonisation of soil after the addition of glucose (a), in topsoil sampled after harvest (b) and in subsoil (c). The scale bar represents 20 µm. From Nunan et al. (2003), with permission.
Gram-positive bacteria, respectively) promised insight into interactions within natural communities, but few studies have assessed their in situ environmental relevance. In heterogeneous natural environments, the impact of signalling compounds will depend on their transport and degradation, as indicated above, but also on specificity of action in complex communities and potential alternative functions (Redfield, 2002; Hense et al., 2007). They therefore illustrate the more important conceptual implications of spatial heterogeneity. The scale at which interactions operate will depend on the mechanisms involved and those mediated by signalling compounds are likely to occur at small spatial scale. Dispersal of soil organisms through motility, bulk flow, belowground transport (soil animals) and above-ground dispersal (birds, wind, aeroplanes) will operate at widely different spatial scales.

The importance of spatial scale is illustrated in a single study of signalling in Bacillus subtilis in which competence and production of surfactins, extracellular enzymes and exopolymers are controlled by quorum sensing through an extracellular pheromone. Bacillus subtilis strains are grouped into pherotypes; members of a particular pherotype can induce competence in members of the same pherotype, but not in members of other pherotypes (Ansaldi et al., 2002; Mandic-Mulec et al., 2003; Stefanić & Mandic-Mulec, 2009), providing a potential mechanism for diversification. It could therefore be predicted that competition between different pherotypes would lead to competitive exclusion and a reduction in richness, but only at the (small) scale at which interactions are taking place. To test this, Stefanić & Mandic-Mulec (2009) found four pherotypes represented in B. subtilis isolates obtained from two closely located 1-cm³ soil samples. Pherotype richness decreased in subsamples of decreasing size, as predicted, but even 1/16-sections often contained more than one pherotype. This gives an indication of the spatial scale at which interactions involving signalling compounds, and other soluble compounds, might operate but community analyses are normally performed at scales (e.g. 1 g, 1 L) that are much greater than those at which communication will take place. Interestingly, the four pherotypes found by Stefanic & Mandic-Mulec (2009) in a Slovenian riverbank soil are the same as those found in North American desert soil samples. This indicates additional large-scale dispersal mechanisms, high migration rates and lack of geographical isolation, as observed previously (Roberts & Cohan, 1995). Further genetic and physiological analysis suggests a strong, but not complete link between pherotypes and ecotypes (Stefanić et al., 2012).

Mechanisms controlling interactions will therefore determine the scale at which they influence community composition and processes. Biogeographical studies and network analysis approaches must consider this, not just when interpreting results, but also when planning experiments. Similarly, biogeographical effects seen at scales of metres or larger are likely to result from larger scale environmental effects, rather than small-scale interactions. Measuring biogeographical patterns of phylotypes without regard to the mechanisms that might be driving interactions and other driving forces associated with those phylotypes is likely to generate descriptive findings, rather than understanding, insight and generic theories. A further important factor is the need for approaches for scaling up small-scale interactive effects in heterogeneous environments to the scales at which we are interested in measuring processes.

Conclusion

The aim of this essay is not to provide an in-depth, comprehensive analysis of the various topics discussed. This is outside its scope and would require major reviews or textbooks. Many of the citations provide evidence that could be countered by evidence from other studies and many of the questions raised do not have ‘correct’ answers. Rather, the essay is intended to encourage reflection on the basis for studies on microbial communities, interactions and ecosystem function. It challenges the rationale for some studies of microbial communities, suggests the need for greater clarity in what we mean by diversity and why we are studying it and suggests the importance of hypothesis-driven approaches, before deciding whether resource-intensive, if ‘cutting-edge’, techniques are relevant. It may be increasingly easy to determine the in-depth community composition through deeper and deeper sequencing, but this may not be required to address many important ecological questions. Certainly, it is unlikely that 1 g soil or 1 L seawater requires 100 000 different 16S rRNA gene-defined phylotypes to function efficiently and understanding of microbial interactions requires much greater consideration of environmental heterogeneity and the spatial scale at which interactions are mediated. Similarly, it is unlikely that a community within 1 g soil or 1 L seawater is exquisitely adapted to and selected for the unique combination of physicochemical and biological characteristics unique to that environment. If niche specialisation and differentiation drive community composition in natural environments, sequence diversity data are redundant without knowledge of the degree to which sequence diversity translates into true phenotypic diversity and without knowledge of temporal changes in communities and links to processes. In the rush to obtain and analyse data, it is easy to sweep difficult issues under the carpet. This applies to both methodological

Processes and interactions in a morass of diversity

© 2012 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved
issues, for example, biases and experimental design, and conceptual issues, for example, neutral theory and relevance of 16S rRNA gene phylogeny to function. Techniques are available to characterise vast diversity within microbial communities. In assessing the links between microbial community composition, interactions and their relevance to ecosystem function, the challenge is to identify key questions and to address them through sound conceptual approaches and with appropriate techniques and experimental design.

**Acknowledgements**

I am very grateful for helpful reviewer comments on this manuscript.

**References**


