Opinion TRENDS in Microbiology Vol.14 No.6 June 2006

Marine microbial diversity: can it be determined?

Carlos Pedrós-Alió

Institut de Ciències del Mar, CMIMA-CSIC, P. Marítim de la Barceloneta 37-42, 08003 Barcelona, Spain

Estimates of the order of magnitude for the total number of microbial species on Earth range from 10^3 to 10^9. Despite global dispersal of microorganisms, this number is probably rather large. The total biodiversity of an ecosystem is composed of two elements: first, a set of abundant taxa that carry out most ecosystem functions, grow actively and suffer intense losses through predation and viral lysis. These taxa are retrievable with molecular techniques but are difficult to grow in culture. Second, there is a seed bank of many rare taxa that are not growing or grow extremely slowly, do not experience viral lysis and predation is reduced. Such taxa are seldom retrieved by molecular techniques but many can be grown in culture, which explains the dictum 'everything is everywhere'.

How many species?

Ecologist Robert May remarked that "we have a catalog of all the celestial bodies our instruments can detect in the universe, but we ignore how many living beings share the Earth with us" (ISME meeting, Barcelona, 1992). This ignorance is particularly large for life in the oceans. Thus, the international initiative 'Census of Marine Life' [1] was launched to "assess and explain the changing diversity, distribution and abundance of marine species from the past to the present, and to project future ocean life." The purpose of the initiative is to make a census of the 'known' diversity in the oceans and to determine what portion of the 'unknown' can be determined to focus research in the most efficient way.

In the case of the diversity of microorganisms, even the right order of magnitude is unknown and the issue is highly controversial [2–4]. The total number of prokaryotic cells in the oceans has been estimated to be $\sim 10^{29}$ [5] but the question is: into how many species are these cells partitioned? This question is not of purely academic interest [6]. First, the unknown diversity hides novel metabolisms (such as the recently discovered photoheterotrophy in the sea) that force a re-evaluation of carbon and energy fluxes in the oceans [7–9]. These processes must be understood so that, for example, precise models of global change can be constructed. Second, the unknown microorganisms are the largest potential reservoir of useful genes for medicine and biotechnology. And, of course, knowledge of microbial diversity is essential to understand evolution [10].

During the past 15 years, molecular studies using the gene sequences that encode the small subunit rRNA (SSU rDNA) have revealed a wealth of new marine microorganisms that belong to the three realms of life: Bacteria [11], Archaea [12,13] and Eukarya [14–16]. The feeling among marine microbiologists is that of living through an age of discovery with no end in sight. However, most of this new diversity has not yet been described because pure cultures of the organisms behind the sequences are necessary to define a species (Box 1). How far has the goal of the Census of Marine Life been fulfilled? First, the known diversity has to be defined; second, the size of the unknown has to be estimated; and finally, I propose a framework within which both can be studied with different strategies.

The known and the unknown

Approximately 6000 species of prokaryotes (http://www.bacterio.cict.fr/number.html) and 100 000 species of protists [17] have been formally described. This is the known

Glossary

Allopatric speciation: the origin of two or more species resulting from divergent evolution of populations that live in geographical isolation.
Biodiversity: the total, specific, taxonomic or genetic richness contained in nature or in any local or taxonomic part of it. The repository of genotypes. The actual richness of 'nature's dictionary' [39,40], which could be also named 'global diversity.'
Biological species concept: two populations are considered to belong to the same species if they can interbreed and their descendants are fertile.
Diversity: the actual distribution of individuals or biomasses into species at a particular time in a given ecosystem, which could also be named 'ecodiversity' [39,40] or 'locally active diversity'.
Ecological species concept: two strains are different species if they occupy different niches and the same species if they occupy the same niche [22].
Evenness: a component of diversity that considers how individuals are distributed among species.
Morphological species concept: a species is defined by having sufficient morphological traits to differentiate it from all other species, dependent on available information and the criterion of experts.
Operational taxonomic units: originally applied to taxa defined by statistical treatment of many characters. Now applied to any taxon defined in a pragmatic way, not necessarily fulfilling the requisites of a formal species definition.
Phylo-phenetic species concept: a species is a monophyletic and genomically coherent cluster of individual organisms that show a high degree of overall similarity in many independent characteristics, and is diagnosable by a discriminative phenotypic property [35].
Richness: a component of diversity that considers the total number of species present.
Speciation: the establishment of reproductive isolation between two or more previously interbreeding populations.
Sympatric speciation: speciation that occurs in the absence of geographical isolation.

Corresponding author: Pedrós-Alió, C. (cpedros@icm.csic.es).
Available online 6 May 2006

 0966-842X/$ - see front matter doi:10.1016/j.tim.2006.04.007

Box 1. What is a microbial species?

There are >20 definitions of species for eukaryotes [53], none of which is satisfactory for all living eukaryotes. It is no surprise that the definition for prokaryotes is also controversial [35]. The morphological species concept is clearly not applicable to prokaryotes and to many of the smallest eukaryotes [35,37,54].

Rosselló-Mora and Amann [35] defend the phylo-phenetic species definition for bacteria as pragmatic, useful, and applicable to all prokaryotes. This definition uses information from SSU rDNA. As a rule of thumb, it is accepted that two strains must show a similarity >97% to be considered part of the same species. However, SSU rDNA does not have enough resolution to delineate species [35,55]. A more powerful technique is DNA–DNA hybridization. Usually, two strains must show >70% hybridization to be considered as the same species. Strains that show >97% similarity in their SSU rDNA show a range of values in DNA–DNA hybridization between <20% and 70%. By contrast, no two strains with >70% DNA–DNA hybridization show a similarity of <97% in SSU rDNA. Obviously, taxa based on a 97% SSU rDNA similarity widely underestimate the number of species if the DNA–DNA hybridization criterion is used. In conclusion, to argue that the number of microbial species is high using the SSU rDNA criterion to count taxa (instead of vice versa) means always underestimating the true number of microbial species.

For the purpose of the present article, however, a definition of a species is not essential. Fortunately, diversity can be calculated with other units besides species. Individuals must be grouped into non-overlapping classes according to a consistent classification criterion [56]. Two descriptors that fulfill these requirements (and are appropriate for microbes) are the number of DNA bands in denaturing gradient gel electrophoresis or the number of clones in a DNA clone library. These units are operational taxonomic units (OTU). The same arguments that are used in the main text for taxa (or OTUs) can be used for species if a satisfactory definition is eventually found [57].

diversity. Most information about these microorganisms comes from the study of pure cultures, at least in the case of prokaryotes, in which a strain deposited in a culture collection is necessary to describe a new species.

The unknown diversity is currently being explored with molecular techniques, particularly cloning and sequencing [6]. A library is constructed with the DNA from a particular sample. Each clone has the SSU rDNA of one member of the original community. As more clones are sequenced, new taxa arise. At the beginning of this process, almost every new clone will belong to a new taxon (Figure 1a) but, as more clones are examined, the curve of taxa versus sequenced clones should approach an asymptote. This would indicate that most diversity in the library has been retrieved. In some cases, the curve does not approach an asymptote but instead increases almost as a straight line. This indicates that many more clones must be sequenced to retrieve the diversity in the library. There are statistical tools to estimate the richness (see Glossary) of a library [18] and different authors have used results from clone libraries and statistical procedures to estimate a global number of microbial taxa in the ocean. Thus, Hagström *et al.* [19] suggested that only a few thousand bacterial taxa were present in the sea, judging from the sequences deposited in GenBank. More recently, these authors estimated that ~360 new bacterial taxa are submitted to GenBank every year. This will increase the total number of taxa but it does not seem to require a major revision of their order of magnitude (between 10^3 and 10^4 taxa) [20].

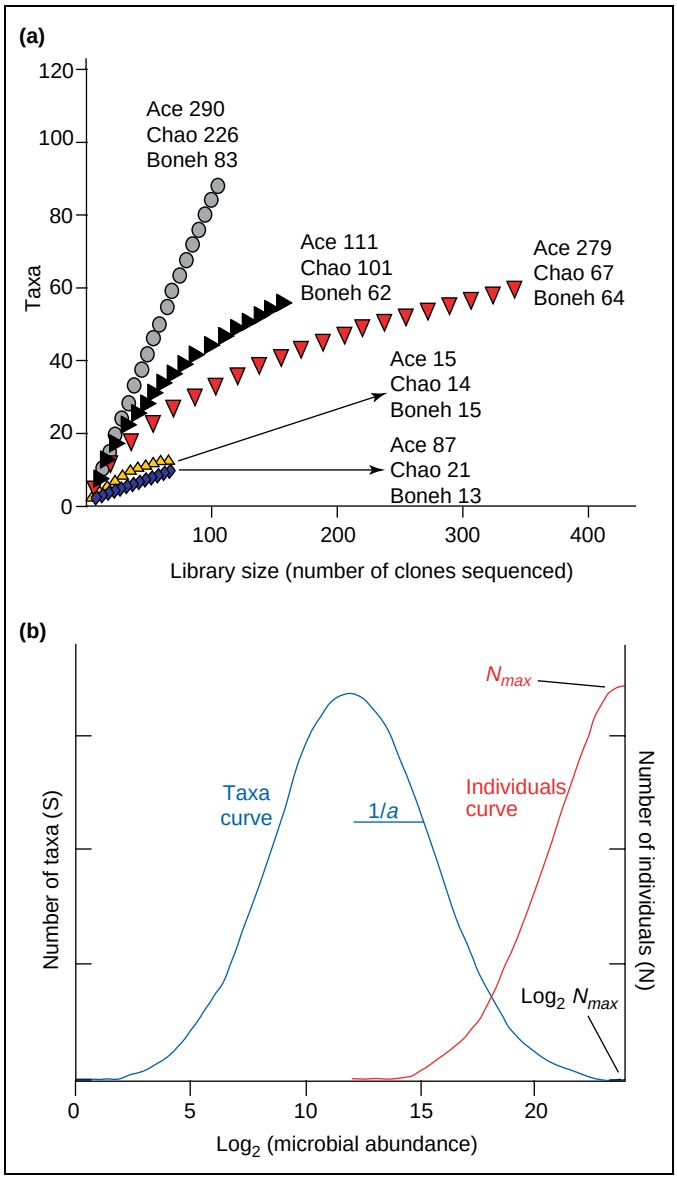

Figure 1. Different ways to estimate the number of microbial taxa in a sample. **(a)** Number of taxa versus clones sequenced in five different libraries. Some curves quickly reach an asymptote whereas others seem to increase continuously. Numbers adjacent to curves indicate estimates of the total number of taxa in each library (Ace, Chao or Boneh) calculated with different statistical methods. Key to colors: French Guiana sediment, gray circles; bioreactor, filled black triangles; hyperthermal spring, filled red triangles; Cariaco anoxic basin 310 m, yellow filled triangles; Cariaco anoxic basin 320 m, blue filled diamonds. **(b)** Number of taxa or number of individuals versus abundance curve. The *x*-axis shows the log_2 of taxon abundance. The blue curve shows the assumed lognormal distribution of taxa versus abundance. The total diversity *St* is the area under this curve. The red curve shows the number of individuals versus abundance. The total number of individuals in the sample is *Nt*, which corresponds to the area under this curve. N_{max} is the number of individuals in the most abundant taxon. N_{max} and *Nt* can be easily measured. This enables calculation of the amplitude (parameter *a*) of the taxon–abundance curve and integration of the area under this curve provides the richness in the sample. Part (a) reproduced, with permission, from Ref. [18]. © (2004) Blackwell Publishing. Part (b) reproduced, with permission, from Ref. [21]. © (2002) The National Academy of Sciences, USA.

These estimates are based on the use of clone libraries but this approach is insufficient to retrieve all of the diversity present in nature (see later). Different authors have provided estimates that would compensate for the shortcomings of the cloning approach. Curtis *et al.* [21] used the taxon–abundance curve (Figure 1b). The assumption is that this relationship follows a lognormal

distribution. In this case, the total number of taxa can be estimated from the total number of individuals (*Nt*), and the abundance of the most common taxon (*Nmax*). The authors' estimate for the whole ocean was 10^6 bacterial taxa. Dykhuizen [22] used a completely different approach and compared differences between microbial communities using total DNA hybridization. The author then assumed a certain number of different microbial communities and estimated that 10^9 different bacterial taxa are present on Earth. These estimates use different definitions of a microbial taxon but the essential argument is that, regardless of the particular definition, the total number is huge (Box 1).

A different school of thought holds that the global number of microbial taxa is low, perhaps in the order of 10^3–10^4 [2,23–25]. These studies refer to morphospecies of ciliates but then extrapolate to all microorganisms, basing their conclusion on a fact and its corollary. The fact is that because microorganisms are small and have many individuals, they are easily dispersed globally. This is an explanation for the classical dictum that 'everything is everywhere, the environment selects' [26]. Most, if not all, microbiologists will agree with this assertion [3,6]. The corollary states that because allopatric speciation will be unlikely, the global number of species will be low [23]. Allopatric speciation takes place when a geographical barrier separates two populations from the same species. The two populations will eventually evolve into separate species if there is no genetic exchange between them (Box 2).

However, this corollary does not hold. First, allopatry is not necessary for speciation [27]. Several mechanisms of sympatric speciation (with the two populations coexisting in the same area; Box 2) have been described for sexually reproducing organisms such as birds [28] or arthropods [29]. Sympatric speciation should be easier for asexual organisms such as bacteria and many eukaryotic microorganisms. Second, studies of bacteria in continuous cultures have shown how fast evolution can proceed in these organisms ([30]; Box 2). And finally, bacterial assemblages harbor microdiversity with an ecological and taxonomical meaning that is still debated [31–33]. If this microdiversity were considered, the coexistence of several closely related but separate clones would make the argument of low number of taxa even less tenable. Thus, global dispersion of microorganisms does not imply a low number of species. It only means that those speciation mechanisms that require allopatry are unlikely to be important for microorganisms. This discussion is frequently tied to the discussion of whether or not there are endemic microorganisms [4]. However, once sympatric speciation is considered, the existence of endemism is not a requirement for a large global number of microbial taxa.

Given that the corollary is contradicted by the repeated appearance of new sequences in molecular surveys, Fenchel [24] advances two arguments about why molecular surveys overestimate microbial diversity. One is that molecular surveys always claim to have found new sequences. However, these sequences might belong to already described organisms that have not been sequenced. This remains a possibility in the case of eukaryotes. However, this is impossible for bacteria or archaea because the SSU rDNA sequence is reported with all descriptions of new species. I think that this question is a minor one that will quickly be solved when the rRNA genes of all pure cultures are sequenced, an almost trivial objective with current sequencing power.

The second point is more interesting. The sequences used in diversity studies are genes that encode rRNA. Because rRNA is essential for the cell, most variation found in the sequence of these genes must necessarily be a result of neutral mutations. According to Fenchel [24], this variability "…does not necessarily reflect phenotypic differentiation in terms of morphological or physiological properties." The implication is that the diversity of sequences found in molecular surveys does not really reflect a diversity of 'true' species. However, in all cases in which the similarity of SSU rDNA from microorganisms has been compared with other taxonomic characters, SSU rDNA from microorganisms has been shown to be extremely conservative [34–37] (Box 1).

The example of the diatom *Pseudo-nitzschia delicatissima* (a marine eukaryotic microbe) is pertinent [36]. Montresor *et al.* (ASLO summer meeting, Santiago de

Box 2. Allopatric and sympatric speciation

Speciation requires that two populations of the same species become reproductively isolated from each other. Then each population can diversify from the other until they become two species. Speciation can occur when the two populations are separated by a geographical barrier (for example, each population colonizes a different island in an archipelago). This is called allopatric speciation. Speciation can also occur when the two populations coexist. This is called sympatric speciation. It is easier to envision how two populations can differ in allopatry. They might find different resources in the two islands or there could be genetic drift in different directions, for example. However, several mechanisms that enable speciation in sympatry have been described [27]. In fact, studies of the mechanisms of speciation in sympatry were considered part of the 'breakthrough of the year 2005' by *Science* [58].

An example among marine microorganisms is that of *Micromonas pusilla*, an abundant and widespread prasinophyte. Isolates from different parts of the ocean are morphologically undistinguishable and, thus, have been characterized as one cosmopolitan morphospecies. Slapeta *et al.* [59] studied a suite of nuclear, mitochondrial and chloroplast genes in *M. pusilla* isolates from different oceans. They found three or four lineages that were sufficiently different to merit species status. Some of these lineages were distributed globally and several of them co-existed in the same places. The authors concluded that these lineages were diversifying "in the absence of apparent geographical barriers (sympatric speciation)."

A case among bacteria was experimentally studied in the laboratory [30]. A culture of *Escherichia coli* was grown on minimal medium in a continuous culture where the only source of carbon and energy was glucose. In this situation, *E. coli* takes up the glucose and degrades it to acetate and some glycerol, which are both excreted into the medium and can be used later. Yet after ~800 generations, three different ecospecies were found to coexist. One was specialized in the use of glucose, another in acetate and the third in glycerol (the ecospecies consisted of 80%, 18% and 2% of the cells, respectively). Thus, three niches had appeared where there was only one and three ecospecies had evolved in one of the most constant and featureless environments possible. It is hard to believe that this phenomenon has not been taking place in the much more complex real world to create uncountable niches for different microbial species to develop.

Compostela, Spain, 2005) used their extensive collection of *P. delicatissima* strains isolated from the Bay of Naples to compare morphology, gene sequences and the ability to reproduce sexually. A strength of this system is that whether two strains interbreed or not determines if they belong to the same biological species. It was found that some strains with the same large subunit rDNA did not mate, whereas others did. However, when the more variable ITS region (interspacer region between the different rRNA genes) was compared in these strains, strains that mated always had the same ITS sequence and strains with different ITS sequences never mated. Thus, rDNA sequences are a conservative estimate of species, not the other way around. Similar studies have been carried out with freshwater green algae [37].

If the number of microbial species is close to the lower estimates proposed by Finlay and Fenchel, a description of a majority of microbial species has already been carried out. If the true numbers are close to the upper estimates, however, the challenge ahead is formidable. From the previous discussion I think it is likely that the number of microbial taxa approaches the upper estimates. The difference between the known diversity (the 10^3 described bacterial species) and the estimated total diversity (10^6–10^9 bacterial taxa) is the unknown. Obviously, the isolation and characterization of 10^6 or 10^9 bacteria are beyond the capabilities of current microbiologists. Thus, one might be inclined to agree with the pessimistic thought that "determining the number of species today is like reaching for the stars; there is no way with the data available today that we are even going to get reasonably close" [38].

Diversity and biodiversity

A way out can be found in the distinction between diversity and biodiversity proposed by Margalef [39,40]. Biodiversity would be the total genetic information on Earth or in any part of it. Diversity, in turn, would be the components that are active and abundant at one particular time and place. A literary analogy would equate biodiversity with a dictionary of all the words in a language, whereas diversity would be the particular set of words (and their frequency of use) chosen for a particular book. This distinction makes a lot of biological sense. Magurran and Henderson [41], for example, analyzed an assemblage of fish in an estuary over a period of 21 years. Some species were present every year and their abundance was high. They called these 'core' species. The remaining species had low abundance and were not found every year. These were 'occasional' species. The authors proposed that core species are responsible for carbon and energy flow in the ecosystem whereas occasional species wander sporadically to areas at their limits of tolerance and eventually disappear.

These ideas can be applied to microorganisms (Figure 2). The most abundant taxa are the core taxa. They are maintained through active growth because they are well adapted to that particular ecosystem. At the same time, they are subject to predation and viral lysis, thus fuelling carbon and energy flow. These taxa constitute the diversity of the ecosystem. The long tail of rare taxa (or 'seed bank') would complete the biodiversity. The taxa in the seed bank would be the equivalent of occasional species. However, in the case of microorganisms, being a member of a rare taxon has fascinating consequences.

Delights of a diluted life

The rare taxa are recruited through immigration, which is dependent on dispersal from other ecosystems. It has already been argued that because of their small size and large numbers, microbes should be easily dispersed everywhere [2,3,6,23,26], thus, immigration rates are expected to be large. By contrast, death and subsequent loss will eliminate taxa from the rare part of the curve, however, these taxa can be expected to have extremely low loss rates. The two main loss factors for microorganisms are viral lysis and predation by protists [42], so by being rare, a microbial taxon is protected from both.

Viruses depend on encounter probabilities to find a prey. If the prey taxon is found below a minimal threshold concentration it will become extremely improbable that the virus finds any of its members. All of the taxa with an abundance lower than this threshold will be protected from this loss factor. Viruses function by a 'kill the winner' strategy [43]. When a microbial taxon grows actively and increases its abundance above the threshold, its specific virus will infect it and reduce its numbers back to the threshold level.

In the case of predation, if the predators are not selective, both abundant and rare taxa will be culled down proportionally and being rare will not be advantageous. My assumption is that rare taxa are rare because they are not growing (or growing slowly) compared with the abundant members of the assemblage. It is well known that bacterivores selectively prey on the largest and most active bacteria [42,44,45]. In fact, a predation resistance strategy is to become small, and actively growing bacteria are known to be bigger than starving bacteria. Rare taxa, therefore, will be represented by smaller cells and will be more resistant to grazing than the abundant, actively growing cells. In effect, the slow-growing taxa will suffer far less predation losses than their more abundant competitors.

As a consequence of the small losses, a rather long tail of rare taxa in microbial assemblages can be expected (Figure 2). An important property of this curve is that rare taxa can be promoted from the seed bank to the core zone of the curve. This might happen in a few days or weeks if conditions become adequate for the growth of the particular taxon. Likewise, a member of the core can be grazed down or reduced by viral lysis beyond the threshold level to become a member of the rare taxa without disappearing from the ecosystem. If it were true that all microbes are easily dispersed globally, the long tail of any ecosystem would include all of the microorganisms on Earth. In effect, the biodiversity of any ecosystem would be identical to the biodiversity on the planet. Ecosystem functions, however, would be carried out by those taxa that make up the diversity of that particular ecosystem at any given time point.

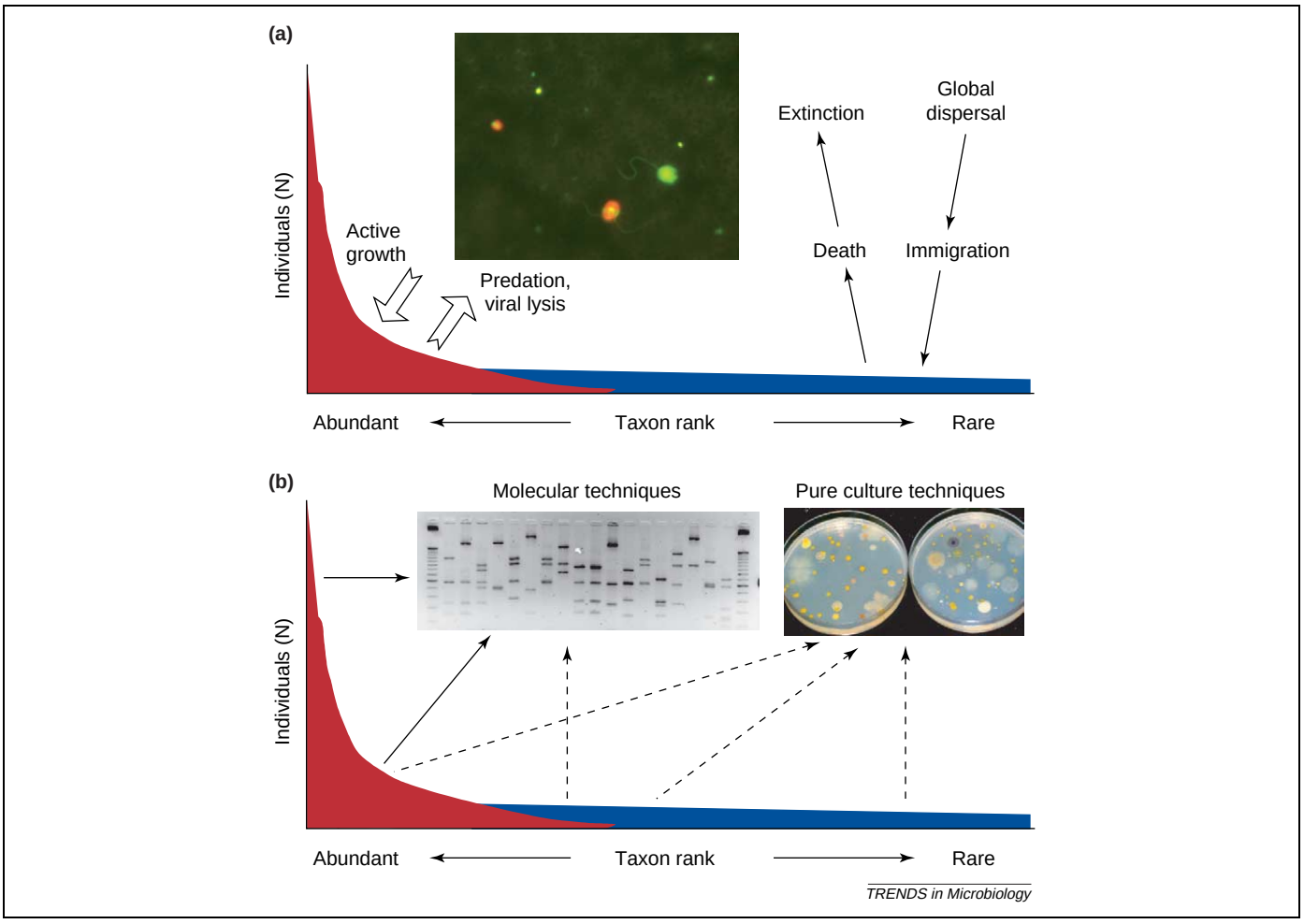

Figure 2. Plots of number of individuals versus taxon, with taxa ranked according to their respective abundance. **(a)** The total curve represents biodiversity and is postulated to be composed of two sections: (i) diversity (red), which contains the abundant taxa that have an active role in carbon and energy flow in a given ecosystem. These taxa are maintained by active growth and suffer losses because of predation and viral lysis. The insert shows two autotrophic flagellates (with red autofluorescence of chloroplasts), one heterotrophic flagellate (large green shape) and bacteria (small green dots), and corresponds to a winter sample from the Beaufort Sea (photograph by Dolors Vaqué). (ii) The blue section of the curve corresponds to rare taxa, which survive in the ecosystem at low abundance, perhaps with resting stages or spores. This 'seed bank' recruits new species through immigration. Because of the easy dispersal of microorganisms, it is likely that essentially all microorganisms reach this ecosystem periodically. By being rare, however, these microorganisms do not experience viral attack and predation is minimized; thus, extinction can only be a result of cell death. Because extinction can be expected to be low while immigration is probably high, this part of the curve might be long and perhaps include the whole biodiversity of microorganisms. **(b)** The same curve illustrates which portions of biodiversity can be retrieved with different techniques. Molecular techniques can presumably access most taxa in the diversity part of the curve (solid arrows) but the biodiversity tail is usually not accessible. The left-hand insert shows RFLP patterns of 18S rDNA clones from a library constructed from a coastal Mediterranean sample of picoeukaryotes (photograph by Beatriz Díez). Pure culture techniques can only retrieve a few of the taxa present in the environment (broken arrows). However, these techniques have the potential to retrieve both abundant and rare taxa. The right-hand insert shows Petri dishes with ZoBell's medium inoculated with a sample from the coastal Mediterranean and incubated in the dark. Yellow and orange colonies correspond to Bacteroidetes, small light-pink colonies are *Roseobacter* species and the large white colonies are γ-proteobacteria (photograph by Jarone Pinhassi).

What portion of the unknown can be determined?

Rare taxa will seldom be retrieved by cloning and sequencing with universal primers. PCR primers will predominantly hybridize with the common taxa and the rare ones will remain unknown for the same reasons that viruses cannot find rare bacterial taxa in a reasonable amount of time. The situation can be improved by using primers specific for some groups of bacteria (such as ammonia oxidizers [46]) by techniques such as nested PCR [47] and by shotgun cloning, in which PCR is not used [48,49]. However, a large portion of the seed bank will remain beyond the reach of current molecular tools. Essentially, only the abundant taxa will be accessible (Figure 2b). Some studies indicate that taxa which make up $\geq 1\%$ of the total cell number can be detected with PCR-dependent techniques, and taxa that comprise $<0.1\%$ are difficult to retrieve [50,51]. I propose that the taxa retrieved by PCR are considered to form the diversity of an ecosystem. Then, by definition, diversity can be determined with current molecular techniques. Abundant taxa such as SAR11 (a group of bacteria that comprise 20–30% of the total count in many oceans) will certainly be found in this group. Taxa that are usually less abundant but important in nutrient cycles, such as ammonia oxidizers, will fall within this 'abundant' group in most instances because their concentrations are close to 1% of the total [52]. Thus, most bacteria that are relevant and active in carbon, energy and nutrient flows will be within the diversity retrievable by molecular techniques.

Concluding remarks and future perspectives

Studies of bacterial diversity within this framework make sense from an ecological point of view. Thus, for example,

latitudinal patterns in the distribution of the bacterial taxa that form the diversity have been found by cloning and sequencing large clone libraries from different oceans (T. Pommier *et al.*, unpublished). If the seed bank from such samples could have been analyzed simultaneously, it is likely that 'everything would be found everywhere'.

Some rare taxa will eventually be discovered with conventional molecular techniques. However, most of the rare taxa in the seed bank will only be accessible if they can be grown in enrichment or, even better, pure cultures. As has been learnt the hard way, this is only possible with a few of the microorganisms that form biodiversity [6].

Within this framework, microbial ecologists who are interested in functional aspects of carbon flow only need to be concerned with diversity (i.e. with the taxa retrievable with a clone library). From the point of view of general ecologists, the comparison of the peculiar situation of the rare microbial taxa with that of rare species of animals and plants is thought-provoking. And for microbiologists who are interested in metabolisms or industrially interesting genes, the realization that a large tail of biodiversity is not currently accessible represents a tremendous challenge to develop ingenious molecular and culturing methods. It is suggestive that in both the cases of stars in an expanding universe and rare bacteria in a diluted microscopic world, the depth of our knowledge depends on the acuity of our tools, whether the tools are a microscope or a telescope.

Acknowledgements

My work is currently supported by grant CTM2004–02586/MAR from the Spanish Ministry of Education and Science and the EU network of excellence 'Marine Genomics Europe' EU-FP6–505403. Ramon Massana, Josep M. Gasol, Jarone Pinhassi, Thomas Pommier and three anonymous reviewers provided useful comments on earlier versions of the manuscript.

References

1 Yarincik, K. and O'Dor, R. (2005) The census of marine life: goals, scope and strategy. *Sci. Mar.* 69(Suppl. 1), 201–208
2 Finlay, B.J. and Esteban, G.F. (2004) Ubiquitous dispersal of free-living microorganisms. In *Microbial Diversity and Bioprospecting* (Bull, A.T., ed.), pp. 216–224, ASM Press
3 Hedlund, B.P. and Staley, J.T. (2004) Microbial endemisms and biogeography. In *Microbial Diversity and Bioprospecting* (Bull, A.T., ed.), pp. 225–231, ASM Press
4 Whitfield, J. (2005) Biogeography: is everything everywhere? *Science* 310, 960–961
5 Whitman, W.B. *et al*. (1998) Prokaryotes: the unseen majority. *Proc. Natl. Acad. Sci. U. S. A.* 95, 6578–6583
6 Pedrós-Alió, C. (1993) Diversity of bacterioplankton. *Trends Ecol. Evol.* 8, 86–90
7 Fenchel, T. (2001) Ecology – marine bugs and carbon flow. *Science* 292, 2444–2445
8 Copley, J. (2002) All at sea. *Nature* 415, 572–574
9 Karl, D.M. (2002) Hidden in a sea of microbes. *Nature* 415, 590–591
10 Baldauf, S.L. (2003) The deep roots of eukaryotes. *Science* 300, 1703–1706
11 Giovannoni, S.J. *et al*. (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* 345, 60–63
12 Fuhrman, J.A. *et al*. (1992) Novel major archaebacterial group from marine plankton. *Nature* 356, 148–149
13 DeLong, E.F. (1992) Archaea in coastal marine environments. *Proc. Natl. Acad. Sci. U. S. A.* 89, 5685–5689
14 Díez, B. *et al*. (2001) Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. *Appl. Environ. Microbiol.* 67, 2932–2941
15 López-García, P. *et al*. (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* 409, 603–607
16 Moon-van der Staay, S.Y. *et al*. (2001) Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* 409, 607–610
17 Margulis, L. *et al.*, eds (1990) *Handbook of Protoctista* (1st edn), Jones & Bartlett Publishers
18 Kemp, P. and Aller, J. (2004) Bacterial diversity in aquatic and other environments: what 16S rDNA libraries can tell us. *FEMS Microbiol. Ecol.* 47, 161–177
19 Hagström, Å. *et al*. (2002) Use of 16S ribosomal DNA for delineation of marine bacterioplankton species. *Appl. Environ. Microbiol.* 68, 3628–3633
20 Pommier, T. *et al*. (2005) Biogeographic analysis of ribosomal RNA clusters from marine bacterioplankton. *Aquat. Microb. Ecol.* 41, 79–89
21 Curtis, T.P. *et al*. (2002) Estimating prokaryotic diversity and its limits. *Proc. Natl. Acad. Sci. U. S. A.* 99, 10494–10499
22 Dykhuizen, D.E. (1998) Santa Rosalia revisited: why are there so many species of bacteria? *Antonie Van Leeuwenhoek* 73, 25–33
23 Finlay, B.J. (2002) Global dispersal of free-living microbial eukaryote species. *Science* 296, 1061–1063
24 Fenchel, T. (2005) Where are all the species? *Environ. Microbiol.* 7, 473–475
25 Fenchel, T. (2005) Cosmopolitan microbes and their 'cryptic' species. *Aquat. Microb. Ecol.* 41, 49–54
26 Quispel, A. (1998) Lourens G.M. Baas Becking (1895–1963). Inspirator for many (micro)biologists. *Int. Microbiol.* 1, 69–72
27 Via, S. (2001) Sympatric speciation in animals: the ugly duckling grows up. *Trends Ecol. Evol.* 16, 381–390
28 Bearhop, S. *et al*. (2005) Assortative mating as a mechanism for rapid evolution of a migratory divide. *Science* 310, 502–504
29 Mendelson, T.C. and Shaw, K.L. (2005) Rapid speciation in an arthropod. *Nature* 433, 375–376
30 Rosenzweig, R.F. *et al*. (1994) Microbial evolution in a simple unstructured environment: genetic differentiation in *Escherichia coli*. *Genetics* 137, 903–917
31 Fuhrman, J.A. and Campbell, L. (1998) Microbial microdiversity. *Nature* 393, 410–411
32 Casamayor *et al.* (2002) Microheterogeneity in 16S ribosomal DNA-defined bacterial populations from a stratified planktonic environment is related to temporal changes and to ecological adaptations. *Appl. Environ. Microbiol.* 68, 1706–1714
33 Acinas, S. *et al*. (2005) Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* 430, 551–554
34 Knowlton, N. (2000) Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia* 420, 73–90
35 Roselló-Mora, R. and Amann, R. (2001) The species concept for prokaryotes. *FEMS Microbiol. Rev.* 25, 39–67
36 Orsini, L. *et al*. (2004) Multiple rDNA ITS-types within the diatom *Pseudo-nitzschia delicatissima* (Bacillariophyceae) and their relative abundances across a spring bloom in the Gulf of Naples. *Mar. Ecol. Prog. Ser.* 271, 87–98
37 Coleman, A.W. (2000) The significance of a coincidence between evolutionary landmarks found in mating affinity and a DNA sequence. *Protist* 151, 1–9
38 Erwin, T.E. (1991) How many species are there? Revisited. *Conserv. Biol.* 5, 330–333
39 Margalef, R. (1994) Diversity and biodiversity – their possible meaning in relation with the wish for sustainable development. *An. Acad. Bras. Ciênc.* 66(Suppl. 1), 3–14
40 Margalef, R. (1997) *Our Biosphere*. (*Excellence in Ecology Series*, number 10), Ecology Institute
41 Magurran, A.E. and Henderson, P.A. (2003) Explaining the excess of rare species in natural species abundance distributions. *Nature* 422, 714–716
42 Pernthaler, J. (2005) Predation on prokaryotes in the water column and its ecological implications. *Nat. Rev. Microbiol.* 3, 537–546
43 Thingstad, T.F. (2000) Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic viruses in aquatic systems. *Limnol. Oceanogr.* 45, 1320–1328

44 Vaqué, D. *et al*. (2001) Dynamics of whole community bacterial production and grazing losses in seawater incubations as related to the changes in the proportions of bacteria with different DNA content. *Aquat. Microb. Ecol.* 25, 163–177
45 Jurgens, K. and Matz, C. (2002) Predation as a shaping force for the phenotypic and genotypic composition of planktonic bacteria. *Antonie Van Leeuwenhoek* 81, 413–434
46 Ward, B.B. (2002) How many species of prokaryotes are there? *Proc. Natl. Acad. Sci. U. S. A.* 99, 10234–10236
47 Benlloch, S. *et al*. (2002) Prokaryotic genetic diversity throughout the salinity gradient of a coastal solar saltern. *Environ. Microbiol.* 4, 349–360
48 Schmidt, T.M. *et al*. (1991) Analysis of a marine picoplankton community by 16S ribosomal-RNA gene cloning and sequencing. *J. Bacteriol.* 173, 4371–4378
49 Venter, J.C. *et al*. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304, 66–74
50 Muyzer, G. *et al*. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* 59, 695–700
51 Casamayor, E.O. *et al*. (2000) Identification and spatio-temporal differences between microbial assemblages from two neighboring sulfurous lakes: comparison by microscopy and denaturing gradient gel electrophoresis. *Appl. Environ. Microbiol.* 66, 499–508
52 Ward, B.B. and O'Mullan, G.D. (2002) Worldwide distribution of *Nitrosococcus oceani*, a marine ammonia-oxidizing γ-Proteobacterium, detected by PCR and sequencing of 16S rRNA and *amoA* genes. *Appl. Environ. Microbiol.* 68, 4153–4157
53 Mayden, R.L. (1997) A hierarchy of species concepts: the denouement in the saga of the species problem. In *Species: the Units of Biodiversity* (Claridge, M.F., Dawah, H.A. and Wilson, M.R., eds), pp. 381–324, Chapman and Hall
54 Potter, D. *et al*. (1997) Convergent evolution masks extensive biodiversity among marine coccoid picoplankton. *Biodiv Conserv* 6, 99–107
55 Stackebrandt, E. and Goebel, B.M. (1994) Taxonomic note: a place for DNA–DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Bacteriol.* 44, 846–849
56 Kolasa, J. and Biesiadka, E. (1984) Diversity concept in ecology. *Acta Biotheor.* 33, 145–162
57 Pedrós-Alió, C. (2005) Diversity of microbial communities: the case of solar salterns. In *Adaptation to Life at High Salt Concentrations in Archaea, Bacteria, and Eukarya* (Gunde-Cimerman, N., Plemenitas, A. and Oren, A. eds), pp. 71–90 (*Cellular Origins, Life in Extreme Habitats and Astrobiology (COLE) Series*, vol. 9), Springer Science
58 Culotta, E. and Pennisi, E. (2005) Evolution in action. Breakthrough of the year. *Science* 310, 1878–1879
59 Slapeta, J. *et al*. (2006) Global dispersal and ancient cryptic species in the smallest marine eukaryotes. *Mol. Biol. Evol.* 23, 23–29