MINIREVIEW

Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity

PHILIP HUGENHOLTZ, BRETT M. GOEBEL,† AND NORMAN R. PACE*
Departments of Plant and Microbial Biology and Molecular and Cell Biology, University of California, Berkeley, Berkeley, California 94720-3102

INTRODUCTION
Our perspective on microbial diversity has improved enormously over the past few decades. In large part this has been due to molecular phylogenetic studies that objectively relate organisms. Phylogenetic trees based on gene sequences are maps with which to articulate the elusive concept of biodiversity. Thus, comparative analyses of small-subunit rRNA (16S or 18S rRNA) and other gene sequences show that life falls into three primary domains, Bacteria, Eucarya, and Archaea (51, 52). Based on rRNA trees, the main extent of Earth’s biodiversity is microbial. Our knowledge of the extent and character of microbial diversity has been limited, however, by reliance on the study of cultivated microorganisms. It is estimated that >99% of microorganisms observable in nature typically are not cultivated by using standard techniques (1).

Recombinant DNA and molecular phylogenetic methods have recently provided means for identifying the types of organisms that occur in microbial communities without the need for cultivation (see references 1, 20, and 35 for reviews). Results from application of these methods to a number of diverse environments confirm that our view of microbial diversity was limited and point to a wealth of novel and environmentally important diversity yet to be studied (34). It is the aim of this review to collate, compare, and incorporate the results of the environmental sequence-based studies into the context of known bacterial diversity. We discuss the sequence data at the taxonomic level of the phylogenetic division because divisions constitute first-order clades for describing the breadth of bacterial diversity. Although we have yet to determine even the outlines of the bacterial tree, common threads are beginning to emerge that revise our current views of bacterial diversity and distribution in the environment.

PHYLOGENETIC DIVERSITY IN THE BACTERIAL DOMAIN
In 1987, Woese described the bacterial domain as comprised of about 12 natural relatedness groups, based mainly on analyses of familiar cultivated organisms such as cyanobacteria, spirochetes, and gram-positive bacteria (all of which, based on rRNA sequence divergence, display greater evolutionary depth than plants, animals, and fungi) (51). These relatedness groups have variously been called “kingdoms,” “phyla,” and “divisions”; we use the latter term. For the purposes of this review we define a bacterial division purely on phylogenetic grounds as a lineage consisting of two or more 16S rRNA sequences that are reproducibly monophyletic and unaffiliated with all other division-level relatedness groups that constitute the bacterial domain. We judge reproducibility by the use of multiple tree-building algorithms, bootstrap analysis, and varying the composition and size of data sets used for phylogenetic analyses. The typical interdivisional rRNA sequence difference is 20 to 25%. For comparison, the 16S rRNAs of Escherichia coli and Pseudomonas aeruginosa, both representatives of the γ group of Proteobacteria, differ overall by about 15%; the 16S rRNAs of E. coli and Bacillus subtilis (“low-G+C gram-positive bacterial” division) differ by about 23%.

At the current stage in the phylogenetic classification of Bacteria, divisions are not consistently named or taxonomically ranked. rRNA-defined divisions are identified by classes (e.g., Proteobacteria [41] and Actinobacteria [42]), orders (e.g., Thermotogales and Aquificales), families (e.g., Chlorobiaceae), generic names such as the Nitrospira group (11), or common names such as the green nonsulfur (GNS) bacteria and low-G+C gram-positive bacteria (51). Division-level nomenclature has not even been consistent between studies, so some divisions are identified by more than one name. For instance, green sulfur bacteria is synonymous with Chlorobiaceae; high-G+C gram-positive bacteria is synonymous with Actinobacteria and Actinomycetales. Indeed, it probably is premature to standardize taxonomic rankings for bacterial divisions at this point when our picture of microbial diversity is likely still incomplete and the topology of the bacterial tree is still unresolved.

In the past decade the number of identifiable bacterial divisions has more than tripled to about 40 due in significant part to culture-independent phylogenetic surveys of environmental microbial communities (21, 34). These analyses rely on sequences of rRNA genes obtained by cloning directly from environmental DNA or, as in the majority of studies, after amplification by the PCR (1, 20, 35). Figure 1 represents the division-level diversity of the bacterial domain as inferred from representatives of the approximately 8,000 bacterial 16S rRNA gene sequences currently available. Although 36 divisions are shown in Fig. 1, several other division-level lineages are indicated by single environmental sequences (9, 21, 37), suggesting that the number of bacterial divisions may be well over 40. Several of the described divisions are well represented by cultivated strains and were the first to be characterized phylogenetically (51). The majority of the bacterial divisions, however, are poorly represented by cultured organisms. Indeed, 13 of the 36 divisions shown in Fig. 1 are characterized only by environmental sequences (shown outlined) and are termed “candidate divisions” to indicate their unsubstantiated status as
new bacterial divisions (21). One of these candidate divisions, OP11, is now sufficiently well represented by environmental sequences to conclude that it constitutes a major bacterial group (see below). Phylogenetic studies so far have not resolved branching orders of the divisions; bacterial diversity is seen as a fan-like radiation of division-level groups (Fig. 1). The exception to this, however, is the *Aquificales* division, which branches most deeply in the bacterial tree in most analyses.

**BACTERIAL DIVERSITY AND DISTRIBUTION IN THE ENVIRONMENT**

Culture-dependent studies indicate that representatives of some bacterial divisions are cosmopolitan in the environment, whereas others appear restricted to certain habitats (39). Culture-independent studies so far conducted reflect and expand this view. Table 1 summarizes the environmental distribution of sequences by habitat type, compiled from most of the available 16S rRNA-based clonal analyses: 86 studies contributing nearly 3,000 sequences. An expanded version of this table that details division-level representation in the individual studies is available at http://crab2.berkeley.edu/pacelab/176.htm. Table 1 includes only divisions for which representatives have been detected in at least two independent studies and for which at least one near-complete 16S rRNA gene sequence is known. Table 1 is, therefore, not an exhaustive listing of potential division-level diversity for all studies.

Sequence representatives of several bacterial divisions have been identified in a wide range of habitats, suggesting the cosmopolitan or ubiquitous distribution of the corresponding organisms in the environment and, potentially, their broad metabolic capabilities. Some of these cosmopolitan divisions are well-known from cultivation studies; however, others are little known or have not yet been detected by cultivation. Figure 2 summarizes the representation of selected cosmopolitan divisions by sequences of cultivated and uncultivated organisms. The *Proteobacteria* (purple photosynthetic bacteria and relatives), *Cytophagales* (*Bacteroides-Cytophaga-Flexibacter* group), and the two gram-positive divisions, *Actinobacteria* and low-G+C gram-positive bacteria, are well represented by cultivat-

![Evolutionary distance tree of the bacterial domain showing currently recognized divisions and putative (candidate) divisions.](http://crab2.berkeley.edu/pacelab/176.htm)
ed organisms and therefore are familiar to us in principle. These four divisions account for 90% of all cultivated bacteria characterized by 16S rRNA sequences and approximately 70% of the environmental sequences collated in Table 1. By contrast, other cosmopolitan divisions revealed by clonal analyses, such as Acidobacterium, Verrucomicrobia, GNS bacteria, and OP11, are poorly represented by sequences from cultivated organisms (Fig. 2) and consequently are little known with regard to their general properties. Although many of the bacterial divisions occur widely, others seem to occupy a more limited range of habitats (Table 1). All cultivated representatives of Aquificales, for instance, are thermophilic hydrogen metabolizers, and all environmental sequences of Aquificales have been obtained only from high-temperature environments. This suggests a specialized habitat niche for this group. Alternatively, the apparently limited environmental distribution may simply reflect a sampling or methodological artifact and representatives of such divisions may be present in a wider range of habitats, but not yet detected.

The database of environmental rRNA sequences is compromised in resolving some phylogenetic issues by a large number of relatively short sequences. More than half of the sequences collated in Table 1 are less than 500 nucleotides (nt) long, which represents only one-third of the total length of 16S rRNA. This is due to an unfortunate trend in many environmental studies of sequencing only a portion of the gene in the belief that a few hundred bases of sequence data is sufficient for phylogenetic purposes. Indeed, 500 nt is sufficient for placement if some longer sequence is closely related (>90% identity in homologous nucleotides) to the query sequence. In the case of novel sequences, however, <85% identity to known sequences is insufficient to place the sequence accurately in a phylogenetic tree and can even be misleadingly.

FIG. 2. Relative representation in selected cosmopolitan bacterial divisions of 16S rRNA sequences from cultivated and uncultivated organisms. Results were compiled from 5,224 and 2,918 sequences from cultivated and uncultivated organisms, respectively.

TABLE 1. Summary of 16S rRNA-based clonal analyses of diversity of uncultivated bacteria

<table>
<thead>
<tr>
<th>Habitats type</th>
<th>No. of studies</th>
<th>No. of sequences</th>
<th>Bacterial division</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geothermal</td>
<td>10</td>
<td>212</td>
<td>Thermotogales</td>
</tr>
<tr>
<td>Soil</td>
<td>16</td>
<td>743</td>
<td>Cytophagales</td>
</tr>
<tr>
<td>Marine</td>
<td>23</td>
<td>687</td>
<td>Archaea</td>
</tr>
<tr>
<td>Freshwater</td>
<td>4</td>
<td>107</td>
<td>Actinobacteria</td>
</tr>
<tr>
<td>Wastewater</td>
<td>5</td>
<td>430</td>
<td>GNS bacteria</td>
</tr>
<tr>
<td>Pollutant associated</td>
<td>7</td>
<td>202</td>
<td>Planctomycetes</td>
</tr>
<tr>
<td>Acid metal leaching</td>
<td>2</td>
<td>21</td>
<td>Thermus/Deinococcus</td>
</tr>
<tr>
<td>Subsurface</td>
<td>6</td>
<td>229</td>
<td>Cyanobacteria</td>
</tr>
<tr>
<td>Symbionts and commensals</td>
<td>10</td>
<td>280</td>
<td>Thermodesulobacterium</td>
</tr>
<tr>
<td>Disease associated</td>
<td>3</td>
<td>7</td>
<td>Thermotogales</td>
</tr>
<tr>
<td>Totals</td>
<td>86</td>
<td>2,918</td>
<td></td>
</tr>
</tbody>
</table>

*An expanded version of this table detailing individual studies is available at [http://crab2.berkeley.edu/pacelab/176.htm](http://crab2.berkeley.edu/pacelab/176.htm).*
Since all but 4 (40, 46, 49, 50) of the 86 studies collated in Table 1 were conducted using PCR to amplify rDNA from extracted environmental DNA, the question arises as to whether molecular analyses accurately reflect the division-level diversity that occurs in the environment. It is well established that PCR-associated artifacts such as differential amplification of different rDNA templates (36, 44), sensitivity to rRNA gene copy number (12), PCR primer specificity (48), sensitivity to template concentration (6), amplification of contaminant rDNA (45), and formation of chimeric sequences (23) may skew our assessment of microbial diversity. Most of the studies collated in Table 1, however, analyzed tens to hundreds of clones, so it seems likely that these studies have sampled the main types of sequences in the communities examined. We believe, acknowledging the caveats of the methodology, that the clonal analyses collated in Table 1 probably include the most abundant (metabolically active) bacterial sequence types in the samples analyzed, likely representing the members of the communities that are involved in the principal metabolic activities, such as carbon cycling.

**ABUNDANT BUT LITTLE-KNOWN BACTERIAL DIVISIONS**

The rRNA sequence studies of environmental organisms probably identify the abundant organisms in the environments studied and, therefore, account for the organisms that participate significantly in the maintenance of the communities. Because of their abundance in the environment, representatives of some poorly studied phylogenetic divisions are predicted to play significant roles in environmental chemistry. Examples of such divisions, which because of their potential environmental significance merit study, are the Acidobacterium division, the Verrucomicrobia, the GNS bacteria, and candidate division OP11.

**Acidobacterium division.** The Acidobacterium group is a newly recognized bacterial division with only three cultivated representatives: Acidobacterium capsulatum (18), Holophaga foetida (26), and Geothrix fermentans (28). Figure 3 is a phylogenetic dendrogram of this group, including selected environmental representatives. The limited physiological information known about these organisms provides few clues to properties that might be general throughout the division. Acidobacterium is a moderately acidophilic aerobic heterotroph; Holophaga and Geothrix are strict anaerobes that ferment aromatic compounds and acetate, respectively. The majority of sequences that make up this division, however, are from environmental clones. At least eight monophyletic subdivisions in the Acidobacterium group are identified by phylogenetic analyses (Fig. 3 [24, 29]). We define a subdivision as a lineage comprised of two or more 16S rRNA sequences within a division that are reproducibly monophyletic and unaffiliated with all other representatives of that division. Acidobacterium subdivisions 1, 2, 3, 4, and 6 are well represented by environmental clone sequences from independent studies, yet no cultivated strains are known with the exception of subdivision 1, represented by A. capsulatum. The widespread occurrence of environmental sequences belonging to the Acidobacterium division (Table 1) suggests that members of this group are ecologically significant constituents of many ecosystems, particularly soil communities. They have been detected in every clonal analysis of soils (with a wide range of chemical properties), as well as in other habitats, including a peat bog, acid mine drainage, a contaminated aquifer, a hot spring, a freshwater lake, and a sample of the Atlantic ocean from a depth of 1,000 m (Fig. 3).

In situ single-cell analyses with fluorescent hybridization probes specific for Acidobacterium subdivision 6 small-subunit rRNA indicate that this subdivision is morphologically diverse (29), as expected for a broad phylogenetic group. Members likely are metabolically diverse as well: the depth of phylogenetic diversity (depth of branching) in the Acidobacterium division is nearly as great as in the Proteobacteria.

**Verrucomicrobia.** Verrucomicrobia is a newly proposed division of Bacteria (17) represented by a handful of isolates: Verrucomicrobium spinosum (after which the division is named) (47), four Prosthecobacter species (17), and three strains of ultramicrobacteria (22). Verrucomicrobia and Prosthecobacter are prokaryotes harvested from freshwater and the ultramicrobacteria, “dwarf-cell” strains only about 0.1 µm in volume, were isolated from a soil habitat. All of these isolates preferentially use sugars as growth substrates. Culture-independent analyses indicate that the Verrucomicrobia, like members of the Acidobacterium division, are widespread in the environment and abundant, particularly in soils (Table 1). Figure 4 shows a dendrogram of representatives of the Verrucomicrobia. Several monophyletic subdivisions are seen, only two of which are represented by the cultivated strains. Clone sequences of this division from soil are predominantly members of the phylogenetically broad subdivisions 2 and 3. The abundance of these two groups suggests their ecological importance. For instance, the abundance of one representative of Verrucomicrobia subdivision 2 (EA25) was estimated by PCR at 10^6 to 10^7 cells per g of a pasture soil sample, 1 to 10% of the total microbial content (25).

In our phylogenetic analyses we consistently find that the division Chlamydia is a specific sister group of the Verrucomicrobia. We find no support for the notion (17, 30, 47) of a specific relatedness of the planctomycetes with the Verrucomicrobia.

**GNS bacteria.** The GNS bacteria have been recognized as a division-level bacterial group for over a decade (51). Even today, however, this division is still represented by only a few isolates. The cultured representatives have a wide range of phenotypes, from anoxicogenic photosynthesis (Chloroflexus) to thermophilic organotrophy (Thermomicrobium). Figure 5 shows the relatedness groups of GNS bacteria detected in the environment. It is apparent from the dendrogram that all of the cultivated representatives except the chlorinated hydrocarbon-reducing Dehalococcoides ethenogenes (31) are related in subdivision 3, together with several clone sequences from a hot spring, a rice paddy, and activated sludge (data not shown). By contrast, most of the environmental sequences described to date fall into a different relatedness group, subdivision 1, with no cultivated representatives. Considering the wide variety of habitats that have contributed GNS sequences (Fig. 5; Table 1), particularly to GNS subdivision 1, members of this division likely play significant roles in the environment.

**Candidate division OP11.** Candidate division OP11 is a recently proposed novel bacterial division for which there is no reported cultivated representative (19, 21). However, several independent clonal studies have reported environmental sequences that together form the OP11 clade. Figure 6 shows a dendrogram of the known environmental sequence representatives of the division, with five subdivisions currently identifiable. OP11 sequences all have highly atypical sequence signatures for the domain Bacteria (51), and they have low sequence identities, only about 80%, to sequences outside the OP11 division. This may be due to higher-than-average mutation rates in OP11 rRNAs, as has been suggested for other groups such as the planctomycetes (27). OP11 sequences have been obtained from a variety of habitats including several different types of soil, freshwater sediments, the deep subsurface, and hot springs (Table 1), suggesting that members of the division
Acidobacterium division

Aacidobacterium capsulatum (DZ6171)

peat bog clone TM29 (X97101)

Acidobacterium capuluiatum

peat bog clone TM29 (X97101)

Acidobacterium division

soil clone MC13 (X85464)

soil clone Ep-T1 152 (Z73936)

soil clone Ep-T2 164 (Z73965)

soil clone Ep-T1 188 (Z73368)

soil clone Ep-T1 153 (Z73304)

soil clone P28 (U66859)

Acidobacterium capsulatum (DZ6171)

peat bog clone TM29 (X97101)

Acidobacterium division

soil clone M26 (U86861)

soil clone DAO52 (Y76746)

peat bog clone TM1 (X97007)

soil clone MC27 (X85468)

soil clone MC26 (X86408)

soil clone P15 (U66848)

soil clone MC29 (X86407)

soil clone M10 (U66697)

soil clone Co2 (AF013515)

hot spring clone OPB3 (AF027004)

soil clone S93 (AF013550)

soil clone RB41 (Z95722)

soil clone 32-11 (Z95711)

soil clone S27 (AF013545)

soil clone RGB3 (Z95702)

soil clone C150 (AF013530)

soil clone 11-25 (Z95709)

soil clone P12 (U66846)

soil clone S111 (AF013560)

soil clone 11-24 (Z95708)

soil clone R8 (U66644)

soil clone M43 (U66825)

soil clone 11-14 (Z95703)

soil clone k24266 (Z95732)

soil clone DAO3 (Y76566)

soil clone C112 (AF013534)

marine clone G81-114 (U86469)

soil clone C28 (AF013527)

soil clone RB24 (Z95717)

soil clone RB24 (Z95721)

soil clone RB24 (Z95717)

soil clone DB2 (Y12997)

soil clone R116 (Z95716)

soil clone S125 (AF013567)

soil clone mb288 (Z95733)

lake clone 420-138 (A010024)

soil clone 52-20 (Z95713)

soil clone i18 (Z95729)

soil clone DB2 (Z95716)

Geothrix fermentans (U41563)

Holophaga foetida (X77215)

CONCLUSION

Phylogenetic trees based on rRNA sequences show that bacterial diversity is represented by natural relatedness groups, the phylogenetic divisions (51). About 36 such divisions are currently identifiable. The final extent of division-level diversity in the bacterial domain is still unknown but clearly will be more than 40 divisions. Culture-independent studies have resulted in multiple hits on the majority of described divisions in different habitat types (Table 1), suggesting that the final number of divisions will be within the same order of magnitude as the present estimate.

The molecular analyses of environmental DNA have revealed substantial phylogenetic diversity with little or no representation among organisms previously studied. Because of their abundance and wide distribution, some of the organisms represented by the sequences likely contribute significantly to the global chemical cycles. Descriptions of newly identified, but apparently important, bacterial divisions such as the Acidobacterium and Verrucomicrobia, are presently confounded by too few cultivated representatives and only rudimentary descriptions of the strains. Cultivation efforts need to be directed at new representatives of the diverse groups for further study. Continued work to sequence the 16S rDNAs of all deposited type cultures (<50% sequenced to date [14]) may also result in
detection of additional cultivated representatives of newly described divisions. It is a challenge to microbial biologists to determine the physiological diversity and environmental roles of these recently articulated divisions of \textit{Bacteria}.

The phylogenetic differences between the bacterial divisions probably are reflected in substantial physiological differences. Some properties, the general properties of \textit{Bacteria}, are expected to be distributed among all the divisions. Division-specific
Green non-sulfur division

contaminated aquifer clone WCHB1-05 (AF050562)
contaminated aquifer clone WCHB1-43 (AF050564)
contaminated aquifer clone WCHB1-62 (AF050566)
sludge clone T76 (Z94009)
contaminated aquifer clone WCHB1-31 (AF050569)
sediment clone RB41 (U62857)
contaminated aquifer clone WCHB1-63 (AF050568)
hot spring clone OPB65 (AF027036)
hot spring clone OP89 (AF027043)
contaminated aquifer clone WCHB1-50 (AF050571)
hot spring clone OPB12 (AF027031)
subsurface clone H1.43.f (AF005749)

subsurface clone H1.4.f (AF005748)
subsurface clone H3.93 (AF005750)
subsurface clone t0.6.f (AF005745)
subsurface clone t0.8.f (AF005746)
Dehalococcoides ethenogenes (AF004928)
peat bog clone TM5 (X97112)

sediment clone RB21 (U62842)

Chloroflexus aurantiacus (M34116)
Chloroflexus aggregans (D32255)
hot spring clone OSV-L-20 (L04703)
Heliothrix oregonensis (L04675)
hot spring clone OS-VI-L-4 (M62774)
Herpetosiphon aurantiacus (M34117)
Sphaerobacter thermophilus (X53210)
Thermomicrobium roseum (M34115)
sediment clone RB03 (U62827)
soil clone M1 (U68589)

marine clone SAR202 (U20797)
marine clone SAR307 (U20798)
subsurface clone H1.2.f (AF005747)

FIG. 5. Phylogenetic dendrogram of the GNS division. Names of cultivated organisms are shown in bold. The habitat source of each environmental sequence is indicated before the clone name. GenBank accession numbers are listed parenthetically. Subdivisions (see the text) are indicated by brackets at the right of the tree. Tree construction and support for branch points was as described for Fig. 1 and 3, respectively. The scale bar indicates 0.1 change per nucleotide.
OP11 division

FIG. 6. Phylogenetic dendrogram of the OP11 division. The habitat source of each environmental sequence is indicated before the clone name. GenBank accession numbers are listed parenthetically. Subdivisions (see the text) are indicated by brackets at the right of the tree. Tree construction and support for branch points was as described for Fig. 1 and 3, respectively. The four MIM clones and F78 clone are unreleased sequences generously made available to us by Pascale Durand (10) and Floyd Dewhirst (8). The scale bar indicates 0.1 change per nucleotide.
novelties are known as well, for instance, endospore formation by the low-G+C gram-positive bacteria or axial filaments (endoflagella) in the spirochetes. Some biochemical properties evidently have transferred laterally among the divisions. For example, the two types of photosynthetic complexes, photosystem I (PSI) and PSI, are each distributed sporadically among the divisions, consistent with lateral transfer (3). Lateral transfer may also have resulted in combinatorial novelty among the divisions; PSI and PSI, for instance, apparently came together in the cyanobacteria to create oxygenic photosynthesis, with profound consequences to the biosphere (3). Many more such division-specific qualities and cooperations should become evident at the molecular level as comparative genomics gives us a sharper phylogenetic picture of bacterial diversity.

ACKNOWLEDGMENTS

We thank John A. Fuerst for providing useful comments on the manuscript and Pascale Durand, Floyd Dewhirst, and Bruce Paster for supplying unreleased sequences. We also thank Michael Tanner and Scott Davis for assistance in establishing the website of additional information. The authors’ research is supported by grants to N.R.P. from the National Institutes of Health and Department of Energy.

REFERENCES

25. Liesack, W., R. Soller, T. Stewart, H. Haas, S. Giovannoni, and E. Stack- 
39. Strunk, O., O. Gross, B. Reichel, M. May, S. Hermann, N. Struckmann, B.
4774 MINIREVIEW J. BACTERIOL.

ERRATUM

Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity

PHILIP HUGENHOLTZ, BRETT M. GOEBEL, AND NORMAN R. PACE

Departments of Plant and Microbial Biology and Molecular and Cell Biology, University of California, Berkeley, Berkeley, California 94720-3102

Volume 180, no. 18, p. 4765–4774, 1998. Page 4767, column 1: Table 1 should appear as shown below.

<table>
<thead>
<tr>
<th>Habitat type</th>
<th>No. of studies</th>
<th>No. of sequences</th>
<th>α</th>
<th>β</th>
<th>γ</th>
<th>δ</th>
<th>ε</th>
<th>Cytophagales</th>
<th>Desulfovibionales</th>
<th>Acidobacteria</th>
<th>Low-G+C gram positives</th>
<th>Acidimicrobiiales</th>
<th>Vermamoebales</th>
<th>Spirochetes</th>
<th>Nitrates</th>
<th>Chloroflexales</th>
<th>Chlorobiiales</th>
<th>TM7</th>
<th>TM6</th>
<th>Thermotogales</th>
<th>Thermodesulfobacterium</th>
<th>Aquificales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geothermal</td>
<td>10</td>
<td>212</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Soil</td>
<td>16</td>
<td>743</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Marine</td>
<td>23</td>
<td>687</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Freshwater</td>
<td>4</td>
<td>107</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Wastewater</td>
<td>5</td>
<td>430</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Pollutant associated</td>
<td>7</td>
<td>202</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Acid metal leaching</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Subsurface</td>
<td>6</td>
<td>229</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Symbionts and commensals</td>
<td>10</td>
<td>280</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Disease associated</td>
<td>3</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Totals</td>
<td>86</td>
<td>2,918</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
</tbody>
</table>

a An expanded version of this table detailing individual studies is available at http://crab2.berkeley.edu/~pacelab/176.htm.
b Incidence of division-level representatives in studies of particular habitat types ranked from most represented to least represented divisions: >75% (☺), 25 to 75% (○), or <25% (□) of studies have representatives of division. No symbol indicates division not detected.
c Excluding organelles.
d Proteobacteria are presented at the subdivision level due to the extensive sequence representation of this division.
e Cannot establish as a monophyletic group in all analyses.