Quantifying Microbial Diversity: Morphotypes, 16S rRNA Genes, and Carotenoids of Oxygenic Phototrophs in Microbial Mats

ULRICH NÜBEL, FERRAN GARCIA-PICHET,* MICHAEL KÜHL,† AND GERARD MUYZER‡

Max Planck Institute for Marine Microbiology, Bremen, Germany

Received 29 July 1998/Accepted 17 November 1998

We quantified the diversity of oxygenic phototrophic microorganisms present in eight hypersaline microbial mats on the basis of three cultivation-independent approaches. Morphological diversity was studied by microscopy. The diversity of carotenoids was examined by extraction from mat samples and high-pressure liquid chromatography analysis. The diversity of 16S rRNA genes from oxygenic phototrophic microorganisms was investigated by extraction of total DNA from mat samples, amplification of 16S rRNA gene segments from cyanobacteria and plastids of eukaryotic algae by phylum-specific PCR, and sequence-dependent separation of amplification products by denaturing-gradient gel electrophoresis. A numerical approach was introduced to correct for crowding the results of chromatographic and electrophoretic analyses. Diversity estimates typically varied up to twofold among mats. The congruence of richness estimates and Shannon-Weaver indices based on numbers and proportional abundances of unique morphotypes, 16S rRNA genes, and carotenoids unveiled the underlying diversity of oxygenic phototrophic microorganisms in the eight mat communities studied.

Due to their physiological diversity, microorganisms play major roles in the cycling of chemical elements within the biosphere, but this relevance for environmental processes is only fragmentarily reflected in our current knowledge about microbial diversity (39, 40) because the small size and morphological simplicity of microorganisms have hampered the study of their diversity. While microbial physiology and genetics can be investigated in great detail in cultivated isolates, the majority of microorganisms have so far resisted cultivation efforts (40). From most habitats studied, with only a few exceptions (47), less than 1% of the microorganisms observed by microscopy have been brought into culture (1). It is clear, then, that current isolation procedures will fail to adequately investigate the microbial diversity extant in natural environments (1, 10). Molecular biological techniques, and particularly the study of small-subunit rRNAs and the respective genes, have provided new insights into the phylogenetic diversity of microorganisms (68). Microbial nucleic acids extracted directly from environmental samples are amenable to comparative analyses of nucleotide sequences (18, 41, 64). Numerous publications based on this approach have reported the exploration of uncultivated microbial diversity in the last decade (10, 40). However, our understanding of forces that shape and sustain microbial diversity in the environment and of the impact that microbial diversity may have on ecosystem processes is as yet very limited (25, 39). Theoretically, empirical investigations of such interdependencies should lead to considerable progress in the field of microbial ecology, but such investigations depend unknowingly on the evaluation of biodiversity in quantitative terms. This quantification has not yet been achieved on the basis of the new molecular methodologies (39), but in principle it is possible and probably desirable (17).

The quantification of diversity requires the grouping of individual elements into nonoverlapping classes according to a differentiating criterion (26). If the study is to be restricted to certain organisms, which usually will be the case, individuals to be excluded from the analysis need to be identified as such. Ecological diversity is considered a function of the number of different classes (richness) and the relative distribution of individual elements among these classes (evenness) (3, 65). Various indices have been proposed as measures of diversity that incorporate both aspects, richness and evenness (30). The Shannon-Weaver index is the most common diversity index used by ecologists (65): it weighs individual classes by their relative abundances. It can be understood as an estimator of the degree of uncertainty attached to the identity of any individual randomly selected from a community, which increases with richness as well as with evenness (29). Optimaly, individual elements in a class should be uniform with respect to their ecology. However, functional diversity, the actual ecologically relevant parameter, cannot be directly determined, and some deviation from this ideal must be expected when single criteria are used as bases for diversity determinations. Exposed to environmental selection, ecological units are also evolutionary units (43, 63), and the use of evolutionarily coherent entities as classes for diversity estimates is desirable. For practical reasons identification procedures should be as little time-consuming as possible, since often large numbers of organisms need to be investigated. Ecologists studying macroscopic plants and animals commonly use taxonomic species as classes for grouping individual organisms and assess species richness and species diversity accordingly (3). The delineation of species on the basis of morphologies is the most common practice (8), but it does not necessarily result in evolutionarily and ecologically coherent entities, particularly when applied to microorganisms. The determination of prokaryotic species richness and diversity in nature is impractical because the current bacteriological species concept applies exclusively to organisms in pure cultures (66). The value of available species concepts for the quantification of diversity probably will depend on the group of microorganisms considered and on the habitats to be studied; it may be necessary to replace species with some other appropriate units of biodiversity (25).
We studied microbial mats from hypersaline waters in evaporation ponds of the saltern in Guerrero Negro, Baja California, Mexico, as well as from salt marshes in its proximity (9, 22). The biomass of these benthic laminated ecosystems is almost exclusively composed of microorganisms, most of which are prokaryotes. We focused our investigations on communities of oxygenic phototrophs, whose activities are the basis for the existence of these mats. These microorganisms, namely cyanobacteria, diatoms, and a small proportion of green microalgae, traditionally have been classified on the basis of their morphologies (2, 7, 49). Diversity studies of microalgae that were restricted to morphotype analyses alone (e.g., reference 60) have been reported in the past. However, the phycological systems of classification do not necessarily reflect evolutionary relationships (35, 46, 67) and may underestimate diversity (11, 69). Therefore, we investigated the diversity of oxygenic phototrophic microorganisms by applying three cultivation-independent approaches in parallel. Classes for grouping the respective elements analyzed were defined by unique cell and colony morphologies, rDNA gene sequences, and carotenoid molecule structures. Cells and molecules from organisms other than oxygenic phototrophs were excluded from the analyses. The relative morphological complexity of oxygenic phototrophic microorganisms and their content of unique pigments allowed us to evaluate the diversity reflected in rDNA genes. The congruence of results obtained for eight different mat communities demonstrated that microbial diversity can be meaningfully quantified. The numbers and proportional abundances of unique morphotypes, carotenoids, and 16S rDNA genes could describe the diversity of oxygenic photosynthetic microorganisms.

**MATERIALS AND METHODS**

**Sampling.** Microbial mats were sampled during the second to fourth weeks of April 1996 (mats P2, P4, P6, NC2, and NC3) and 1997 (mats P3/4, P5, and NC2). Sampling sites were located in evaporation ponds of the saltern in Guerrero Negro, Baja California Sur, Mexico, and in the salt marsh of Ojo de Liebre Lagoon. Detailed descriptions of these sites can be found elsewhere (9, 22). The characteristics of the microbial mats investigated are summarized in Table 1. Measurements of photosynthesis were performed in a field laboratory at time of collection. Salinities in evaporation ponds typically vary by less than 5% per year. The salinities and water depths in tidal channels vary with tidal cycles.

**TABLE 1.** Characteristics of the microbial mats studied

<table>
<thead>
<tr>
<th>Mat</th>
<th>Field conditions</th>
<th>Photic zone characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>P2/4</td>
<td>Evaporation pond</td>
<td>Soft</td>
</tr>
<tr>
<td>P4</td>
<td>Evaporation pond</td>
<td>Cohesive, rubbery</td>
</tr>
<tr>
<td>P5</td>
<td>Evaporation pond</td>
<td>Cohesive, rubbery</td>
</tr>
<tr>
<td>P6</td>
<td>Evaporation pond</td>
<td>Jellyatinous, translucent</td>
</tr>
<tr>
<td>NC2</td>
<td>5–8</td>
<td>Soft</td>
</tr>
<tr>
<td>NC52</td>
<td>0.2–0.5</td>
<td>Soft</td>
</tr>
<tr>
<td>NC3</td>
<td>0.3–0.6</td>
<td>Very cohesive, rubbery, compact</td>
</tr>
</tbody>
</table>

* Measured in the field site at time of collection. Salinities in evaporation ponds typically vary by less than 5% per year. The salinities and water depths in tidal channels vary with tidal cycles.

**Carotenoid extraction and analysis.** Frozen samples of mat layers (100 to 400 mg, representing approximately 60 mm² of mat surface) corresponding to photic zones were ground in a mortar while cooled by liquid nitrogen. The ground samples were extracted in 10 to 15 ml of degassed acetone for 24 h in the dark at 4°C. Extracts were clarified by filtration on Whatman GF/F glass fiber filters and subsequently concentrated under a stream of N₂ gas. Concentrated pigment extracts were separated and analyzed by high-pressure liquid chromatography (HPLC), with on-line detection by diode array-based spectroscopy between 350 and 700 nm, allowing for the determination of typical carotenoid spectra. The details of chromatographic conditions and equipment were essentially as previously described (24).

**DNA extraction.** The layers corresponding to photic zones were aseptically cut from mat cores (100 to 400 mg, representing approximately 60 mm² of mat surface) and homogenized in Dounce tissue homogenizers (No 3.0, KM, Germany). Cell lysis and DNA extraction were performed as described previously (38). Briefly, the suspensions were repeatedly frozen and thawed and subsequently incubated in the presence of sodium dodecyl sulfate and proteinase K. Cell lysis was controlled microscopically. DNA was extracted by applying hexadecymethylammonium bromide, phenol, chloroform, and isomyl alcohol and precipitated by the addition of isopropl alcohol.

**PCR.** The oligonucleotide primers CYA359F and CYA781R were selected to amplify 16S rDNA gene segments from cyanobacteria and plastids (38). The numbers in the primer designations refer to the 5’ end of target signature sites in 16S rRNA genes (Escherichia coli numbering [5]). A 40-nucleotide GC-rich sequence was attached to the 5’ end of the primer CYA359F to improve the detection of sequence variation in amplified DNA fragments by subsequent denaturing-gradient gel electrophoresis (DGGE [38]). As templates for amplifications, 10 ng of DNAs extracted from mat samples was added to each 100-µl reaction mixture.

**DGGE and digital image analysis.** Amplification products generated by duplicate PCRs with the same template DNAs were pooled and subsequently purified and concentrated by using the QIAquick PCR purification kit (Diagen, Bethesda, Md.). To enable the transformation of fluorescence values into ratios, the resulting gels were run in the gel documentation system and photographed with a digital image gel documentation system (CyberTech, Berlin, Germany). The intensities of gel band fluorescence were measured on digital images by using the gel-plotted macro implemented in the NIH-Image software package (version 1.62 [National Institutes of Health, Bethesda, Md.]). To enable the transformation of fluorescence values into amounts of DNA in individual bands, we designed and applied on each gel a DNA mass calibration standard for DGGE, which was a mixture of PCR products with known concentrations (see Fig. 4).

**Microscopy and morphotype quantification.** The layers corresponding to photic zones were cut from formaldehyde-fixed mat samples with scalpel blades and sectioned vertically into sublayers of approximately 0.5 to 0.3-mm mat surface area. These pieces were placed on glass slides in 1 drop of water and chopped and stirred to achieve even distribution. Sketches of morphotypes were prepared for illustration (Fig. 1). For each subcore, 25 to 40 randomly chosen phase-contrast microscopic fields were photographed at 400-fold magnification. Counts and size measurements were performed on projections of the resulting slides, exclusively taking into account focused cells of oxygenic phototrophs. Diatom valves without plastids were omitted from the analyses. For unicellular organisms, cell numbers were counted. For filamentous organisms, total filament lengths were divided by the respective mean cell length to calculate cell numbers. Final cell numbers or total filament lengths were converted into biovolumes with geometric formulae (15). Rounded cells were considered spheres, and rod-shaped and filamentous organisms were considered cylinders. Biovolumes were determined by applying three cultivation-independent approaches in parallel. Classes for grouping the respective elements analyzed were defined by unique cell and colony morphotypes, rDNA gene sequences, and carotenoid molecule structures. Cells and molecules from organisms other than oxygenic phototrophs were excluded from the analyses. The relative morphological complexity of oxygenic phototrophic microorganisms and their content of unique pigments allowed us to evaluate the diversity reflected in rDNA genes. The congruence of results obtained for eight different mat communities demonstrated that microbial diversity can be meaningfully quantified. The numbers and proportional abundances of unique morphotypes, carotenoids, and 16S rDNA genes could describe the diversity of oxygenic photosynthetic microorganisms.

**DNA extraction.** The layers corresponding to photic zones were aseptically cut from mat cores (100 to 400 mg, representing approximately 60 mm² of mat surface) and homogenized in Dounce tissue homogenizers (No 3.0, KM, Germany). Cell lysis and DNA extraction were performed as described previously (38). Briefly, the suspensions were repeatedly frozen and thawed and subsequently incubated in the presence of sodium dodecyl sulfate and proteinase K. Cell lysis was controlled microscopically. DNA was extracted by applying hexadecymethylammonium bromide, phenol, chloroform, and isomyl alcohol and precipitated by the addition of isopropl alcohol.

**PCR.** The oligonucleotide primers CYA359F and CYA781R were selected to selectively amplify 16S rDNA gene segments from cyanobacteria and plastids (38). The numbers in the primer designations refer to the 5’ end of target signature sites in 16S rRNA genes (Escherichia coli numbering [5]). A 40-nucleotide GC-rich sequence was attached to the 5’ end of the primer CYA359F to improve the detection of sequence variation in amplified DNA fragments by subsequent denaturing-gradient gel electrophoresis (DGGE [38]). As templates for amplifications, 10 ng of DNAs extracted from mat samples was added to each 100-µl reaction mixture.

**DGGE and digital image analysis.** Amplification products generated by duplicate PCRs with the same template DNAs were pooled and subsequently purified and concentrated by using the QIAquick PCR purification kit (Diagen, Bethesda, Md.). To enable the transformation of fluorescence values into amounts of DNA in individual bands, we designed and applied on each gel a DNA mass calibration standard for DGGE, which was a mixture of PCR products with known concentrations (see Fig. 4).

**Microscopy and morphotype quantification.** The layers corresponding to photic zones were cut from formaldehyde-fixed mat samples with scalpel blades and sectioned vertically into sublayers of approximately 0.5 to 0.3-mm mat surface area. These pieces were placed on glass slides in 1 drop of water and chopped and stirred to achieve even distribution. Sketches of morphotypes were prepared for illustration (Fig. 1). For each subcore, 25 to 40 randomly chosen phase-contrast microscopic fields were photographed at 400-fold magnification. Counts and size measurements were performed on projections of the resulting slides, exclusively taking into account focused cells of oxygenic phototrophs. Diatom valves without plastids were omitted from the analyses. For unicellular organisms, cell numbers were counted. For filamentous organisms, total filament lengths were divided by the respective mean cell length to calculate cell numbers. Final cell numbers or total filament lengths were converted into biovolumes with geometric formulae (15). Rounded cells were considered spheres, and rod-shaped and filamentous organisms were considered cylinders. Biovolumes were determined by applying three cultivation-independent approaches in parallel. Classes for grouping the respective elements analyzed were defined by unique cell and colony morphotypes, rDNA gene sequences, and carotenoid molecule structures. Cells and molecules from organisms other than oxygenic phototrophs were excluded from the analyses. The relative morphological complexity of oxygenic phototrophic microorganisms and their content of unique pigments allowed us to evaluate the diversity reflected in rDNA genes. The congruence of results obtained for eight different mat communities demonstrated that microbial diversity can be meaningfully quantified. The numbers and proportional abundances of unique morphotypes, carotenoids, and 16S rDNA genes could describe the diversity of oxygenic photosynthetic microorganisms.
Estimation of richness and diversity. The richness and diversity of morphotypes, 16S rRNA genes, and carotenoids of oxygenic phototrophic microorganisms were estimated. Classes for grouping organisms and molecules were defined by unique cell and colony morphologies, nucleotide sequences, and pigment molecule structures. Cells and molecules from organisms other than oxygenic phototrophs were excluded from the analyses.

Morphotype analysis was performed by light microscopy. Diatoms and cyanobacteria could be distinguished from most other microorganisms due to their sizes and characteristic morphologies. Epifluorescence microscopy was used to identify putative members of the family Chloroflexaceae. These green filamentous bacteria occur in marine microbial mats (45) and may be mistakenly identified as cyanobacteria, but they were distinguished by their lack of visible fluorescence. Proportional abundances of morphotypes were calculated on the basis of cell numbers and, alternatively, cell volumes. 16S rRNA gene fragments were amplified by PCR from cyanobacterial and plastid DNA after nucleic acid extraction from mat samples (38). Phylum-specific amplification enabled the exclusion from the analyses of DNA from organisms other than oxygenic phototrophs (38).

Numbers and proportional abundances of the unique rRNA gene segments amplified were estimated after DGGE analysis of PCR products. Carotenoids were extracted directly from mat samples, and their analysis by HPLC enabled the determination of the numbers and abundances of unique pigments. Quantification of the relative abundance of specific carotenoids was based on area integration of peaks resulting from absorption at 470 nm, and no attempts were made to introduce corrections for differences in extinction coefficients. Peak delimitation and area integration were carried out automatically by the instrument’s software. Absorption peaks in these chromatograms corresponding to pigments other than carotenoids (chlorophyll a, pheophytins, diverse bacteriochlorophylls, and scytonemin) were identified from the corresponding absorption spectra (350 to 700 nm) and deleted from the automated analysis a posteriori (for an example of this procedure, see Fig. 5). The assumption was made that the contribution to total carotenoids of pigments from heterotrophic and chemotrophic bacteria was negligible, as phototrophs make up the bulk of the biomass in the photic zone and they contain much higher specific amounts of carotenoids than nonphototrophic bacteria. In those samples displaying measurable amounts of bacteriochlorophylls, a correction for carotenoids stemming from anoxygenic phototrophic bacteria was carried out. These were identified by the comparison of retention times and spectra to carotenoid standards from purple sulfur bacteria and green filamentous nonsulfur bacteria. However, these corrections were seldom necessary. The majority of carotenoids detected could be assigned to typical cyanobacterial and diatom standards from cultivated strains.

Using the approaches described above for triplicate samples from each oxygenic phototrophic mat community, the numbers, $D$, and proportional abundances, $a_i$, of classes $i$ were determined. Shannon-Weaver indices ($H'$) (30, 52) were calculated as
Because HPLC. Accordingly, crowding was likely to affect the results if values of PS-lating expectation values for PS taken as corrected estimates of richness.

Once the values of PS are computed, a simple correction consists of calculating expectation values for PS directly as an estimate of the richness of the morphotypes, PS. Arithmetic means and standard errors were calculated from the results of triplicate analyses. Statistical analysis was performed by using SPSS version 6.1 software (SPSS, Erkrath, Germany).

In the following discussion, subscripts refer to the respective analyses of morphotypes (M), rRNA genes (R), and carotenoids (C). Corrections for crowding of bands and peaks in electrophoretic and chromatographic analyses transformed Ds and Ds into richness estimates Sm and Ss (see below). Ds was used directly as an estimate of the richness of the morphotypes, Sm. Arithmetic means and standard errors were calculated from the results of triplicate analyses. Statistical analysis was performed by using SPSS version 6.1 software (SPSS, Erkrath, Germany).

Correction for crowding of the number of classes measured by electrophoresis and chromatography. With increasing numbers of bands or peaks detected in electrophoretic or chromatographic analyses, respectively, the probability increases that classes cannot be discerned because they run at identical positions in the gel or chromatogram. In the following paragraph, we describe a way to approximate this probability distribution and, on that basis, to estimate how many classes are likely to have been missed in a given analysis due to crowding.

Let k be the number of specific classes present in a sample, D the number of classes actually detected in the analysis, and Dmax the maximum number of classes that the analytical procedure can detect. Due to crowding within the chromatogram or electrophoretic gel, for any given S there exists a certain probability to detect only D classes, where D = 5 and D = Dmax. A distribution, P(D), describing this probability can be used to correct measured values for crowding. Assuming that there are no preferred sites of occurrence for the classes within the chromatogram or electrophoretic gel, i.e., that crowding is homogeneous, and that all classes can be detected with equal resolution, then the probability, p, that two classes in a sample with S = 2 cannot be discerned is p = P2(D = Dmax) = (Dmax)2. The probability that we measure D classes in a sample with S + 1 specific classes is P2(D + 1) = P2(D)Dp + P2(D + 1) [1 - (D - 1)p]. Because P2(1) = 1, if p is known, one can inductively calculate any value of P2(D). Alternatively, P2(D) can be computed directly as

\[ P2(D) = \left( \frac{1}{Dmax} \right)^2 \sum_{i=0}^{Dmax-1} \left( \frac{Dmax}{D} \right)^{D-i} \left( \frac{1}{D} \right)^i \]

Once the values of P2(D) are computed, a simple correction consists of calculating expectation values for D, Dexp, for any given S by using P2(D), as

\[ Dexp = \sum_{D=1}^{Dmax} D P2(D) \]

Values of S that yield a Dexp corresponding to a measured value of D can be taken as corrected estimates of richness.

Values for Dmax were estimated, on the basis of the average peak width and total distance of separation for our experiments, to be 200 for DGGE and 96 for HPLC. Accordingly, crowding was likely to affect the results if Ds was >20 and Ds was ≥14, respectively. None of our DGGE results exceeded this threshold, but values of carotenoid richness in most cases needed to be raised by one to three classes (i.e., by up to 12%). We did not attempt to take into account the effect of varying class frequencies on crowding or to correct the calculation of Shannon-Weaver indices, as such corrections can be extremely complicated. Thus, our values for HC tend to underestimate carotenoid diversity.

### RESULTS

#### Diversity of morphotypes.

Morphotypes that were distinguished are illustrated in Fig. 1. In total, 36 different morphotypes were observed, of which 25 were cyanobacteria, 10 were diatoms, and 1 was a green alga. The number of morphotypes detected in each of the specimens is considered an estimate of the morphotype richness, S M, of the respective oxygenic-pho-totrophic community. Proportional abundances based on cell counts and, alternatively, estimated cell volumes were used to calculate the Shannon-Weaver indices, H M (Table 2). Figure 2 illustrates the proportional abundances of morphotypes in two of the eight mats on the basis of cell counts performed on triplicate randomly drawn subcores.

A marked patchiness in mat community structure was observed at the scale of cyanobacterial colonies and filaments (10 to 100 μm). Therefore, a significant dependence of S M and H M on the number of microscopic fields investigated had to be expected. The minimum number of cells needed to achieve representative subsamples was determined for each of the mats by randomly removing slides (microscopic fields) one by one from the respective analyses and subsequently determining richness and Shannon-Weaver indices for the resulting samples of reduced sizes. For a low number of cells included in an analysis both parameters increased (not necessarily monotonically) concomitantly with increasing sample size and then, provided that a large enough number of cells was investigated, leveled off. A subsample was assumed to be sufficiently representative if its further enlargement caused no further increase.

### Table 2. Numbers of classes, richnesses, and Shannon-Weaver indices determined

<table>
<thead>
<tr>
<th>Mat</th>
<th>S M</th>
<th>H M</th>
<th>H V</th>
<th>S R</th>
<th>H R</th>
<th>D C</th>
<th>S C</th>
<th>H C</th>
</tr>
</thead>
<tbody>
<tr>
<td>P2</td>
<td>10.67 ± 0.33</td>
<td>1.45 ± 0.09</td>
<td>1.82 ± 0.12</td>
<td>11.33 ± 0.66</td>
<td>1.38 ± 0.08</td>
<td>20.67 ± 0.66</td>
<td>22.67 ± 0.66</td>
<td>2.28 ± 0.03</td>
</tr>
<tr>
<td>P3/4</td>
<td>7.33 ± 0.33</td>
<td>0.98 ± 0.09</td>
<td>1.42 ± 0.17</td>
<td>4.00 ± 0.00</td>
<td>0.35 ± 0.02</td>
<td>9.33 ± 0.88</td>
<td>9.33 ± 0.88</td>
<td>1.62 ± 0.03</td>
</tr>
<tr>
<td>P4</td>
<td>10.33 ± 0.88</td>
<td>1.10 ± 0.14</td>
<td>1.07 ± 0.17</td>
<td>10.00 ± 0.00</td>
<td>1.75 ± 0.02</td>
<td>22.00 ± 1.15</td>
<td>24.00 ± 1.16</td>
<td>2.28 ± 0.02</td>
</tr>
<tr>
<td>P5</td>
<td>7.00 ± 0.58</td>
<td>1.02 ± 0.12</td>
<td>1.31 ± 0.03</td>
<td>6.67 ± 0.33</td>
<td>1.31 ± 0.09</td>
<td>17.00 ± 0.58</td>
<td>18.00 ± 0.58</td>
<td>1.97 ± 0.02</td>
</tr>
<tr>
<td>P6</td>
<td>9.67 ± 0.33</td>
<td>1.39 ± 0.05</td>
<td>1.72 ± 0.10</td>
<td>10.00 ± 0.00</td>
<td>1.90 ± 0.02</td>
<td>14.67 ± 0.33</td>
<td>15.67 ± 0.33</td>
<td>2.20 ± 0.03</td>
</tr>
<tr>
<td>NC2</td>
<td>12.00 ± 0.58</td>
<td>1.41 ± 0.21</td>
<td>1.68 ± 0.08</td>
<td>12.00 ± 0.00</td>
<td>2.09 ± 0.01</td>
<td>22.00 ± 1.00</td>
<td>24.00 ± 1.00</td>
<td>2.44 ± 0.02</td>
</tr>
<tr>
<td>NC52</td>
<td>12.00 ± 0.58</td>
<td>1.18 ± 0.13</td>
<td>1.65 ± 0.06</td>
<td>12.33 ± 0.33</td>
<td>1.74 ± 0.05</td>
<td>13.67 ± 0.66</td>
<td>14.00 ± 1.00</td>
<td>1.90 ± 0.03</td>
</tr>
<tr>
<td>NC3</td>
<td>12.67 ± 0.33</td>
<td>1.67 ± 0.04</td>
<td>1.29 ± 0.03</td>
<td>16.00 ± 1.00</td>
<td>2.08 ± 0.11</td>
<td>24.00 ± 1.00</td>
<td>26.67 ± 1.34</td>
<td>2.53 ± 0.02</td>
</tr>
</tbody>
</table>

* D, number of classes detected; S, richness; H M, Shannon-Weaver index. The subscripts refer to the respective analyses of morphotypes (M), volumes of morphotypes (V), 16S rRNA genes (R), and carotenoids (C). The means and standard errors from triplicate analyses are shown.
of the determined diversity (44, 60). Figure 3 shows plots of these analyses for two specimens. The sample sizes needed to detect all rare and localized morphotypes (Fig. 3A) and to weight them according to their actual abundances (Fig. 3B) were estimated from such plots to be 2,000 to 3,000 cells. Significant heterogeneity of the distribution of organisms at the scale of millimeters caused major differences among microscopical analyses of triplicate preparations from the same mat. For example, the surface of the mat from pond 5 consisted of a loose film of diatoms (Nitzschia sp.; morphotype 30) forming irregularly scattered tufts up to 1 mm in height and diameter. Concomitantly, depending on the exact site of sampling, the proportion of diatoms in a randomly drawn subcore, representing only a 0.5- by 0.5-mm mat surface, varied from 1.4 to 9.5% of the total cell numbers of oxygenic phototrophs. This level of patchiness resulted in large variances of estimates of both abundances of individual morphotypes (Fig. 2) and morphotype diversity as reflected in $H^M$ (see Fig. 7), whereas $S_M$ was less affected. In all specimens, the rarest morphotypes detected accounted for less than 1% of the total number of cells and total biovolumes. Richness estimates and Shannon-Weaver indices varied almost twofold among mats (Table 2). An exceptionally low diversity was detected for Shannon-Weaver indices, varied between 1.62 and 2.53, almost twofold. The crowding of peaks was likely to have affected the results of HPLC analyses (Fig. 5). While estimates of carotenoid richness could be corrected accordingly, this was not attempted for Shannon-Weaver indices. Thus, the latter may underestimate diversity. There may have been some misidentification events with carotenoids, particularly with the less abundant ones, as the spectra from small peaks were not always clear. These may have been degradation products of original photosynthetic carotenoids, or they may have stemmed from other bacteria.

**Diversity of 16S rRNA genes.** Total nucleic acids were extracted from triplicate mat samples, and 16S rRNA gene segments were amplified from cyanobacterial and plastid DNA by applying a phyllum-specific PCR that had recently been developed on the basis of published nucleotide sequences (38). The sequence-dependent separation of the resulting amplification products by DGGE (36) generated band patterns that were characteristic for each of the mats (Fig. 4). The number of bands visible in denaturant gradient polyacrylamide gels (corrected for crowding) provides an estimate of the richness, $S_R$. Fluorescence measurements and comparisons to a DNA mass calibration standard (Fig. 4) allowed the calculation of the amounts of DNA in individual gel bands, proportional abundances of sequence-defined populations, and Shannon-Weaver indices, $H^R$ (Table 2). An exceptionally low diversity was detected in mat community P3/4. Among the other communities the variation of $S_R$ was more than twofold, and Shannon-Weaver indices varied from 1.31 to 2.09, with coefficients of variation derived from triplicate analyses smaller than 7%. The lower detection limit for ethidium bromide-stained DNA in these gels was approximately 10 ng per band, which is 2% of the total amount of PCR product applied to each gel lane. To date, we have been able to extract, reamplify, and sequence DNA from the majority of gel bands. This analysis should allow the determination of phylogenetic relationships among organisms forming the investigated mats and cultivated reference strains, which is outside the scope of the present paper. However, it may be significant to note that no “heteroduplex bands” (12) could be detected and that all determined sequences were derived from cyanobacteria or plastids, so that most probably, none of the bands represents any undesired amplification product.

**Diversity of carotenoids.** According to carotenoid pigment analysis, more than twofold differences in richness could be observed among mats. Richness $S_C$ was estimated to range from 9.33 $\pm$ 0.88 in mat P3/4 to 24.00 $\pm$ 1.00 in mat NC3 (Table 2). The corresponding estimates of the Shannon-Weaver diversity index, measured from the relative abundance of each carotenoid, varied between 1.62 and 2.53, almost twofold. The crowding of peaks was likely to have affected the results of HPLC analyses (Fig. 5). While estimates of carotenoid richness could be corrected accordingly, this was not attempted for Shannon-Weaver indices. Thus, the latter may underestimate diversity. There may have been some misidentification events with carotenoids, particularly with the less abundant ones, as the spectra from small peaks were not always clear. These may have been degradation products of original photosynthetic carotenoids, or they may have stemmed from other bacteria.

**DISCUSSION**

**Units of biodiversity.** Biologists studying the ecology of communities of plants or animals usually choose species as basic units for diversity estimates. For prokaryotes this is currently impossible because of the bacteriological practice of delineating species on the basis of cultivated strains (66). At present, probably far less than 5% of extant cyanobacterial species are successfully cultivated (6) and only a minority of these cultures are axenic (7). This disproportion reflects general problems of obtaining, within a scientist’s lifetime, a collection of strains that represents the microbial richness extant in an environment.

**FIG. 4.** Composite figure of ethidium bromide-stained DGGE separation patterns of PCR-amplified segments of 16S rRNA genes. Mixtures of PCR products derived from five cyanobacterial strains were applied to each gel as standards (in lanes 1 to 3 [top to bottom]) are Synechocystis sp. strain B-77-Scyjav, Synechococcus elongatus SAG 1402-1, Microcoleus chthonoplastes MPI-NDN-1, Geitlerinema sp. strain PCC 9452 ("Microcoleus" sp. strain 10 mfx), and Cyanophyceae sp. strain PCC 7418). The standard in lane 1 allows gel-to-gel comparisons. The DNA mass calibration standard in lanes 2 and 3 enables the transformation of measured band fluorescence values into amounts of DNA (in lane 2, the amounts of DNA in individual bands are (top to bottom) 528, 176, 59, 20, and 7 ng; in lane 3, half of the amount of the standard in lane 2 was applied). The gels labelled P4 and NC3 show separation patterns of PCR products derived from triplicate sampling cores of those microbial mats. The arrowheads indicate the bands included in the subsequent analyses.
genes are strongly conserved in function and structure, and their information content is therefore limited. Strains of bacteria with considerably different physiologies were reported to contain identical 16S rRNA genes (43). On the other hand, slightly different rRNA gene sequences detected in an environmental sample do not prove the presence of several microbial populations but may instead be derived from a single organism (37). Similarly, a single microalga or cyanobacterium usually produces a multiplicity of carotenoid types, and conversely, identical pigments may be extracted from several different algae.

None of the approaches we applied enables the indisputable determination of absolute numbers of ecologically distinct populations forming a community. Yet the congruity of the results obtained indicates that all of the characteristics analyzed provide valuable information on the diversity of the organisms of interest. Therefore, their simultaneous investigation allows a meaningful comparison of the different communities in relative terms. Richness data obtained for the eight mat communities by applying three independent approaches are positively linearly correlated (Fig. 6), meaning that, on average, an increase in the number of unique 16S rRNA genes is accompanied by an increase in the number of morphotypes and types of carotenoids. The correlation of morphotype and 16S rRNA richness is particularly strong and highly significant. Interestingly, for six of the eight mats the ratio of \( S_M \) to \( S_R \) is 1.0 (mean values of triplicate analyses). However, in the communities with the highest (NC3) and lowest (P3/4) richnesses according to carotenoid and genetic data, these ratios are 0.8 and 1.8, causing the slope of the respective regression curve to deviate from one. The richness determined on the basis of carotenoids for all mats is higher than the number of either morphotypes or rRNA gene sequence types, which is due to the fact that all phototrophs produce more than a single type of carotenoid (19). The ratio of \( S_C \) to \( S_R \) ranges from 1.1 to 2.7, causing a less significant positive correlation of these data and possibly indicating that the number of pigments synthesized depends on the specific identity of the organism. Thus, for some communities richness estimates based on different approaches can be rather contradictory, which strongly confirms the notion that the analysis of any single characteristic can be misleading. But, very importantly, the simultaneous application of three independent approaches unveils an underlying trend of richness among the various communities investigated, which may now be related to further mat characteristics or other environmental parameters.

**Proportional abundances.** The determination of abundances of microorganisms depends on meaningful units of
counting, the choice of which is difficult and may to some extent depend on research objectives. Since microorganisms are clonal organisms, some authors have argued that colonies or trichomes instead of cells should be considered individuals (23, 60). However, in the densely packed communities within microbial mats, the boundaries of colonies often could not be recognized. Therefore, we simply counted cells and, as an alternative measure, weighted them by their respective individual volumes. When we applied these alternative approaches, drastically variable cell sizes caused significantly divergent estimates of the proportional abundances of certain morphotypes, seriously affecting Shannon-Weaver indices. While the abundances of cell components may reflect the relative proportions of organisms, the number of marker molecules per cell depends on the cell’s specific identity and physiological state. The production of carotenoids depends on environmental conditions, with the quantity and spectral quality of incident irradiation being especially important (42). The number of rRNA genes per cell varies with the number of copies per genome as well as the number of identical chromosomes (and replicated parts thereof) per cell. The latter may change considerably with growth conditions (4, 62).

Because of practical problems with all of the methodologies applied, the relative abundance of any population detected may deviate from that actually present in a sample. Morphotype-based abundance determinations are biased in favor of large organisms, since they are more likely to be encountered in focused microscopic planes. Procedures for extracting nucleic acids and pigments from mat samples may be selective in principle. However, the application of alternative protocols employing mechanical disruption (bead beating) or enzymatic digestion (lysozyme treatment) of cells resulted in no detectable differences in DGGE band patterns (data not shown), and microscopic observations indicated that complete lysis was achieved by the protocol that was finally applied. The quantitative use of PCR may be compromised when the reactions reach a plateau phase of amplification. This typically occurs at a product concentration of 10⁻⁶ M (50, 56), which is approximately the final molarity yielded in our experiments. By performing amplification reactions with a series of template dilutions, we could confirm that above this limit the integrated amplification efficiency decreased; however, the proportional abundances of DNA molecules with different sequences remained unaffected (data not shown). Other potential sources of bias cannot be excluded, however. Primer degeneracies and greatly varying G+C contents of amplified DNA molecules have been suspected to cause differential amplification efficiencies (61). Unknown target organisms may exist, the 16S rRNA genes of which do not contain the signature sites necessary for efficient amplification (38). Hybridization techniques may allow us to determine abundances of environmental nucleic acids more accurately than by PCR. However, their application depends on the availability of suitable nucleic acid probes, which either require prior knowledge of the sequences to be detected (synthetic oligonucleotide probes [1, 53]) or need to be prepared from available reference DNA (polynucleotide probes [21, 28]).

Despite the inherent difficulties of population abundance determinations, overall congruent estimates of diversity were obtained based on the analyses of 16S rRNA genes, morphotypes, and carotenoids. Shannon-Weaver indices calculated for the eight mat communities on the basis of the different approaches are positively linearly correlated (Fig. 7). Mainly because of the relatively low 16S rRNA gene diversity detected in community P3/4, the slopes of the regression lines deviate from one. Similar to the richness results, carotenoid analyses for all mats yielded higher absolute values for Shannon-Weaver indices than did other measurements. Shannon-Weaver indices based on proportional volumes of morphotypes (not shown in Fig. 7) are exceptional, since they are only weakly and insignificantly correlated with those based on cell number proportions (r = 0.35; P = 0.39) and not at all correlated with indices calculated on the basis of 16S rRNA genes (r = 0.09) and carotenoids (r = 0.02). Since individual cell volumes differed by up to 3 orders of magnitude, the transformation of cell counts into biovolumes significantly changed the proportional abundances calculated for certain morphotypes. As was the case for estimates of richness, binary comparisons of Shannon-Weaver indices for particular communities are also contradictory when based on different methodologies. This again reflects the limitations inherent in all of the identifying traits and their analyses.

Quantification of microbial diversity. Facing the numerous limitations inherent in the various methodologies currently available, many authors conclude that an ecological evaluation of microbial diversity has not yet been convincingly reported (39, 57) or that it would be—at present—impossible (55). The identification of basic units of microbial diversity is difficult, and the determination of the proportional abundances of microbial populations can be very ambiguous. However, the latter may be of interest because, generally, rare species in a community have little effect on the overall flux of energy and matter but may instead become important under changing environmental conditions (51). Therefore, if any current activities of communities or ecosystems are investigated, diversity indices weighting populations by their proportional abundances may be more relevant than the number of distinct populations. Such indices also are less sensitive to the detection limits of the
respective methodologies applied. This feature may be especially useful for the study of habitats such as soil, with difficult-to-determine and tremendously high microbial richness.

Many of the basic problems discussed in this paper are not specific to the exploration of the microbial world. For the majority of ecological collections the diversity can be estimated and expressed only in relative terms. A comprehensive census usually is not achievable, and in many cases even random samples cannot be drawn (44). Species concepts of larger organisms may themselves be as controversial as is their relevance for diversity estimates (32). Depending on research objectives, it may be more fruitful to take into account the organisms’ specific identities and their ecologically relevant properties (31, 59). However, diversity is an inherent aspect of community structure, and it has been reported to be related to properties (31, 59). However, diversity is an inherent aspect of community structure, and it has been reported to be related to properties (31, 59).

Acknowledgments

We gratefully acknowledge S. Koch and C. Wafer for numerous extractions of carotenoids and nucleic acids and Exportadora de Sal, S. A. de C. V., Baja California Sur, Mexico, for their continued logistic support. We thank E. Clavero for her introduction to the classification of the diatoms, R. Amann for use of the facilities, and T. Richter for the explicit formula for calculating $P_{D}(D)$.

The research described in this paper was financially supported by the Max Planck Society and the Deutsche Forschungsgemeinschaft.

References


