(2)

ANNUAL AND FINAL PROGRESS REPORT

Grant # N00014-87-K-0813 **R & T Code:** 41

Principal Investigator: Norman R. Pace

Institute: Indiana University, Bloomington

Grant Title: Phylogenetic Analysis of Marine Picoplankton Using rRNA Sequences

Period of Performance: July, 1990 - Feb., 1991

Objective:

To analyze phylogenetically the dominant constituents of oligotrophic marine picoplankton at selected sites using ribosomal RNA gene cloning and sequencing.

Overview:

This report is a summary of the overall project period. Formal funding for the program ended 02/91, although a small component was extended until 9/91. The overall achievements of the program during the main performance period were outstanding. The essential goals of the program were fulfilled. The activities of the laboratory in natural population analysis during the project period have had good impact on the fields of microbial ecology and microbial phylogenetics.

Report Narrative:

Phylogenetic identification of picoplankton. Using tangential flow filtration, we collected bulk amounts of marine picoplankton from the central Atlantic Ocean (Sargasso Sea, Hydrostation S) and northern Pacific Ocean (Aloha Station). DNA prepared from both populations was analyzed by hybridization using kingdom-specific probes complementary to 16S rRNA, and determined to be largely (>90%) eubacteria: little eukaryote-specific and little or no archaebacterial-specific DNAs were detected. DNA from the Aloha Station was suitable for cloning experiments and a random library of 20 kb *Sau*3a fragments in phage λ EMBL3 was established. This library is available to other investigators upon request. During these past two years, screening that library for 16S rRNA-containing clones, sequencing the genes and carrying out phylogenetic analysis occupied substantial effort. Approximately 50 rRNA gene-containing clones were identified by hybridization and the unique types determined by sequencing using single dideoxynucleotide reactions. Several hundred nucleotides of sequence was determined for each of the unique types.

The Aloha picoplankton proved to be a fairly simple population. Only 15 different types of rRNA genes have been identified, 14 eubacteria and one eukaryote (a dinoflagellate). The eubacteria are all restricted to well-known phylogenetic groups, the α- and γ- groups of "proteobacteria" ("purple" bacteria, *sensu* Woese), and the cyanobacteria. The identified cyanobacterial picoplankton are all closely related to *Synechococcus* spp. Considerable is known about the proteobacteria and cyanobacteria phylogenetic groups that can be extrapolated to the picoplankton, since properties that are common to a phylogenetic group are expected to occur in any particular representative of that group. Phylogenetic trees of picoplankton clones are shown on the *Highlight* page. Clone numbers are shown for the cyanobacteria; asterisks indicate the proteobacteria.

Quantitative analysis of clones. We currently are working toward a quantitative assay of cloned rRNA genes in mixed populations. This should approximate the distribution of the organisms in the population. We find, however, that simple hybridization tests using probes based on rRNA sequences are not optimal for quantitative analysis. Longer probes (> ca. 50 nucleotides) typically are problematic because of cross-species hybridization due to sequences shared among related organisms. Oligonucleotide probes (< ca. 30 nucleotides) can discriminate between targets on the basis of only one or a few mismatches, but the stringency of hybridization required for such precision also results in loss of much or most of the homologous signal. (Theoretical reasons for this are based on the high dissociation rates of oligonucleotide hybrids.) The necessity to achieve a balance between specificity and quantitation is difficult to control for, experimentally, and it casts uncertainty on many oligonucleotide hybridization results, particularly in the case of low-abundance sequences. Because of this uncertainty, we are developing a method for quantitative analysis of clone distribution based on a nuclease protection assay to discriminate perfectly matched hybrids from mismatched ones. The method should be generally useful for studies in microbial ecology.

Continued development of single-cell phylogenetic stains. We previously showed that fluorescently labeled oligonucleotides complementary to phylogenetic group-specific sequences can be used to identify single cells. Probes used initially had a single fluor molecule, fluorescein or rhodamine, attached to the oligonucleotide. Consequently, detection of probe-binding to individual cells requires the presence of ca. 5000 ribosomes (the probe targets). In some cases of organisms (e.g. phototrophs) probe-binding also is compromised by intrinsic fluorescence that overwhelms probe-fluorescence. A method is needed for bleaching cells that does not interfere with subsequent hybridization results. In order to increase the sensitivity of the fluorescent probes, we have experimented with increasing the number of fluor molecules attached to oligonucleotides. However, addition of multiple probes to one end of the oligonucleotide results in reduction of fluorescence (due to quenching); addition of fluors to multiple nucleotides in a given oligonucleotide results in destabilization of the hybrid. Nearly two-fold gain in sensitivity is achieved by attaching fluors at both 5'- and 3'- ends of oligonucleotides. However, further sensitivity through adding multiple fluors will require use of longer probes so that fluor molecules can be sufficiently well-separated to minimize quenching.

Use of PCR for population analysis. A random recombinant library in a phage λ vector is the least selective method available for retrieving rRNA genes from a mixed population of organisms, but it requires substantial DNA (>100 μg) and, consequently, substantial biomass. In order to accommodate small amounts of natural DNA, we (and others) are using "polymerase chain reaction" (PCR), coupled with primers complementary to universally conserved rRNA sequences, to amplify rRNA genes for cloning and sequence analysis. We are interested in using micromanipulation to obtain cells from the environment for phylogenetic analysis. Substantial experience has now been gained with PCR. In general the method works well, but some noteworthy problems have been encountered.

A significant problem in the use of universally conserved rRNA sequences as PCR primers is that all commercially available Taq DNA polymerase is contaminated by organismal DNA. This contaminating DNA presumably derives mostly from chromatography columns, etc., used in preparing the enzyme. Although (usually) not a problem with other primers, the universally applicable rRNA primers detect the DNA and the resulting signal can eclipse that due to a low amount of added template. In an effort to reduce the contaminating DNA, we obtained an expression clone of Taq DNA polymerase and developed an enzyme purification protocol that reduces, but does not yet eliminate,

contaminating DNA. In-house production of Taq polymerase has been economically useful, as well: we obtain $>10^6$ units of this expensive enzyme from 1 l. of induced cells. We are still working at suppressing residual contamination, but the improved purification method results in enzyme preferable for population analysis.

Another potential problem with PCR is that rRNA genes of different organisms in a mixed population may not amplify equally well. We have preliminary evidence that this may be a problem, and are currently investigating the issue.

Training:

During the course of the program the laboratory served as host to training a number of investigators from other institutions in the use of rRNA-based technology. These visitors recognized the utility of the methods to their own ecological or phylogenetic studies. Examples of visitors include F. Azaur (S.I.O.), T. Burger-Wiersma (U. Amsterdam), C. Cavanaugh (Harvard), D. Distel (S.I.O.), J. Fuhrman ( ), R. Rossen (Center for Great Lakes Studies), J. Waterbury (W.H.O.I.), and others.

Future Directions:

The program to characterize picoplankton and develop rRNA-based microbiology methods has been successful. The work with picoplankton continues with former colleagues: E. DeLong (now at W.H.O.I.), S. Giovannoni (Oregon State), and T. Schmidt (U. Miami). Future efforts of the laboratory will continue with development of methods for population analysis without cultivation of target organisms. However, main focus will be on high-temperature ecosystems.

Inventions

None

Recent Publications Acknowledging ONR Support

1. Pace, N.R. and A.B. Burgin. (1990). Processing and evolution of the ribosomal RNAs. In "The Structure, Function and Evolution of Ribosomes," W. Hill (ed.). Am. Soc. Microbiol., pp. 417-425.

2. DeLong, E.F., T.M. Schmidt, and N.R. Pace. (1990) Analysis of single cells and oligotrophic picoplankton populations using 16S ribosomal RNA sequences. In "Recent Advances in Microbial Ecology," T. Hatori, Y. Ishida, Y. Maruyama, R. Morita, and A. Uchida (eds.). Japan Scientific Societies Press, pp. 697-701.

3. Giovannoni, S.J., E.F. DeLong, T.M. Schmidt and N.R. Pace. (1990). Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. Appl. Environ. Microbiol., 56:2572-2575.

4. Schmidt, T.M., E.F. DeLong and N.R. Pace. (1991). Phylogenetic identification of uncultivated microorganisms in natural habitats. In "Rapid Methods and Automation in Microbiology and Immunology." Springer-Verlag, in press.

5. Eden, P.A., T.M. Schmidt, R.P. Blakemore, and N.R. Pace. (1991). Phylogenetic analysis of *Aquaspirillum magnetotacticum* using PCR-amplified 16S ribosomal RNA-specific DNA. Internal. J. System. Bacteriol. 41:324-325.

6. Schmidt, T.M., B. Pace and N.R. Pace. (1991). Detection of DNA contamination in *Taq* polymerase. BioTechniques (in press).

7. Schmidt, T.M., E.F. DeLong and N.R. Pace. (1991). Analysis of a marine picoplankton community using 16S rRNA gene cloning and sequencing. J. Bacteriol. 173:4371-4378.

8. Pace, N.R. (1991). Origin of life--Facing up to the physical setting. Cell 65:531-533.

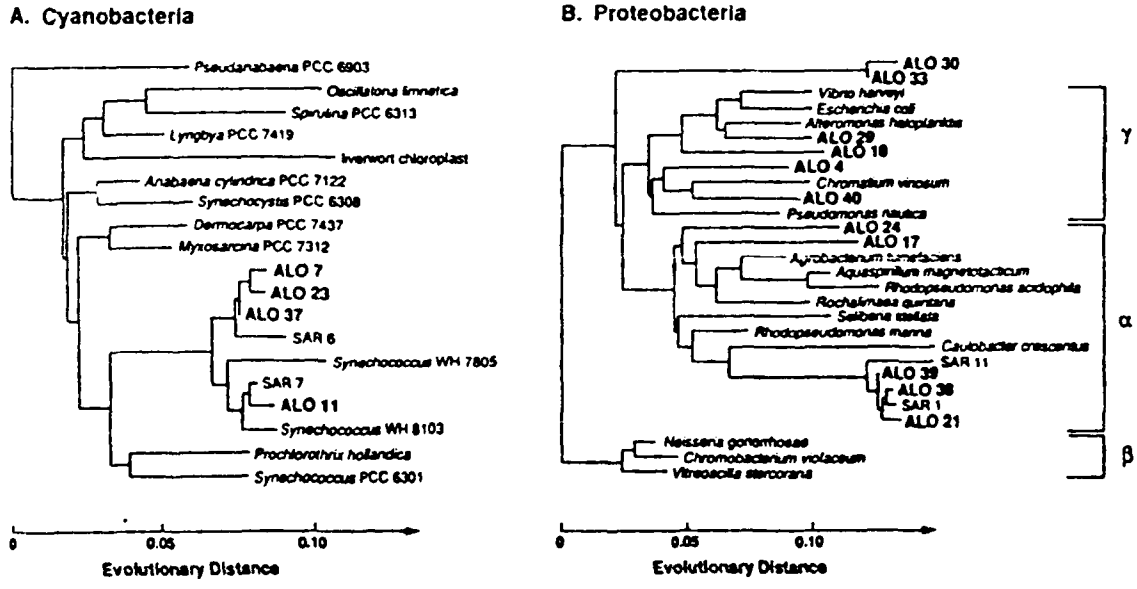
Highlight Page (following)

Objectives

- Analyze marine picoplankton using rRNA sequences.
- Develop probes for single-cell phylogenetic analysis.

Accomplishments

- Tangential flow filtration used for bulk collection of picoplankton.
- Hybridization analysis shows picoplankton from Atlantic and Pacific sites are dominantly (>90%) eubacteria. Few eukaryotes, no archaebacteria detected (at low resolution).
- Fluorescently labeled phylogenetic group-specific oligonucleotides allow analysis of single cells.
- Representative, recombinant library of Pacific Ocean picoplankton established (20 kb *Sau*3A fragments in l EMBL3). This library is available to the community.
- Numerous 16S rRNA genes identified, isolated, sequenced, and phylogenetically analyzed; mostly "proteobacteria" ("purple bacteria" *sensu* Woese), and cyanobacteria. Trees show "coordinates" for the picoplankton genotypes.



Future Directions

- Quantitative analysis of cloned rDNAs in mixed population.
- Correlating cloned (natural population) rDNAs with cultivated organisms.
- Characterization of high-temperature ecosystems

Norman Pace
Indiana University