

A.H. Treusch<sup>1</sup>, K.L. Vergin<sup>1</sup>, L.A. Finlay<sup>1</sup>, R.M. Burton<sup>2</sup>, C.A. Carlson<sup>3</sup>, S.J. Giovannoni<sup>1</sup>

<sup>1</sup> Department Of Microbiology and <sup>2</sup> Department of Mathematics, Oregon State University, Corvallis, OR, USA;

<sup>3</sup> Department Of Ecology, Evolution And Marine Biology, University Of California, Santa Barbara, CA, USA

## Abstract:

Marine ecosystems and their bacterioplankton communities play an important role in global biogeochemical cycles. Despite progress in the cultivation and metagenomic analyses of marine microorganisms, many questions about spatial and temporal patterns in water column microbial communities are unresolved. Long-term studies showing trends in community composition and their dynamics are rare, because of the difficulty of obtaining large sets of samples collected in similar fashion over several years.

In this study we used samples collected during monthly cruises ranging from 1991 to 2004 at the Bermuda Atlantic Time-Series (BATS) study site. To profile microbial community composition we used a terminal restriction fragment length polymorphism (T-RFLP) approach with multiple enzymes. The resulting large dataset was analyzed employing a bioinformatic pipeline of tools developed in house for noise filtration, normalization, aligning of samples and the putative identification of groups of organisms represented by the different T-RFLP peaks. To verify the identification of fragments, small subunit ribosomal RNA gene clone libraries were constructed and clones representing interesting T-RFLP peaks were sequenced.

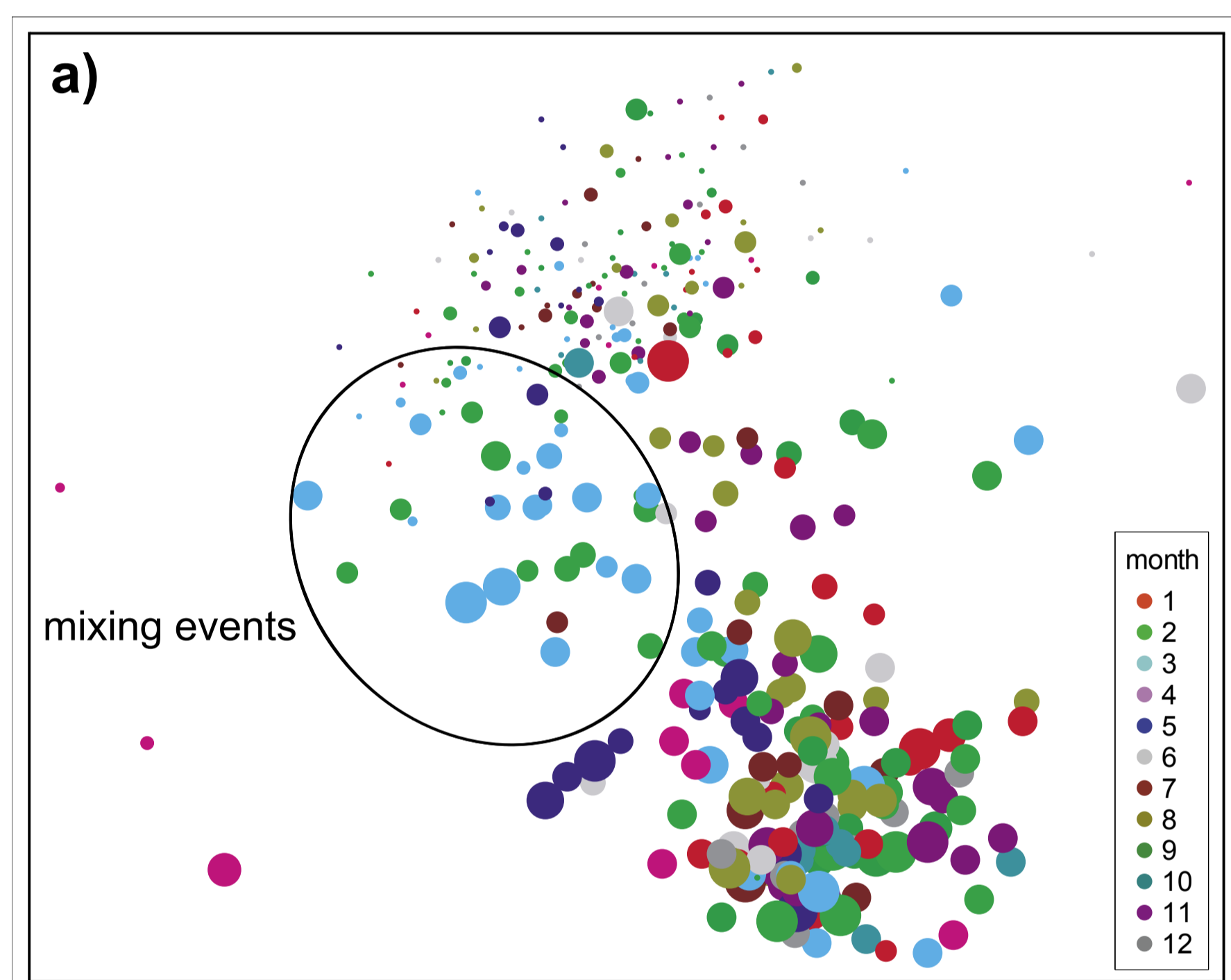
The results of the community profiling showed that over 100 of the observed T-RFLP patterns had seasonal distributions and close to 150 had distinct spatial distributions in the water column. Organisms represented by these T-RFLPs were identified as belonging to the bacterial clades SAR11, SAR86, OCS116 and SAR406, among others.

Further analyses are focusing on correlating the microbial distributions with environmental factors such as temperature, depth, salinity and organic and inorganic nutrients.

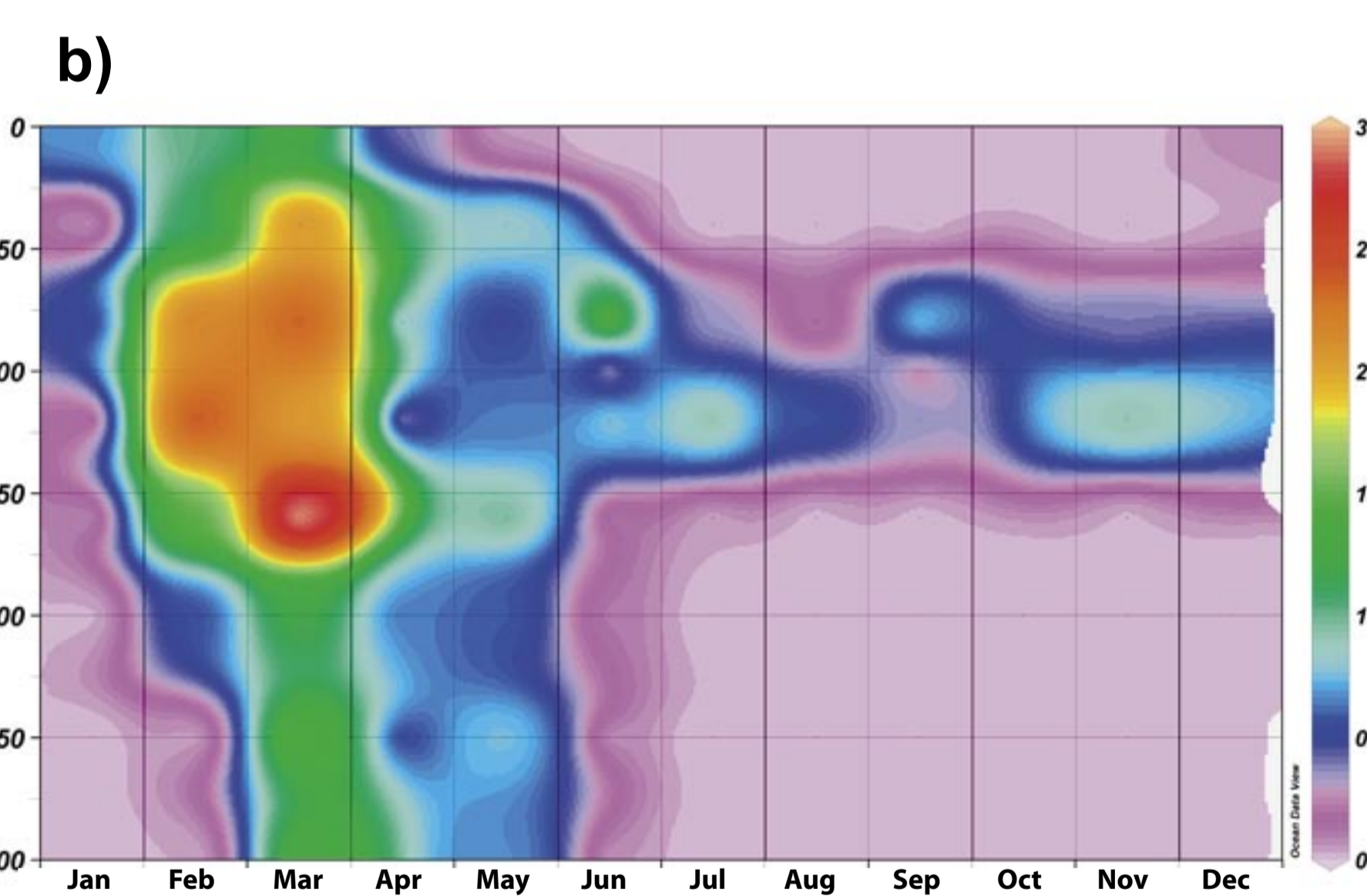
## Figure 1:

### Spring mixing of water column at BATS.

a) Nonmetric multidimensional scaling (NMS) ordination of the complete T-RFLP dataset. Samples are color coded for the month they have been collected. Sampling depths are indicated by the diameter of the dots (larger dots represent deeper samples). Shallower and deeper water samples cluster in distinct groups. Samples collected during spring time cluster independent of their depth, reflecting the convective overturn events.



b) An example: bloom of the alpha-proteobacterial group OCS116 in response to mixing events. The plot shows a composite year generated from all available datapoints.



## Results and Conclusions:

During ordination analysis, it became obvious that trends described earlier for a subset of the data (Morris et al., *Limnology and Oceanography*, 2005) could be corroborated. The complete dataset showed a clustering of the samples coming from the shallower surface layer (0-120 m) and those from the deep surface layer (160-300 m) (Figure 1 a). Spring time convective overturn events in the Sargasso Sea (Hansell and Carlson, *Deep-Sea Res. II*, 2001) were also reflected in the composition of the microbial community, as a subset of samples taken from February to May in different years and depths clustered more close together than to samples from the respective depths (Figure 1 a). This massive change in the community structure is caused by the bloom of several groups of organisms during the mixing, most notably phytoplankton. Heterotrophic bacterial groups (e.g. OCS116, Figure 1 b) also bloom during this period, presumably using the resources provided by the primary producers.

Looking at the distribution of terminal restriction fragments of different lengths revealed trends with depth and environmental factors at other times of the year. SAR11 subgroup 1a was present in shallower waters throughout the year but showed a distinct bloom in the summer with abundances up to 15% of the total detected fluorescence (Figure 2 a). SAR86 (Figure 2 b) could only be detected in the upper 100 m of the water column and reached the highest abundances during the summer. SAR406/Marine Group A, SAR202 and SAR324 (Figure 2 c,d,e) were mainly found in the mesopelagic part of the water column with blooms during the summer (SAR202, SAR324) and blooms in spring and late summer/fall (SAR406).

Further work is focusing on identifying the organisms associated with triples that were predicted with high probability and relating these groups of organisms to environmental factors to identify long term trends.

## Figure 2:

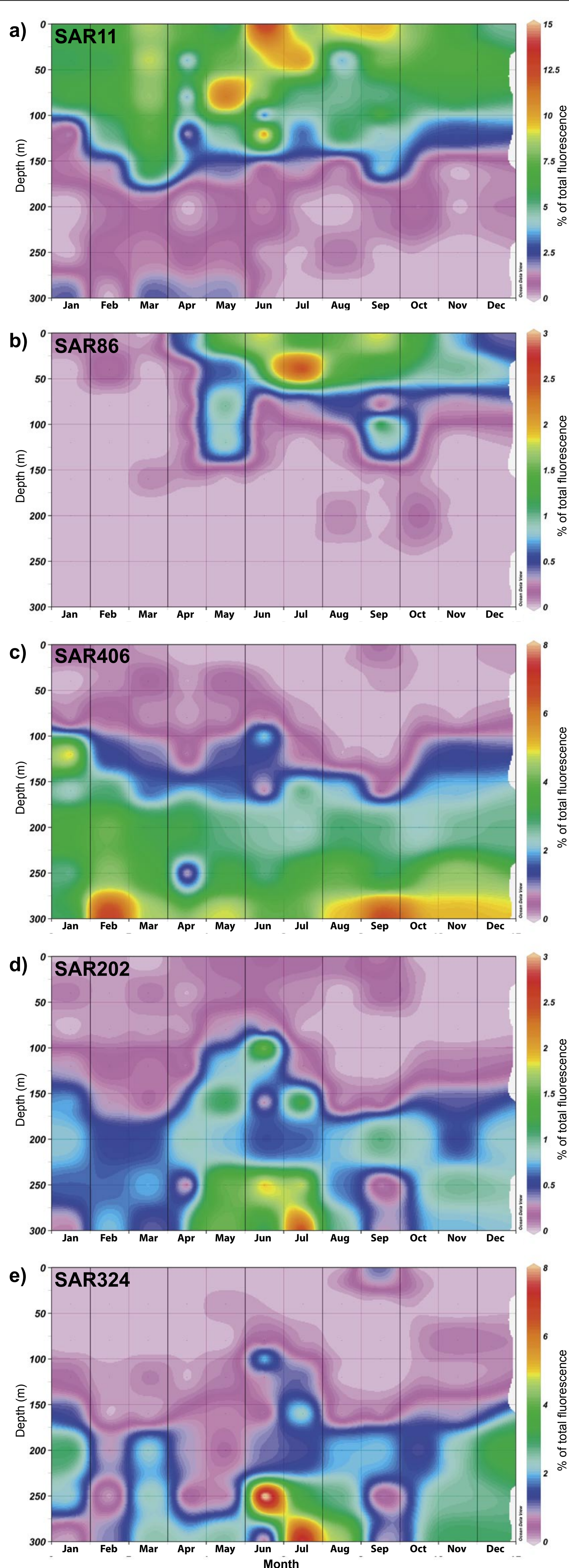


Figure 2: Spatial and temporal distribution of selected organisms.

Plotted are composite years of combined and normalized fluorescence data of triples (see Table 1). ODV (Schlitzer, *Ocean Data View*, <http://odv.awi.de>, 2006) was used for plotting.

a) SAR11 1a, b) SAR86, c) SAR406/Marine group A, d) SAR202, e) SAR324

## Methods:

Monthly samples (0 m and 200 m) were collected at the BATS site (32°N, 64°W) from August 1991 to February 1994 and September 1997 to January 2004. Additionally, depth profile samples (0, 40, 80, 120, 160, 200, 250, 300 m) were collected on 35 sampling dates. From the altogether 402 samples, nucleic acids were extracted and 16S rRNA genes amplified using the bacterial primers 27F-FAM/519R. The T-RFLP analysis was performed using the restriction enzymes *BsuRI*, *Bsh1236I* and *MspI* in parallel.

Most organisms were identified by one fragment from each enzyme (a triple) resulting in a dataset with a higher resolution of the community structure. To analyze the dataset we developed a workflow in which we combined readily available and custom programmed tools to adjust for systematic errors in the data and to identify the organisms present. One algorithm was developed for the *de novo* prediction of putative triples based on presence/absence and abundance of the respective DNA fragments in the dataset (Burton, Treusch et al., in preparation). For the verification of predicted triples, six small subunit ribosomal RNA gene clone libraries were constructed from representative samples. Selected clones were sequenced and their phylogeny reconstructed using the software package ARB (Ludwig et al., *Nucleic Acids Research*, 2004). To reveal non-linear trends in community structure we performed a non-metric multidimensional scaling (NMS) analysis using the PC-ORD software package (McCune and Grace, *MJM Software*, 2002).

Table 1:

	<i>BsuRI</i>	<i>Bsh1236I</i>	<i>MspI</i>
OCS116	302	104	160
SAR11 1a	113	363	149
SAR86	187	106	140
SAR406	230/ 264	106/399/ 400	142/ 150
SAR202	155	394	209
SAR324	405/ 407	105	130/ 161

Table 1: Observed triples for selected organisms.

## Acknowledgements:

We thank the crew of the *Weatherbird II* for their assistance and support, Rachel Parsons for technical assistance and Robert Morris for his input. We further want to thank the BATS chief scientists for assistance in water collection and accommodating wire time requests. Research was funded by the *Marine Microbiology Initiative of the Gordon and Betty Moore Foundation* and a *NSF Microbial Observatory* grant. AHT is supported by a *Feodor Lynen Fellowship of the Alexander von Humboldt Foundation*.