New Biotechnologies for sampling the ecological diversity of the oceans: The Informatics Challenge

Carmen Palacios (*1*,*2*), Bertil Olsson (*3*), Antje Boetius (*1*), Philippe Lebaron (2)
1 Max Planck Institute for Marine Microbiology. Celsiusstr. 1. 28359 Bremen Germany
2 Marine Biological Laboratory. 7 MBL St. Woods Hole 02543 MA USA
3 Observatoire Océanologique Banyuls, Laboratoire de Microbiologie, 66651Banyuls-sur-mer, France

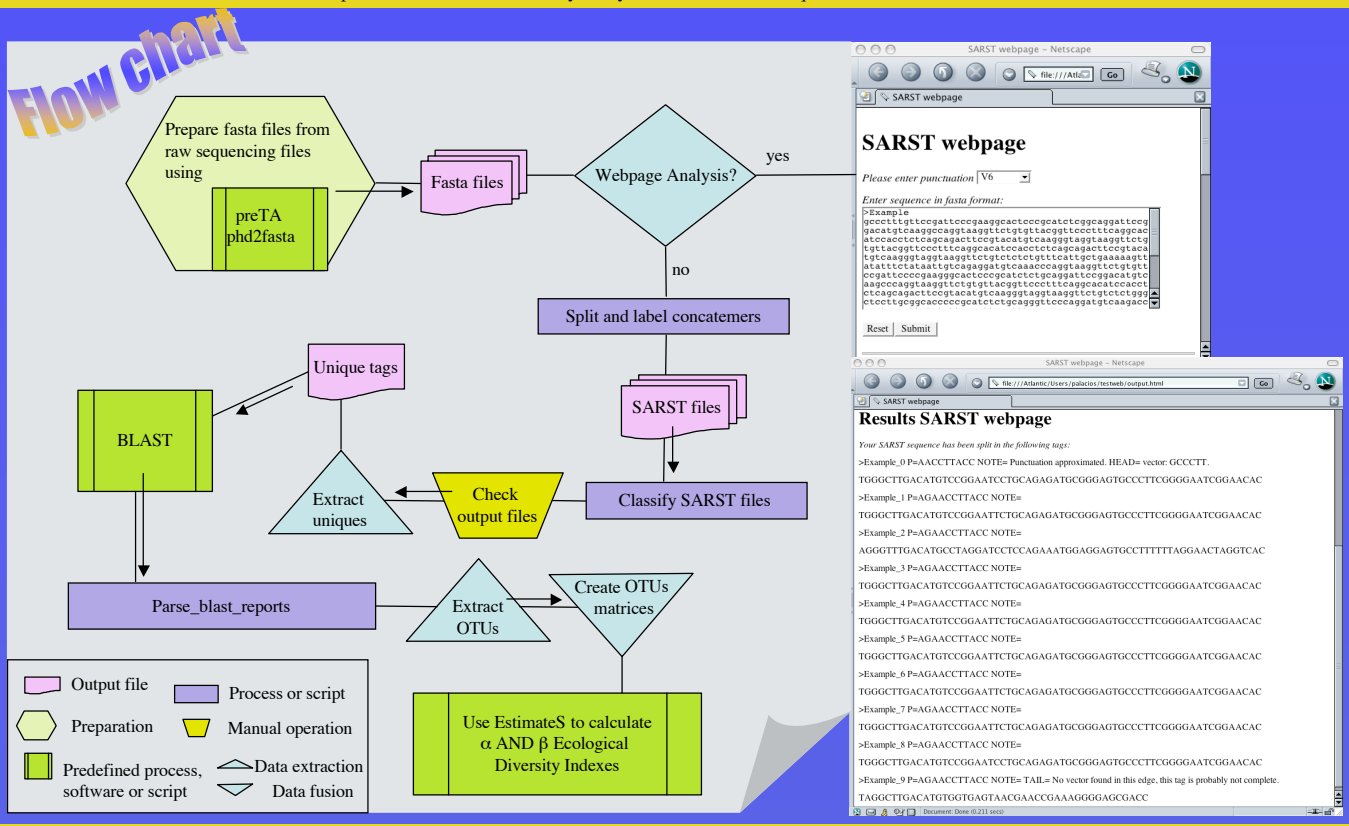
Background information

Exploring the ecological diversity of natural environments is taking the study of biodiversity one step further by unravelling the relationships among the organisms and with their habitat. But ocean biodiversity seems so large and our knowledge of it too small. New technologies might help to decrease the gap helping to better understand the ecology of the vast oceans. Molecular methods continue revolutionizing biodiversity studies, particularly challenged in the field of microbiology. For instance, microbes are the primary engines of the biosphere and may represent as much as 90% of biomass in marine systems.

A new set of technologies are being developed that might overcome the technical difficulties that microbiologists still face when studying ecological diversity. The key feature of these new methods is the use of the intrinsic phylogenetic information contained in genetically *hypervariable* regions of the 16S ribosomal RNA molecule to extract phylotype information directly from sequencing (Bertilsson et al 2002, Neufeld et. al. 2004, Kysela et. al. 2004). In other words, the obvious advantage of these newly developed technologies for ecological diversity studies is that they allow detecting all organisms present in natural samples providing simultaneous direct counts of its relative abundances, avoiding the difficulties and assumptions underneath band profiles in the fingerprinting methods (DGGE, TRFLP) most widely used to characterize microbial communities until today. The method we have developed is SARST-V6 (Serial Analysis of V6 Ribosomal Sequence Tags, Kysela et al 2004). Its strategy is to render concatemers of the PCR amplified V6 rDNA hypervariable region from natural samples increasing by at least 6-fold the yield of a single sequence compared to a regular16S rDNA library study.

Why this communication?

SARST-V6 imposes a new challenge to the study of biodiversity; it implies processing large amount of sequences. Computerization is necessary in order to treat huge datasets like this. We have built a series of scripts to facilitate the analysis of SARST sequences. The flow chart below outlines the **SARST-V6 pipeline.** It comprises the stages that involve from getting your sequences out of the chromatograms until estimating species richness, evenness and β indexes of ecological diversity. Violet and blue panels in this flow chart correspond to the scripts and programs contributed by this communication. Furthermore, a web page will be posted soon for user friendly analysis of SARST-V6 sequences over the internet.



What is next? Ecological diversity of sunken woods in deep ocean

We have successfully applied SARST-V6 to several environments (Kysela *et al.* in press, Palacios *et al.* in preparation). Our next challenge will be to characterize the microbial community composition of sunken woods in deep marine waters using all these helpful newly developed tools. Sunken woods are very interesting yet unexplored habitats from a biodiversity point of view. Degradation most likely occurs by specific microbes and animal communities adapted to the use of woods or other high carbon content materials as substrates. At least for metazoans, these environments might act as stepping-stones for chemosynthetic communities that inhabit hydrothermal vents and cold seeps. However, to the best of our knowledge, there is no published work that characterizes the microorganisms that dwell sunken woods in oceans, neither anybody has demonstrated the development of anaerobic communities in them.

Bibliography

Bertilsson S, Cavanaugh C.M., Polz M.F. 2002 Sequencing-Independent method to generate oligonucleotide probes targeting a variable region in bacterial 16S rRNA by PCR with detachable primers. Applied and Environmental Microbiology 68:6077-6086.

Kysela D, Palacios C, Sogin M.L. 2004. Serial analysis of V6 ribosomal sequence tags (SARST-V6): a method for efficient, high-throughput analysis of microbial community composition. Environmental Microbiology (in press)

Neufeld JD, Yu Z, Lam W, Mohn WW. 2004. Serial analysis of ribosomal sequence tags (SARST): a high-throughput method for profiling complex microbial communities. Environmental Microbiology 6:131-144

Acknowledgements

We thank you Dave Kysela, Andrew MacArthur and Laura Shulman for inputs on the code of different scripts. We are also thankful to Mitchell L. Sogin for financial support through a NAI grant.