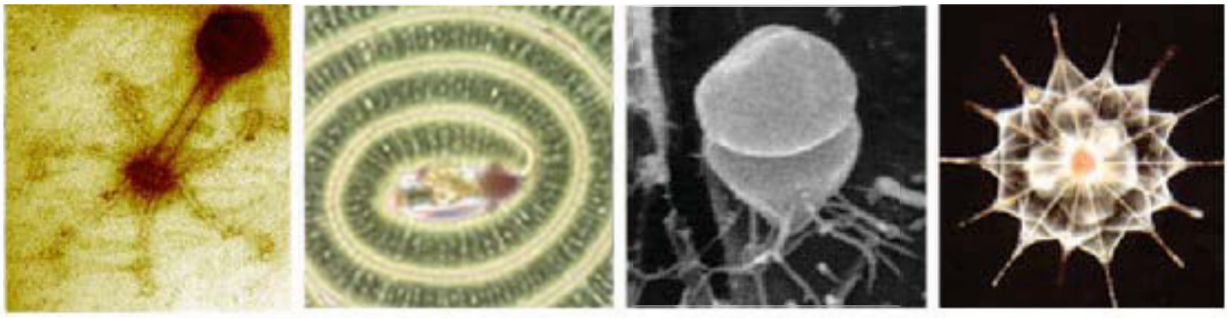


International Census of Marine Microbes
1st Annual Meeting

June 12th-15th, 2006

NH Leeuwenhorst Noordwijkerhout, The Netherlands

Meeting Organizers

Principal Investigators

Mitchell L. Sogin Marine Biological Laboratory (MBL)

Jan W. de Leeuw The Royal Netherlands Institute for Sea Research (NIOZ)

Secretariat

Linda Amaral-Zettler (MBL)

Scientific Organizing Committee

Gerhard Herndl (NIOZ)

David J. Patterson (MBL)

Stefan Schouten (NIOZ)

Lucas Stal (NIOO)

Staff Coordinator

Trish Halpin (MBL)

Table of contents:

Programme and Logistics

International Census of Marine Microbes 1st Annual Meeting Agenda
June 12th-15th, 2006
NH Leeuwenhorst Noordwijkerhout, The Netherlands

Monday, June 12th, 2006

All day Arrival at NH Leeuwenhorst, Noordwijkerhout
1700: ICoMM Scientific Organizing Committee (SOC) Meeting
1700-1900: Poster Set-up
1900-2100: Icebreaker party/reception with drinks and food.

Tuesday, June 13th, 2006

0830: Welcome, logistics, purpose of the meeting (**Baross, Sogin, de Leeuw, Amaral Zettler**)
0900: **Sogin/de Leeuw**: Meeting the ICoMM Challenge, Unfathomable microbial diversity in the deep sea: an unexplored "rare biosphere"
1000: Coffee Break
1030 **Carles Pedrós-Alió:** Marine microbial diversity: can it be determined? (30 minutes)
1100: **Slava Epstein:** Rarefaction (30 minutes)
1130: **Gerhard Herndl**: Integrating prokaryotic microdiversity into ecosystems function theory (30 minutes)
1200: Lunch Break
1300: **Steve Giovannoni:** High throughput cultivation of microbes (30 minutes)
1330: **John Heidelberg** Sorcerer II (30 minutes)
1400: Discussion of working group objectives
1430: Working Group Session I
1700-1900: Poster Session
1900: Dinner

Wednesday, June 14th, 2006:

0830: **Peg Riley**: Lateral gene transfer (1 hour)
0930: Plenary Session
1130: Working Group Session II
1215: Lunch Break
1330: **Katrina Edwards/Julie Huber:** Seamounts (1 hour)
1430: Working Group Session III
1600: Plenary Short Reports
1700: End of second day
1715: ICoMM Scientific Advisory Council (SAC) meeting
1900: Dinner

Thursday, June 15th, 2006:

0830: **Steve D'Hondt** : IODP (1 hour)
0930: **John Baross** – SAC report to meeting
1030: Working Group Session (Synthesis)
1215: Lunch Break
1330: **Antje Boetius**: Microbial diversity in marine sediments – what do we know about vertical, horizontal and temporal distribution patterns? (30 minutes)
1400: Complete writing
1500: John Baross to present SAC recommendations / **Plenary Session** to make decisions and planning the future led by SAC.
1700: End of Meeting

Meeting Locations

Registration desk	**Lobby**
Plenary Sessions	**Erasmus 149**
Working Group Sessions	**Erasmus 103**
	Erasmus 105
	Erasmus 107
Icebreaker party	**Alegria**
Congress Dinner	**Alegria**

Dear ICoMM meeting participants:

We are looking forward to your participation in the **International Census of Marine Microbes (ICoMM) June Meeting** on 13th-15th of June, 2006 at the NH Leeuwenhorst Noordwijkerhout, The Netherlands.

ICoMM (http://icomm.mbl.edu) is part of the Alfred P. Sloan Foundation's Census of Marine Life (http://www.COML.org/). The most general statement of **ICoMM**'s goal is to develop a highly-resolved biodiversity database for marine microbes and to understand how these populations evolve and redistribute on a global scale. Participants in ICoMM seek to (1) catalogue all known diversity of single-cell organisms inclusive of the Bacteria, Archaea, Protista and associated viruses, (2) to explore and discover unknown microbial diversity, and (3) to place that knowledge into appropriate ecological, biogeochemical and evolutionary contexts. Examples of questions that ICoMM will address include but are not limited to:

- What governs the evolution of marine microbial lineages within complex marine communities?
- Why do marine microbial consortia retain functionally equivalent but genetically distinct lineages?
- Is there a marine microbial biogeography and if so, what are the principal drivers or restrictors?
- How does genotypic diversity shape phenotypic diversity, and how does this diversity influence the biogeochemistry and the functioning of marine ecosystems?

ICoMM serves the international scientific community through its efforts to coordinate research activities and secure funding for studies of marine microbial diversity. The web site http://icomm.mbl.edu/ contains important information about the organization and objectives of ICoMM.

We have attached a final agenda (both as Word and PDF documents) and a list of participants at the end of this document. We plan to distribute a meeting booklet that includes abstracts of presented talks, posters, and/or descriptions of each participant's research interests. **If you have not yet done so, please submit an abstract by the 5th of June otherwise we will not be able to include it in the meeting booklet. Also, if you will be bringing a poster, please notify either Linda Amaral-Zettler (amaral@mbl.edu) or Trish Halpin (thalpin@mbl.edu) of your poster title and authors so that we can include these in the meeting booklet as well.**

We have identified at least four working group topics for the meeting participants to tackle new questions of relevance to the International Census of Marine Microbes. The primary focus of the meeting will be on the use of high-throughput technology to extensively sample marine microbial communities including the low-abundance taxa that account for the *rare biosphere*. We recognize the importance of thinking about community microbial structures in terms of population and community evolution, ecological function, biogeochemical processes and metagenomics. We will use the elements from our discussions for developing new competitive research initiatives that will further ICoMM's overall

objective of *defining the range of genetic diversity and relative numbers of different microbial organisms at sampling sites throughout the world's oceans.*

This is the first community-wide meeting for the ICoMM initiative which is intended to function as a research coordination network. Please consider whether you would like to formally affiliate with ICoMM by providing information about your research programs including objectives, research goals, and results. ICoMM's principle coordination objective is to assist in securing resources for a global census of marine microbial diversity. We will achieve this through a series of workshops and support of grant writing activities. We are also requesting participants who are active in field programs and willing to share protocols with other members of the community**, to bring sufficient information about their recommended sampling procedures and experimental protocols with them to the meeting (preferably in electronic format).** In addition to successful protocols, we are also interested in garnering information about protocols/techniques that have been tested but have not worked well for your applications. This information will then be summarized and made available on the ICoMM website.

We anticipate that this will be both an enjoyable and productive workshop.

We look forward to welcoming you in The Netherlands,

Mitchell Sogin

Jan de Leeuw

Linda Amaral-Zettler

Logistics:

Meeting Site and Accommodations:

<u>You are not required to reserve a hotel room</u>. Jan W. de Leeuw, our host in the Netherlands, has made arrangements for meeting space and hotel accommodations at the NH Leeuwenhorst hotel.

NH Leeuwenhorst
Langelaan 3
2211XT Noordwijkerhout
The Netherlands
Tel.: +31 (0)252 378428
Fax: +31 (0) 252 378891
http://www.nh-hotels.com

Ground Transportation:
For those of you with similar arrival times at the Schiphol airport (see list at the end of this document), we have arranged for shared taxis to the hotel. Please come to the meeting point in the arrivals hall at the Schipol Airport and look for colleagues and/or a representative of the Brouwers Taxi Company. If you have sent your itineraries to Trish (thalpin@mbl.edu) or Linda (amaral@mbl.edu) we will pass the information on to the taxi company. They will monitor the arrival times of your flights and collect multiple people for shared taxi rides (at a cost of 51.50 euros/taxi). We've arranged this option for June 12th and 13th. On other days, people will have to take the train to Leiden Central Station and arrange for their own taxi to the hotel. In the event that there are no other colleagues at the meeting point or absence of a Brouwer Taxi representative, please take the train to Leiden station. Use the bus service (the NH Leeuwenhorst Express-bus runs every 15 minutes between 0800 and 0915 Monday-Friday- **after 0915 you will need to take a taxi**) or a taxi to the hotel. We will not be able to reimburse you for rental cars during your stay.

The ICoMM registration desk in the entrance hall of the hotel will be open on June 12th from 10 am until noon on June 13th.

Meals:

There will be an icebreaker party with drinks and light fare on the evening of June 12th. Breakfast, lunch and dinner (plus coffee during break-times) will be provided on Tuesday and Wednesday, June 13th and 14th. Breakfast and Lunch (plus coffee) will be provided Thursday, June 15th. We will also provide up to $40 per diem for travel days. Please let us know if you have dietary restrictions.

Expense Reports: After you return from the workshop, you will need to submit an expense reimbursement form (which will be available at the workshop) to:

Trish Halpin
Staff Coordinator
Josephine Bay Paul Center for
Comparative Molecular Biology and Evolution
Marine Biological Laboratory
7 MBL Street
Woods Hole, MA 02540
USA
thalpin@mbl.edu
Tel: 508-289-7282
Fax: 508-457-4727

RECREATION FACILITIES:

AQUA FITNESS CENTER

Fancy a refreshing swim? A workout? A healthy tan? In the Aqua Fitness Center, you can combine exercise with relaxation.
The opening times of the fitness and sauna are:

Monday - Friday	07.00 a.m. – 10.00 p.m.
Saturday	09.00 a.m. – 09.00 p.m.
Sunday	09.00 a.m. – 05.00 p.m.

SWIMMING POOL

In the Aqua Fitness Center (ground floor, G wing), there is a swimming pool with whirlpool.
A visit to our swimming pool and the whirlpool is free for our hotel guests.
The opening times of the swimming pool are:

Monday	07.00 a.m. – 09.00 a.m. & 06.00 p.m. – 10.00 p.m.
Tuesday	07.00 a.m. – 09.00 a.m. & 11.30 a.m. – 10.00 p.m.
Wednesday	07.00 a.m. – 09.00 p.m. & 06.00 p.m. – 10.00 p.m.
Thursday	07.00 a.m. – 10.00 p.m.
Friday	07.00 a.m. – 09.00 a.m. & 06.00 p.m. – 10.00 p.m.
Saturday	09.00 p.m. – 09.00 p.m.
Sunday	09.00 a.m. – 05.00 p.m.

Naturally, there are changing-rooms with showers available but you'll also find a hot-whirlpool, bubble-bath, waterparasol, jetstreamers, a turbo-jet and a solarium.
Towels and swimming costumes can be hired at the front-office desk where you can also collect the key to your locker. The solarium costs € 8.50 for 20 minutes.

SAUNA

A visit to our sauna costs € 8.50 per person per session. Deposit for the key is € 5.00. You can then make unlimited use of the entire sauna, including Turkish bath, plunge bath, foot bath, massage showers and relaxation room. For the opening times of the sauna we refer you to our main reception, the Aqua Fitness Center and/or the notice board in the lobby.

FITNESS STUDIO

Use of the fitness studio with exercise apparatus and weights will cost you € 8.50.
From Monday to Friday, group training is conducted at fixed times. The training times are posted up at various places in the building. The maximum size of the groups is 15 persons per training session.

CHEERS SPORTS CAFÉ

Near the Rotunda in the basement, you'll find an American-style sports café. The Cheers Sports Café is a real meeting-place. Why? Because it's located at the centre of a complete range of sporting facilities:

SQUASH

The two squash courts are open from 7.00 a.m. to 01.00 a.m. Hire of a court costs € 5.50 for half an hour and € 10,00 for an hour. You can hire rackets and tennis balls at the bar, or when closed, at the front-office desk. Shoes with black soles are not permitted. Shoes are available at the front-office desk. Reservations are accepted only on the same day as you wish to use a court.

TABLE TENNIS

Table tennis is free. Reserving is not possible. Balls can be purchased on the spot for € 0.50.

BILLIARDS

Use of the billiard tables is € 0.50 for 20 minutes. Balls and cues etc. can be found in the area where the tables are set up. Balls for American pool are released when € 1.00 is inserted in the slot. Another billiard table stands at the end of the F-corridor (free of charge). It is not possible to reserve.

DARTS

Use of the dartboard is free of charge. Reserving is not possible. Darts are available at the front-office desk and in the café.

BOWLING

There are 4 bowling lanes. Hire of a lane costs € 12.00 for an hour. Special bowling shoes are obtainable free of charge. You can reserve at the bar, or in case closed, at the front-office desk. The lanes are open all days from 07.00 p.m. until 01.00 a.m. and all day Saturday and Sunday. Other times by arrangement with the front-office staff.

OUTDOOR

Once you've seen all that NH Leeuwenhorst has to offer on the inside, you might well forget that we also have attractive outdoor sports facilities:

TENNIS

Our all-weather outdoor tennis courts are open the whole year round from 07.00 a.m. to sunset. You can find them in the wood near the sports field. The use of our tennis court is free of charge for our guests. You can obtain rackets, balls and tennis shoes at the front-office desk. You may only reserve a court on the same day as you wish to use it.

CYCLING

You can rent a bicycle for € 1.25 per hour. If you wish to use the bicycle for a day, we will charge you € 7.50. Deposit € 10.00.
From 8 persons we can rent bicycles for you. We are pleased to make you an offer.
Information and/or reservation at the front-office desk. Cycling routes are also available there.

WALKING

Take a walk through the bulb fields, the dunes or the wood. You'll find a walking route posted on the notice board at the tower. The red posts will guide you to the wood and the country side of South Holland. You can get information about other routes at the front-office desk.

GOLF

Guests at NH Leeuwenhorst can use the facilities of the Golf Center Noordwijk offering a 9 holes par-3 course, driving range, pitch and putting green.
Reservations via the front-office desk. You must have a golfing proficiency certificate.

The above mentioned tariffs are valid up to and including 31 December 2006.
Times and prices subject to change.

Sport facilities are used at your own risk.

List of Participants

Silvia G. Acinas
NIOO-KNAW The Netherlands
Institute of Ecology
Department of Marine Microbiology
Korringaweg 7, P.O. Box 140
4400 AC Yerseke, The Netherlands
Email: s.acinas@nioo.knaw.nl

Linda Amaral-Zettler
The Josephine Bay Paul Center for Comparative Molecular Biology
and Evolution, MBL
7 MBL Street, Woods Hole, MA 02543 USA
Email: amaral@mbl.edu

Felipe Artigas
Université du Littoral (ULCO)
UMR 8013 ELICO, MREN -ULCO 32
av. Foch, 62930 Wimereux, France
Email: Felipe.Artigas@mren2.univ-littoral.fr

John Baross
University of Washington
Oceanography Department
260 Marine Sciences Bldg., Seattle Washington 98175-7940, USA
Email: jbaross@u.washington.edu

Judith v Bleijswijk
The Netherlands Institute for Sea Research
Texel, Den Burg, The Netherlands
Email: judith@nioz.nl

Antje Boetius
Max Planck Institut
für Marine Mikrobiologie
Celsiusstr. 1
D-28359 Bremen, Germany
Email: aboetius@awi-bremerhaven.de

Henry Boumann
The Netherlands Institute for Sea Research
Texel, Den Burg, The Netherlands
Email: boumann@nioz.nl

Peter Burkill
Southampton Oceanography Centre
Biogeochemistry & Ecosystems Research Group
Southampton, United Kingdom
Email: p.burkill@noc.soton.ac.uk

D. Chandramohan
The National Institute of Oceanography
19/3, 3rd St.,Ratnapuri Colony
J.N.Salai Koyambedu
Chennai-600107 (T.N.), India
Email: drd.chandramohan@gmail.com

Guiseppe D'Auria
Universidad Miguel Hernandez de Elche
ES 03550 San Juan, Spain
Email: gdauria@umh.es

Steven D'Hondt
University of Rhode Island
Graduate School of Oceanography
Narragansett Bay Campus
South Ferry Road
Narragansett, Rhode Island 02882, USA
Email: dhondt@gso.uri.edu

Jan de Leeuw
The Royal Netherlands Institute
for Sea Research
P. O. Box 59 1790 A B den Burg
Texel, The Netherlands
Email: deleeuw@nioz.nl

Virginia Edgcomb
Woods Hole Oceanographic Institution Marine Chemistry and
Geochemistry Department
Woods Hole, MA 02543, USA
Email: vedgcomb@whoi.edu

Katrina Edwards
Woods Hole Oceanographic Institution Marine Chemistry
and Geochemistry Department
Woods Hole, MA 02543, USA
Email: katrina@whoi.edu

Slava Epstein
Northeastern University
Biology Department
134 Mugar Life Sciences
360 Huntiington Avenue
Boston, MA 02115, USA
Email: s.epstein@neu.edu

Carola Espinoza
Center for Oceanographic Research
in the Eastern South Pacific
P.O. Box 160C
Casilla 160-C, Concepción, Chile
Email: carespin@udec.cl

Isabel Ferrera
Biology Department
Portland State Unversity
P.O. Box 751
Portland, Oregon 97207, USA
Email: iferrera@pdx.edu

Victor Ariel Gallardo
Center for Oceanographic Research in the Eastern South Pacific
P.O. Box 160C
Casilla 160-C, Concepción, Chile
Email: vagallar@udec.cl

Steve Giovannoni
Oregon State University
Corvallis, Oregon 97331-4501, USA
Email: steve.giovannoni@orst.edu

Frank Oliver Gloeckner
Max Planck Institut
für Marine Mikrobiologie
Celsiusstr. 1
D-28359 Bremen, Germany
Email: fog@mpi-bremen.de

Martha Liliana Gómez García
INVEMAR - Instituto de Investigaciones Marinas y Costeras José Benito
Vives de Andréis
A.A. 1016 Cerro Punta Betín
Santa Marta, Colombia
Email: mlgomez@invemar.org.co

Maria-Judith B.D. Gonsalves
National Institute of Oceanography
Dona Paula, Goa 403004, India
Email: mjudith@nio.org

John Heidelberg
The Institute for Genomic Research
9712 Medical Center Drive
Rockville, Maryland 20850, USA
Email: jheidel@tigr.org

Gerhard Herndl
The Netherlands Institute for Sea Research
Texel, Den Burg, The Netherlands
Email: herndl@nioz.nl

Julie Huber
MBL and the NASA Astrobiology Institute
7 MBL Street, Woods Hole, MA 02543 USA
Email: jhuber@mb.edu

Elena Ivars-Martinez
Evolutionary Genomics Group
Universitas Miguel Hernandez de Elche
Avenida de la Universidad s/n.
Elche 03202, Spain
Email: eivars@umh.es

Renzo Kottman
Max Planck Institut für Marine Mikrobiologie
Celsiusstr. 1
D-28359 Bremen, Germany
Email: rkottman@mpi-bremen.de

Alexandra Kraberg
Stiftung Alfred-Wegener
Institut für Polar and Meeresforschung in
der Helmholtz-Gemeinshaft
Kurpromenade, (Building C-46)
D-27498 Helgoland, Germany
Email: akraberg@awi-bremerhaven.de

William Li
Bedford Institute of Oceanography
P. O. Box 1006
Dartmouth, Nova Scotia B2Y 4A2, Canada
Email: LiB@mar.dfo-mpo.gc.ca

Debbie Lindell
Civil and Environmental Engineering
Masschusetts Institute of Technology
15 Vassar Street, MIT 48-424
Cambridge, MA 02139, USA
Email: dlindell@mit.edu

Jose Lopez
Division of Biomedical Marine Research
Harbor Branch Oceanographic
5600 US North
Ft. Pierce, Florida 34946, USA
Email: lopez@HBOI.edu

Connie Lovejoy
Université Laval, Département de biologie
VCH 4042-A, Pavillon Alexandre-Vachon Sainte-Foy Québec G1K 7P4, Canada
Email: connie.lovejoy@bio.ulaval.ca

Alexander Loy
Department of Microbial Ecology
University of Vienna
Althanstrasse 14
A-1090 Wien, Austria
Email: loy@microbial-ecology

Ana Belén Martín Cuadrado
Evolutionary Genomics Group
Universitas Miguel Hernandez de Elche
Avenida de la Universidad s/n.
03202 Elche, Spain
Email: amartin@umh.es

Alison Murray
Desert Research Institute
Division of Earth and Ecosystem Sciences
2215 Raggio Parkway
Nevada 89512-1095, USA
Email: alisonemurray@gmail.com

Phillip Neal
The Josephine Bay Paul Center for Comparative Molecular Biology
and Evolution, MBL
7 MBL Street, Woods Hole, MA 02543 USA
Email: pneal@mbl.edu

Rodolfo Paranhos
Departamento de Biologia Marinha, Universidade do Brasil
UFRJ - Laboratório de Hidrobiologia Avenida Pau Brasil 211, Prédio do CCS, bloco A, sala A1-071, Cidade Universitária, Ilha do Fundão, Rio de Janeiro,
RJ 21941-590, Brasil
Email: rodpar@biologia.ufrj.br

David Patterson
The Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, MBL
7 MBL Street, Wood Hole, MA 02543, USA University of Sydney
Heydon Laurence Building
Sydney, Ausralia
Email: dpatterson@mbl.edu

Carles Pedrós-Alió
Institut de Ciències del Mar
CMIMA
Passeig Marítim de la Barceloneta 37-49
E-08003 Barcelona, Spain
Email: cpedros@cmima.csic.es

Wim Pool
The Royal Netherlands Institute
for Sea Research
P. O. Box 59 NL
1790 A B den Burg, Texel, The Netherlands
Email: pool@nioz.nl

Alban Ramette
Max-Planck-Institut for Marine Mikrobiologie, Microbial Habitat Group
Celsiusstrasse 1
D-28359 Bremen, Germany
Email: aramette@mpi-bremen.de

Thomas Reinthaler
The Royal Netherlands Institute
for Sea Research
P. O. Box 59 NL
1790 A B den Burg, Texel, The Netherlands
Email: reinthal@nioz.nl

Anna-Louise Reysenbach
Portland State University
Department of Biology
SB2 Rm 246
1719 SW 10th Avenue
Portland, Oregon 97201, USA
Email: reysenbacha@pdx.edu

Margaret Riley
Biology Department
University of Massachusetts
221 Morrill Science Center
Amherst, MA 01003, USA
Email: riley@bio.umass.edu

Francisco Rodriguez-Valera
Universidad Miguel Hernandez de Elche
ES 03550 San Juan, Spain
Email: frvalera@umh.es

Stefan Schouten
Nederlands Instituut voor
Onderzoek der Zee
P.O. Box 59, NL-1
790 AB Den Burg, Texel, The Netherlands
Email: schouten@nioz.nl

Mitchell Sogin
The Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, MBL
7 MBL Street, Woods Hole, MA 02543, USA
Email: sogin@mbl.edu

Lucas Stal
Netherlands Institute of Ecology
(NIOO-KNAW)
P.O. Box 140
NL-4400 AC Yerseke, The Netherlands
Email: l.stal@nioo.knaw.nl

Mike Taylor
Department of Microbial Ecology
University of Vienna
Althanstrasse 14
A-1090 Wien, Austria
Email: wagner@microbial-ecology.net

Marcel van de Meer
The Royal Netherlands Institute for Sea Research
Texel, Den Burg, The Netherlands
Email: mvdmeer@nioz.nl

Edward vanden Berghe
Flanders Marine Institute
Oostende, Belgium
Email: wardvdb@vliz.be

Michael Wagner
Department of Microbial Ecology
University of Vienna
Althanstrasse 14, A-1090 Wien, Austria
Email: wagner@microbial-ecology.net

Tian Xiao
Key Laboratory of Marine Ecology & Environmental Sciences
Institute of Oceanology
Chinese Academy of Sciences
Nanhai Road, Qingdao
Shandong, 266071, P. R. China
Email: txiao@ms.qdio.ac.cn

Bess Ward
Princeton University
Department of Geosciences
M51 Guyot Hall
Princeton, New Jersey 08544, USA
Email: bbw@princeton.edu

Shi Ning Zhou
Sun Yat-Sen (Zhongshan) University
College of Life Sciences
Sun Yat-Sen (Zhongshan) University
135 XinGangXi Road
Guangzhou, 510275 Guangdong
P.R. China
Email: lsszsl@zsu.edu.cn

Research Summaries

Silvia G. Acinas

Exploring genomic micro-diversity and dynamics of two types of cyanobacteria: *Synechococcus* and *Pseudanabaena.*

Department of Marine Microbiology, Netherlands Institute of Ecology NIOO-KNAW, Yerseke, The Netherlands.

The combination of isolation by culturing and molecular approaches such as genome sequencing and quantitative PCR (QPCR) have been used recently as a successful strategy to explore genomic micro-diversity and distribution of six known marine cyanobacterial *Prochlorococcus* ecotypes. Despite recent progress however, little is known about other cyanobacterial populations, which represent a relevant fraction of the phototrophic microorganisms and therefore a crucial topic for the International Census of Marine Microbes.

This study focuses on structure-function relationships of two cyanobacterial populations: ***Synechococcus*** and ***Pseudanabaena.*** Our extensive culture collection containing more than 100 strains of these cyanobacteria, isolated from diverse aquatic environments, will provide insights into their genomic heterogeneity, structure and dynamics. Currently we are sequencing the 16S-23S rDNA internal transcribed spacer (ITS), the large ribosomal subunit 23S rRNA gene and part of the phycocyanin operon (cpc) to explore the phylogenetic structure of these cyanobacterial groups. Our results will extent the phylogenetic resolution and in combination with specific phenotypic traits such the differentiation of photosynthetic pigments phycoerythrin (PE) or phycocyanin (PC) rich isolates, will help to reveal coexisting genomes that may represent different ecotypes within these populations. Further molecular approaches will include the quantitative PCR (qPCR) and Amplified Fragment Length Polymorphism (AFLP) to examine cyanobacterial population dynamics and genomic microheterogeneity respectively. Finally, our main goal would be tracking the distribution of such coexisting genomes (ecotypes) over spatial and temporal scales and exploring the mechanisms driving selection and diversification within and among different ecotypes.

Linda Amaral-Zettler

A microbial diversity survey of a sewage- and thermally-impacted estuary: Mt. Hope Bay, Massachusetts
Amaral-Zettler, LA, Laatsch A, Rocca J, Dennett MR, and Gast, RJ

Coastal marine environments have been impacted by human activity for several centuries, including shoreline alteration, nutrient introduction, sedimentation, toxic compound release, and thermal modification. Mt. Hope Bay, Massachusetts is an ideal site to base a study of human pathogen presence and distribution because it has several important sources of human impact, including sewage disposal sites and the thermal outfall of a power plant within a mile of each other. The Bay is currently undergoing limited monitoring for several different parameters, including fish populations, river runoff, meteorological forcing, tidal cycles and water chemistry as part of the Mt. Hope Bay Natural Laboratory (MHBNL) program, a 5-year interdisciplinary project at the University of Massachusetts Dartmouth School of Marine Science and Technology (SMAST). While phytoplankton and zooplankton communities in the water column have been fairly well monitored, although not at a molecular level, microbial communities remain relatively uncharacterized. We report the first comprehensive (eukaryal, bacterial, archaeal) data from small-subunit ribosomal RNA gene clone libraries for samples collected near the thermal plume and underlying sediments of the Brayton Point Power Plant. We have partial sequences of nearly 4,000 clones from 2 different sites and have further sequenced 1,000 unique clones from these to full-length. Not surprising, our findings reveal a highly diverse consortium of the three domains including relatives of sludge bacteria, polyaromatic hydrocarbon-degrading bacteria, and representatives related to the genera *Staphylococcus*, *Streptococcus*, and *Clostridium*. Phylogenetic analyses will further unveil the relationships of many of these clones and determine whether they are related to known pathogens and may possibly represent undescribed taxa.
It is clear that even limited knowledge about the overall microbial community composition can lead to important observations about the ecosystem as a whole. Furthermore, understanding whether free-living pathogens participate in relationships with other members of the microbial community will be important in understanding their distributions and persistence.

Antje Boetius

Microbial diversity in marine sediments – what do we know about vertical, horizontal and temporal distribution patterns?

Antje Boetius and Alban Ramette (MPI for Marine Microbiology, Microbial Habitat Group, Celsiusstr. 1, 28359 Bremen, Germany; aboetius@mpi-bremen.de; aramette@mpi-bremen.de)

This presentation will deal with the known, unknown and knowable factors influencing microbial diversity in the seabed. Investigations of animal diversity on the ocean floor have resulted in some general principles. For example, animal diversity is higher at bathyal compared to abyssal depths, at vents and seeps compared to surrounding regions, at habitats under medium disturbance compared to low and high disturbance, in oxygen rich compared to oxygen depleted habitats, in tropical compared to polar settings. We still do not know if any of these factors are relevant to explain variation of microbial diversity. Exploration of microbial diversity in the seabed so far has been mostly limited to comparisons of numbers of 16S rDNA sequences of various phylogenetic units, which is commonly used as "species richness" definition in microbial ecology, despite known problems with PCR biases (e.g. Bowman et al. 2003). This approach has at least helped to show that there are distinct differences in community composition between sites, and that certain microbial clusters appear endemic to certain habitats (e.g. to hydrate-containing sites (Inagaki et al. 2006); to hydrogen-rich hydrothermal vents (Reysenbach and Shock 2002)). A few works have also tackled the question of relative abundances of species using FISH or quantitative PCR to obtain in situ abundances of phylogenetic groups. Due to the considerable time and expertise needed to implement those methods, the amount of data available for statistical analyses has however been a limiting factor (e.g. Knittel et al. 2003; 2005; Bowman et al. 2005). Some of the first works in microbial ecology of the seabed showed a correlation of the abundance and activity of microbial communities with increasing water depth and decreasing flux of organic matter to the seabed (Deming and Baross 1993, Lochte 1992), or increasing sediment depth into the deep biosphere realm (Parkes et al. 1993), but it remains unknown if such trends are accompanied by significant changes in species richness or relative abundance of species. Biogeochemical investigations have established the image of the "electron tower" (the vertical sequence of electron donors and acceptors available for energy gain by microbes) as a main structuring factor of microbial metabolism. This concept has influenced most recent investigations of microbial diversity in the seabed, which concentrate on the distribution of functional groups in their potentially preferred sediment horizons (scales of cm to hundreds of m). On these scales, almost nothing is known about horizontal variation in microbial diversity, which may be caused by sediment mixing, small-scale variability introduced by animal burrows, seabed morphologies, etc. However, such investigations are important to test the applicability of ecological principles to microbial communities such as the taxa-area relationship (Horner-Devine et al. 2004). Furthermore, we do not have any knowledge about temporal variation of microbial diversity in the seabed, neither about seasonal patterns, nor succession of communities after disturbances. Using a combination of community fingerprinting techniques (e.g. Hewson et al. 2003, 2006) and multivariate gradient analyses, we show that it is now possible to identify the main factors at stake in very complex situations involving hundreds of genotypes, tens of environmental variables, along with vertical, horizontal and temporal information of the samples. Based on examples from our work on coastal sediments and from deep-sea ecosystems, which

correlate community composition with habitat variability, this presentation will discuss emerging ecological trends of microbial diversity in the seabed.

Henry Boumann

Biophysical properties of newly-discovered membrane lipids: Insights into the functioning of cell membranes of marine micro organisms;

HENRY A. BOUMANN, JAAP S. SINNINGE DAMSTÉ, STEFAN SCHOUTEN;
Royal Netherlands Institute for Sea Research (Royal NIOZ), Dept. Marine Biogeochemistry and Toxicology, P.O. Box 59 , 1790 AB Den Burg – Texel, The Netherlands;
boumann@nioz.nl

Biological membranes play an essential role in life by serving as selective barriers between aqueous milieus. The basic architecture of these biomembranes has been extensively studied in many living cells and is composed of a wide variety of lipids that are generally composed of a polar head group and two apolar tails. Recently, we have discovered two particularly important groups of marine microorganisms that synthesize extraordinary membrane lipids in nature: Anammox bacteria biosynthesize linearly concatenated cyclobutane lipid structures, and non-thermophilic crenarchaeota produce unique glycerol dibiphytanyl glycerol tetraether lipids containing a cyclohexane ring. In the current work, we aim to unravel the head group moieties of these lipids and to study the biophysical features of the intact lipids. To achieve this goal, we will apply several techniques, including liquid chromatography coupled to (tandem) mass spectrometry, surface plasmon spectroscopy and micropipette aspiration in membrane-like environments. Furthermore, computer simulations of the dynamics of the membranes will be performed. The results will shed light on the impact of the membrane lipids on the structure and functioning of the cells of these environmentally important organisms.

Peter Burkill

Peter Burkill: description of research interests

I am Prof of Ocean Science at the National Oceanography Centre, University of Southampton, UK. As Head of the George Deacon Division, I lead a team of 44 physicists, biologists, chemists and modellers who study the biogeochemical controls and fate of biological production in the pelagic and benthic open ocean (www.noc.soton.ac.uk/GDD). With respect to ICOMM, we have broad interests in microbial communities in GDD. These range from biophysical interactions at the mesoscale and microscale, through understanding microbial functional biodiversity to biodiversity of protists in the abyss (Todo et al 2005, Foraminifera flourish at the ocean's deepest point. *Science* 307: 689). In a recent study (Martin et al, 2005. Extreme spatial variability in marine picoplankton and its consequences for interpreting Eulerian time series. *Biology Letters doi* 10.1098/rsbl.2005.0316), we studied the distribution of bacteria, picophytoplankton and protists demonstrating community patches with 10^3 variance in concentration over km spatial scales. We use flow cytometry routinely to tease apart bacterial communities and to address "who is doing what?" We have gone some way to addressing how *Prochlorococcus* thrives in the oligotrophic ocean (Zubkov et al, 2003. High rate of uptake of organic nitrogen compounds by *Prochloroccus* cyanobacteria as a key to their dominance in oligotrophic oceanic waters. *Applied & Environmental Microbiology* 69: 1299-1304; Zubkov et al, 2004. Depth related amino acid uptake by Prochlorococcus cyanobacteria in the Southern Atlantic tropical gyre FEMS Microbiology Ecology 50 (3): 153-161) using shipboard flow sorting. Recently, we have got fascinated in the functional role of low and high DNA content bacteria in the ocean (Zubkov et al, 2006. Bacterioplankton of low and high DNA content in the suboxic waters of the Arabian Sea and the Gulf of Oman: abundance and amino acid uptake. *Aquatic Microbial Ecology* 43: 23-32; Mary et al, In Press. Bacterioplankton with low nucleic acid content: metabolically active and dominated by SAR11 along an Atlantic Meridional Transect. *Aquatic Microbial Ecology*). Future research will include these approaches linked with genomics, proteomics to develop better understanding of how microbes respond to their environment, and better cytometric approaches to quantifying microbes (Zubkov & Burkill In Press. Syringe pumped high-speed flow cytometry of oceanic phytoplankton. *Cytometry;* Zubkov & Burkill submitted. Flow cytometric enumeration of DNA-stained planktonic protists in contrasting regions of the Atlantic Ocean. *J Plankton Research*) as well as understanding the microbial biogeochemistry of the N and S Atlantic oligotrophic gyres using the Atlantic Meriodional Transect Programme.

D.A. Caron

IcoMM Abstract – D.A. Caron Research Program

Our work is focused on the diversity, trophic interactions, population dynamics, physiological adaptations and biogeochemical significance of protistan assemblages (microalgae and protozoa) in aquatic ecosystems. Areas of on-going research:

-Field and laboratory studies of the protistan communities of sea-ice, water column and benthic habitats in the Ross Sea, Antarctica: Molecular biological characterizations of microbial eukaryote diversity (small subunit ribosomal RNA gene sequences obtained from environmental samples) have been conducted, and complemented with the establishment of a protistan culture collection consisting of a wide variety of taxa. Trophic studies have examined herbivory and bacterivory by microzooplankton (largely protistan) during austral spring, summer and fall in the Ross Sea. Physiological studies have employed cultures of Antarctic taxa to study the effect of low environmental temperature on growth rate, growth efficiency and nutrient remineralization of protistan species.

-Broad geographical surveys of microbial eukaryotic diversity: Genetic studies of microbial eukaryotes in environmental samples (small subunit ribosomal RNA gene sequences, and fragment analyses of rRNA genes) have been conducted on samples collected from various locations/depths in the world ocean in order to characterize the diversity of natural protistan assemblages, and to examine the distribution of these species on a global scale. Overall estimates of species diversity have been high, with large numbers of relatively rare phylotypes present at any given locale.

-Studies of the ecology of harmful algae: Field-based and laboratory-based observational and experimental studies have been conducted to understand and predict blooms of specific harmful and nuisance algae in U.S. coastal waters. Observational work has documented the spatial and temporal extent of bloom events, and correlated physical/chemical parameters (e.g. nutrient concentrations) with abundances of specific phytoplankton taxa and/or toxin concentrations. Experimental (meosocosm) approaches have focused on factors that might cause (and prevent) the development of these blooms.

-Development and application of sensor networks for conducting novel observations in freshwater and marine ecosystems: A prototype wireless network consisting of static sensor nodes (buoys) with basic chemical/physical sensors (e.g. temperature, chlorophyll) and an autonomous robotic boat equipped for sensing and sampling have been constructed and deployed in freshwater ecosystems for making novel measurements of plankton dynamics. This system has been employed to document small-scale spatial and fine-scale temporal changes in phytoplankton abundance, and to follow the migratory behavior of phytoplankton assemblages in-situ.

Guiseppe D'Auria

Searching for a marine bacteria biogeography. The ribosomal operon ITS as a tool

Giuseppe D'Auria and Francisco Rodriguez-Valera

The Intergenic transcribed spacers (ITS) between 16S and 23S genes are highly variable within short evolutionary time. This property increases enormously the discrimination power to detect clusters below the species clade. It can also be amplified by PCR using conserved primers in the 16S and 23S rRNA genes what facilitates enormously the task to accumulate large amounts of sequences for this region. One of the problems to establish biogeographic patterns in bacteria has been precisely the lack of resolution of the 16S rRNA sequence to identify clonal lineages due to its slow variation in spite of the large numbers of sequences available in databases. This work aims to identify clusters of identical (100%) or nearly identical (99%) ITS sequences from several marine environments. We have carried out the sequencing of a large set of amplicons containing this region and gathered all the sequences available in public databeses. This way significant numbers of sequences from the Antarctic and Arctic Oceans, Pacific off-shore California, Sargasso Sea, West and East Mediterranean Sea and North Pacific Ocean were obtained and compared. Globally represented ITS sequences related to Archaea and Alphaproteobacteria were identified.

In Archaea clusters containing identical sequences could retrieved from the Antarctic, Arctic and Western Mediterranean Sea. Also some highly related clusters of Alphaproteobacteria were found in the Antarctic and Arctic Ocean, off-shore California. Apparent endemisms were also observed in all locations and phylogroups strictly associated to the sampling depth were also found.

Garcia-Martinez, J., and F. Rodriguez-Valera. 2000. Microdiversity of uncultured marine prokaryotes: the SAR11 cluster and the marine Archaea of Group I. Mol Ecol 9: 935-48.

Gurtler, V., and V. A. Stanisich. 1996. New approaches to typing and identification of bacteria using the 16S-23S rDNA spacer region. Microbiology 142 (Pt 1): 3-16.

Gurtler, V., and B. C. Mayall. 1999. rDNA spacer rearrangements and concerted evolution. Microbiology 145 (Pt 1): 2-3.

Jan de Leeuw
Royal Netherlands Institute for Sea Research (NIOZ)

Why ICoMM?

Recent discoveries such as the Anoxic Oxidation of Methane (AOM) through sulfate and AOM through nitrate and nitrite by consortia of Bacteria and Archaea, the massive existence of nitrifying chemoautotrophic marine Crenarchaeota and the anoxic oxidation of ammonia to dinitrogen by Anammox bacteria indicate that:
a) our understanding of biogeochemical cycling of elements such as C, N and S is still fragmentary and that
b) the biogeochemical cycles are fully dominated by microbial processes.

A more systematic approach to improve our understanding of microbial processes and their evolution determining the functioning of biogeochemical cycles now and in the past starts with a much better qualitative and quantitative knowledge of microbial diversity per sé. In other words we have to know who is there to begin with in order to find out who is doing what, where and when.
To qualitatively and quantitatively determine (marine) microbial biodiversity is a phenomenal task which can only be achieved by concerted and coordinated actions worldwide using highly advanced and well-documented sampling strategies, sequence and other analyses, bio-informatics and data management.

To improve our understanding of (functional) microbial diversity in the past and to improve our knowledge regarding the speed of the "molecular clock" we have to rely on the presence of species-specific lipids (biomarkers) in well-dated lacustrine and marine sediments since the most specific biochemicals like DNA, RNA or proteins are ice creams for bacteria and other microbes and are thus very rapidly degraded and mineralized during transport to the sediments and/or in the sediments, whereas lipids are very stable in that respect. Thus, a "lipidomics" approach to establish microbial palaeo-diversity and the evolution of biosynthetic pathways and changes thereof over time can be seen as another (very modest) feature of ICoMM to reconstruct biogeochemical cycles and thereby palaeoclimate changes through well-documented analyses of sediments and appropriate archiving of sedimentary species-specific lipid data.

Jan W. de Leeuw

Maria-Judith De Souza e Gonsalves

Abstract: My research work deals with habitats and fields of study on the biogeochemical roles of bacterioplankton, and adds to our knowledge of ecosystem dynamics throughout the aquatic realm, from freshwaters of the Antarctic, to the coastal and open Indian ocean. A multi-scale level investigation of the diversity and status of bacterial adaptations in natural and extreme environments, their biochemistry, functional aspects of their interaction with their environment and products of biotechnological applications have given new insights to marine microbiology.

The work on Antarctic bacteria describes how under supersaturation of oxygen these bacteria shutdown their activity and express it once a reductant is added to their milieu. It also investigated the physiological adaptations in bacteria to extreme conditions. It suggests that anaerobiosis outcompete aerobiosis in the Antarctic lacustrine bacteria. It was a landmark finding that a great percentage of Antarctic bacteria are viable and therefore could participate in various activities.

My work in the estuarine ecosystem on the dynamics of particle-associated bacteria (PAB) on Mandovi - Zuari estuary delineates the role of PAB in tropical waters where there is a paucity of information. It is the first of its kind in India, which stresses the importance of PAB in the interconversion of particulate organic matter to dissolved organic matter and vise versa and their role in the food web. Moreover, it also deals with the effect of different environmental parameters on their production and activity. Interestingly, this study suggests that estuarine bacteria are nutritionally less resilient than off shore forms.
Oil degradation in the marine environment is often hampered by the non-availability of phosphorus. My research on phosphate solubilizing bacteria (PSB) from general oceanic regions has shown how these organisms could be useful for bioremediation in phosphate starved environments that can be used as a novel and natural strategy for future bioremediation of oil spills
Recently I participated in the deep drilling cruise and will now be investigating the different bacterial communities present in the deep sediments (300m).

Virginia Edgcomb

An Intensive Study of Eukaryotic Microbial Diversity Across the Oxic/Anoxic Interface of the Cariaco Basin, Venezuela

Virginia Edgcomb[1], Sunok Jeon[2], Vladimir Bondarev[2], Slava Epstein[2]

[1]Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA 02543 USA
[2]Department of Biology, Northeastern University, Boston, MA 02115 USA

We have conducted an in-depth 18S rDNA sequence analysis of eukaryotic microbial diversity in the water column of the Cariaco Basin, Venezuela. The Cariaco Basin is the world's largest permanently-anoxic marine environment. To date we have surveyed over 10000 clones obtained from 3 different stations within the basin at four discrete water depths corresponding to positions within, above, and below the oxycline, and from 900m depth (significantly below the oxycline). As such, the near completion of this phase of the project represents the most comprehensive survey of this type of marine environment to date. The phylogenetic analyses of the surveyed community are in progress, but preliminary results reveal surprising insights into the extent of novel protistan lineages at all taxonomic levels. One goal is to obtain complete, or nearly complete coverage of protistan diversity in at least selected samples. Our aim is to obtain experimentally determined – as opposed to statistically calculated – figures of protistan local species richness, and use these as benchmarks to compare the performance of several competing statistical approaches to estimating microbial diversity. The second phase of the project will begin this fall when we will use the information we have gathered to direct the design of fluorescent in situ hybridization (FISH) probes to target particular groups/OTUs of interest for studies that combine FISH and quality SEM ultrastructure examination. FISH studies will also allow us to assess the ecological importance of selected recorded protistan types recovered in our molecular surveys, by quantitatively describing their dynamics over time and across environmental gradients. We will also make every attempt to enrich for and cultivate the above species for physiological, biochemical, and morphological investigation.

Katrina J. Edwards

Katrina J. Edwards, Cara M. Santelli, Erin Banning, Brandy Toner, Dan R. Rogers
Geomicrobiology Group
Department of Marine Chemistry & Geochemistry
Woods Hole Oceanographic Institution
katrina@whoi.edu

Geomicrobiology in oceanography: Mineral-microbe interactions in the deep sea

At mid ocean ridge (MOR) spreading centers and seamounts, the ocean crust is exposed to oxygenated seawater and microbes. On a global basis the MOR systems represents a ~600,000 km^2 continuous rock-colonizing ('epi- and endo-lithbiontic crusts') community of microbes, ranking it the largest rock-hosted microbial ecosystem on Earth. Fluid circulation at ridge flanks and seamounts may "inoculate" the subsurface, extending this ecosystem hundreds of meters into porous basement. In theory, microbes can propagate where water, nutrients, and energy are available, and temperatures permit. However, few studies have empirically examined the nature of deep-sea hock hosted microbial ecosystems in any detail. Consequently, fundamental questions remain to be addressed, such as, what is the abundance and activity of these deep-sea microbial communities? What 'types' of microbes – from a phylogenetic and physiological standpoint – comprise deep-sea rock-hosted communities? And importantly, what role, if any, does this biosphere play in important geochemical processes such as rock weathering and elemental exchange between ocean and crust?

To answer these questions, we are studying deep-sea microbial communities at the Lo'ihi seamount, Hawaii, and at the East Pacific Rise 9°N (a RIDGE 2000 Integrated Study Site) on a long-term, time-series basis. A multi-disciplinary approach for studying microbial communities colonizing rocks and minerals at these sites will be discussed. Our work applies microbiological and geochemical methods for 'full-circle, spatially-resolved geomicrobiology': a suite of integrative, iterative approaches designed to link specific microbes and microbial communities with biogeochemical processes. The microbiological techniques involved include culturing, physiological studies, and both 16S rDNA and genomic sequencing to determine the composition and function of these communities. In-situ experiments with rock substrates, in tandem with examination and analysis of environmental samples, are being assessed using 16S rDNA phylogenetic approaches, fluorescence and electron microscopy, and synchrotron-based mineralogical and chemical methods. Our specific goals are to establish definitive roles for microbes in ocean crust weathering in the deep sea; however, in principal these types of approaches can be applied in an integrative fashion to study mineral-microbe interactions in a variety of Earth environments.

Slava Epstein

Northeastern University, Boston, MA U.S.A.

Research in S.Epstein's lab falls into two categories: culture-dependent and culture-independent. The first focuses on the "Great Plate Count Anomaly" – and reasons why so many (prokaryotic) microbes refuse to grow in the lab. Synergistic projects aim at design alternative cultivation approaches, ways to domesticate uncultivated species, and explore their biotechnological potential. The second group of projects is more relevant to the goals of the ICoMM meeting. This (culture-independent) direction is developing in three venues:

1. Nature of microbial species. Two projects in the lab aim at determining the level of genetic variations within classical taxonomic units vs among such units, looking at reconciliation of taxonomic schemes produced by alpha-taxonomists and molecular phylogenetists. The focus is on microbial eukaryotes. At a minimum, we would like to find a practical measure of rRNA gene divergence that separates organisms into morphologically and ecologically meaningful clusters. At this time, 1% divergence seems to be a good candidate for a species cut-off value at least in marine ciliates. A separate project studies ecological differences and similarities between strains identical – or nearly identical – in their conserved genes' sequences.

2. Microbial discovery. Three projects in the lab, mostly in collaboration with other scientists (V.Edgcomb, D.Patterson, G.Taylor, T.Stoeck, M.Yakimov, L.Giuliano), aim at surveying, discovering, describing, and eventually cultivating representatives of novel lineages of microbial eukaryotes. The largest of the three is represented by a Microbial Observatory in the Cariaco Basin off the coast of Venezuela. The other two focus on extreme environments of Arctic (Greenland) and high salinity/deep sea (Mediterranean). Not surprisingly, each of the three has uncovered an enormous diversity and richness of novel protistan forms, but going beyond a simple detection of novel sequences proves challenging.

3. Patterns of diversity. Two projects, both in collaboration with John Bunge, focus on local and global diversity of both pro- and eukaryotes. The long-term goals are a) to understand if microbial communities are indeed composed of thousands of interacting species – or most of them are essentially inconsequential in the given community, and b) if there are meaningful patterns in global microbial distribution. The immediate objectives are more modest: to develop statistically valid tools to measure and predict microbial diversity based on small samples of this diversity (as all our clone libraries are). Parametric distributions appear to be good candidates for analyses of local diversity. Global diversity and its patterns are more difficult to handle; these seem to require completely new statistical approaches (under development).

Isabel Ferrera

Diversity and distribution of *Aquificales* from deep-sea vents in the Lau Basin

Ferrera I,[1] Banta AB,[1] Webb J,[1] Kelly SM,[1] Kirshtein JD,[2] Voytek MA,[2] Reysenbach AL[1]

[1]Portland State University, Biology Department, 1719 SW 10th Ave, Portland, OR, 97201

[2]US Geological Survey, Water Resources Division, 12201 Sunrise Valley Dr., Reston, VA 20192

The distribution and diversity of the thermophilic bacterial lineage, the *Aquificales*, was studied along a newly explored deep-sea hydrothermal system (the Valu Fa Ridge and Eastern Lau Spreading Center, ELSC) and compared with *Aquificales* diversity from other mid-ocean ridge systems. Using a combination of molecular approaches based on the 16S rRNA gene (DGGE and cloning), and enrichment culture and isolation, we show that the diversity of the *Aquificales* at the ELSC is much higher than any other deep-sea vent site studied. Abundance of this group was quantified using QPCR. Using multidimensional scaling (MDS) and analysis of similarity (ANOSIM) measures for exploring patterns in diversity we show that there is a significant difference between *Aquificales* diversity between sites on the ELSC. Additionally, the *Aquificales* genera showed patterns of co-occurrence and were not distributed randomly. Furthermore, unlike most other deep-sea vent sites where the most frequently obtained culture and environmental *Aquificales* sequence is from strains of *Persephonella marina*, this species was not detected at Lau, but instead, *P. hydrogeniphila*, *Desufurobacterium spp.*,*Thermovibrio spp.* and newly described *Hydrogenivirga spp.* were detected. Additionally, clone libraries showed a greater microheterogeneity within the *Persephonella* sequences than previously observed. These results demonstrate the importance of *Aquificales* in deep-sea hydrothermal ecosystems, but also provide good models for exploring global patterns of microbial diversity and biogeography.

Victor Ariel Gallardo

INITIAL PROGRESS ON THE NEWLY DISCOVERED BENTHIC FILAMENTOUS BACTERIA OF THE HUMBOLDT CURRENT SYSTEM

Carola Espinoza,[1] Jeppe L. Nielsen[2] and V. A. Gallardo[1,3]

[1] **Center for Oceanographic Research in the Eastern South Pacific (COPAS), Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepcion, Concepción, Chile.**
[2] **Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Denmark.**
[3] **Departamento de Oceanografía, Universidad de Concepción, Chile.**

ABSTRACT

The shelf sediments under the Humboldt Current System have provided new opportunities to study big marine bacteria which were supposed to be scarce in the world's ocean. Since the discovery of a diverse assemblage of big filamentous bacteria off Concepción, central Chile and their confirmation off northern Chile, Perú and Costa Rica, investigations which were more of the field survey type and based exclusively on morphology through phase contrast microscopy and photography, have moved into the molecular realm. Present efforts on the most conspicuous of the big filamentous bacteria, the so-called 'anaconda', a pigmented, non-vacuolated filamentous organism, extracted from field samples by means of micromanipulation, have included so far their 16S rRNA genes subjected to PCR amplification and screening by DGGE. The ecophysiology of the organism is being studied by micro-auto-radiography (MAR) in combination with FISH (fluorescence *in situ* hybridization).

Gallardo Talk

BIG FILAMENTOUS BACTERIA ARE MORE COMMON THAN PREVIOUSLY THOUGHT. NOTES ON THE FINDINGS IN SHELF SEDIMENTS THE OF THE EASTERN SOUTH PACIFIC AFTER A LONG NON-EL NIÑO PERIOD

Victor Ariel Gallardo[1, 2], Carola Espinoza[2], and Jeppe Lund Nielsen[3]

[1]Department of Oceanography; [2]Center for Oceanographic Research in the Eastern South Pacific (COPAS), Universidad de Concepción, Concepción, Chile; [3]Department of Biotechnology, Chemistry and Environmental Engineering, University of Aalborg, Aalborg, Denmark.

An unusually long non-El Niño period after the largest of the last century (the El Niño 1997-98), has provided the conditions in the shelf sediments under the oxygen minimum zone off central Chile, for the development of an abundant and diverse community of giant filamentous bacteria, among other microbial organisms. Until now we have been recording the succession of the most conspicuous morphs in this community and surveying their occurrence in the eastern Pacific i.e., Golfo Dulce, Costa Rica; Callao and Independence Bay, Peru; Iquique and Mejillones, northern Chile. Samples from sediments under salmon culture pens of southern Chile have also been analyzed. The method throughout has been phase contrast microscopy and microphotography. Investigations have now moved into the molecular realm and are beginning to show results at least from one of the 'flag ship' species of this new assemblage, the so-called 'anaconda'. In this talk emphasis is placed on the need to support these studies on the larger and most conspicuous end of the microbial biosphere continuum.

Abstract for a talk at the Annual Meeting of the International Census of Marine Microbes (ICoMM), June 12-15, 2006, NH Leewenhorst, Noordwijkerhout, The Netherlands.

Stephen Giovannoni

Domesticating Bacterioplankton: New Methods Produce a Surfeit of Cultured Oligotrophs

Stephen J. Giovannoni, Ulrich Stingl and Joshua Kitner
Oregon State University, Corvallis OR, USA 97331

Recently introduced high throughput culturing (HTC) procedures have led to a dramatic increase in the number of important bacterioplankton groups that are being cultivated and studied in a laboratory setting. This approach is based on Button's method for isolating cells by extinction culturing in natural seawater. Approximately 2500 cultures of pelagic marine have been isolated by this approach. Up to 14% of cells from coastal seawater were cultured using this method, a number that is 1400 to 140-fold higher than obtained by traditional microbiological culturing techniques. Among the cultured organisms are many unique cell lineages that have been named as new phyla, families, and genera. Ninety percent of the cells recovered by the HTC project do not replicate in Petri dishes of agar media. A majority of the isolates obtained are obligate oligotrophs that display logistic growth curves in seawater. These strains are being used to study microbial metabolic processes at natural substrate concentrations and cell densities, which are typically about three orders of magnitude less than in common laboratory media. Twenty-seven genomes from HTC isolates are now sequenced or in ques for sequencing. It is apparent that some abundant bacterioplankton groups do not replicate in the current generation of HTC media, and will require further innovation before isolation can be achieved.

Frank Oliver Gloeckner

Megx.net - database resources for marine ecological genomics

Thierry Lombardot[a], Renzo Kottmann[a], Michael Richter[a], Hanno Teeling[a], Christian Quast[a] and Frank Oliver Glöckner[a, b],

[a]Microbial Genomics Group - Max Planck Institute for Marine Microbiology (D-28359 Bremen, Germany) and [b]International University Bremen (D-28759 Bremen, Germany)

Marine microbial genomics and metagenomics is an emerging field in environmental research. Since the completion of the first marine bacterial genome in 2003, the number of fully sequenced marine bacteria has grown rapidly. Concurrently, marine metagenomics studies are performed on a regular basis, and the resulting number of sequences is growing exponentially. To address environmentally relevant questions like adaptations of organisms to oceanic provinces and regional differences in the microbial cycling of nutrients, it is necessary to couple microbial sequence data with geographical information and supplement them with contextual information like physical, chemical and biological data. Therefore, new specialized databases are needed to organize and standardize data storage as well as centralize data access and interpretation. We have implemented Megx.net, a set of databases and tools that handle genomic and metagenomic sequences in their environmental contexts.

Megx.net includes: i) the Genomes Mapserver - a geographic information system (GIS) to systematically store and analyse marine (meta)genomic data in conjunction with contextual information (metadata); ii) an environmental genome browser with fast search functionalities; iii) a database with precomputed analyses for selected complete genomes and iv) a database and tool to classify metagenomic fragments based on oligonucleotide signatures.

The **Genomes Mapserver** is furthermore part of the EU-project **MetaFunctions** which has two main objectives:

1. to visualize the geographic distribution of genes from marine prokaryotes including geographic browsing on a world map and specific search functionalities (Geographic-BLAST)
2. to assist in assigning potential functions to hypothetical genes by the integration of genomic data with metadata about the environment.

The Genomes Mapserver contains currently 20 sampling sites of completely sequenced marine genomes and 30 samples for marine metagenome studies with around 58,000 and 1.2 million genes, respectively.

Megx.net and the Genomes Mapserver are publicly accessible at http://www.megx.net and www.metafunctions.org.

Reference: Lombardot, T., R. Kottmann, H. Pfeffer, M. Richter, H. Teeling, C. Quast, and F. O. Glöckner. 2006. Megx.net - database resource for marine ecological genomics. Nucleic Acid Res. 34:D390-D393.

Martha Liliana Gómez

SELECTION AND APPLICATION OF MARINE BACTERIA WITH DEGRADING ABILITY OF PERSISTENT ORGANIC COMPOUNDS (POC) IN THE PACIFIC AND COLOMBIAN CARIBBEAN [1]

Martha Liliana Gómez and Jenny Dussán

Instituto de Investigaciones Marinas y Costeras-INVEMAR. Punta Betin. Santa Marta. Colombia. mlgomez@invemar.org.co jdussan@uniandes.edu.co

ABSTRACT

The incidence of persistent organic compounds (POC) such as chlorinated pesticides and hydrocarbons on the terrestrial surface and marine water of Colombia, are originated by the agricultural, industrial and marine activities; it generates economical losses and damaging environmental impact. This perspective has required the production of new studies in order to obtain the knowledge about POC, subsequently for take up procedures necessaries to estimate the pollution on the environment and then to mitigate or to eliminate the effect of the polluting agents. An alternative for such aim is in the use of biorremediación processes by means of the use of strains of native bacteria with capacity to degrade POC.

In Colombia some researches have been made in the field of the biodegradation and bioremediation, have been modest in the marine area, by this INVEMAR with the financial support of COLCIENCIAS, within the framework of the project "Selection and application of native marine bacteria with degrading ability of persistent organic compounds (POC) in the Pacific and the Colombian Caribbean" made the recovering of native bacteria strains of marine and estuarine Colombian environments with POC - degrading ability to make the first collection (58 strains) registered at the Museum of Marine Natural History of Colombia (MHNMC) (Record N°. 082 IavH), these represent a great potential for the future to prepare microbial pooles with the purpose of use them in bioremediation.

Was possible the isolation and identification of 136 native morphotypes obtained in Caribbean and Colombian Pacific sediments. These strains were put under tests of selective pressure to different concentrations of POC, 22 from Caribbean and 9 from Pacific were chosen by high tolerance of DDT and Aldrin between 1,600 ng/L and 2,000 ng/L. For hydrocarbons, 9 strains from Caribbean and 9 from Pacific grew between 1 to 8 % v/v of ACPM. Based on the individual abilities from these bacteria, were produced mixed bacterial cultures from Pacific and Colombian Caribbean to estimate their abilities of degrading POC. The results showed a degrading percentage of 68,8 % for the mixed bacteria culture from Caribbean and 68 % from Pacific in presence of DDT (1600 ng/L) for 60 days. In test with hydrocarbons (2% v/v ACPM) there was 68,61 % degradation of n-alkane in 21 days for the mixed bacteria culture from Caribbean, and 54,8 % from Pacific.

24 bacterial strains from mixed cultures were identified by 16S ribosomal RNA gene sequencing amplified by PCR, where *Bacillus* and *Pseudomonas* were predominant. The results showed a great bacterial diversity in these areas, and 3 new strains never cultured (CCBM132, CCBM133 and CCBM145).

---

[1] Part of the project with COLCIENCIAS N° 2105-09-13524, presented as closing report. Researchers: Gómez, M., Hurtado, C., Casanova, R., Lozano, J., Campos, N., Marín, B., Narváez, S., Ruíz, R., Reyes, V. y Cortés G. (INVEMAR-Colombia). Dussán, J. (Universidad de los Andes). Nuñez, R. (CEBIMAR, -Cuba).

Gerhard Herndl

Integrating prokaryotic microdiversity into ecosystems function theory

Gerhard J. Herndl, Dept. of Biological Oceanography, Royal Netherlands Institute for Sea Research (NIOZ), 1790 AB Den Burg, The Netherlands, email: Herndl@nioz.nl

Sufficient information on the phylogenetic diversity of marine prokaryotes have been collected over the past 15 years to allow assessment whether ecological theories on diversity and ecosystem functioning largely established on findings in plant communities also hold for marine prokaryotic communities. From the plethora of ecosystem functions of prokaryotes, examples are given in this presentation from two main functions, the remineralization of dissolved organic carbon to CO_2 and the production of methane by anaerobic bacteria. Both functions have been measured over a wide range of different marine subsystems. The responsible bacterial consortia are well defined in terms of electron acceptors and donors used and moreover, exhibit considerable richness. Thus, these two major prokaryotic consortia are ideal to exemplify and discuss functional versus phylogenetic variability in marine bacterial communities. As evident from these two examples, there is substantial complementarity among prokaryotic species, i.e., species with overlapping ecological niches. Species-rich communities are therefore more productive because more of the overall resource is used. Species-rich communities are also functionally more stable since niche overlap leads to parallel processing of a given substrate and synergistic effects within a network of species. Thus, this 'complementary mechanism' seems to be more important than the 'selection mechanism', i.e., species-rich communities are more productive because they are more likely to contain individual species with a large effect on ecosystem functioning. The recently emerging view that there is substantial microdiversity beyond the 97% sequence similarity level in marine prokaryotes has not been incorporated into general ecosystems theory yet. In this presentation an attempt is made to highlight some approaches to accomplish this. Based on recent major advances using a novel pyro-sequencing approach it becomes obvious that 1) the 'rare but closely related genotypes' are never attaining substantial abundance. Thus, Beijerinck's 'everything is everywhere' postulate probably does not hold for this rare biosphere. 2) Next to the few highly abundant genotypes present in all the water masses, several very closely related genotypes are present, decreasing in abundance with decreasing sequence similarity, leading to a bell-shaped distribution pattern of genotypes around the dominant one. 3) The rare genotypes all exhibit a certain level of sequence dissimilarity to the abundant genotypes, however, the sheer richness of the rare genotypes is enormous.
This 'rare but closely related genotypes' are probably originating from divergence from the relatively few but abundant genotypes present. The genotypes most closely related to the parent genotype reach abundances almost as high as their parent genotype. Therefore, they likely contribute to parallel substrate processing and hence, to the complementarity network of prokaryotic activity and therefore, add to stability in ecosystem function of prokaryotic plankton.

Julie A. Huber

Microbial ecology of subseafloor communities at deep-sea hydrothermal seamounts

Josephine Bay Paul Center, Marine Biological Laboratory, Woods Hole, MA, USA 02543; E-mail: jhuber@mbl.edu; Telephone: 508-548-3705, x6616; Fax: 508-457-4727

Circulation of hydrothermal fluids and seawater occurs within the upper 500 m of porous oceanic crust and provides a rich environment for microbial growth in the subseafloor. Enrichment cultures, geochemical indicators, and sequence analyses of PCR amplicons of ribosomal RNA genes demonstrate that these crustal fluids host a microbial community composed of organisms indigenous to the subseafloor, including anaerobic thermophiles and mesophilic sulfur oxidizers, and organisms from other deep-sea habits, such as seawater. However, the subseafloor microbial communities remain undersampled and our knowledge of what microbes are present and how they are distributed in this dynamic geochemical environment over time and space is fragmentary.

This work focuses on determining the microbial diversity and genomic content of the subseafloor microbial community at geographically and geochemically distinct deep-sea hydrothermal seamounts. The approach uses a combination of methods, including the application of 454 tag sequencing, metagenomics, culturing, and geochemical methods to diffuse fluids from deep-sea hydrothermal seamounts. Diffuse vent fluid samples are being collected from three locations: Axial Seamount, an active submarine volcano on the Juan de Fuca Ridge in the Northeast Pacific Ocean (45.92° N, 130° W); seamounts along the Mariana Arc (14-22° N, 143-146° E) in the Western Pacific; and Loihi Seamount (18.92° N 155.27° W), located 30km southeast of the big island of Hawaii. All three locations, Axial, Loihi, and the Mariana Arc, are recently eruptive seamounts located above 2000 m and host diffusely venting fluids with high concentrations of carbon dioxide. However, their geological and chemical setting differs greatly; Axial is a mid-ocean ridge seamount with fluids dominated by high concentrations of hydrogen sulfide, Loihi is a mid-plate hotspot seamount with extremely high concentrations of dissolved iron (FeII), and the Mariana seamounts are at a convergent plate boundary and host a variety of fluids, including those with very low pH and high concentrations of particulate sulfur.

Questions to be addressed include, what is the distribution and relative abundance of microbial lineages in the subseafloor environment? Is there a core set of lineages or genes specific to the subseafloor environment? How do genes and lineages compare in different environments and are there unique lineages or genes at individual sites? If so, how are they linked to distinct parameters at that site? Finally, what is the extent of horizontal gene flow in the subseafloor and how does this gene flow shape microbial lineages and their adaptation to this unique environment? Pilot studies at Axial Seamount have begun with the construction of fosmid libraries and 454 sequencing of PCR amplicons that span the V6-hypervariable region of ribosomal RNAs from two diffuse flow vents. The collection of this suite of samples and application of a combination of methods will allow us for the first time to look at the distribution and abundance of subseafloor microbial communities at geographically and geochemically distinct deep-sea hydrothermal seamounts across the Pacific Ocean.

Elena Ivars-Martinez

Multilocus sequence typing of the marine copiotrophic bacterium *Alteromonas macleodii*

Ivars-Martínez, E.; D´Auria, G. and Rodriguez-Valera, F.

A. macleodii is a common marine bacterium found in temperate latitudes mostly as part of the aggregate associated bacterioplankton. We have gathered a collection of 23 isolates from different geographic locations, Black Sea, Eastern and Western Mediterranean, the English Chanel, Pacific Ocean (off-shore Hawaii) , Andaman Sea (Thailand) and depths. Primers were designed for conserved regions of eight housekeeping genes (Pmg, PntA, gyr B, rpoB, met G, gly A, Dna k, Suc C) PCR amplified and and sequenced for all 23 strains. We formerly detected the presence of two ecotypes, one restricted to the deep Mediterranean while the other was found in surface temperate waters around the world. The new data show that also isolates from Black sea could form another ecotype presumably specialized in the brackish waters typical of this water body. On the other hand we have now evidence showing the presence of the deep ecotype in a sample from the English Channel albeit a surface one. We have found before deep ecotype representatives in up-welling areas in the Mediterranean, but its presence in the North Atlantic shows that this organism could be much more widespread than we originally thought. In any case, we have evidence of recombination events involving all three ecotypes showing that they are not genetically isolated.

Alexandra Kraberg

Current Research Interests:

The main theme of my research at the moment is related to intraspecific diversity in different components of the phytoplankton and microzooplankton and most importantly the possible ecological implications of such intraspecific diversity. Within this framework I am involved in/ supervising several projects:

1. the intraspecific morphological/ genetic variation in *Thalassiosira rotula*. Such variation can hamper taxonomic studies per se, as well as assessments of the biodiversity of aquatic communities. This work is carried out with many different clones from culture collections and with seasonal clones isolated at the Helgoland Roads long term monitoring site. Genetic, physiological and morphological studies are carried out in the context of a large ongoing foodweb study at the Biological Station Helgoland.
2. The biology of heterotrophic dinoflagellates, particularly their role as consumers in marine food webs and again the impact of intraspecific diversity on these assessments. The possible impacts of heterotrophic dinoflagellates as consumers along with other microzooplankton and metazooplankton have hitherto only been very poorly studied. Reasons for this might be the difficulties associated with culturing these species. We have therefore recently established a small culture collection for heterotrophic dinoflagellates, and are supervising several students who are carrying grazing experiments, morphological studies etc. in species such as *Protoperidinium bipes*. The backdrop for these studies is again the Helgoland food web project and I am particularly interested in the extent to which heterotrophic dinos can influence/ control spring bloom succession. The cultures will also be used for studies of intraspecific morphological variability in the cultured dinoflagellates and for studies investigating the ingestion process in selected dinoflagellates.

I remain interested in the development of tools for the presentation, analysis and dissemination of data online. With PLANKTON*NET (http://www.awi.de/PlanktonNet) we have so far restricted ourselves to the presentation of taxonomic data but pending funding proposals would attempt to link this to the analysis of numerical data from various European long-term data series

Debbie Lindell

Transfer of genes between marine cyanobacteria and phages and their expression in recipient genomes.

Debbie Lindell, Matthew B. Sullivan, Jessica A. Lee, Luke R. Thompson, Claudia Steglich, Andrew C. Tolonen, Erik Zinser, Sallie W. Chisholm

Massachusetts Institute of Technology, Cambridge, MA 02139, USA

The transfer of genetic material between organisms is an important mechanism in evolution leading to both genetic and functional diversification. Microbial genome sequencing over the past decade has implicated phages (viruses that infect bacteria) as one of the key agents mediating genetic exchange. One of the more surprising recent findings is the presence of photosynthesis genes, *psbA* and *psbD*, encoding the two core photosystem II reaction center proteins, in the genomes of phages that infect marine cyanobacteria. During a survey of over thirty cultured cyanophages in our collection we found that 88% of the cyanophage genomes contain *psbA* and 50% contain both *psbA* and *psbD*. Phylogenetic clustering of these genes suggest that they have been transferred from host to phage in a discrete number of events over the course of evolution (4 for *psbA* and 2 for *psbD*) followed by horizontal and vertical transfers between cyanophages. Analysis of *psbA* and *psbD* sequences from the viral fraction of seawater revealed significant sequence diversity, much of which is represented in cultured cyanophage.

Using *Prochlorococcus* MED4 and the T7-like podovirus P-SSP7 as a model host-phage system we found that the *psbA* gene is expressed during infection. Other bacterial-like genes that have been acquired by this phage are also expressed. These include *hli* – encoding a cyanobacterial stress response protein, *nrd* – encoding ribonucleotide reductase and *tal* – encoding transaldolase. We hypothesize that the expression of these genes enhances phage fitness by helping the phage obtain the energy and/or nucleic acid substrates necessary for maximal phage genome replication. In addition to phage acquired host genes, we found evidence for the transfer of numerous *hli* stress response genes back to *Prochlorococcus* after a period of evolution in the phage. The resulting expansion of this gene family in high-light adapted *Prochlorococcus* strains may contribute to niche expansion among *Prochlorococcus* in the high-light, low nutrient waters of the open oceans. Indeed these acquired *hli* stress response genes have undergone specialization in *Prochlorococcus* and are differentially expressed in response to light and nitrogen stress, as well as over a diel cycle. These findings strongly suggest that the acquired genes have attained functional importance in the recipient genomes and that the swapping of genes between hosts and phages has likely played a role in the development of genetic and physiological variants within *Prochlorococcus* and the phages that infect them.

Jose V. Lopez

Marine Sponge Hosts Are Oases of Microbial Diversity

Jose V. Lopez , Dedra Harmody, Cheryl L. Peterson , Angela Ledger, Karen Sfanos, Shirley A. Pomponi, Peter J. McCarthy
Harbor Branch Oceanographic Institution, Ft. Pierce Florida USA

For the past several years, our laboratory in the Division of Biomedical Marine Research (DBMR) at Harbor Branch Oceanographic Institution has been conducting a biotic survey and inventory of cultured eubacterial and fungal isolates and the culture-independent microbial consortia of several different marine invertebrate species (mostly demosponges) from various oceanic regions using molecular genetics techniques (16S rRNA). The primary goal was to establish a database and characterize patterns of specific microorganismal associations (or "symbioses") within these unique microcosms. The collection is distinguished by a high proportion of marine invertebrate samples (mostly Porifera), collected from >150 fsw (feet seawater) down to 3000 fsw using the Johnson Sea-Link submersibles. Overall, invertebrate hosts varied widely according to taxonomy (>40 invertebrate taxa, 26 sponge families), ecology, depth and geographic location. A major conclusion is that marine sponges offer an oasis for microbial biodiversity rather than biomass at various depths.

In the first phase of the inventory, the DNA fingerprinting technique of amplified rDNA restriction analysis (ARDRA) and DNA sequencing of approximately 750–800 base pairs (bp) encompassing hypervariable regions in the 5' portion of the small subunit (SSU) 16S rRNA gene was applied to a subset (~15%) the total Harbor Branch Oceanographic Marine Microbial Culture Collection (HBMMCC). The collection consists of approximately 17,000 microbial isolates, with 11,000 from a depth of greater than 150 fsw, making the HBMMCC a unique, viable collection for microbiological diversity. More than 2273 heterotrophic bacterial isolates were inventoried for 16S SSU taxanomy, showing that this subset of the HBMMCC contains at least 224 different phylotypes from six major bacterial clades (Proteobacteria (Alpha, Beta, Gamma), Cytophaga, Flavobacteria, and Bacteroides (CFB), Gram+high GC content, Gram+ low GC content). Notably, 11 phylotypes were < 93% similar to the closest sequence match in the GenBank database even after sequencing the complete 16S rRNA gene, indicating the likely discovery of novel microbial taxa. Furthermore, previously reported ''uncultured'' microbes, such as sponge "symbionts", appear in the HBMMCC. Lastly, we are currently adding over 150 unique marine fungal 18S sequences to the inventory.

Since culturing techniques can still only retrieve 0.1 -1.0% of the microbial diversity found in nature, we applied a metagenomic approach to about 20 distinct culture- independent 16S rRNA clone libraries derived from 8-10 different sponge taxa. Duplicates of each host taxon were included to determine any possible species specific associations. For the culture-independent 16S rRNA dataset, the detection of *Spirochaetes*, *Planctomycetales*, *Cyanobacteria*, *Riftia*, *Arcobacter*, *Chloroflexi*, *Thermales* , *Nitrospira*, *Acidobacter*, *Deltaproteobacteria*, *Myxobacteria*, *Epsilonproteobacteria*, *Archaea* and numerous matches to "uncultured" lineages (SAR, JAWS series etc.) differed considerably from the cultured isolate collection. Many sponge – associated microbes also diverged from the surrounding environmental flora sampled from sediments and ambient seawater based on our sequencing or from the recent marine microbial literature. Rarefaction analyses of specific libraries indicated that this survey of sponge-derived uncultured microbial symbiont diversity was not exhaustive, and that many more uncharacterized microbial taxa will be harbored by untested sponge species. In both datasets, ubiquitous or "cosmopolitan" eubacterial taxa (*Beijerinckiaceae*, soil clone C083, green non-sulfur bacteria) occur among several of the host sponges. The results of this research are now available online as a searchable taxonomic database (www.hboi.edu/dbmr/dbmr_hbmmd.html).

Connie Lovejoy, PhD
Professeure adjointe
Dépt de Biologie, Université Laval, Québec QC, Canada
Email: connie.lovejoy@bio.ulaval.ca

Recent Background
Since obtaining a tenure track position in June 2004, I have been building a research team and a functional laboratory to study Arctic microbial diversity. In August and September 2005 we sampled up to six depths at 21 stations from the North Water Polynya (Canada and Greenland) to the Alaska US boarder. In collaboration with the Institut Ciences del Mar, my group has also been analyzing pelagic microbial DNA collected as part of the Canadian Arctic Shelf Exchange Study (CASES) an international 3 year field program that included a one year icebreaker deployment. We are in the process of constructing SS rRNA clone libraries for archaea, bacteria and small eukaryotes from different regions and seasons including the overwinter period.
Our sampling strategy includes collecting all ancillary CTD data as well as, nutrients, chl *a* (and limited HPLC pigments) concentrations, epifluorescence counts of eukaryotes, prokaryotes and viruses.

Research Focus:

The major objective of my research is to identify microbial species and ecotypes, to define factors that determine protist species dominance and associated protist and bacterial consortia under specific hydrological regimes, and to relate these findings to ecosystem function. Specifically, our aim is to define the relationships between physical conditions and microbial assemblages that differ in their biogeochemical and ecological influences towards carbon, energy and nutrient fluxes (sedimentation, versus mineralization and recycling, versus transfer to higher trophic levels). A major outcome to date is that we have confirmed the notion that the Arctic is a marine microbial province.

Current projects and objectives: Because of the high cost of genomic work, so approach has by necessity been much directed. Our priorities are as follows. 1) Describe and evaluate the taxonomic affinities and distributions of pico-eukaryotes, archaea and bacteria over synoptic scales and relate these data to specific physical regimes in cold oceans and seas. 2) Define community associations and microbial consortia within fine scale physical regimes over vertical sampling scales. 3) Develop fluorescent in situ hybridization (FISH) probes for novel protist groups. 4) Discover mechanisms that may promote stable associations among microbial groups. We are testing several hypotheses that arose from my earlier work, for example; that diatom blooms are terminated by other protists via intersize-class grazing and that bacterial species drive eukaryotic species outcome (e.g., via metabolite accumulation or pathogenesis). The results of these studies will allow us to develop and test more refined hypotheses on species distributions, functions and effects on nutrient recycling and remineralization.

International Polar Year involvement. The Canadian Government has announced two avenues for participation in the upcoming International Polar Year, one via NSERC and the second funded within the Canadian Government, Department of Indian and Northern Affairs. I am currently a co- demander on three Canadian NSERC proposals investigating different aspects of microbial biodiversity in the Arctic (PAME, MERGE and the overwintering Canadian Flaw Lead study, Part I-Science). Within the second category I am the lead researcher on the Canadian CoML IPY proposal Climate Change and Arctic Marine Diversity. It is predicted that there will be a ca. 20% success rate for these projects, my current funds and collaborations mean that we will continue collecting samples in the Arctic throughout the IPY, however funding the analysis is not guaranteed.

Alexander Loy

Ribosomal RNA-targeted oligonucleotide microarrays for highly parallel structure-function analysis of complex microbial communities: PhyloChips and Isotope Arrays

Alexander Loy, Michael Taylor, and Michael Wagner

Department of Microbial Ecology, University of Vienna, Austria, loy@microbial-ecology.net

The overwhelming extent of microbial diversity on our planet necessitates the development of novel techniques for rapidly and efficiently screening the dynamics and distribution of microbes in the environment. In principle, DNA microarrays fulfil all requirements in this regard. Large 16S rRNA sequence collections form the basis of our current perception of this microbial diversity and offer an ideal source for developing phylogenetic oligonucleotide probes targeting microorganisms at different hierarchical levels. Hundreds to thousands of such probes can be immobilized on a microarray (PhyloChip) and used for the detection of either native 16S rRNA or PCR-amplified 16S rRNA genes recovered from an environmental sample. Among the PhyloChips (1) that we have developed are two which are of particular relevance for the "International Census of Marine Microbes" agenda. The SRP-PhyloChip targets all known sulfate-reducing microorganisms, key players in the sulfur and carbon cycles of the sea floor, while the NIT-PhyloChip (2) was designed for identifying all cultivated and yet uncultivated anaerobic ammonium oxidizing (ANAMMOX) and nitrifying microorganisms, including the recently identified crenarchaeal ammonia oxidizers.
Furthermore, the recently developed Isotope Array approach employs PhyloChips for monitoring the consumption of a radioactively-labelled substrate by defined members of a microbial community (3). Combination of the IsotopeArray approach with the above-mentioned PhyloChips, or with PhyloChips newly customized for a habitat of interest, will allow comparative, phylogenetic and functional profiling of complex microbial assemblages over time and space in a highly efficient manner.

1. Loy, A., Lehner, A., Lee, N., Adamczyk, J., Meier, H., Ernst, J., Schleifer, K.-H. and Wagner, M. (2002) Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. Appl. Environ. Microbiol., 68, 5064-81.
2. Taylor MW, Adamczyk J, Stoecker K, Salkowitsch P, Radax R, Loy A, and Wagner M. (2006) 16S rRNA-based oligonucleotide microarray for the simultaneous detection and identification of all known lineages of chemoautotrophic nitrifiers. In preparation.
3. Adamczyk, J., Hesselsoe, M., Iversen, N., Horn, M., Lehner, A., Nielsen, P. H., Schloter, M., Roslev, P. and Wagner, M. (2003) The isotope array, a new tool that employs substrate-mediated labeling of rRNA for determination of microbial community structure and function. Appl. Environ. Microbiol., 69, 6875-87.

Allison Murray

Bacterioplankton diversity and genomic signatures of coastal Antarctic bacterioplankton.

Alison E. Murray, Joseph J. Grzymski
Earth and Ecosystem Sciences, Desert Research Institute, 2215 Raggio Parkway
Reno NV, 89512, USA
alison@dri.edu and joeg@dri.edu

Western Antarctic Peninsula is experiencing one of the most rapid rates of climate warming on Earth, with an increase of 5ºC in the mean winter temperature in 50 years. Concomitant with this warming trend, mean annual sea ice extent has declined by ~30%. Impacts on upper trophic levels are evident, though there have been few, if any studies that have considered the impacts on bacterioplankton in the Southern Ocean. A major limitation in predicting an impact is that we have very little information concerning the bacterioplankton (both archaea and bacteria) in terms of diversity, ecosystem roles and adaptive capabilities. These features are all encoded in the Antarctic bacterioplankton environmental genome. We are studying the diversity and environmental genomes of winter and summertime Antarctic bacterioplankton to determine the diversity, community structure, genomic content and adaptations that result in their livelihood in this high latitude environment. Genomic sequencing is supported through an award from the Joint Genome Institute Community Sequencing program which will produce 2,400 SSU rRNA gene sequences sampled from surface waters collected every other month through the year (6 libraries in total), 34,000 fosmid end sequences and 300 40Kb clones from two (winter and summer) bacterioplankton libraries.

Phylogenetic analysis of SSU rRNA genes in coastal picoplankton sampled off the Antarctic Peninsula using PCR amplification and shotgun cloning approaches show that representatives of common marine groups (α and γ-proteobacteria, Bacteriodetes) often fall into clusters with other high latitude sequences, suggesting a common evolutionary history of cold-adapted microorganisms. We have sequenced genome fragments from 11 different representatives of this late winter picoplankton assemblage in order to get a broader perspective on the genomics of cold adaptation. The results of this genomics survey have identified ~ 450 open reading frames (ORFs) coding for a full complement of metabolic, stress response, and biogeochemically interesting functions. These ORFs are now being used as gene targets on DNA microarrays for use in genomic DNA and mRNA hybridizations to relate presence of organisms and expression of a subset of their genomes.
Following gene-finding and sequence annotation, we have analyzed properties of the coding regions from a large number of predicted proteins. Analysis has included comparing amino acid composition, identified substitutions and bias in composition as well as regions of disorder in genome sequences representing diverse Antarctic marine bacteria spanning four phyla (Proteobacteria, Actinobacteria, Bacteriodetes, and Gemmatimonadetes) in comparison to homologous proteins from mesophilic bacteria. The most significant changes in these Antarctic bacterial protein sequences include a reduction in salt-bridge forming residues like arginine, glutamic acid and aspartic acid, reduced proline content and a reduction in stabilizing, hydrophobic clusters. Stretches of disordered amino acids were significantly longer in the Antarctic sequences than in the mesophile sequences. Overall our results suggest that underlying genotypic and biochemical adaptations to the cold are inherent to life in the permanently subzero waters of the Antarctic.

Ana Belén Martín Cuadrado

Abstract: "Metagenomic study of the bacterioplankton in the deep Mediterranean Sea"

Martin-Cuadrado, A.B.,, Alba J.C., Moreira, D., López-García, P., Rodriguez-Valera, F.

Most of the biomass and biological/biochemical activities of the deep ocean correspond to microbes that are nearly totally unknown. Metagenomics or microbial environmental genomics and high throughput sequencing have opened an alternative avenue to classical approaches to study them. We have begun a systematic exploration of the genetic reservoir represented by prokaryotic microorganisms of the deep Mediterranean Sea. The Mediterranean basin is unique in being very deep (average depth 2000 m) but warm throughout the water column (14ºC), in comparison to most of the areas of Atlantic or Pacific become very cold below 200 m (4-2ºC). The adaptation of microorganisms (and their proteins) to high pressure environments remains very poorly known. We believe that the deep Mediterranean could provide a special habitat in which it would be possible to differentiate the effects of the high pressure characteristic of the abyssal regions from those due to the low temperatures normally associated with large depths.

We have generated about 8,000 sequences derived from the ends of the fosmids from a library of picoplanktonic biomass obtained at a depth of 3000m in the Eastern Mediterranean. We report here a comparative genomic analysis of these data and other originated from studies carried out around the world´s oceans and particularly, with the ones derived from the water column at the North Pacific Subtropical Gyre (DeLong, 2006) and the Sargasso Sea metagenome database (Venter, 2004). We categorized sequences from each depth in BLAST searches against each database, and identified the corresponding proteins. We also show a distribution in COG categories that were represented in one sample versus another.

As expected from previous studies, the Mediterranean data were remarkably different and underline the uniqueness of this land-locked sea. A number of studies have shown that the bacterial community structure in the open ocean varies significantly with depth, but interestingly, Mediterranean Sea deep-water microbial sequences shared many more homologous sequences with the ones from the North Pacific Subtropical Gyre found at moderate depths (200-500-700 m) than with those from the deepest samples (4000 m). This may be due to several factors, such as the fact that both habitats are similar in their average temperature.

Phillip Neal

Abstract of Research
ICoMM Meeting
Amsterdam, Holland
June 12-15, 2006

Phillip Neal is a senior research assistant at the Marine Biology Laboratory in Woods Hole. He does research on the following topics:

1. Database development for 454 sequence data
2. Parametric estimates of OTU abundance
3. Multiple alignment software for large datasets
4. Mapping taxonomies to sequences
5. Taxonomic classification of sequences via compression
6. Taxonomic classification of sequences via naïve Bayesian classifiers

Furthermore, Phillip Neal does the technical development of the 'MICROBIS' database and supports the interface of ICoMM data with OBIS.

Rodolfo Paranhos

BACTERIOPLANKTON DIVERSITY IN A TROPICAL POLLUTED COASTAL BAY.

R. P. Vieira[1], A. S. M. Gonzalez[1], M. B. M. Clementino[1,2], A. M. Cardoso[1], R. M. Albano[3], O. B. Martins[1] & R. Paranhos[1]. Universidade Federal do Rio de Janeiro[1], Fundação Osvaldo Cruz[2], Universidade do Estado do Rio de Janeiro[3]

It is well known that heterotrophic prokaryotes play important roles in the structure and dynamics of aquatic ecosystems, degrading organic matter and making nutrients available to the environment. Guanabara Bay is a eutrophic estuarine system located in a humid sub-tropical region surrounded by the second largest metropolitan area in Brazil, that we studied with a microbial observatory approach since 1998. Sub-surface water was collected at a well-characterized gradient pollution transect. Bacterial abundance (measured by flow cytometry), metabolic activity (^{3}H-leucine incorporation) and chemical variables were determined to characterize the transect stations. Prokaryotic diversity was studied by PCR amplification of 16S rRNA gene and analyzed by cloning and comparative sequence analysis. Prokaryotic counts varied from $5.86x10^4$ to $2.34x10^6$ cells.mL^{-1}. Bacterial activity varied from 44 to 2804 ngC.L^{-1}.h^{-1}. Nutrient analyses showed a highest concentration of ammonia in inner bay waters, while in oligotrophic oceanic waters the main form was nitrate. We observed a increase in microbial diversity from polluted (low div) to oligotrophic (high diversity) waters The identification of prokaryotic groups responsible for the nitrification process is relevant to understand the nitrogen biogeochemical cycle in Guanabara Bay. The majority of bacterial phylotypes were affiliated with Proteobacteria taxa. The analyses of archaeal groups are under analysis. These results point to the existence of a great diversity of eubacterial and archaeal communities in eutrophic Guanabara Bay water. The increasing pollution is posing a serious problems treat to this important coastal bay. We are creating a site to hold the **MORio** – the ***M***icrobial ***O***bservatory at ***Rio*** de Janeiro to further study microbial trends in tropical environments, using Guanabara bay as a model ecosystem.

David Patterson

Management of information about organisms

Primary activities relevant to the ICoMM initiative include:

- The development of micro*scope (http://microscope.mbl.edu), an innovative communal website for descriptive information about marine microbes. This is based on original content management software which can be used as a template for other websites. The web site is designed for participation in a layered informatics world, such that the contents are very portable through web services.
- Micro*scope software has been adopted by the European-based plankton*net project.
- Development of the data model for MICROBIS, the data management environment for ICoMM.
- Development of taxonomically intelligent software and services for biodiversity informatics as part of the uBio project (http://www.ubio.org). Taxonomic intelligence refers to the incorporation of taxonomic practices and expertises within systems that manage information about organisms.

Carles Pedrós-Alió

Marine Microbial Diversity: is it knowable?

Carles Pedrós-Alió

Estimates of the order of magnitude for the total number of microbial species on Earth range from 103 to 109. Despite global dispersal of microorganisms, the number is probably rather large. The total biodiversity of an ecosystem is composed of two elements. First, a set of abundant taxa that carry out most ecosystem functions, grow actively and suffer intense losses through predation and viral lysis. These taxa are retrievable with molecular techniques, but difficult to grow in culture. And second, there is a seed-bank of many rare taxa. These taxa are not growing or grow very slowly, do not experience viral lysis and predation is reduced. These taxa are seldom retrieved by molecular techniques, but many can be grown in culture, thus explaining why "everything is everywhere".

Alban Ramette

TITLE
Bacterial Activity and Community Structure in Sandy Sediments

AUTHORS
Alban Ramette[1], Simone Böer[1], Jed A. Fuhrman[2] and Antje Boetius[1]

AFFILIATION
[1]Max Planck Institute for Marine Microbiology, Bremen, Germany.
[2]Department of Biological Sciences and Wrigley Institute for Environmental Studies, University of Southern California, Los Angeles, California, USA

ABSTRACT
The microbial ecology of coastal zones of the oceans is still poorly understood, despite the fact that sandy sediments represent a major proportion of coastal zones and play a key role in the global cycles of carbon and nitrogen. Here we described bacterial enzymatic activities and community shifts of such complex ecosystems as influenced by sampling time (four seasons over two years), depth (0-5, 6-10, and 11-15 cm), and interactions with other organisms (diatoms and microalgae). Sampling was done in sandy sediments of the long-term ecological research station of Sylt (North Frisian Wadden Sea, Germany). Bacterial enzymatic activities in the sandy sediments were found to be unexpectedly high, with potential rates at the sediment surface up to 5-6 μM per hour for commonly investigated enzymes such as β-glucosidase or chitobiase. These rates, however, followed distinct seasonal changes with 3-4 fold decreases between summer and winter, corresponding to a temperature drop of more than 20°C. Both bacterial enzymatic activities and bacterial growth rates were closely correlated with the distribution of microalgae within the sediments. Those variations were further correlated with changes in bacterial diversity using the Automated rRNA Intergenic Spacer Analysis (ARISA) method according to Hewson et al. (2003 Microb. Ecol. 46:322-336), which offered a high-resolution description of the variation in diversity of these microbial populations. Bacterial community shifts were thus posited to be the results of the interactions of seasons, depth and diatom blooms.

Anna-Louise Reysenbach
Portland State University, Portland OR 97201, USA

Patterns of diversity microbial diversity at deep-sea vents

My research focuses on global patterns of bacterial and archaeal diversity associated with high temperature hydrothermal vent chimneys. In some cases, these patterns can be linked to geochemical, geological and physical constraints. For example, the archaeal and bacterial community structure associated with different types of sulfide structures (flanges versus chimneys) appear to be significantly different. Understanding what factors influence and constrain microbial diversity at deep-sea vents will provide a more informed and ecological approach when considering a census on deep-sea vent microbes.

Furthermore, my lab is interested in growing previously uncultured microbes from deep-sea vents. For example, we detected the Aquificales in clone libraries in the late 90's and then successful cultured *Persephonella marina* from deep-sea vents on the East Pacific Rise. We have subsequently explored the distribution and patterns of diversity of the Aquificales from numerous deep-sea vent systems, and noted that although strains of *P. marina* are widely distributed (identified in the Mid-Atlantic, in the Indian Ocean and Eastern Pacific), it is noticeably absent from clone libraries and cultures from samples from the western and south western Pacific vents. Additionally, we recently isolated and characterized a widespread archaeal lineage which is the first true acidophile from deep-sea vents.

Peg Riley

Biology Department
University of Massachusetts

Phylogenetics, Genomics and a Bacterial Species Concept.

The goal of this work is to evaluate the core genome hypothesis, which posits that there is a core set of shared genes that define a bacterial species. Although it is clear that mechanisms exist for abundant and widespread genetic transfer between microbial lineages, the observation of phenotypic clustering argues for genomic stability and cohesion. To evaluate the importance of genomic and evolutionary stability versus genomic flux, we employ population and comparative genomic methods. Such analyses suggest that, for at least *E. coli* and *S. enterica*, there is a core genome that is shared within, but not between, these two related species. If the core genome hypothesis holds for many bacterial lineages, then it may be possible to revise the existing Biological Species Concept originally proposed by Ernst Mayr such that is can be usefully applied to bacteria.

Francisco Rodriguez-Valera

Micro-Mar: A database for dynamic representation of marine microbial biodiversity.

Giuseppe D'Auria, Ravindra Pushker, Jose Carlos Alba-Casado and Francisco Rodríguez-Valera

Prokaryotes represent a major component of all marine ecoystems. Their reluctance to grow in pure culture brought during the last decades a massive amount of PCR based data about these microbial communities. To analyse these huge amount of data from a global point of view new tools are required that would allow the scientific community to approach problems from **bioprospecting to development of evolutionary and speciation models.**

Micro-Mar database was born to gather molecular data from several oceanic locations, previously published in generic databases. The sequences included are stored with their geographic origin and, where provided by the authors, with physicochemical parameters. A scrupulous attention was also given to the taxonomical description of the stored sequences. In that way the users can easily compare their own data with the database looking also into the ecosystem properties of the close related hits for ecological inferences more effective than the simple phylogenetic.

Micro-Mar provides a friendly interface that allows the users to search sequences within the database by free text search, by geographic locations or by BLAST analysis of unknown sequences. The results coming from the various search systems are displayed in a classic tabular format or by Dynamic Map providing a clear representation of the best hits geographic origin. Up to date the Micro-Mar database stores 16S, ITS, 23S and several CDS sequences for a total of 1241 sequences related to Archaea and 8875 to Bacteria superkingdom. The database is online available at http://egg.umh.es/micromar/. However, due to lack of funding, it could disappear any time soon.

Stefan Schouten

Royal Netherlands Institute for Sea Research
Department of Marine Biogeochemistry

My general research subject is the organic biogeochemistry of marine sediments, i.e. the reconstruction of present and past microbial communities, biosynthetic pathways, biogeochemical cycles, environments and climates by structural and stable isotopic analysis of organic compounds in micro-organisms, marine waters and sediments. Detailed research subjects include:

- Biosynthetic pathways, structures and isotopic compositions of lipids in algae, bacteria and archaea
- Combination of lipid and genetic information for analysing past and present biodiversity
- Assessing the functioning of microbial communities by labelling experiments
- Development of new organic proxies for paleoenvironmental reconstructions
- Determination and interpretations of ^{13}C, ^{15}N and ^{2}H in sedimentary and cellular compounds
- Isolation and structural identification of biomarker lipids in micro-organisms and sediments
- Identification, isolation and synthesis of (novel) sedimentary organic compounds

I have worked with a number of microbiologists and microbial ecologists looking at diverse micro-organisms, e.g those involved in the methane cycle and those present in microbial mats from Yellowstone park. Over the last few years I specifically focussed on the microbial ecology of non-extremophilic archaea and bacteria which produce tetraether lipids. I also am involved in work studying the microbial ecology of planctomycetes performing the anaerobic oxidation of ammonia. However, I have a broad interest in marine microbes producing characteristic lipids which can be traced back as fossil molecules.

Mike Taylor

Marine benthic eukaryotes as reservoirs of microbial diversity

Mike Taylor & Michael Wagner
Department of Microbial Ecology, University of Vienna, Austria

The existence of substantial microbial communities living on and within marine eukaryotes ("hosts") is widely known, yet host-associated microbes are frequently overlooked in estimates of marine microbial diversity. There are more (potentially many more) than 200,000 species of marine eukaryotes, many of which comprise phyla not found outside of the oceans. However, with few exceptions (e.g. sponges, corals, hydrothermal vent worms) their associated microbial communities remain largely uncharacterised. This leaves a substantial gap in our knowledge of marine diversity, and signals the urgency of greater research efforts in this area. We suggest that these efforts should be focussed in two main areas: (1) diversity surveys of host-associated microbial communities – a so-called "discovery phase", and (2) directed studies to better understand the biology of these systems and the implications for diversity at a broader scale. For example, the extent to which microbes are specialists or generalists (in terms of host-use patterns) has crucial implications for diversity estimates. If the majority of host associates are confined to a specific host species, then this would be indicative of high overall diversity; conversely, if the same microorganisms are found on a wide range of hosts then total diversity will be much lower. These arguments also influence our perception of symbiont biogeography: if "everything is everywhere" then these organisms must be generalists which do not from exclusive associations with one host plant or animal. Such considerations are critical if we are to gain a more complete understanding of microbial diversity in the oceans. Here, we focus on marine sponges (one of the better understood host taxa) to illustrate this point. We performed a census of all (~1500) publicly available, sponge-derived 16S rRNA gene sequences, and contributed a further 200 sequences from three hitherto unstudied sponges. At least 15 bacterial phyla, as well as both major archaeal lineages and an assortment of eukaryotic microorganisms, are now known from sponges. The biogeography of these organisms is of considerable interest, as certain symbiont lineages are found in a wide range of distantly related, geographically disparate hosts, but are (so far) not found in seawater or any other environments. We consider these results from a global diversity perspective, and call for a better understanding of host-associated microorganisms in general.

Xiao Tian

Dynamics of bacterioplankton in the East China Sea and Yellow Sea

Zhao San-Jun, Li Hong-Bo, Wang Chen-Yang , Zhang Wu-Chang , XiaoTian*
(Key laboratory of Marine Ecology and Environmental Science
,Intitute of Oceanology,The Chinese Academy of Sciences,Qingdao,266071)

Abstract

Heterotrophic bacteria was investigated in the East China Sea and Yellow Sea (ECSYS) from 1999 to 2005. The spatial distribution of heterotrophic bacteria abundance, bacterial secondary production and the contribution of heterotrophic bacterial biomass to phytoplankton primary biomass were different in different seasons. Depth profiles of individual bacterial abundance (BA) were 5.98(3.24-15.69) ¥10^5 and 6.21 (2.37-13.62) ¥10^5 cells ml^{-1} in spring and fall. Individual bacterial production (iBP) was 0.78 (0.001-2.04) mgC m^{-3} d^{-1} and 2.27 (0.27-7.77) mgC m $m^{-3}d^{-1}$ respectively in spring and fall. Meanwhile the averaged turnover rates (m=iBP/iBB) in fall (average was 0.21 d^{-1}) was 3 times as high as it was in spring (average was 0.07d^{-1}). Value of IBB: IPB ratio in eutrophic-zone were 4-101% (average was 35%) in spring and 24-556% (average was 121%) in fall. Correlation between biological variables and environmental factors (temperature, nitrate) indicated temperature controlled dynamic of bacterioplankton during spring. Temperature and supplement controlled the dynamic of bacterioplankton in different hydrographic conditions in fall.

There was different distribution of bacterioplankton in the different hydrographic conditions and areas. Tidal front is an important factor determining the distribution of bacterioplankton in late spring and early summer of southern Yellow Sea. (1) The biomass of Synechococcus ranged from 7.62 to 22.06 mgC m^{-3} in May(2001), from 8.53 to 27.52 mgC m^{-3} in June(2001), from 0.69 to 55.90 mgC m^{-3} in June(2002). Synechococcus distributed more often in frontal area and middle-surface layer of a stratified zone; (2) The average ratio of cyanobacterium (Synechococcus spp.) biomass (CB) to phytoplankton biomass (PB) was 0.58 (May, 2001), 0.77 (June, 2001), 0.31 (June, 2002) along the transect in the three cruises, respectively. And the highest value occurred especially in the surface and bottom layer of stratified region; (3) The heterotrophic bacteria biomass was 7.56-51.82 mgC m^{-3} in May(2001), 8.54-24.77 mgC m^{-3} in June(2001) and 3.12-10.05 mgC m^{-3} in June(2002) respectively. Maximal abundance of heterotrophic bacteria occurred in stratified and mixied zone.

In the Yellow Sea Cold Water Mass (YSCWM), CB was 0.78-33.49 mgC m^{-3} (average was 6.26 mgC m^{-3}) and heterotrophic bacteria biomass was 1.58-21.25 mgC m^{-3} (average was 5.79 mgC m^{-3}). CB/PB ranged from 2% to 99% (average was 42.5%), but the ratio of heterotrophic bacteria biomass to phytoplankton biomass ranged from 0.05 to 6.37 (average was 0.85) in summer. The distribution of bacterioplankton was accorded with temperature and salinity, the minimum value of bacterioplankton biomass occurred in YSCWM. At the same time, the microzooplankton (20-200 m) was the grazer who prey on Synechococcus in the area, and the ingesting rate was about 0.20-0.42 d^{-1}.

* XIAO Tian email:txiao@ms.qdio.ac.cn

Marcel T.J. van der Meer
Royal Netherlands Institute for Sea Research

Microbial mats such as those in hot spring are thought to be modern analogs of the Earth's earliest and most long-lived fossils, stromatolites. Modern hot spring mats can be formed by two types of photosynthetic microorganisms, cyanobacteria and green nonsulfur bacteria. It is important to distinguish which type of microorganism constructed stromatolites, as only the former could have introduced oxygen into Earth's atmosphere. Even though some green nonsulfur bacteria can have an autotrophic metabolism, it has always been assumed that in cyanobacterial mats they are (photo)heterotrophic. Based on the analysis of several cultivated bacteria and different hot spring microbial mats we can now demonstrate that each type of phototroph has a unique spectrum of lipids, with typical stable carbon isotopic signatures. These stable carbon isotopic signatures suggest green nonsulfur bacterial autotrophy also in cyanobacterial mats.

An alternative explanation for the ^{13}C enriched green nonsulfur bacterial lipids could be the transfer of ^{13}C enriched cyanobacterial poly-glucose fermentation products. Therefore, we have studied the diel fluctuations in, and stable carbon isotopic composition of storage product pools, like poly-glucose, and lipids. Using percoll gradient separation we were able to separate green nonsulfur bacteria from cyanobacteria to a large degree and we were able to measure the stable carbon isotopic composition of glucose from both groups separately. The results from these measurements and ^{13}C bicarbonate labeling experiments suggest that green nonsulfur bacteria are the main autotrophs in the morning, during the low light period, and that cyanobacteria are the main photoautotrophs during the afternoon. ^{13}C bicarbonate pulse labeling and ^{13}C acetate labeling experiments suggest that green nonsulfur bacteria in these mats are also heterotrophic growing on organic matter produced by cyanobacteria.

Two "relevant" *Chloroflexus*-like bacteria have been isolated from Yellowstone hot spring microbial mats in order to study their metabolic properties and lipid content.

Developing, and compiling data for, a microbial lipid database in the context of the International Census of Marine Microbes.

Marcel T.J. van der Meer
Royal Netherlands Institute for Sea Research
PO Box 59
1790 AB Den Burg, Texel
The Netherlands
Telephone: (+31) (0)222-369584
Fax: (+31) (0)222-319674
E-mail: mvdmeer@nioz.nl
http://www.nioz.nl

Bess B. Ward

Princeton University

Research Relevant to International Census of Marine Microbes

My research programs focus on the nitrogen cycle in aquatic environments, and as one important theme, address the relationships between microbial community structure/diversity and ecosystem function. We have assembled some of the largest available datasets on diversity of functional genes in the nitrogen cycle (ammonium monooxygenase genes from ammonia-oxidizing bacteria, nitrite reductase genes from both nitrifiers and denitrifiers, nitrate reductase genes from both eukaryotic phytoplankton and heterotrophic bacteria). With collaborators, we are developing methods for evaluating species abundance relationships on the basis of gene diversity information. We find that different functional groups exhibit quite different ranges of overall diversity, which has implications for the resilience and redundancy of natural assemblages. We also find that some groups show distinct bloom dynamics under naturally varying environmental conditions, very much like patterns that are commonly observed among phytoplankton, but difficult to detect among microbes. These databases have been used to develop functional gene microarrays that can be used to quantify both relative abundance and activity of microbial groups defined on the basis of sequence similarity. We also maintain a large bacterial culture collection (especially nitrifiers) and are involved in the analysis of several marine microbial genomes.

Shi Ning Zhou

Denatured gradient gel electrophoresis (DGGE) analysis of bacterial community composition in deep-sea sediments of the South China Sea

Xin-tian Lai, Xiao-fan Zeng, Shu Fang, Ya-li Huang, Shi-ning Zhou*
State Key Laboratory for Biocontrol, College of Life Sciences, Zhongshan University,
Guangzhou, 510275 P. R. China

Abstract
The South China Sea, which is one of the largest marginal seas in the world, is predicted to have suitable accumulation conditions and exploration prospect for natural gas hydrate. The aim of this study was to explore the bacterial community of deep-sea sediments from such an ecosystem. DNA was extracted by five different methods and used as templates for PCR amplification of V3 regions of 16S rDNA gene. Denaturing gradient gel electrophoresis (DGGE) was used to separate the amplified products and analyse the 16S rRNA gene diversity of sediment samples. The results of DGGE indicated that the bacterial community composition is influenced by DNA extraction methods. Sequencing dominant bands demonstrated that the major phylogenetic groups identified by DGGE belong to *Proteobacteria*, *Bcteroidetes*, gram-positive bacteria and Achaea. Integrating different DNA extraction procedures are needed to analyse the actual bacterial diversity from environment when the amplification of 16S rRNA gene and construction of representative clone library were adopted.
Key words: Deep-Sea Sediments, Bacterial community, DGGE, Phylogenetic Analysis

---

Fungal Communities in Deep Sea Sediments of the South China Sea

Xin-tian Lai, Li-xiang Cao, Shu Fang, Ya-li Huang,
Yong-cheng Lin□ Shi-ning Zhou*
State Key Laboratory for Biocontrol, College of Life Sciences, Sun Yat-sen□Zhongshan□ University,
Guangzhou, 510275 P. R. China

Abstract
To describe fungal community in methane hydrate-bearing deep marine sediments, fungal internal transcribed spacer (ITS) regions of rRNA genes amplified from five sediments at different depth were analyzed. Five ITS libraries of rRNA genes were constructed and four hundred and thirteen clones were restricted analysis and grouped into 25 restriction patterns. ITS sequences of 44 representatives clone were determined and compared with the GenBank database by using gapped-BLAST. The phylogenetic analysis indicated that there is a broad spectrum of fungi in the deep sea sediments, including *Phoma*, *Lodderomyces*, *Malassezia*, *Cryptococcus*, *Cylindrocarpon*, *Hortaea*, *Pichia*, *Aspergillus* and *Candida*. Several rDNA sequences determined in this study like operational taxonomic units □OUT□#2 and #24 are not specifically affiliated with any currently documented fungal sequences in the public database. This is the first report on fungal communities in methane hydrate-bearing deep-sea sediments of the South China Sea.
Keywords: Methane Hydrate□ eep sea sediments□fungal community□ nternal transcribed spacer regions of rRNA genes

---

***Author for Correspondence**
Mailing address: College of Life Sciences, Zhongshan(Sun Yat-sen) University,
135 XinGangXi Road, Guangzhou, 510275, P. R. China
Fax: 86-20-84036215
E-mail: LSSZSL@zsu.edu.cn; or LSSZSL@mail.sysu.edu.cn;

The ICoMM Science Plan

(ICoMM)

Science Plan

Mitchell L. Sogin
Josephine Bay Paul Center for Comparative Molecular Biology and Evolution
The Marine Biological Laboratory at Woods Hole
7 MLB Street
Woods Hole MA 02543 USA

J.W. de Leeuw
The Royal Netherlands Institute for Sea Research (NIOZ)
P.O. Box 59
1790 AB Den Burg
Texel, The Netherlands

June 22, 2004

TABLE OF CONTENTS

1. Executive summary:

Microbes of untold diversity in marine environments are the primary catalysts of energy transformation, and are responsible for > 98% of the carbon and nitrogen cycling [1]. An estimated 3.6 x 10^{30} microbial cells with cellular carbon of ~3 x 10^{17} grams may account for more than 90 percent of the total oceanic biomass [2]. The number of bacteriophage and viruses may be one hundred-fold higher. With such enormous populations, the accumulation of mutations should lead to very high levels of genetic diversity and phenotypic variation. Yet, traditional microbiological methods have described only 30,000 protists [3-5] and fewer than 5000 kinds of prokaryotes [6].

Today we are witness to a revolution in microbiology. Just as the first microscopes unveiled an unseen microbial world, the use of molecular techniques to enumerate different kinds and numbers of single-cell organisms has forever changed perceptions of the natural world. Microbial diversity is at least 100-1000 times greater than estimates based upon cultivation-dependent surveys [7]. Comparisons of genome sequences from cultivated and naturally occurring microbial populations reveal unanticipated levels of metabolic diversity and suggest new modes and mechanisms for evolutionary change. Microbes account for the preponderance of life's genetic and metabolic variation, but our understanding of microbial diversity and the evolution of its population structures in the oceans is only fragmentary.

To develop a description of biodiversity in the oceans, the Census of Marine Life (CoML) must look beyond metazoans and plants. It must develop a strategy to (1) catalogue all known diversity of single-cell organisms inclusive of the Bacteria, Archaea, Protista and associated viruses, (2) to explore and discover unknown microbial diversity, and (3) to place that knowledge into appropriate ecological and evolutionary contexts. Several existing or proposed CoML field projects including CeDAMar, ChEss, MAR-ECO, GoMA, NaGISA, CMarZ, Reefs, Arctic, Antarctic, Sea Mounts etc. either have microbial initiatives or the potential to develop microbial-based projects. Yet, there is no global effort to acquire information about diversity and distribution of microbes and associated viruses from the three domains of life in the World's oceans. This proposal describes an **I**nternational **C**ensus **o**f **M**arine **M**icrobes (**ICoMM**). It will advocate for and coordinate investigations of microbial diversity (Bacterial, Archaeal, Protistan and Viral) and their population structures in marine environments. **ICoMM** will have five major activities. The first is to support scientific working groups. These will focus on (1) open ocean and coastal systems, (2) benthic systems, and (3) technology that is specifically required for a microbial census. The second is to develop the database resource **MICROBIS**, which will organize morphological, molecular and contextual information for marine microbial diversity within a framework that integrates into OBIS. The third is to provide resources that can facilitate and coordinate requests for research support from governmental and private foundations. The fourth is to facilitate education and outreach of **ICoMM** to make it visible to the general public and raise awareness of its goals. Finally, **ICoMM** will support pilot projects that have the potential to shape larger-scale research initiatives in marine microbial diversity.

To be successful, **ICoMM** must promote international cooperation and forge linkages with existing and new CoML field projects for collecting samples, contextual information and new technologies. At the same time, **ICoMM** must engage the broader community of

microbiologists with collateral interests in microbial diversity, evolution, biogeography and their functional roles in marine systems.

2. Uncharted Diversity of Marine Microbes: *The Known, Unknown and Unknowable.*

Communities of Bacteria, Archaea, and Protists account for greater than 90 percent of oceanic biomass and 98 percent of primary production [1, 2]. Stable isotopes studies reveal that for more than three billion years, these microscopic factories –initially anaerobic and later aerobic– mediated biogeochemical processes that shaped planetary habitability [8]. Today the oceans world-wide are teeming with microscopic and macroscopic life forms. Rich, chemosynthetic microbial communities thrive at deep-sea hydrothermal vents [9]. Abundant Archaea populate oceanic midwaters [10]. Very large populations of picoplankton including diatoms, dinoflagellates, picoflagellates and cyanobacteria are the primary catalysts in carbon fixation [11], orchestrate the cycling of nitrogen [12] and form the base of the traditional marine food web. Heterotrophic SAR11 represents the dominant clade in communities of ocean-surface bacterioplankton [13] while nonphotosynthetic protists of unknown diversity control the size of picoplankton populations and regulate the supply of nutrients into the ocean's food webs.

Amazing advances in microbiology over the past fifty years force us to think in terms of ever shifting boundaries between what is known, unknown and unknowable about single-cell organisms. In the late 60's, microbiologists had lost hope of constructing a robust natural system for microbial taxa. New molecular techniques developed during the 1970's opened pathways for establishing microbial phylogenetic relationships that were unknowable using traditional techniques (comparisons of phenotypic characters such as morphology, staining properties, metabolic capabilities, and physiology). Modern technologies (molecular techniques, automated fluorescence cell sorting, etc.) have demonstrated the great abundance and diversity of microbial life forms in the oceans, and DNA sequencing of environmental genomes (metagenomics) provides evidence of hitherto unrecognized physiological categories among the planktonic microbes. With the acceptance of the significance of microbial food webs in the 1980s [14, 15] and discoveries of microbial mediated biogeochemical cycles, oceanographers recognized the pivotal role of microbial communities as catalysts in oceanic processes. Biologists reached the profound conclusion that the continued survival of all multi-cellular life is contingent upon complex microbial communities of under-described and possibly unknowable diversity.

If we are to assemble a comprehensive description of marine biodiversity and the processes that shape habitats for multi-cellular life, we must determine what kinds of microorganisms occur in benthic and planktonic open ocean and coastal systems. For the traditional alpha taxonomist, a "kind" of organism is comparable to the concept of OTUs (Operational Taxonomic Units) for describing animal and plant species. Based upon traditional methods, the number of recognized microbial OTUs is almost trivial when compared to estimates of 10^6 to 10^8 species for marine fauna. This modest assessment of microbial diversity is not consistent with a 3.5 billion-year evolutionary history during which microbes have developed an enormous metabolic repertoire to cope with Earth's dynamic environment. In contrast, culture-independent descriptions for the microbial world, which rely upon comparisons of homologous genes (phylotypes), reveal a much richer diversity. Sequence comparisons of polymerase chain reaction products (PCR amplicons) that target phylogenetically conserved regions of ribosomal RNA (rRNA) coding regions, demonstrates that microbial diversity ranges from 10^5 to greater than 10^7 kinds of organisms [7].

Traditional microbiology has failed to culture more than 99.9 percent of these newly discovered "phylotypes" from marine environments. Using this powerful technology, the microbiologist can also make distinctions between cells with identical morphologies and enumerate differences in community structure between microbial populations. Despite the impact of new information provided by the molecular biology toolbox, traditional techniques must not be abandoned since it is within this context that our understanding of marine microbial ecology has developed.

The hallmark of microbial diversity is biochemical innovation that single-gene studies cannot fully describe. Within the next few years, molecular biology will allow us to incorporate a definition for functional capacity or inducible phenotype in descriptors of microbial diversity [16, 17]. Microbiologists are able to identify the occurrence of a particular functional or structural gene and exploit it as a marker of diversity within an isolate or for members of a naturally occurring microbial population. In a similar manner, post genomic technology can measure gene expression patterns as a means to differentiate between "kinds" of microorganisms. As a direct consequence of increased activity in marine metagenomics, the combination of high-throughput DNA sequencing, expression profiling and proteomics can describe new traits, novel functions, and unusual enzymes in microbial populations. In some cases, entirely new phyla with novel functions are being discovered [18]. These aid in understanding the evolution of life in this ancestral habitat and lead to sounder descriptions of new communities and species. Sequencing data will also be wedded to newly emerging molecular assays that incorporate automated sampling technologies and which will lead to finer temporal and spatial resolution of molecular diversity. If advances in genome technology and bioinformatics continue on the current trajectory, sequence scans of entire genomes or communities of genomes [19] coupled with high-throughput gene expression or proteomic profiles may become the standard for defining diversity and monitoring distribution patterns for microbial species.

To fully understand microbial marine diversity it is important to integrate sequence-based studies with phylogenetically-rich information from isotopic analyses and characterizations of metabolic and biosynthetic products. For example, isotopic analyses have pinpointed lipids produced by novel Archaea that oxidize methane anaerobically [20]. Follow-up investigations at sites rich in these products have revealed abundant new phylotypes that are related to methanogens [21]. The abundance of carbon-14 and carbon-13 in lipids produced by planktonic Archaea [22] proves that those organisms are assimilating large amounts of inorganic carbon from the ocean's midwaters and must be growing as autotrophs. Unprecedented lipid structures have been traced to previously unknown planctomycetes and the long-sought capability for anaerobic oxidation of ammonia. These are just a few examples of the novel insights that can be achieved when molecular and biochemical information are combined.

3. International Census of Marine Microbes:

3.1 Objectives.

This proposal implements recommendations that are relevant to CoML objectives as outlined in the document **Unveiling the Ocean's hidden majority: a roadmap**. The most general statement of **ICoMM**'s goal is to develop a highly-resolved biodiversity database for marine microbes and to understand how these populations evolve and redistribute on a global scale. Beginning with Haeckel's reports from the Challenger expedition of over 100 years ago [23], traditional microbiological approaches have made important contributions to our

knowledge of microbial eukaryotes too numerous to recount here, but little about Bacteria or Archaea. Most of what we must learn about microbial diversity in the oceans will depend upon the application of molecular techniques. Early molecular studies of marine microbial diversity only considered the Archaea and the Bacteria [24-27]. Recent molecular-based searches have already identified novel eukaryotic lineages in the water column and in warm anoxic sediments [28, 29]. Combined with fluorescence *in situ* hybridization technologies (FISH), it is already possible to associate novel, molecular-based lineages with specific morphologies. Efforts should be made to bring newly discovered key taxa into culture for more detailed investigations. One of our challenges is to create a bridge to expertise of the past.

Knowing what "kinds" of organisms exist within a marine microbial population and how community structure changes in response to environmental shifts are high priorities for **ICoMM**. Sampling strategies and the collection of contextual information will be important elements of this census. For example, culture – independent surveys reveal unanticipated numbers of distinct phylotypes in the benthos and plankton of open ocean and coastal waters. In contrast, deep-sea vents separated by thousands of miles sometimes display lower levels of diversity [27] but often harbor anaerobic thermophiles that have nearly identical rRNA sequences, even though these organisms have not been detected in open ocean waters. Mechanisms that might explain this biogeographical distribution will require studies of chemically-similar vent environments and strategically located, intermediate stations. The high-throughput DNA sequencing of environmental shotgun libraries from an oligotrophic, low diversity environment [19], provides another lesson about the importance of sampling strategies. This landmark study shows that current de-facto standards of a few hundred to a few thousand sequences for PCR amplicons of conserved genetic elements e.g. rRNA coding regions- cannot fully describe microbial diversity. A more complete accounting of diversity will dictate significant increases in data collection. But this comes at a considerable cost both in terms of reagents and in analytical efforts. To maximize the science return from such costly, high-throughput studies, marine microbiologists must identify the most important questions to be addressed and the best study sites and strategies for obtaining unambiguous answers.

The historical events and underlying mechanisms that led to contemporary microbial diversity are mostly uncharted (exceptions might include the marine foraminifera). The goals of **ICoMM** include cataloguing and discovery, but must extend to an understanding of the processes by which marine microbial diversity has been created and is maintained. Genome-based studies suggest that large-scale genetic exchange corresponding to tens of thousands of base pairs from unknown genetic sources can occur over timescales required by microbes to adapt to shifts in environmental chemistry. Stunningly, we have only scratched the surface of marine environments but already learned that the correct conceptual framework for describing the dynamics of metagenome evolution and shifts in diversity might not yet be known. Some of the fundamental questions that we must address and molecular approaches make this possible include:

1) How many kinds of microorganisms exist in marine environments and what governs the evolution of microbial lineages within complex microbial communities?

2) Why do complex microbial consortia retain functionally equivalent but genetically distinct lineages rather than selecting for a single "winner" with an optimal suite of metabolic activities?
3) Does the diversity of a microbial guild relate to the stability of its functioning?
4) Is there a biogeography for distinct microbial lineages and, if so, what are the principal drivers or restrictors? What genomic changes, if any, are associated with relocation of dormant organisms over large distances?
5) How widespread is horizontal gene transfer and does it completely obliterate phylogenetic patterns for microbes? Do viruses mediate this process?
6) Do chemical environments select for lineages endowed with particular metabolic capabilities, or does the unit of selection correspond to individual genes that can transfer particular metabolic functions between lineages?
7) What accounts for large-scale genetic variation in microbial genomes that share a very recent common ancestry? Is there a cryptic source of genetic information that selectively invades microbial genomes, or are there undocumented mechanisms that can rapidly generate novel coding capacity within a bacterial chromosome?
8) How does genotypic diversity shape phenotypic diversity, and how does this diversity influence the functioning of ecosystems?

When coupled with a larger genomic context, the interpretation of data from molecular-based field studies will challenge even the most advanced genetic algorithms and evolutionary theory. This enterprise will demand interdisciplinary efforts to explore the dynamics of microbial population biology, genome diversity, and the metabolic basis of biogeochemical processes.

3.2 Strategy

Unlike CoML initiatives that focus upon geographical locations (e.g. Arctic, Antarctic, GoMA, MAR-ECO, NaGISA, POST, TOPP etc.), or restricted environments (e.g. ChEss, Seamounts, CeDAMAR, etc.), **ICoMM** will embrace a world-wide strategy to explore the diversity and distribution patterns of all kinds of single-cell organisms in marine environments. Understanding the diversity of marine microbes is a mega-science problem that requires new approaches to mapping diversity, grand strategies, integration of diverse communities, and enabling studies that will explore processes – whether ecological or evolutionary. The community of marine microbiologists that must participate in this enterprise is diverse but they do not yet form a unified community. A problem of this magnitude requires careful planning and international cooperation. Because we know so little about the limits of microbial diversity or whether biogeographical distribution patterns exist for microorganisms, major advances will occur by 2010 albeit complete descriptions may require decades of research.

To address the key scientific questions outlined above (3a. Objectives), ICoMM must seek community consensus about research priorities and an integrated experimental plan. Unification of this discipline will require the development of shared, enabling technologies and standardized measurements in the same way that DNA sequencing and "bar coding" has provided a common means to index metazoan and plant biodiversity. Constituents of ICoMM must agree upon sampling regimes and mechanisms for sharing samples, contextual information and new data with the scientific community. We must determine how to bring together the existing molecular data into a single framework/synthesis or establish coding

standards that promote electronic exchange of information including close ties with OBIS. An important goal will be to make data from ICoMM readily accessible to process oriented interest in microbial oceanography. It will be especially important to form alliances with relevant CoML and other marine microbiology initiatives. For example, ICoMM's advisory board and working groups include participants from ChEss, CeDAMar and GoMA. Because of overlapping interests in certain protist groups, ICoMM has agreed to cooperate with CMarZ in development of programmatic infrastructure. Preliminary discussions are also underway to establish a Protistan focus Group at the interface of both programs. Other collaborative activities will include participation in the European Union projects BASICS (Bacterial single—cell approaches to the relationship between diversity and function in the sea coordinated by J. Gasol, CSIC, Barcelona, Spain), MIRACLE (Microbial Marine Communities Stability: from Culture to Function, coordinated by Francisco Romero, Inst Biomar, Spain), PICODIV (Monitoring Biodiversity of Pico-Phytoplankton in Marine Waters – coordinated by Daniel Vaulot, Brest France), ALIENS (Algal Introductions to European Shores, (coordinated by Jose M. Rico Ordas, Univ. Oviedo, Spain), MARBEF (Marine Biodiversity and Ecosystem Functioning, coordinated by Carlo Heip), EurOcean (coordinated by Paul Treguer, IFREMER Brest and Louis Legendre, Lab Oceanographique Villefranche sur mer, France) and participation in the several U.S. programs including the NSF Research Coordination Network "Seamount Biogeosciences Network" submitted by Scripps Institution of Oceanography, the NIH/NSF funded Center of Oceans and Human Health at Woods Hole (organized by John Steggeman at the Woods Hole Oceanographic Institution), the NSF RIDGE 2000 initiative, and international collaborations i.e. the MBL and the Alfred Wegener Institute joint effort to develop the WEB resource plankton*net. Finally, ICoMM must set an agenda to guide the development of funding strategies and provide support for pilot projects that have the potential to generate additional support from governmental agencies and private foundations. Upon receipt of initial funding in the Fall of 2004, ICoMM's first task will be to formalize collaborative relationships with ongoing CoML programs, relevant European and US initiatives including the Sorcer II expedition, and other existing projects that contribute towards ICoMM's objectives. The scope of ICoMM's activities by 2010 will be proportional to available resources from foundations and governmental funding agencies. An attached supplement provides cost estimates based upon different kinds of measurements applied to different sampling regimes. The dynamic range of these cost estimates is admittedly enormous and it is clear that the marine microbiology community must establish priorities for initial funding. The total resource requirement ranges from tens of millions of dollars to NASA-size efforts costing multiple billions of dollars. None of these projections takes into account efficiencies that we should expect from advances in technology. It is entirely reasonable to expect that the cost of molecular analyses will drop by one or two orders of magnitude over the next decade.

3.3 Organization of ICoMM

The MBL will be the lead organization and will support a Secretariat, a small administrative staff, and a computational biology group charged with development of the ICoMM data base MICROBIS (see below). NIOZ & NIOO-CEME in The Netherlands will fund a European coordinator and will employ a data base specialist who will integrate data from our international collaborators into MICROBIS. The MBL and NIOZ & NIOO-CEME formed a partnership in the preparation of this proposal. The Secretariat will coordinate

ICoMM activities including setting agendas, developing a community-driven database, and providing support (financial and organizational) for meetings of ICoMM's constituency.

ICoMM will coordinate scientific activities through a multi-tiered interface that will engage the general marine microbiology community, ICoMM's specialized working groups and its Scientific Advisory Committee (SAC). Three working groups (Open ocean and coastal systems, Benthic systems, and Technology) will consider the science questions posed under 3.1 Objectives as they develop a plan to address the challenges outlined under 3.2 The working groups for Benthic systems and for the Water column will consider the current status of the field, the most promising approaches for exploring marine microbial diversity, sampling requirements and potential obstacles. The Technology working group will be cross-cutting and will consider issues that overlap with the other two working groups. Their primary charge is to determine what kinds of methods and which targeted genes will be most appropriate for meeting ICoMM's scientific objectives. They will also evaluate alternative methods for sample processing, standards for data collection and data sharing.

Collectively, the three working groups will propose objectives, agendas and resource requirements for consideration by the SAC, which will guide and monitor development of ICoMM activities. These interactions will provide guidance for a broader community of representative marine microbiologists who will meet at least annually in order to move ICoMM's agenda forward. Members of the ICoMM Secretariat and the SAC will review funding requests associated with the preparation of research proposals including either financial support or DNA sequencing support for small-scale pilot projects. Examples of four such projects are provided in the Appendix.

The division of labor into the three working groups allows us to be inclusive of the taxa to be studied and addresses fundamental differences between the benthos and the water column that will impact experimental design and processing of data. Separate working groups for the Benthos and the Water Column face different challenges in surveys of microbial diversity. The communities of organisms that inhabit these environments have different compositions and structures. The physical environments are dissimilar and different nutrient and energy pathways drive each of these systems. Chemosynthetic energy and heterotrophy dominate the Benthos, whereas photosynthesis drives Open ocean and coastal water systems. There are fundamental differences in the physical stability, scale and patchiness and therefore sampling protocols for the two types of habitats will be different. Even the extraction of biopolymers requires alternative technologies for samples collected from the benthos versus open ocean and coastal waters (water column samples). In general, we have a clearer understanding of the microbiology and physical parameters of open ocean and coastal waters, where the systems complexity is lower and the technology demands are better developed. The evaluation of benthic diversity poses special problems associated with differentiating between organisms that are endemic versus the introduction of cells that normally live closer to the surface via sedimentary processes.

3.4 Membership of Secretariat, SAC and Working Groups.

Secretariat			Scientific Advisory Council (SAC)	
PI:	**Mitchell L. Sogin**	**MBL**	**John Baross**	**Univ. Wash.**
Co PI:	**Jan W. de Leeuw**	**NIOZ**	**Robert Anderson**	**Bigelow**
Secretariat /EPO	**Linda Amaral-Zettler**	**MBL**	**Edward DeLong**	**MIT**
Co-I	**Stefan Schouten**	**NIOZ**	**Victor Ariel Gallardo**	**Univ. of**

				Conc.
Co-I	**Gerhard Herndl**	**NIOZ**	**Antje Boetius**	**MPI**
Co-I	**Lucas Stal**	**NIO**	**Carlos Perdros-Alio**	**ICM**
Co-I	**David J. Patterson**	**MBL**	**Francisco Rodriguez-Valera**	UMH

Working Groups:

Open ocean and coastal systems	**Benthic systems**	**Technology**
David Karl	Andreas Teske	Rudi Amann
Steve Giovanonni	Katrina Edwards	Chris Scholin
Daniel Vaulot	Steve D'Hondt	Eric Mauther
Curtis Suttle	David M. Patterson	Robert Friedman
Peter Burkhill	Jim Prosser	Michael Kuhl
Penny Chisholm	Anna-Louise Reysenbach	

4. Data base development: MICROBIS

ICoMM will support the development and maintenance of MICROBIS, which is a distributed knowledge resource that provides systematic and biogeographic information for marine viruses, archaea, bacteria, photosynthetic eukaryotes and heterotrophic protists. The design of MICROBIS allows it to integrate seamlessly with OBIS and it takes advantage of the MBL's development effort for construction of the image-rich WEB resource, micro*scope (http://www.mbl.edu/microscope). Using the MBL's star*model for sharing distributed information about microbial diversity between different WEB portals, micro*scope currently integrates information from plankton*net, a network of distributed information that includes collaborators in Japan, Australia, Germany, France, Norway, Denmark and the US. Plankton*net seeks to develop encyclopedic knowledge resources for marine phytoplankton (e.g. http://e-bck.rd.awi-bremerhaven.de/protist/baypaul/microscope/general/page_01.htm or http://www.sb-roscoff.fr/baypaul/microscope/general/page_01b.htm). Web sites using the star*template derived from micro*scope are assembled quickly and allow distributed teams to work co-operatively to create resources of a grand scope and scale. The Data Model meets inter-operability requirements of OBIS and of other major databases (e,g., TreeBase, GenBank, the Ribosomal Database Project, the European RISSC, MIRACLE etc.). Records will include names and latitude and longitude information, will be annotated with Dublin core, ISO and TDWG-SDD - metadata standards, and incorporate DiGIR and SOAP-based protocols to promote cross-resource indexing, search and retrieval.

MICROBIS will employ a Distributed Workgroup Environment to enable a diverse community of users to manage unprecedented volumes of largely molecular data; as well as developing scaleable and flexible internet services that will allow many users to contribute to, access, organize and package information to suit the needs of a diverse community of users. Integration relies heavily on the TNS system developed at the MBL/WHOI library to emulate taxonomy within internet services. TNS exploits the universal system of metadata – the names and the classification of organisms – that has been applied to most biological information, and uses this to organize and index information locally and remotely, to create taxon-specific links between data sources, to promote inter-operability by standardizing the names in previously independent databases, or to provides services that will mark up documents with taxonomic metadata and catalogue the resources. TNS is developed in close

compliance with the International Union of Biological Sciences Taxonomic Database Working Group (TDWG) (http://efgblade.cs.umb.edu/twik/bin/view/SDD/WebHome).

To enable the international community to contribute descriptive information into a communal knowledge repository about marine microbes, the repository will include, the names, synonyms, taxonomic authorities, descriptions, images, references, web sites, distribution, ecology, dynamic links on all marine microbes. This system will share resources with micro*scope and plankton*net.

5. Education and Outreach

The outreach and education components of ICoMM are important. The lack of familiarity with the diversity and significance of microbial communities demands that we make a strategic and targeted commitment to education and outreach. Our proposed Education and Outreach activities include two objectives: 1) to raise community awareness of ICoMM; 2) to provide resources that will underpin the education of marine microbiology in schools and universities. We will work closely with the Office of Marine Programs at the University of Rhode Island (URI_OMP) and draw on their experience of existing CoML projects to implement the ICoMM education and outreach strategy. That strategy will take advantage of new informatics initiatives. We will use MICROBIS to open up access to resources across the ICoMM program narrowing the gap between researchers and consumers of knowledge. Working with the MLER (Microbial Life Educational Resources) project that has been funded through the NSF National Science Digital Library program, we will generate a library of digital educational resources with models for how those resources may be embedded in K-12 and undergraduate educational packages. This will be based on the model already developed for the geosciences (http://serc.carleton.edu/introgeo/index.html).

We will provide to URI_OMP the necessary imagery, content, text, and out-link bundles for CoML portal subprojects. Our outreach liaison officer (Linda Amaral-Zettler) will become a member of the CoML Education and Outreach network and has already developed contacts with Sara Hickox. Our budget will ensure attendance at annual meetings. We will add customized access to the resources of the micro*scope, plankton*net, and MLER web-sites to each area of the CoML portal. The web-based knowledge environments micro*scope and plankton*net are discussed above while MLER is summarized below. We will add educational resources and special navigational pathways to MICROBIS to the 'Partner Resources' page. Finally we will hold our own facilitation workshops, and/or link with existing workshops being developed at the MBL in the context of other programs.

We are well positioned to do this. The team is committed to outreach and education includes participation in the Microbial Diversity course at the MBL, teacher education workshops (MBL) and the Astrobiology Education and Outreach program. We have biodiversity informatics initiatives that will improve access to resources; and we are actively involved in educational research programs funded by the NSF.

6. ICoMM's Pogress (June, 2004).

The process of preparing an application to the Alfred P. Sloan foundation to support ICoMM activities has already influenced funding decisions. Several of ICoMM's organizers (Sogin, Schouten, Patterson, Stal) participated in an international meeting of marine microbiologists that was organized by M.L. Sogin and supported by Alfred P. Sloan and the Gordon B. Moore Foundations. This meeting produced the white paper: *Unveiling the Ocean's Hidden Majority: a Roadmap, December 17, 2003, Guidelines from a November, 2003 strategic planning workshop MBARI / Moss Landing, CA*. The Agouron and Gordon B.

Moore Foundations used our guidelines to make strategic decisions about funding activities in microbial oceanography that will impact activities relevant to ICoMM's primary objectives. In the case of The Agouron Foundation, the document served as a framework for a meeting held in January 2004 in San Diego. Although they have yet to determine the exact areas of marine microbiology to fund, the Agouron Foundation will support an annual course in microbial oceanography similar to their offering in geo-microbiology. At least one significant module of their course will cover molecular techniques that ICoMM considers to be important for a census of marine microbes. David Kingsbury of the The Gordon B. Moore Foundation used the planning document to convince his Board about the wisdom of investing in microbial oceanography. In April, the Moore foundation decided to invest approximately 143 million dollars in microbial oceanography over the next ten years. The program will initially fund twelve fellows in marine microbiology with direct cost awards that will range from ~$300,000 to $800,000/year. Dr. David Karl of the University of Hawaii is the first recipient of the Moore funding, but other fellows to be announced in the near future have interests in marine microbial diversity issues that are directly relevant to ICoMM's objectives. Several have agreed to participate in our working groups. The Moore foundation has also committed nearly five million dollars to support Craig Venter's ongoing expedition that will sample globally distributed marine environmental genomes. This project will revolutionize what we know about marine microbial diversity, but it is only the tip of the iceberg. The Moore foundation will consider other large-scale projects over the next ten years and one of ICoMM's tasks will be to develop competitive proposals. Finally, several other activities have commenced at the MBL that converge with the objectives of ICoMM. The first is a molecular based microbial population and genome study of microbes from effusive flows of hydrothermal vents. The National Research Council has awarded a grant to Julie Huber and she will carry out the work in M.L. Sogin's laboratory. The second is the recent expansion of MICROBIS to support plankton*net as described above. Finally, M.L. Sogin and L. Amaral-Zettler are Co-investigators in a new project that will study marine microbial population diversity in the context of a Center for Oceans and Human Health. We anticipate that the activities of ICoMM will stimulate additional funding.

7 Literature Cited:

1. Atlas, R.M. and R. Bartha, *Microbial Ecology. Fundamentals and Applications*. 1993, Redwood City: The Benjamin/Cummings Publishing Company, Inc.
2. Whitman, W.B., D.C. Coleman, and W.J. Wiebe, *Prokaryotes: the unseen majority.* Proc Natl Acad Sci U S A, 1998. 95(12): p. 6578-83.
3. Corliss, J.O., *The kingdom Protista and its 45 phyla.* BioSystems, 1984. 17: p. 87-126.
4. Patterson, D.J., *The Diversity and Diversification of Protists.* American Naturalist, 1999. p. S96-S124.
5. Lee, J.J., S.H. Hutner, and E.C. Bovee, eds. *An illustrated guide to the protozoa.* 1985, Society of Protozoologists: Lawrence, Kansas. 629.
6. Guerrero, R., *Bergey's manuals and the classification of prokaryotes.* Int Microbiol, 2001. 4(2): p. 103-9.
7. Pace, N.R., *A molecular view of microbial diversity and the biosphere.* Science, 1997. 276(5313): p. 734-40.

8. Des Marais, D.J. and M.R. Walter, *Astrobiology: exploring the origins, evolution, and distribution of life in the Universe.* Annu Rev Ecol Syst, 1999. 30: p. 397-420.
9. Jannasch, H.W. and M.J. Mottl, *Geomicrobiology of deep-sea hydrothermal vents*, *Science*. 1985. p. 717-725.
10. Karner, M.B., E.F. DeLong, and D.M. Karl, *Archaeal dominance in the mesopelagic zone of the Pacific Ocean.* Nature, 2001. 409(6819): p. 507-10.
11. Garcia-Pichel, F., *Cyanobacteria*, in *Encyclopedia of Microbiology*, J. Lederberg, Editor. 2000, Academic Press: San Diego.
12. Capone, D.G., et al., *Trichodesmium, a globally significant marine cyanobacterium*, *Science*. 1997. p. 1221-1229.
13. Morris, R.M., et al., *SAR11 clade dominates ocean surface bacterioplankton communities.* Nature, 2002. 420(6917): p. 806-10.
14. Azam, F., et al., *The ecological role of water-column microbes in the sea.* Marine Ecology Progress Series, 1983. 10: p. 257-263.
15. Pomeroy, L.R., *The ocean's food web, a changing paradigm.* Bioscience, 1974. 24: p. 499-504.
16. Zehr, J.P. and B.B. Ward, *Nitrogen cycling in the ocean: new perspectives on processes and paradigms.* Appl Environ Microbiol, 2002. 68(3): p. 1015-24.
17. Steward, G.F., et al., *Development and Testing of a DNA Macroarray To Assess Nitrogenase (nifH) Gene Diversity.* Appl. Environ. Microbiol., 2004. 70(3): p. 1455-1465.
18. Béjà, O., et al., *Unsuspected diversity among marine aerobic anoxygenic phototrophs.* Nature, 2002. 415(6892): p. 630-3.
19. Venter, J.C., et al., *Environmental Genome Shotgun Sequencing of the Sargasso Sea.* Science, 2004. 304(5667): p. 66-74.
20. Hinrichs, K.U., et al., *Methane-consuming archaebacteria in marine sediments.* Nature, 1999. 398(6730): p. 802-5.
21. Boetius, A., et al., *A marine microbial consortium apparently mediating anaerobic oxidation of methane.* Nature, 2000. 407(6804): p. 623-6.
22. Sinninghe Damste, J.S., et al., *Distribution of membrane lipids of planktonic Crenarchaeota in the Arabian Sea.* Appl Environ Microbiol, 2002. 68(6): p. 2997-3002.
23. Haeckel, E., *Report on Radiolaria collected by H. M. S. Challenger during the years1873-1876*, in *The Voyage of the H. M. S. Challenger*, C.W.T.a.J. Murray, Editor. 1887, Her Majesty's Stationary Office: London. p. 1-1760.
24. DeLong, E.F., *Archaea in coastal marine environments.* Proc Natl Acad Sci U S A, 1992. 89: p. 5685-5689.
25. Fuhrman, J.A., K. McCallum, and A.A. Davis, *Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific Oceans.* Applied and Environmental Microbiology, 1993. 59: p. 1294-1302.
26. Giovannoni, S.J., et al., *Genetic diversity in Sargasso Sea bacterioplankton.* Nature, 1990. 345(6270): p. 60-3.
27. Teske, A., et al., *Microbial diversity of hydrothermal sediments in the Guaymas Basin: evidence for anaerobic methanotrophic communities.* Appl Environ Microbiol, 2002. 68(4): p. 1994-2007.

28. Lopez-Garcia, P., et al., *Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton.* Nature, 2001. 409(6820): p. 603-7.
29. Edgcomb, V.P., et al., *Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment.* Proc Natl Acad Sci U S A, 2002. 99(11): p. 7658-62.

Working Group Reports

Working Group Report
Technology Working Group

The following is the agenda and notes taken during the ICOMM technology meeting held in Bremen Jan 31-Feb1, 2005. The outline identifies topics discussed and makes two major recommendations.

1. The first is to add another working group of database experts plus biologists and biostatisticians. This might fit with the Moore Foundations goals.
2. The census should consider putting together support for a pair of proper "Flagship projects" that would focus all of our technology on detailed sampling. It should correspond to a well-studied site and include high throughput microbial population structure studies, metagenomic investigations, high throughput culturing, etc.

Agenda for the ICoMM Technology Group meeting
Max Planck Institute of Marine Microbiology, Bremen
January 31st-February 1st 2005

Meeting participants:

Rudolf Amann, Max Planck Institute (MPI), Bremen, Germany (ramann@mpi-bremen.de)
Guiseppe D'Auria, Universidad Miguel Hernandez de Elche (UMH), Alicante, Spain (gdauria@umh.es)
Jan W. de Leeuw, The Royal Netherlands Institute for Sea Research (NIOZ), The Netherlands (deleeuw@nioz.nl)
Frank Oliver Glöckner, Max Planck Institute (MPI), Bremen, Germany (fgloeckn@mpi-bremen.de)
John Heidelberg, The Institute for Genome Research (TIGR), MD, USA (jheidel@tigr.org)
Gerhard Herndl, The Royal Netherlands Institute for Sea Research (NIOZ), The Netherlands (herndl@nioz.nl)
Michael Kühl, Marine Biological Laboratory, Helsingor, Denmark (mkuhl@bi.ku.dk)
Eric Mathur, Diversa Corporation, San Diego, CA, USA(emathur@diversa.com)
Mitchell L. Sogin, Marine Biological Laboratory, (MBL), Woods Hole, MA, USA (sogin@mbl.edu)

Meeting Location: MPI Bremen, Celsiusstrasse 1, D-28359 Bremen, room 4020

Background:
The International Census of Marine Microbial Life (ICoMM) is part of a larger Census of Marine Life (CoML) funded by the Sloan Foundation. There will be three "microbial" working groups on Benthos, Plankton, and Technology.
The technology group should tackle issues regarding **how** the microbial oceanography community should pursue a census of microbes. We should attempt to outline the best strategy for carrying out the census of marine microbes while at the same time considering key scientific questions like: What mechanisms guide the evolution of microbial populations?

How do microbial population structures shape planetary processes? etc etc. etc.
Major topics for our discussions could be

1) The kinds of data to be collected and challenges in collecting the data
2) What is the most important contextual data to be collected?
3) Access to samples, preservation of samples, density of sampling, opportunities for sampling, drivers for new technology development
4) Database issues.

Our group could suggest standards for collection of molecular data and their archiving. This immediately brings to mind the following more specific questions:
Should we develop a set of standard operating procedures?
What genes should we encourage for sequence analysis?
How can we engage or develop high throughput capabilities?
Is there a need for vouchering samples and how would we accomplish such a task?
What other technologies should be developed for building a census of marine microbes?
Considering resource limitation, what is the role of genomics in the census?
How should data be processed?
How should data be represented in MICROBIS? What other data elements should be brought into MICROBIS? Which other databases can be linked to MICROBIS?
Are there opportunities to forge alliances with ARB or RDP or other databases?
How can we tie the molecular data back to what is known about microbiology?

Preliminary Schedule
Sunday, January 30th

19:00 a.m.	Dinner at Hotel Atlantic
Monday, January 31st	
9:15 a.m.	Welcome & Logistics: Rudi Amann
9:30 a.m.	ICoMM overview + Report from first meeting of Benthos group in Southampton: Mitch Sogin
10:15 a.m.	General discussion on tasks of the Technology working group/ agenda of the first meeting
11:15 a.m.	Discussion 1: Kinds of data to be collected in ICoMM/priorities
12:15 pm	break for lunch
1:15 p.m.	Topic 1 continued focusing on technological challenges/guidelines and needs for the development/access to new technologies
2:15 p.m.	Coffee Break
2:30 p.m.	Discussion 2: Collection of contextual data, which and how?
4:30 p.m.	Summary of first day discussions, start with report
5:30 p.m.	End
6:00 p.m	move to downtown Bremen, dinner at restaurant “Schröter’s Leib und Seele” starting at 7 pm
Tuesday, February 1st	
9:00 a.m.	Discussion 3: Sampling techniques, sample preservation, access to preserved samples
10:30 a.m.	coffee break

10:45 a.m. Discussion 4: Database issue
12:15 p.m. break for lunch
1:15 p.m. Database issues continued with relationships to other programs
2:30 p.m. Break
2:45 p.m. How to proceed? Funding strategies and white papers
4:00 p.m. Official end of meeting

Issues to be discussed:

1. Data Water column
    a. contextual information
        i. **Location -**
        ii. **Latitude, longitude,**
        iii. **depth,**
        iv. **Sampling methodology – see section on Sampling**
        v. **Important parameters**
            1. **Chemical**
                a. **CTD – Conductivity, Temperature and Salinity in the Water column (Density), pH,**
            2. **Physical data**
                a. **Depth**
                b. **GPS,**
                c. **time**
                d. **light intensity/depth**
        vi. Desirable parameters
            1. fluorescence
                a. In situ fluorescence – Chlorophyll measurement.
                b. Variable fluorometry – fluorescence kinetics-quantum yield measurements. Maps in situ the activity of primary productions. (Falkowski)
            2. Oxygen
            3. Alk.
            4. DOC
            5. POC
            6. sulphides
            7. suphates
            8. nitrates
            9. phosphates
            10. turbidity to estimate particulate matter.
            11. Current – Relates to time series measurement. If sample drifts with current, will collect from same populations. If fixed, the current will sweep in different populations.
            12. Atmospheric data: Temp, humidity, wind strength and direction, solar radiation
    b. Biological related data.
        i. sequences - see below
        ii. proteomics – see below

iii. metabolomics – see below
iv. lipidomics –see below
v. staining data – acridine orange staining, Propidium iodine.
    1. Microscopy
    2. FACS
vi. Particular size distributions – in situ camera for marine snow – not common
vii. Need quantitative measurements – number of species and their relative abundance – One concern is how to correct for multiple operons. The operon problem is big and not solved easily by sequencing. Is there a technical solution? Most organisms may have one or two operons. The worst case scenario is that we will off by a factor of 2 or 10. We recognize there is noise in all molecular techniques but it may be no worse than noise afforded by differential lysis, PCR bias but this is better than what we have learned by cultivation.
    1. qPCR
    2. SARST (but sensitive to operon number in organism)
    3. specific probes
    4. TRFLP,
    5. FISH which is cumbersome is only way to get absolute relative numbers but difficult to carry out in high throughput-a technology challenge. Currently can only multiplex three dyes.
    6. DNA microarrays to monitor hypervariable regions using labeled PCR amplicons plus control probe outside of rRNA hypervariable region to measure accessibility of probe on array.
    7. Explore the potential of using entire operon including spacer region to handle the operon problem.

2. Soft sediment data (Part of Benthic systems)
    a. contextual information
        i. Location -
        ii. Latitude, longitude,
        iii. depth,
        iv. Sampling methodology – see section on Sampling
        v. Important parameters – need in situ data when possible*.
            1. **Chemical**
                a. **Oxygen***
                b. pH*
                c. H_2S*
                d. Porewater chemistry*gel techniques available.
            2. **Physical data – will vary depending upon environment e.g. sediments vs. sea floor vs. vents etc.**
                a. **Depth**
                b. **GPS,**
                c. **time**
                d. **porosity - sediments**
                e. **grain size distribution- sediments**

f. **permeability - all**
g. **flows and currents – vents,**
h. **incident light - is it always dark?**
i. **Temperature***
j. **Boundary layer characteristics**
k. **2 dimensional Imaging data***
l. Structural characterization of data – gradients and physical structures.

vi. Desireable parameters – most by probe technology
1. fluorescence
a. In situ fluorescence – Chlorophyl measurement.
b. Variable flourometry – fluorescence kinetics
2. CO_2*
3. Ca*
4. Nitrate, Nitrite, Ammonia*
5. Carbonate*
6. Hydrogen*
7. H_2O_2*
8. Sulphides*
9. SO_3*
10. Alk. *
11. DOC
12. POC
13. phosphates
14. turbidity to estimate particulate matter.
15. two dimensional O_2 pH imaging*
16. in situ time sampling

b. Biological related data.
i. sequences - see below
ii. proteomics – see below
iii. metabolomics – see below
iv. lipidomics
v. staining data – acridine orange staining, Propidium iodine.
1. Microscopy
a. Sectioning of frozen sediment samples
b. Cryro SEM

3. Technical Biological Capabilities
a. Nucleic acids – All sequencing studies should have a pilot phase in which sample sequencing will guide high throughput sequencing.
i. rRNA sequencing- phylotypes – current gold standard but it is not perfect technology.
1. Concerns about efficient lysis and extraction from all cells.
a. Cell lysis control – monitor by microscopy
2. Concerns about differential amplification
3. Close to full length sequences are desirable –but there is a trade off between throughput and information content.

4. Development of SARST technology for generating reads appropriate for counting different phylotypes and for detecting novelty.
5. Normalization may be needed to eliminate major players in what appears to be monolithic communities – alternative is to sequence more deeply
6. Reconditioning PCR –2nd and 3rd round PCR amplifications to increase sensitivity and minimize cell number requirements.
7. tRNA carrier to swamp nucleases and RNAase inhibitors to minimize degradation
8. Need multiple primer sets – ICoMM can provide published list of primers - .
9. Lowest samples sizes are approximately 2000 cells
    a. Can increase scale of sampling in water column samples to get larger number of cells.

ii. Single gene vs. MLST-
1. MLST can only correlate appearance of different gene families i.e. major phylotypes with DSR clades or MCR clades etc. vs.genomic context –
2. genomic context gets around the problem of interpreting MLST but can't control which MLST are in same large DNA insert library. MLST in environmental samples is a very different game than MLST in cultured isolates. For both MLST and Metagenomics in community sequencing approaches the classification of the sequenced fragments into "organism bins" is still an unsolved problem in case they do not overlap. The overlay of two trees based on different phylogenetic markers is problematic. Intrinsic DNA signatures mainly based on oligonucleotide frequencies as well as G+C content and codon usage have shown that they can guide the bining process.

iii. Environmental or metagenomic
1. What is overlap between rRNAs from PCR and rRNA from metagenomics. – WARNING – assembler may be constructing chimeric sequences with metagenomic datasets.
2. Need to make rRNA PCR libraries from same DNA samples as we use for construction of small insert and larger insert metagenomics libraries.
3. Need short insert and intermediate insert libraries and archive sufficient DNA and library material for future analyses. This implies some sort of archiving activity for key libraries and samples.

iv. Large Insert libraries:
1. Large inserts with rDNA can be toxic. – causes bias in libraries – problem is likely presence of efficient rRNA promoters from cloned gene.

2. Fosmids – prescreening using FISH to identify clones with rRNAs.

v. Benthic environments have high variability:
    1. Time series need to be array based – too dynamic.

vi. COGs based upon 25 million base pairs may be sufficient to differentiate between sites.

vii. Nested Combined approach –
    1. **Developed concept of a small number of flagship sites** for heavy sampling and detailed workup.
    2. Start with initial pilot project screen study to decide where to do the big study. This would avoid aberrant results in an unfortunate or contaminated site.
    3. Use rRNA to look at microbial diversity. rRNA survey – deep levels multiple examples of a biotype. –Deep would be in terms of extinction curve up to 10,000 reads.
        a. SARST or analogous to deep level everywhere.
        b. FISH, 5000-1000 clones,
    4. Number of examples of biotypes appropriate to study might be driven by informed advice by the community or by experience. This uncertainty is part of exploration. BATS and HOTS – oceanographic moderately deep sites, long time series. Infrastructure exists. Jan identified four sites. May need time samples as part of the time series.
    5. Metagenomic at high coverage for a few flagship sites –

viii. Single cell genomics. **100 cell** genomics is current reality
    1. micro capsules on flow cytometry. Public domain technology.
    2. extrachromosal DNA elements must be absent – kiss of death to rolling circle methods.
    3. use epicenter kits. Obtain 70kb fragments.
    4. Need at least 100 cells.
    5. Won't have cultivar to work with.
    6. how do we decide which single cell genome to work with.

b. DNA microarrays
    i. Even random arrays can be valuable for differentiating between DNA populations. – DNA microarrays may be the future of pattern recognition or fingerprinting of community composition.
    ii. Also useful for rDNA analyses – can be developed for determining relative numbers. We recognize the operon problem.
    iii. Gene expression arrays although capable of providing functional descriptions are too complex and unreliable for functional surveys.

c. Fractionation of population- This is a strategy to obtain subpopulations that are less complex than the total population. Can be done at nucleic acid stage or even according to the organisms life style – i.e. attached vs. non-attached cells etc.
    i. Fractionation according to GC using bisbenzymide.

ii. Differential filtration can also fractionate population but do not throw anything away. –Particle fractionation can resolve different kinds of communities but it does eliminate the ability to generate data about relative numbers of all different kinds of organisms in a sample.

4. Cultivation
    a. Should we increase cultured representatives for physiological studies?
    b. One method is cultivation by dilution into low nutrient medium.
    c. Other technique is using beads and forming micro-colonies after running sea water through bead column.
        i. Takes 3-4 weeks to get colony of 100 cells. Look for 100 cell beads vs. 1 cell on a micro-droplet bead. 60% of time the 100 cell guys grow to 10^8.
        ii. Should use environmental conditions to increase efficiencies of successful cultivation. For example, reproduce gradients.
        iii. Perhaps use of longer initial incubations to get microgel wells to grow. Instead of 40% success of microgels growing after 3-4 weeks, it may go higher.
        iv. Can use FTIR to sort microbial diversity in 10,000 differ cultures.
        v. When used to study Sargasso Sea, obtained most of the clades in the Venter data set but the exact overlap was minimal. Unclear whether the Culturing effort or the sequencing effort had failed to approach a plateau in terms of discovery curves. It should also be noted that these studies were on totally different samples.
        vi. This suggests an experiment that could inform us about percent of populations that reported by molecular studies,can be cultured.
            1. Carry out high throughput culturing using the micro-colony technique but incubate for several weeks under different conditions e.g. within and without light or under different oxygen conditions before sorting by FACS. It might be good to culture under steep gradient conditions- might be technically challenging.
            2. Carry out molecular surveys to very deep levels i.e. 10,000 clones from the original water sample and 10,000 sorted colonies.
            3. Compare rRNA populations looking for extent of overlap between rRNA amplicons from the cultivars versus from the environmental isolates.
    d. Use of cultivation in genomics
        i. Discussed Moore sequencing by Ventor. Anticipate 5X coverage of 120 genomes –80 in progress. Phylogenetically disperse. Should we push for full coverage? -
        ii. At some level in the phylogenetic tree we would like closed genomes. Closely related genomes benefit quickly from closed genomes. Divergent genomes don't benefit to the same extent from closure.
    e. Use of cultivation in proteomics
        i. Can use on cultivars

ii. Not useful on environmental genomes because of complexity
iii. On pure cultures, proteomics is powerful.
iv. Environmental samples: proteomics may not be able to reproducibly identify functional genes at this time.

5. Metabolomics and Functional genomics
   a. One strategy is to select clones for analyzing genes in a genomic context is to screen large insert libraries according to function rather than homology.
   b. Nitrolase tree presented from clones isolated according to functional characteristic. There may be many functions of interest to marine biologists for which there already exist functional assays for recombinant clones containing fosmid or BAC sequences that can be screed in medium-high thruput manners. I.e. selection, agar plate color metric assay, MT plates, FACS, etc.– ICoMM should provide Diversa with a list.
   c. Bio markers in general have potential to be informative
   d. Intermediate metabolites
   e. Single genome metabolism vs. environmental metabolism
      i. Discussed idea that the active metabolism and hence functional diversity of an organism may reflect the microbial population context. In other words activity may reflect communication with other organisms in the population.

6. Database for ICoMM – General comment: This is a special area where ICoMM's activities can make an important contribution. We need a working group to handle the technical details but it should also include biologists who drive the database construction with scientific questions. An important issue will be making sure that standards allow database interoperability.
   f. Databases that we want to link
      i. MICROBIS
      ii. ARB
      iii. MicroMar
      iv. IMCG (Integrated microbial community genomes)
      v. GMOD – type databases for large insert libraries
      vi. Database of GMOD like databases – see http:/www.tigr.org/tdb/MBMO or http://www.mbl.edu/giardia
      vii. Proteomics and metabolomics.
      viii. Recommend meeting for database experts.
   g. Database capabilities today. **Must establish standards for sharing data between participating databases. Must push for high-level initiative to organize marine microbial databases that can talk to each other. … Will need funding. – ARB is slow because of lack of resources. Funding for ARB has been all but impossible. Is this a job for the UN???? The following are issues that need to be addressed in database construction**
      i. Overlay contextual data
      ii. Identification of sequences (broadly defined) in GIS context.
      iii. Data input – high thru put
      iv. Phylogenetic excellence –e.g. ARB and/or CIPRES
      v. Super tree capability

vi. Distribution of enzymatic and bio-activities i.e. function in marine environments.
vii. With ties to geochemical processes.
viii. With ties to genomics.
ix. Kinds of additional database that may be constructed.
    1. Functional database
    2. Proteomic database
    3. Lipid database
    4. Metabolomics

h. Database capabilities of the future. – Will require standards
    i. Highly integrated, cross referencing databases of metagenomics, metabolomics, proteomics
    ii. Interfaces with high throughput phylogenomics –possibly CIPRES
    iii. Big issue of capturing old data – Base line information.
        1. older molecular studies
        2. need information about GIS, location, depth etc.
        3. Biogeochemical information from study sites – may drive new molecular or census-based studies.
        4. Links to Prokaryotes and Bergeys databases?

i. Money for maintenance is a big issue to resolve
    i. Emphasize concept of working databases and necessary link to curation vs. archival databases.

7. Other issues:
    a. Archival samples – who gets the DNA and who stores it?
    b. Metagenomic clones from John Heidelberg
        i. Small insert clones not saved
        ii. Medium to larger inserts were saved
        iii. Filters stored in sucrose lysis buffer – as in Giovanonni
        iv. Sediment samples require 100's of grams for biogeochemical science.
        v. Cores kept at –70°C for nucleic acids
        vi. Need to have anaerobic reducing environments. Immediate treatment. Anaerobic hoods ship board. Counting device on ship board is important to know where you are but could use NASA technologies i.e. LAL or ATP assays.
        vii. Phage mentioned but don't know best practice. Disagreement about best practice – Wolmark freezing in glycerol, Foeherster DNAse followed by 4C, Suttle other in between.

    c. Sampling strategy – replication? Density? Best sites? Depth of sampling in terms of numbers of reads. HOTS or BATS do deep sampling.
    d. **Developed concept of flagship sites** for heavy sampling and detailed workup. Start with initial pilot project screen study to decide where to do the big study. This would avoid aberrant results in an unfortunate or contaminated site.
        i. Use rRNA to look at microbial diversity.

ii. FISH, 5000-1000 clones, BATS and HOTS – oceanographic, moderately deep sites, long time series. Infrastructure exists. Jan identified four sites. May need time samples as part of the time series.

8. Sampling scales
9. Sample sharing – who makes decisions about precious samples?
10. Attached sampling note from Chris Scholin who could not attend the meeting.

FROM CHRIS SCHOLIN
Greetings -

Again, my apologies for not being able to participate in this important workshop. I'd hoped to write something up on water samplers that might be of use as the group considers how it might acquire samples as part of ICoMM's mission. In that regard, we have precious little time at most locations (as in having a human presence) to sample microbial populations and relate changes in their abundance/gene expression to short and long-term alterations of the chem/phys environment. I think this reality, coupled with the value of having physical samples from different environments (even if they're just sitting in the freezer), speaks to need to automate sample collection and archival.

There are a # of commercially available instruments out there that may help in this regard. I've collected and attach here some information about samplers that can be combined with other in-water observatory systems to enable collection of physical samples in the context of many other measurements. Indeed, depending on the observatory in question, it is possible for several sampling modes: pre-programmed (e.g., time-dependent); human in the loop adaptive sampling (i.e., event-dependent - e.g., with respect to out put of an optical sensor); autonomous adaptive sampling (as in previous). Some of the samplers provide minimal chemical processing of material collected (e.g., fixation for nucleic acid recovery or whole cell analyses). In any case, it's worth considering use of such devices in conjunction with on-going efforts to establish local, regional and global observatory backbone (incl. fresh waters as well). Enormous investments are being made along those lines, and it seems ICoMM would be well advised to take advantage of that infrastructure as you gain not only infrastructure (operations/platforms) but contextual data products as well .

Beyond 'simple' samplers, new systems are emerging that are making it possible to conduct molecular analytical analyses remotely, in situ, even subsurface. Two examples are the "autonomous microbial genosensor" (AMG; John Paul et al., U of S Florida, Center for Ocean Technology; http://www.onr.navy.mil/sci_tech/ocean/reports/docs/om/03/omfries2.pdf; http://www.marine.usf.edu/microbiology/genosensor.shtml) and the "environmental sample processor" (ESP, Scholin et al., MBARI, web site still under construction - http://www.mbari.org/microbial/ESP/). The AMG is designed around NASBA isothermal amplification analyses and has no capacity for sample archival. The ESP emphasizes the

interface with the environment and sample processing (archival, extraction, fractionation), and supports a variety of molecular probe analyses (e.g., probe arrays, antibody assays, and many other analyses possible too). There are other in-water sensors that may make it possible to conduct whole cell
(e.g., FISH) analyses autonomously, such as the in situ flow cytometer that WHOI researchers are working on (there are other in-water flow cytometer systems out there, but none I'm aware of also allow for probe application).

I hope this very brief bit of information will be of use. I look forward to participating more in the future --

Chris Scholin
Monterey Bay Aquarium Research Institute (MBARI)
7700 Sandholdt Rd.
Moss Landing, CA 95039-0628
USA
phone: 831-775-1779
fax: 831-775-1620
web: www.mbari.org

Sampler Name	Max # of samples	Max sample vol	Sample vol range	Design	Size	Weight	Sampling rate	Power	Cost	Notes
EnviroTech, AquaLAB Gas-tight deep ocean water sampler	50	1L	75mL - 1L	Syringe, multi-port valve	Length = 26" Diam = 7"	Air = 39.6 lbs (plus 11 lb controller)	200cc/min	V = 13.5 VDC Ipeak = 740 mA, I (motors still) = 140 mA, Isleep = 125uA	$37k	Gas tight titanium foil bags) 2 stroke acquisition (inhalation and pump to bag) Rated to Full ocean depth Slow Reusability of bags? Cost?
EnviroTech, AquaMonitor Smart Water Sampler	50	3L	100mL - 1L	Syringe, multi-port valve	Length = 24.5" Diam. = 5.7"	Air = 17.6 lbs Water = neutrally buoyant	200 cc/min	V = 12 VDC Ipeak = 740 mA, Imean = <2mA	$23-30k	Similar issues to Aqualab Handling issues potentially problematic
McLane Remote Access Sampler	48	500 mL	100-500mL	Pump with multi-port valve	Length = 65" Width = 17" Height = 17"	Air = 240 lbs (containers full- >325 lbs) Water = 125 lbs	75 mL/min (also have 150 mL/min pump)	31.5 VDC <1A when pumping	$27k	Heavy Expansion system - clean Collects particulates Bag reusability issues ($13/bag) Full ocean depth 25 port valve available Configurable for smaller package
Challenger Oceanic Multi-sampler	10	1L	1L	rotary and linear actuators. Flow sensor.	Length = 11.8" Diam. = 15.7"	Air = 33 lbs	4min/1L max (syringe)	12-24 VDC at 2A max	~$10k	Hard bottles (not as clean) Fewer samples (may need >1) Small Single stroke acquisition More mechanically complex Simple control input Could put together bag/valve/pump system Much faster sampling if there is no filter
Chelsea Technology group. Autonomous Plankton Sampler (Collects particles)			Uses standard 270 um mesh silk or 80 um nylon gauze	Updated Hardy mechanism	5.5" x 9.25" x 11.5"	28.1 lbs	Can operate in towed vehicle at 25 knots	6 x 1.2v cells	$18k	Just collects particles May be good solution for those wanting large volumes

A list of questions relevant for all ICoMM working groups

Diversity:

1) What metric can be used to describe microbial diversity?
2) For molecular measures, what are the strengths and weaknesses of single-gene, genomic, and populations-level perspectives? Which of these will have the greatest long-term benefits?
3) How should the dynamics of diversity be handled, and should the approach be based on species (phylotypes) or populations?
4) How will approaches to microbial diversity differ from those used in the Census of Marine Life, which focuses mainly on metazoans?

Integration:

1) What level of biodiversity is necessary to interpret ecological, physiological and process-related observations?
2) How can process and ecological data inform us about diversity and what levels of information are required?
3) What are the key scientific questions that a census can address?
4) Where are the gaps in the investigative framework?
5) How is diversity related to process stability? Can predictive frameworks be defined?

Sampling, prioritization and coordination with other programs:

1) Schedules, locations and priorities.
2) Relationship to sampling program in the Census of Marine Life and related subprojects.
3) Are there mileposts that will logically define phases of the project?
4) What observations are needed at each sampling site?
5) How should we address temporal variations?
6) How should we address spatial heterogeneity, particularly with regard to commensal populations and chemosynthetic environments (e.g. seeps, whale falls, wood falls)?

Databases:

1) What is the structure of the information that will be produced?
2) What are the preferred techniques for carrying out a census?
3) How can databases be structured to facilitate communication?

Relationships with other programs:

1) RIDGE, ODP, Genomes to Life, Microbial Observatories,
2) Biodiversity Organization issues: Centralized, Coordinated, Distributed, Combinational Models

How to Proceed:

1) Funding Strategies
    i. What are the highest priorities?
    ii. Where could ICoMM seed monies be most effective towards initiating programs and collaborations in benthic systems?

iii. What are the key programs for benthic studies? How can benthic diversity research be better advocated within these or other programs for multi-institutional and international programs?

2) White Papers

i. What audiences should be targeted?

What are the most important messages?

Working Group Report
Benthic Systems Working Group

ICoMM Benthic SystemsWorking Group Meeting
Southampton Oceanography Center (SOC) January 14th-15th 2005
Meeting participants:

Dr. Katrina J. Edwards, WHOI (katrina@whoi.edu)
Dr. Mitchell L. Sogin, MBL (sogin@mbl.edu)
Prof. Steven D'Hondt, GSO (dhondt@gso.uri.edu)
Prof. David M. Paterson, Univ St. Andrews (dp1@st-andrews.ac.uk)
Prof. James Prosser, Univ Aberdeen (j.prosser@abdn.ac.uk)
Prof. Andreas Teske, Univ N Carolina (teske@email.unc.edu)
Prof. Bo Barker Joergensen, MPI (bjoergen@mpi-bremen.de)
Prof. Anna-Louise Reysenbach, Portland State Univ (reysenbacha@pdx.edu)
Prof. Paul Tyler, Univ Southampton PI of ChEss (pat8@soc.soton.ac.uk)
Dr. Stefan Schouten, NIOZ (schouten@nioz.nl)
Dr. Eva Ramirez Llorda – Secretariat of ChEss

ICoMM Benthos Working Group Agenda

Friday, January 14th

10:00 a.m.	Welcome: Katrina Edwards
10:15 a.m.	Logistics:
10:30 a.m.	Overview of ICoMM: Mitch Sogin
11:00 a.m.	Overview of ChEss: Paul Tyler – Eva Ramirez Llorda
11:30 a.m.	The Benthos working group charge: Katrina Edwards

--break for lunch—

1:15 p.m.	Diversity discussion; led by James Prosser; Andreas Teske rapporteur
3:15 p.m.	Break
3:30 p.m.	Biodiversity Integration; discussion led by Anna-Louise Reysenbach; Stefan Schouten, rapporteur
5:30 p.m.	Diversity summary; Andreas Teske & Stefan Schouten; Katrina Edwards rapporteur
6:00 p.m.	End

Saturday, January 15th

9:00 a.m.	Sampling, Prioritization, & Program integration; Edwards discussion lead; Paul Tyler rapporteur
10:15 a.m.	Break
10:30 a.m.	Databases: Bo Jorgensen discussion lead; David Patterson rapporteur

12:00 p.m.	Break for lunch
1:15 p.m.	Relationships to other programs; Steve D'hondt lead; Mitch Sogin rapporteur
3:15 p.m.	Break
3:30 p.m.	How to proceed; Andreas Teske discussion lead; Edwards rapporteur
5:00 p.m.	End

ICoMM Benthic Systems Working Group discussions: Southampton Oceanography Center meeting Jan 14-15.

General comment:

The ICoMM **Benthos working group** initiated their discussions using questions posed by the secretariat to each of the ICoMM working groups (see below and indicated by italics). Within the context of these discussions, the **Benthos includes everything from sedimentary systems, seamounts, hydrothermal ridge axes, and bare rock ridge flank environments.** A secondary circumscription of **Benthic organisms might be: "everything that is driven by "Dark energy".** Early discussions reflected the working groups collective view that an international census of marine microbes must include more than the counting of microbes by some genetic or physiological criteria. The census must gather contextual information in order to interpret the significance of diversity measurements and to maximize the return of scientific information from investments in the census. As a corollary, if ICoMM defines diversity through measurements of active genes, the environmental context will radically alter microbial expression patterns. For example, the same organism can have unlike functions and physiologies in different geochemical contexts. We also express the view: -- A census of what is there is not always a census of what is active. --- screening diversity does not lead to self-evident physiological categories.

Diversity:

The discussion about diversity was challenging because there are many ways to describe genotypic and phenotypic similarities and differences between kinds of microbes. We reject the use of the term species in the classical sense used to describe sexually compatible multicellular organisms. There are no self-evident "species" criteria, with a unified level of resolution that is equally applicable to questions in microbial ecology, biogeochemistry, and evolution. We are working with different metrics at different scales of resolution; it is unavoidable at present. Comparisons of genes, genomes and function offer different modes for describing microbial diversity. **The Technology working group should carefully address existing and future technical capabilities for measuring microbial diversity.**

1) *What are the most timely and important questions regarding benthic microbial diversity?*
   a. The working group discussed the recurrent theme: **Is everything everywhere?** Are there evolutionary lineages of microorganisms in marine environments and is it possible to identify biogeographical distribution patterns?

b. The working group explored the question: **Should ICoMM be concerned with genetic, genomic diversity or functional diversity?**
    i. Genetic diversity as defined by a single or small number of genes may not always be sufficient for a census for a variety of reasons including: 1) inadequate information to resolve location or phylotype in molecular trees, 2) conflict between relationships inferred from different genes (either due to horizontal gene transfer or altered rates of evolutionary change). 3) technical problems associated with different efficiencies of DNA extraction or design of primers for amplifying highly diverged homologues.
    ii. Estimates of genomic diversity in microbes are complicated by the potential for large-scale movement of genetic elements between microbial genomes and the technical challenge and expense of collecting genome sequences from microbial populations. Without large scaffolds, the assembly of shotgun environmental sequences from complex microbial populations may be beyond our grasp. **The technology group should explore optimal and alternative strategies for gathering information about the genomic context for a gene that is highly conserved or that specifies a particular function.**
    iii. Since ecosystem parameters will influence functional diversity, contextual information e.g. water chemistry, temperature, etc. will be essential for interpreting functional diversity of microbial populations.
c. How does diversity relate to function and ecosystem processes?
d. How does the choice of gene influence diversity assessments and inference about presence or absence of functional groups in a complex community?
e. What scales of heterogeneity – spatial, temporal are most appropriate for the census?
f. How can we link diversity at different scales?
g. What is the optimal measure of microbial diversity

2) *What metric can be used to describe microbial diversity?*
    a. Functional diversity –molecular techniques e.g. expression profiling, qPCR can serve as proxy for activity of specific functions.
    b. Phylogenetic Diversity- rRNA, other conserved genes, or multiple conserved genes
    c. Genomic Diversity- Genomic context of genes associated with a conserved gene, or with a cluster of genes that can be identified within large DNA insert libraries by functional assays. Includes sequence analysis of BAC's, Cosmids and Fosmids.
    d. Depending on the research question under investigation, we may need different metrics. For example, if we seek to focus on the genetic diversity of a particular phylotype, it will be necessary to use rapidly evolving sequences. In contrast, a general survey of all organisms in a microbial population will rely upon more slowly evolving gene sequences. In some cases the census will seek information about relative or absolute numbers of different kinds of organisms.

3) *For molecular measures, what are the strengths and weaknesses of single-gene, genomic, and populations-level perspectives? Which of these will have the greatest long-term benefits.*
    a. Problems associated with single-gene and genomic level perspectives are summarized above. One potential option would be the sequencing of genomes from representative single cells from different environments or several genomes from a single environment. This will require advances in cloning technology and at least a 100 fold reduction in the cost of DNA sequencing
4) *How should the dynamics of diversity be handled, and should the approach be based on species (phylotypes) or populations?*
    a. The working group was less concerned about temporal diversity than diversity at different sites. Temporal diversity is important but may not be possible to monitor in the benthic environment because of challenges of obtaining samples. **The technology group should address time-series sampling capabilities in the benthos.**
5) *How will approaches to microbial diversity differ from those used in the Census of Marine Life, which focuses mainly on metazoans?*
    a. **The metric of diversity**
        i. Critical discussion of this thorny issue will be necessary. There are no self-evident "species" criteria, with a unified level of resolution. We are working with different metrics at different scales of resolution; it is unavoidable at present. **Can the technology group define metrics using different techniques to guide ICoMM participants?**
        ii. Depending on the research issue, we may need to use different metrics. E.g. DNA sequence diversity, DNA arrays, expression profiling, qPCR, FISH technology, etc.
        iii. Evaluate a number of selected key environments in terms of microbial diversity. Multiple examples should be studied for each selected key environment. Here the key question is to determine whether or not a particular environment selects for a specific microbial population structure.
        iv. Single cells genomics are the ultimate screening tool to evaluate the biogeography issue. This will of course require enormous reductions in cost of DNA sequencing and annotation as well as development of technology for producing libraries from single cells. **The technology group should address how to encorporate diversity measurements based upon genomics into a census that also reports on the absolute and/or relative numbers of different kinds of microbes in a sample.**

6) **Other issues:**

The working group discussed experimental strategies and how they could translate into fundable projects. Getting support for "exploration" will be very difficult in a climate that explicitly encourages hypothesis driven research. The discussions on Integration outlined a number of experimental paradigms that ICOMM should formulate in terms of testable hypotheses.

Integration:

The following questions about Integration were addressed in the context of outlining experimental strategies that could attract funding under **Links between divrsity measurements, environmental conditions and functional diversity.**

1) *What level of biodiversity is necessary to interpret ecological, physiological and process-related observations?*
2) *How can process and ecological data inform us about diversity and what levels of information are required?*
3) *What are the key scientific questions that a census can address?*
4) *Where are the gaps in the investigative framework?*
5) *How is diversity related to process stability? Can predictive frameworks be defined?*

a, Links between diversity measurements, environmental conditions and functional diversity.

i. Strategy 1) By linking genetic and physiological diversity ("function") it is possible to formulate hypotheses about links between genotype and phenotype.
ii. Strategy 2) Looking the other way around, an analysis could start with metabolic function and diversity in a given environment, and extrapolate from these data towards potential diversity. What types of organisms can be expected in a given environment? Can the genes discovered in a certain environment account for all functions in this environment? Imagine new gene functions and classes. Do not start with screening diversity and try to fit it into a priori functional categories – the categorical framework may not fit.
iii. Strategy 3) The continuum of environmental conditions and niches may correspond to a continuum of sequence diversity. Several thousand genes may have to be sequenced, to document this link.
iv. Strategy 4): Global diversity patterns can be observed in a single species; these may be a consequence of distinct environmental conditions and pressures, or biogeographic separation.
v. Strategy 5) Complement genomics with proteomics. Proteomics today gives genomic-level information, which indicates high complexities. Getting good, pure protein in sufficient quantities is the limiting factor. However, in some cases, there is useful information to be obtained if a key component is dominating (ANME-1 mcrA protein from Black Sea ANME mats) or biomass is sufficiently high to enable processing and purification.
vi. Strategy 6) Focus on groups with a well-developed toolbox that Couples function and diversity. SRBs are an excellent example where the coupling is tight --- excellent molecular equipment, primers, arrays, extremely complex and pervasive fermentations – not much work done (hydrogenase as devil's advocate idea). Methanogenesis as a simpler example

b. **Biogeography and diversity:**

The working group also discussed the concept of biogeography of microbes or Microbial distribution

Principal problem 1). Can we reject "everything is everywhere" with absolute certainly? A strict rebuttal is not possible, since we don't have a 100% complete census of any microbial habitat. A microorganism may remain undetected by various tools although other datasets indicate their existence. On the other hand, if the hypothesis is correct it eliminates the need to sample every environment to absolute completeness.

Principal problem 2). The speed of evolutionary gene flow vs. environmental exchange determines whether biogeography is possible at all; genetic speciation must be faster that exchange mechanisms. Strength of dispersal mechanisms that override genetic isolation and evolution of separate lineages. Examples are the phylogenetic separation of freshwater/saltwater bacteria as shallow, but separate clusters (Zwaart and Crump 2003).

i. **Strategy 1**: Fosmid libraries from geochemically similar environments, to differentiate changes in communities that are not necessarily affected by geochemistry.

ii. **Strategy 2**: Look for organisms in the wrong place. If organisms show up out of place again and again, that is an indication for an effective environmental dispersal mechansisms. Examples are thermophiles in arctic sediments.

iii. **Strategy 3**: Use same field campaigns for multiple investigators with multiple targets

iv. **Strategy 4:** Compare three (ten) very similar or almost identical habitats (geochemically) for microbial community structure and genomic repertoire

Sampling, prioritization and coordination with other programs:

7) *Schedules, locations and priorities*. **This topic should be addressed by SAC and/or Technology working group as well as by participants at the general meeting.**
8) *Relationship to sampling program in the Census of Marine Life and other ongoing or currently planned sampling efforts.*
This topic should be addressed by SAC and/or Technology working group
9) *Are there mileposts that will logically define phases of the project?* **This topic should be addressed by SAC and/or Technology working group**
10) *What observations are needed at each sampling site? Density of sampling: A key question or hypothesis "Is everything everywhere?" If we could demonstrate unequivocally that everything is everywhere, it will reduce the requirements for sampling*. This topic was addressed under ***Integration***
11) *How should we address temporal variations?*
12) *How should we address spatial heterogeneity, particularly with regard to commensal populations and chemosynthetic environments (e.g. seeps, whale falls, wood falls)?*
This topic should be addressed by SAC and/or Technology working group as well as by participants in general meeting. The Benthos systems working group did not define detailed spatial sampling strategies for particular sites but the discussions

captured the idea that several seeps, vents, the sea floor and geographically distributed sites including sediments with similar water chemistries should be sampled and compared.

Databases: The following questions guided the Benthos working group's discussions of data bases. The technology working group may wish to address these issues in greater detail.

1) *What is the structure of the information that will be produced?*
2) *What are the specific database needs for benthic systems?*
3) *What are the preferred techniques for carrying out a census in benthic systems?*
4) *How can databases be structured to facilitate communication?*

US databases funded by NSF (*Note: NSF data base contact Peter Cornellon URI*) are generally judged according to their technical merits rather than utility to the investigator. ICOMM will emphasize content and a format that facilitates access. Currently, ICOMM relies upon MICROBIS which provides image-rich information about protists and a microbial name serving capability. It also functions as a "traffic cop" in the sense it provides links between its content and relevant information residing on other data bases accessible through the WEB.

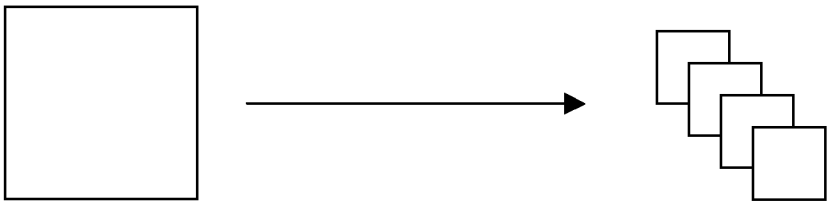
Within CoML there is a primary data base:: OBIS plus (Eurobis), as well as web sites for each of the field projects. The Chess database is undergoing revision and will be integrated to OBIS. ICoMM must be compatible with OBIS and other CoML web sites. It should strive to share taxonomic information and naming capability with other COML databases – particularly for zooplankton which houses some of the protists.

Micro*scope:

ICoMM website: MICROBIS based on Micro*scope: (Contact Paddy Patterson), this has received diverse funding (principally through NASA), and a system is developed and growing, currently with approximately 3 million names

- Non-molecular at present, image rich database of microorganisms and a detailed compilation of names that can be organized according to alternative taxonomic schemes.
- Access by name, habitat, but at present focused on Protists
- Image linked to other relevant sites on the web

Database web-resources



Different classifications can be added at a later stage (see later phased approach)
NOTE: Star site concept very useful: software can be exported to other sites, new data base synchronized back to micro*scope, however there remains danger in accuracy if there is not a single gatekeeper in charge of the information.
Volunteer submissions of images and data via a webmaster was also discussed.
The Engine for Micro*scope is UBIO, OBIS is considering using UBIO.
The inclusion of basic biochemical information for cultured representatives of certain groups (along the lines as available in papers describing newly named taxa) was discussed.

MICROBIS could take data such as phylogentic trees, possibly via a link with ARB to allow search for nearest relative to new sequence data. It was suggested that only the most simple, core information for research purposes should be provided in MICROBIS, rather than detailed analysis such as would be used for publication quality data because of the dangers associated with automated alignment of sequences (though it is noted that this is now becoming standard practice due to the increased data coming on line with genomic sequencing efforts). **The technology group should address which molecular data base(s) can optimally provide services and links to microbis.**

Important questions:

What is the depth of information to include for a census?

For example, the database could search by sequence and provided nearest relative orphylogenetic position. Lead (Link) to page and information available on related types. Includes biochem, **Technology group should address optimal way to integrate with other molecular data bases and interoperability.**

How do you judge relatedness……standard probabilistic methods
If sites (databases) already have information, use them in the framework, eg meta-site concept don't duplicate in MICROBIS use existing resources.
Eg Tree of life, Miracle (biogeochem linked to organisms), also biogeography sites
Can set up database to search for occurrence of sequence by geographic or site data with contextual information:
Structure of MICROBIS: phylotype and contextual info and links
Important information: What are the contextual data required:
Where is the quality control (A)…not possible to be too precise
eg Can database hold controversial (arguable) information?….yes, but its always up to the individual to make an assessment.

FISH data and molecular probes could also be added at later date, the database may be developed in phases
As new information is added then new searches will be required , this is part of the dynamic use of the data.
Options 1) rebuilt its own phylo trees
2) use another data base to update phylogenetic basis
the latter is preferred if formal link can be established

(A) census data should be simple and not duplicate effort, use of links better,
MICROBIS would ship sequence to ARB for analysis, Note: ARB always behind genbank
MICROBIS should be user friendly and census driven..

Technical hierarchy: GENBANK: ARB : MICROBIS (latter more contextual info, central repository for census information, images, interpretative, search tools, points to ARB and GENBANK

Data entry:

Data entry should establish a common set of minimal identifiers and contextual information. Mechanism in place but not public yet, password protected system will be in place to prevent spurious information, comment possible but this leads to some problems
Need expert reviewers to insure integrity of data check on how other databases handle this, but does slow down progress
Is the database going to include environmental parameters? Are databases with such info in existence?
Seamount database (through GERM) can be used to design this…links to information can be included
We should ask for information on sample sites and relevant conditions, and agree format, e.g ODP system , little in other eg ChEssThis is a problem with GENBANK contextual info often lacking, only optional and not encouraged
However, IODP can be asked to ask authors to include this information, RIDGE2000 is an example of good practice
Try and insure contextual information is included in as many places as possible IODP can insist on this information future cooperation can be based on this to help compliance
Steve D'Hondt asked that people let him know what is required and he will forward the request…

Coffee discussion. Environmental data much more variable and difficult to collect, yet very useful in understanding the distribution and likely occurrence of organisms/sequences in specific environments.

Ask CHeSS what OBIS is capturing in terms of environ data: (A-not much)

MICROBIS : **Key environmental parameters that should be added for context and linked to the sequence/isolate information**

1. Location: latitude, longitude, and depth.

2. Aerobic/ anaerobic (Oxygen concentration)
3. pH
4. Temperature
5. Salinity
6. Conductivity
7. Broad characteristics: sediment, rock, water

Other databases with supporting information should be identified and linked, cross-referenced. MICROBIS or ICoMM could serve as a portal to information about potential sampling opportunities that might be afforded by planned cruises and the associated investigators (i.e., provide an International BFG–**SAC working with secretariat should identify those opportunities**

Example programs that provide partial information:
RIDGE2000http://ridge2000.bio.psu.edu/HURL:
http://www.soest.hawaii.edu/HURL/hurl.html
IODP: http://www.iodp.org/
UNOLS: http://www.unols.org/
EUROCEAN: http://www.eurocean.org/categories.php?category_no=90

Relationships with other programs:

1) RIDGE, ODP, Genomes to Life, Microbial Observatories, RCN
2) What are the future opportunities that are currently funded for sampling? How can sample access be coordinated?
3) Biodiversity Organization issues: Centralized, Coordinated, Distributed, Combinational Models

How to Proceed:

4) Funding Strategies
    a. What are the highest priorities?

There are many priorities that should be addressed depending on the benthic environment of interest that are detailed below. But as a premise it should be underscored life in extreme environments is a major existing driver for scientific research. In the deep ocean that is not always fully tapped. For example, IODP has one of the 3 organizing themes microbiology there have not been any microbiological legs besides leg 201
Benthic Priorities:
Subseafloor Priorities:
E.g., 1: ponded sediment/young ridge flank at ~30°N MAR. permits study of sedimentary communities (~30 m sediment) and recovery of young (7 Ma) hydrologically active crust.
E.g., 2: low-activity (in upper water column) (30°S) vs. high activity sedimentary systems (in UWC) (~45-50°S) as a function of crust age in Southern Pacific Ocean sites.
E.g. 3: Black sea: comparing the modern microbial community with the sedimentary community, looking for evidence of degradation in-situ vs. preservation of water-column processes occurring in the present and past.
E.g., 4: Cariaco basin: green varved sediments that record an excellent paleoceanographic history and have been studied intensively for this but never been drilled for microbiology.

Coastal priorities:
Estuaries - sites of high primary productivity and potentially autotrophic pathways. E.g., Venice estuary, highly impacted, productive, historical.
Stromatolites – give us a window into the early evolutionary processes. Communities of functionally interactive microbes that display low diversity. High spatial gradients in chemisty (O2, salinity, biomass) that can be quantified in-situ. Enables functional as well as phylogenetic diversity to be assayed. Shark bay, Bahamas
Deep Sea priorties:

Exploration – vents and beyond. Current NSF funding is largely focused on long-term monitoring of select sites but does not support exploratory research. NOAA exploration is currently the only US platform that is engaged in significant exploratory research. ICoMM could help facilitate this research by emphasizing an exploratory program aimed at contributing to a microbiological survey.

Rock communities at ridge axes in bare cold rock habitats. Studies are needed to determine the succession of endolithic microbial communities on bare rock off axis of ridges as a function of age.

Bentic sediments in traps to address the issue of "sediment drift" for the purposes of looking at benthic dispersal are needed. Currently there are only sediment traps for looking at either planktonic sedimentation or deep-ocean dispersal of vent larva. No studies have yet looked at the dispersal of benthic dispersal of prokaryotes in off-axis environments.

Arctic/Antartic priorities:
Sea ice communities: the floating benthos.

Where could ICoMM seed monies be most effective towards initiating programs and collaborations in benthic systems?
Promoting and supporting finacial endeavors between international collaborators is a must. ICoMM should advocate for funding programs that support international efforts between different programs, for example, for a European anologue to Microbial Observatories to be developed.

What are the key programs for benthic studies? How can benthic diversity research be better advocated within these or other programs for multi-institutional and international programs?

In the US RIDGE, Ocean Sciences, Microbial observatories. Internationally, IODP is the only effective one – and it supports little post-cruise science. NASA supports many post-cruise studies but has no mechanism for funding cruises.

How to proceed:
5) White Papers
    a. What audiences should be targeted?

Societies (AAM, ASM, ASLO, EGS, AGU), funding agencies, sea-going organizations (WHOI, Scripps).

b. What are the important questions?

Is everything everywhere? How are microbes dispersed in the deep ocean? What is the extent of microbiology in the sub-surface ocean? What is the total biomass supported in the deep and what supports it?

c. What are the important messages?

The microbiology of the benthos is nearly entirely uncharted except in coastal systems, and even these habitats are under sampled. Because of issues related to obtaining deep-ocean samples there is an absolute critical need to have contextual information for benthic surveys of any sort and this needs to be a integral component of any integrated census effort.

Questions for ICoMM working groups

Below are questions that will guide each of the working groups. It is not exhaustive and discussion leaders are encouraged to address additional issues that might be relevant to the goals of ICOMM.

Diversity:

2) *What are the most timely and important questions regarding benthic microbial diversity?*
3) *What metric can be used to describe microbial diversity?*
4) *For molecular measures, what are the strengths and weaknesses of single-gene, genomic, and populations-level perspectives? Which of these will have the greatest long-term benefits*
5) *How should the dynamics of diversity be handled, and should the approach be based on species (phylotypes) or populations?*
6) *How will approaches to microbial diversity differ from those used in the Census of Marine Life, which focuses mainly on metazoans?*

Integration:

6) *What level of biodiversity is necessary to interpret ecological, physiological and process-related observations?*
7) *How can process and ecological data inform us about diversity and what levels of information are required?*
8) *What are the key scientific questions that a census can address?*
9) *Where are the gaps in the investigative framework?*
10) *How is diversity related to process stability? Can predictive frameworks be defined?*

Sampling, prioritization and coordination with other programs:

13) Schedules, locations and priorities.
14) Relationship to sampling program in the Census of Marine Life and other ongoing or currently planned sampling efforts.
15) Are there mileposts that will logically define phases of the project?
16) What observations are needed at each sampling site? Density of sampling: A key question or hypothesis "Is everything everywhere?" If we could demonstrate unequivocally that everything is everywhere, it will reduce the requirements for sampling.
17) How should we address temporal variations?
18) How should we address spatial heterogeneity, particularly with regard to commensal populations and chemosynthetic environments (e.g. seeps, whale falls, wood falls)?

Databases:

4) What is the structure of the information that will be produced?
5) What are the specific database needs for benthic systems?
6) What are the preferred techniques for carrying out a census in benthic systems?
7) How can databases be structured to facilitate communication?

Relationships with other programs:

6) RIDGE, ODP, Genomes to Life, Microbial Observatories, RCN
7) What are the future opportunities that are currently funded for sampling? How can sample access be coordinated?
8) Biodiversity Organization issues: Centralized, Coordinated, Distributed, Combinational Models

How to Proceed:

1) Funding Strategies
    a. What are the highest priorities?
    b. Where could ICoMM seed monies be most effective towards initiating programs and collaborations in benthic systems
    c. What are the key programs for benthic studies? How can benthic diversity research be better advocated within these or other programs for multi-institutional and international programs?
2) *White Papers*
    a. What audiences should be targeted?
    b. What are the important questions?
    c. What are the important messages?

Working Group Reports
Open Ocean and Coastal Systems Working Group

International Census of Marine Microbes:

***Open Ocean and Coastal Systems Workshop Discussions*:** University of Hawai'i at Manoa, May 10th and 11th, 2005

Summary of Workshop:

The Open Ocean and Coastal Systems Working Group discussed options for a community-based, international program to explore marine microbial diversity. A major challenge will be to develop strong consensus about the optimal sampling strategies for short-term, mid-term and long-term objectives. Through extended dialogues between members of the Open Oceans and Coastal Systems working group and leaders of the breakout groups for sampling, technology and databases, the ICoMM secretariat will seek better-defined short and mid-term strategies for consideration by meeting attendees and the general community of microbial oceanographers. Long-term strategies will require more extensive planning efforts that engage the broader community of marine microbiologists and significant increases in funding. A more detailed accounting of the workshop proceedings follows this list of ideas and concepts that most of the workshop participants strongly supported.

1. ICoMM's scientific objectives as it pertains to the open ocean and coastal environments will require high degrees of coordination and significant increases in funding.
2. ICoMM should promote international collaborations. A short-term project would be to map where groups around the world are engaged in ICoMM relevant activities, particularly in coastal waters where time series measurements are tractable.
3. To obtain maximum scientific return on resource investments, ICoMM activities must integrate studies of microbial populations with contextual information (depth, geospatial information, temperature, luminocity, a suite of biogeochemical parameters, etc. from temporal samplings at specific sites) that will inform us about the interplay between microbial mediated activities and oceanic processes.
4. A functional or physiologic census is just as important as a taxonomic census.
5. ICoMM should promote the use of common protocols and techniques that can be calibrated across different laboratories.
6. Scientific questions will drive the sampling strategy and measurement requirements. There was enthusiasm for globally distributed microbial population surveys and for intensive studies of localized ecosystems. The workshop did not reach a clear consensus about optimal sampling strategies, but there was general recognition that single point (as opposed to temporal) samplings are insufficient.
7. The differing ideas about optimal sampling strategies argue for a "nested approach" in which information is sought at multiple scales. Sampling density will depend on the efficiency of measurement technologies and available resources.

8. The ICoMM community must convince the database community about the importance of including contextual data with annotations of gene sequences. Minimal information includes latitude, longitude, time and depth.
9. Rather than a "monster database", ICoMM should encourage the development of specialized yet interdependent databases that capture phylogenetic, molecular, physiological and contextual information. These should be federated databases capable of sharing information. If the current MICROBIS strategy of "synchronizing" databases were to be adopted, information would be both shared and maintained in a redundant fashion thus ensuring its availability on the internet for future investigators.
10. There is a need to develop experimental and predictive modeling capabilities.
11. The next generation of microbial oceanographers must cross-train in marine microbiology, molecular biology, biogeochemistry and bioinformatics.
12. There is a need for an informatics-based workshop on molecular ecology.
13. ICoMM should play a pivotal role in linking programs such as the NSF's ORION and NOAA's Integrated Ocean Observing Initiative.
14. ICoMM should use the Marine Microbe Forum: http://www.sb-roscoff.fr/marine_microbes/ to exchange ideas about Marine Microbiology.

Invited Workshop Presentations:

Thirty-five marine microbiologists attended a two-day workshop at the University of Hawai'i at Manoa sponsored by the International Census of Marine Microbes' – ICoMM (http://icomm.mbl.edu). The objective was to explore options for developing an international, community-based census of marine microbes in open ocean and coastal systems. On the first day of the meeting, **David Karl** (the Chair of the Open Ocean and Coastal Systems working group) discussed the importance of knowing the genetic and functional diversity of microorganisms in the oceans and **Mitchell L. Sogin** (Co-PI of ICoMM with **Jan de Leeuw** of NIOZ) provided information about the history of ICoMM and its objectives. A series of fifteen-minute talks by eleven other participants offered descriptions of ongoing marine microbiological studies at several different sites using a range of standard and advanced technologies. (see attached Open Ocean and Coastal Systems Workshop schedule for speakers and titles).

John Heidelberg of TIGR headed off the talks with a presentation on what TIGR and the Venter Institute are doing towards marine metagenomics. Metagenomics is very much an exploratory science and does not test theories or models. John described the results of their shotgun sequencing work off of Bermuda that consisted of 4 samples and 1.6 million sequence reads. In this work, they detected 1,100 16S sequences. There is a second expedition going on called the Global Ocean Survey that in part aims to trace the steps of the Beagle and Challenger Expeditions. It began in Bermuda and is currently in Australia. John and colleagues Bill Nelson and Ed DeLong are developing a database to make the data from these voyages available to the scientific community. They are searching for ways to make the data they have collected useful to other researchers. Researchers would probably be happier with metabolic maps but

these are still a long way from becoming readily available. So far they have 7.7 million reads, 3 million assemblies and 4.5 billion base pairs of DNA.

Forest Rohwer shared his work on uncultured marine viruses that typically occur at 10^7/ml in the surface seawater. Most of these are eating bacteria – bacteriophage. They play a major role in global carbon cycles and can affect microbial diversity by killing off particular strains of microbes. Forest detailed some of the methods his lab is using to study viruses in the sea. He presented the concept of a "metagenomic species" as being defined by assembly parameters of metagenomic data. He has been using Monte Carlo Analysis to predict the numbers and relatedness of viral genomes. For example, a marine sediment sample has around 10,000 viral genotypes for 1 kg of sediment and 200 liters of seawater contains 5,000 genotypes. Forest has been working on making dynamical programming available to everyone. He has also been exploring how the Power Law appears to be a better predictor of the number species in a sample than other models. Forest concluded by sharing some his work on corals where he has discovered a lot of host-specific microbial populations.

Heidi Sosik's research took the workshop beyond the traditional tools of bottles, nets and filters to *in situ* flow cytometric methods that allow her to ask questions about what regulates coastal phytoplankton populations. She and colleagues Rob Olson and Alexi Shalapyonok have developed the FlowCytobot – a flow cytometer that detects individual particle properties such as fluorescence and light scattering and works underwater in an autonomous mode to provide real-time access to data. FlowCytobot is deployed at the Martha's Vineyard Coastal Observatory, a facility that is available to any users who want to test new sensors in a setting with high power and data bandwidth. Most recently, Sosik and colleagues have developed a second-generation instrument, Imaging FlowCytobot, that combines aspects of FlowCytobot capabilities with in-flow single cell imaging techniques. Using these tools, they have been asking questions like: What causes inter-annual variability and what processes lead to this kind of variability? Some cell identifications can be made down to the species level. She is also interested in what is the ecological entity, how does it change and when does it matter.

William Li reviewed some of the recent reports in the literature that address issues surrounding microbes and spatial scaling. He pointed out that a compilation of extant microbes listed in textbooks might fall short of a true census because the concept of species is problematic. Bill outlined recent studies on soil fungi and salt marsh bacteria in which the taxa-area relationships were used to extrapolate from local to regional scales. The slopes of these relationships were low, indicating that taxonomic richness is not greatly dissimilar at different scales, suggesting a ubiquitous distribution of many microbes. However, more recent studies of bacteria in water-filled treeholes and of phytoplankton in limnetic and marine systems indicate that the slopes are much higher in non-contiguous habitats. In other words, diversity at local scales may not be easily extrapolated to the global scale. The taxa-area issue remains unresolved for marine microbes. Bill suggested that Alan Longhurst's concept of the

biogeochemical provinces in the ocean might be one way to focus our census efforts. This approach has been used to scale up primary production from the regional to the global scale. For the census of marine microbes, it therefore seems worthwhile to understand the patterns and mechanisms that relate microbial diversity to primary production.

Carlos Pedrós-Alió described ways that we might use remote sensing to ask questions relevant to the census. In illustrating this point, Carlos summarized some of the work of Rafel Simó and colleagues who are using remote sensing to test some parts of the CLAW hypothesis formulated by Charlson, Lovelok, Andreae and Warren (Nature326:655-661, 1987): that dimethylsulfide (DMS) plays a role in regulating the temperature of the planet. Dimethylsulfoniopropionate (DMSP) gets converted to DMS (a volatile compound), the main source of biologically formed sulfur in the atmosphere above the oceans. Phytoplankton produce DMS that escapes into the atmosphere where it is oxidized to sulfuric acid, acts as a nucleus for the condensation of water and ultimately contributes to the albedo of the planet. When albedo increases, less solar radiation reaches the microbial plankton populations resulting in less photosynthesis and less DMS production creating a feedback loop that contributes to the regulation of the Earth's temperature. These scientists found that if the mixed layer depth is very shallow, then almost 100% of DMSP is converted into DMS, and as the mixed layer depth increases this value goes down. Using the mixed layer depth, chlorophyll concentrations and the DMS relationship, these investigators were able to show that predicted DMS concentrations were nicely correlated with the real DMS concentrations. Carlos challenged the workshop participants to see if there are other scientific questions that can answered using remote sensed data beyond the chlorophyll and DMS examples. In summary, Carlos emphasized that you don't have to be an expert to take advantage of remote sensing data and that even though what you are measuring may not be directly related to remote sensing, there may be ways of correlating your data with remotely sensed data to enhance your area of research.

Osvaldo Ulloa presented his work on low oxygen minimum zones off of Peru and Chile. Oxygen minimum zones (OMZ) are defined as those zones with oxygen concentrations of 0.5ml O_2/L or 22 μM. They are typically distributed off the eastern coasts of the ocean. Some reports indicate that they are areas of low diversity but no one has looked at microbial populations in detail. Some of the questions facing this environment include: How do intermediate waters vary in OMZ's, and how stable are they? Osvaldo's group has found that OMZ's are not stable and that they vary over large scales and geological time scales. OMZ's need to be studied in relation to the physical forces that shape them. Osvaldo's group has been examining the nitrogen cycle in OMZ's. He has been detecting high rates of nitrification, a high diversity of denitrifiers, and the presence of anaerobic ammonia oxidizing bacteria that may be a significant sink for oceanic nitrogen.

Gordon Taylor presented his work on Microbial Processes in the Cariaco Basin. The Cariaco Time Series is a collaboration between three U.S. institutions (U South Florida, U South Carolina and Stony Brook U) and three Venezuelan institutions (Fundacion La Salle de Ciencias, U de Oriente and U de Simon Bolivar) that has been ongoing since 1995. The Cariaco's setting is along a productive coastal margin, prone to strong seasonal upwelling that makes it a very dynamic portion of the coastal ocean. Annual primary production is three times that found at the subtropical BATS and HOTS stations and carbon fluxes exported from the epipelagial are twice as large. This is the world's largest truly marine anoxic basin and the only US-sponsored time series in the tropics. The Cariaco Basin is considered a natural laboratory for studying the biogeochemistry and microbiology of an anoxic system, where organisms may have novel metabolisms and physiologies. It may also contain evolutionarily significant phylotypes. Two Microbial Observatory programs have recently been added to the Time Series Program, representing collaborations between scientists from Stony Brook University, University of Louisiana at Lafayette, Northeastern University, Universidad de Simon Bolivar (Caracas, VE), WHOI and MBL. These projects are examining both prokaryotic and eukaryotic (protist) community dynamics using a variety of molecular, cultivation and manipulative experimental techniques. Their goal is to better understand how geochemical gradients organize microbial communities in this sulfidic, oxygen-depleted environment. Methods being exploited include CARD-FISH, MICRO-FISH, DGGE, T-RFLP, SSU rDNA libraries and oligo-FISH combined with SEM for protistan molecular and α-taxonomies (Stoeck, Fowle & Epstein 2003). They are finding many prokaryotes related to organisms isolated from hydrothermal vents, cold seeps and sulfur-dominated habitats. Many novel protistan 18S rDNA sequences have been recovered from anoxic waters, including protozoa related to known anaerobes from animal guts and shallow anoxic habitats as well as a novel deeply-branching clade with no known close relatives. In addition to novel thiosulfate disproportionating prokaryotes, cultivation studies have yielded the first cultivable thiosulfate-oxidizing manganese oxidizing chemoautotrophic bacterium.

David Caron spoke of his joint Microbial Observatory (MO) with Jed Fuhrman that has worked cooperatively with SPOTS (1997). SPOTS has provided a wealth of contextual information for the MO. The MO started in 2000 and emphasizes prokaryotic and eukaryotic discovery based on diversity studies sampled on short and long time scales. The San Pedro basin is hypoxic below 150 meters. Samples are taken at four depths from the surface through this hypoxic zone. They have used a combination of flow cytometry, epifluorescence microscopy, and a combination of SSU clone libraries and ARISA, TRFLP to explore aspects of microbial diversity from viruses to eukaryotes. They have discovered 800 distinguishable types of the SAR 11 bacterial clade. They are exploring the relationship between prokaryotic and eukaryotic presence in relation to predator/prey and symbiosis relationships or other associations. Although many protistan taxa are morphologically defined, they still represent a large diversity. Dave and his group have been trying to determine a proxy

of similarity that can be used to separate most species of protists in GenBank – they have settled on 95%. Of 1200 clones they found 488 OTUs with 95% similarity cut-off – so most of the species are present but rare. It may be that what's out there is present but is rare. Dave pointed out a couple of myths that are not true, 1) we have a good estimate of protist diversity, 2) we can forget about the species concept. We have to keep in mind that species are the unit of evolution. The definition of species is a matter of perspective.

Mike Zubkov presented work on the Atlantic Meridional Transect (AMT, NERC UK funded consortium) that extends from the British Isles to the Falkland Islands. Zubkov and colleagues have been using flow cytometric methods to sort cells, followed by PCR, sequencing and subsequent phylogenetic probe construction to return to the environment and identify the organism of interest to couple their identity with biogeochemical processes. They have been using tracers followed by cell-sorting to estimate the rates of amino acid uptake at ambient concentrations – this allows evaluation of microbial population growth in situ. Questions being asked include: Do *Prochlorococcus* populations in the surface waters rely on organic nutrients (such as amino acids) more than those at depth? On the contrary they found that *Prochlorococcus* living in the deeper waters was consuming 50% the amount of methionine tracer given compared to 25% of the methionine given to those *Prochlorococcus* living in the surface waters. Deep cells are taking up more of the methionine than average prokaryotic cells in these waters. This approach allows us to learn about the physiology of these cells in situ. They also found higher *Prochlorococcus* cell activity in temperate waters than inside the Gyres. In a different study they examined the mesoscale spatial variability in picoplankton in the Celtic Sea and found that cell concentrations of *Synechococcus* and heterotrophic bacteria vary up to 50-fold over distances as short as 12 km. Advection of such spatial variability through a time-series site would therefore constitute a major source of 'error'. Consequently, attempts to model and to investigate the ecology of globally important microorganisms in situ must take into account and quantify the hitherto ignored local spatial variability as a matter of necessity.

Matt Church reviewed the on-going efforts and progress at the Hawaii Ocean Time series (HOT) program. Based on the nearly 17 year record of monthly observations at the field site for the HOT program, Station ALOHA, investigators are beginning to assemble information on temporal variability in plankton processes and biogeochemical cycling spanning diurnal to decadal time scales. Some of the questions that have guided research at Station ALOHA include: "What are relevant time scales for us to study phytoplankton diversity?" and "Are these time and space scales the same as ones that control biogeochemical cycles?" In 1988, Dave Karl and Roger Lukas initiated observations at Station ALOHA under the auspices of the JGOFS (Joint Global Ocean Flux Study). The central objectives of the HOT program are to characterize time-dependent dynamics at the Station ALOHA. The following points summarize the findings to date: 1) prokaryotes dominate the system, 2) oceanic biology regulates nutrient stoichiometry and carbon export, 3) the surface ocean

appears chronically oligotrophic, 4) nitrogen fixation plays an important role in nutrient dynamics and carbon export, and 5) the organization of plankton populations seems to be controlled by longer-term oceanic teleconnections. Church urges the microbial oceanographic community to consider time series stations in future census of marine microbe efforts. The existing data sets available from time series stations provide us with a framework for a census. If we are interested in examining how diversity maps onto biogeochemical cycling we need to sample at the appropriate time scales to capture both short term population dynamics and longer term (decadal and inter-decadal time scales) ecosystem transitions.

Craig Carlson reviewed progress on the Oceanic Microbial Observatory, a microbial observatory associated with the Bermuda Atlantic Time Series (BATS) station. The BATS program, on their 200th cruise this year, provides the relevant biogeochemical backdrop for this observatory. The work represents a collaborative effort between Carlson's group at UC Santa Barbara, Steve Giovannoni's group at Oregon State University and some of the research staff lead by Rachael Parsons at the Bermuda Biological Station for Research (BBSR). Objectives of this observatory were to 1) identify spatial and temporal patterns in specific bacterioplankton and prokaryotic populations and 2) to initiate experiments to investigate potential linkages between microbial processes, community structure and biogeochemical processes and events (mixing events, changes in nutrient fields, etc.) with an emphasis on discovery. Giovannoni's role has been to bring some of the uncultured bacteria into culture through low-nutrient, high-throughput extinction culturing methods. Giovannoni and Rappe recently succeeded in bringing SAR11 into culture. There are also ongoing outreach efforts geared at education in the form of summer courses at the BBSR – John Heidelberg and Steve Giovannoni offer a course in Marine Genomics and Bob Morris and Craig Carlson offer a course in Marine Microbial Ecology. BATS is located 80 km southeast of the island of Bermuda – in the Northwestern Sargasso Sea and is characterized by seasonal oligotrophy and annual patterns of temperature availability and mixing. Carlson and his colleagues are using the BATS long time series data to provide focus to questions on microbial diversity. Their carbuoy mesocosm experiments wherein deep water was inoculated into surface water media produced interesting trends of significant bacterial production and removal of DOC (3 to 5 μM). This suggested a physical separation of a zone of DOM production from a zone of DOM remineralization that may be related to the microbial community present. They are also getting quantitative data using FISH for specific clades of bacteria – SAR11, *Cytophaga* and *Roseobacter*. Differences exist– SAR11 makes a major contribution and *Cytophaga* can be up to 10 to 15 % depending on the time of the year and position in the water column. Giovannoni has >2,000 strains of bacteria that won't grow on agar – 18 were selected for genomic sequencing by the Moore Foundation.

Comments about ICoMM goals and relevant issues:

The second day of the meeting provided opportunities for participants to raise additional questions and issues that potentially impact ICoMM and its goals. **Daniel Vaulot** discussed the formation of an electronic Marine Microbe Forum at the address: http://www.sb-roscoff.fr/marine_microbes/. **Lucas Stal** presented data about distribution and measurement of N_2 fixation by marine microbes and **Ricardo Letelier** discussed the role of remote sensing and the need to integrate very large data sets of remote sensing of physical and biological data. In the near future we can look forward to even richer satellite data sets with hourly observations. **Oscar Scholfield** summarized a series of new initiatives that would benefit from the microbiological activities within the ICoMM community. Most are tied to the Ocean Observing Initiative and recognition that we need to develop a nested observing network strategy to look at overlapping scales. New initiatives such as the ORION program will provide support for development of infrastructure. A partner program NOAA's IOOS Integrated Ocean Observing systems together with ORION may have as much as 500 million dollars/year to invest and ICoMM could play an important role in linking these communities together. Some of these programs will support research on high power moorings distributed on a global scale. There will be regional Arrays i.e. fiber optic nodes. Neptune will provide access to the deep ocean and there are proposals for Coastal networks that will facilitate long-term time series measurements. Schofield will provide to ICoMM a more detailed summary of new opportunities in the Ocean Observing Initiatives including contact information for lead scientists who might facilitate collaborations and new research opportunities in microbial oceanography. Finally, Dave Karl reminded the workshop that we would soon be celebrating the International Polar Year and that ICOMM should be involved. This could be the year of the microbe!

Before adjourning to individual working groups, **Grieg Steward** challenged the meeting participants to think in terms of a real census of microbes so that we might understand their ecology and interpret genomic data etc. Rather than an exercise in stamp collecting, the census represents a phase of discovery. We need the census to fully take advantage of new information from genome sensors, gliders etc. The new tools coming on line are very powerful but sometimes inefficient. Until recently, surveys were limited to 16S rRNA sequence comparisons. Yet, if we want to address functional diversity, this is not enough to make progress. Metagenomics and the hope of inferring metabolic potential is both promising but inefficient science. One day single cell genome projects may partially circumvent this inherent inefficiency, but it will still be necessary to first "tease apart" microbial community structures before we study their DNA at the genomic level. This initial characterization must include morphology, biochemistry, analysis of lipids and detailed information about microbial population structures. We must be able to sample molecular diversity at levels required to detect minor members of populations. This will require technology developments that reduce the cost of molecular diversity surveys. The minor population members must be important for survival of the community otherwise they would not survive over the long term.

Specific science questions will dictate the details of sampling and measurement methods. It would be highly desirable to carry out time course studies of spatial grids at a large number of sites, but given available resources, such a scenario would not be attractive for funding. Difficult decisions must be made through workshops and community meetings, ICoMM can contribute its collective wisdom towards identifying the most important

priorities. For example, questions about global change will require more ocean sampling on global scales. Alternatively, if the goal is to understand how microbes are catalyzing transformations, it becomes necessary to study a particular system at greater levels of detail in order to understand likely emergent properties.

Goals for working groups in the workshop:

The participants formed three break-out groups including: Group 1 - **Sampling**, Group 2 - **Measurements**, and Group 3 - **Data analysis and training**. In addition to the general questions assigned to all of the ICoMM working groups (Benthic, Open Ocean and Coastal Systems, Technology and Scientific Advisory Council) each breakout group addressed a specific set of questions. Dave Karl and Mitchell Sogin emphasized the importance of focusing discussions on "**HOW**" to collect, organize and analyze relevant data, as opposed to arguing more philosophical questions such as "What is a microbial species?" The organizers also reminded the participants that Big questions in science demand community-based efforts similar in scope to those used in the Physics, Astronomy and Genetic community for tackling seemingly impossible tasks such as the construction of high energy Accelerators, large earth-based and orbiting observatories, and the full sequencing of the human genome. Microbial Oceanography is likely more complex than any of these disciplines because it requires massive amounts of data (genetic, physiological, biogeochemical), major computational capabilities, extensive modeling and BIOLOGY. Any strategic plan developed through this workshop should consider resource constraints available in the short term (immediately available capabilities within the community), the mid-term (resources that can be competed for from existing funding sources) and long term – (large scale resources that will require broad-based support from an international constituency of scientists and policy makers at the highest levels of government and the private sector.) Finally, ICoMM's goal extends far beyond the mere identification and counting of different kinds of microbes in the sea. To obtain maximum scientific return on resource investment, ICoMM activities must integrate studies of microbial populations with contextual information that will inform us about the interplay between microbial mediated activities and oceanic processes.

Plenary Discussions:

The general plenary discussion prior to reports by the breakout groups asked the question **"What is the goal of the census?"** Is it a description of all organisms in all marine environments? (Genotypic and phenotypic diversity) Is it a description of microbes at a selected latitude, longitude and depth? (Descriptions of microbial populations at one or more geographical locations) or Is it a description of organisms that are associated with a particular process? (Functional diversity). Should the census include direct counts of microbes and measurements of chlorophyll? Does it correspond to enumerating subgroups and clades of *Prochlorococcus*? All of these issues are important topics for marine microbiology and each will contribute to an International Census of Marine Microbes. There was strong agreement that the science questions will drive the sampling strategy and requirements for specific measurements. There was enthusiasm for both a globally distributed survey of microbial populations and for intensive studies of localized ecosystems. Key questions include "Are all

microbes everywhere? How is the microbial world connect to circulation patterns? What roles do microbes or microbial activities as defined by the genome play in major biogeochemical processes? One of the strongest rationales for supporting ICoMM research initiatives relates to the development of tools for understanding emergent properties of marine ecosystems. The task that this workshop should address is to set realistic priorities for allocating existing and future resources.

Breakout Group 1. Sampling. (William Li - Group Leader, David Kirchman - Rapporteur, Matthew Church, David Caron, Gordon Taylor, John Waterbury, Osvaldo Ulloa, Daniel Vaulot, James Cowen, David Karl, and Lucas Stal).

The scale of sampling (spatial and temporal) and the ability of analytical techniques to discriminate between different kinds and numbers of microbes will define the resolving power of an International Census of Marine Microbes. The "Sampling" working group considered the following issues:

1) *To address questions of diversity and distribution of different kinds of microbes: An important goal of these investigations is to address the question* ***Is everything everywhere?***

Should we explore many sites in a superficial manner?

or

Should we focus upon two or three intensely studied oligotrophic sites?

2) *What is the optimal sampling strategy for an intensely studied site?*

3) *What is the optimal global sampling strategy?*

4) *How shall we address the issue of temporal sampling?*

5) *What is the relevance of quantitative measurements and what is the best available methodology?*

6) *What is the known role and what might be the role of viruses in shaping microbial population structures? – What measurements are needed and how can they be obtained?*

Discussions in the sampling group initially considered the enormity of the task at hand both in terms of total volume of the oceans (estimated to be ~1-4 X 10^{18} m^3) with a potential population of 10^{29} microbial cells. Strategic efforts to sample microbial populations in the water column according to a selected measurement criteria must consider tradeoffs between high density studies with many data elements from a small number of sites versus the cursory analysis (fewer data elements) of many samples that are either globally distributed or that span smaller scales at one or a few sites. The group also considered the effect of collecting large sample volumes versus a larger number of small samples that can provide increased levels of spatial and or temporal resolution. If the goal is to understand the relationship between microbial populations and global models, investigations of local processes orchestrated by smaller scale distribution patterns of microbes may be of minor importance.

Importance of temporal measurements: The working group used *Synechnococcus* studies to examine some of these tradeoffs. For these organisms, transects from South American to the North Atlantic were not informative because snap-shot measurements inherent in transect sampling will miss temporal variation. The studies of seasonal cycles off Woods Hole were more instructive and diel cycles provided a great deal of information about *Synechnococcus* biology and ecology. These results argue for the temporal sampling of a limited number of sites instead of broad-scale transects.

Limitations of single site studies. Analysis of *Synechococcus* populations only in Woods Hole waters provides a counter example because such a geographically restricted temporal sampling scheme will miss *Prochlorococcus*. This suggests that a global description of a few key microbes would be of value, but it must include a temporal component. Comparisons of time-series studies of microbial populations at multiple sites are important and if a sufficient number of sites are included it becomes easier to establish international collaboration.

These differing points of view argue for a "nested sampling approach" in which information is sought at multiple scales. Sampling density will depend on the efficiency of measurement technologies and available resources. For example, cytometric analyses can be adapted to time series investigations whereas the expense of metagenomics restricts the application of this technology to snap shot studies.

An important role for ICoMM would be to "help establish international collaborations". This would mitigate regulatory problems that foreigners encounter when working in territorial waters. A short-term project would be to map where groups around the world are engaged in ICoMM relevant activities, particularly in coastal waters where time series measurements are more tractable. As part of this survey a questionnaire could determine where and the frequency of sampling, the target organisms, the methodology and willingness to share data and engage in collaborations. This could be the first step in carrying out a global census of marine microbes. ICoMM should lead this effort and it should take advantage of existing organizations and networks such as Antarus for South America, and POGO. This effort would benefit from a contact person in different geographical regions. ICoMM would collect this information and collate results on its web site. It could also promote pilot projects to prime the pump for mid-term and long term projects.

ICoMM should also promote the use of common protocols and techniques that can be calibrated across different laboratories. For example it is possible to calibrate methods for obtaining direct counts, DOC and nutrient measurements in different experimental environments.

Mid-term objectives: The working group endorsed a mid-term objective of using commercial shipping for obtaining survey data from surface waters and the development of centers that could carry out technologies not available in some marine biology laboratories. The community should also take full advantage of the International Polar Year (2007-2008) to work in the Arctic and Antarctic waters.

Long-term objectives: The working group discussed long term projects and the need for remote "microbial sensors" which currently have very limited capability for monitoring biological parameters. There is also a need to develop experimental and modeling capabilities that are predictive. Current models define rates but lack ability to differentiate activities of

different kinds of microbes. They also explored how many biomes would be suitable for detailed analyses. The number of recommended sites ranges from 2 to 200 with approximately 20 representing sites that already have historical information including biogeochemical data. A list of candidate sites includes: 1. HOT, 2. BATS, 3. KNOT, 4. Station P, 5. EAST, 6. 48 N, 7. Canary Island, 8. COPAS, 9. SEATS, 10. CATS, 11. Antarctic, 12. Leo, 13. ORION - part of the coastal ocean observatory work, South Pacific - off Easter Island-currently not under investigation. Cost estimates for each site includes on the order of $1 million/site for ship time (assumes monthly sampling) and at least $1 million for laboratory work but potentially much more if the studies include large scale molecular investigations. Each of these sites would operate for ten years but at least 10 of the sites would continue on a longer term.

The needs for such an ambitious program would include developing an appropriate labor force and securing resources for ship time. The working group did not recommend specific sampling frequencies beyond setting a minimum of biannual collections at each of these international sites.

Challenge to the Open Ocean and Coastal Systems Working Group plan for sampling. The sampling group recommended a tractable, long-term strategy that will require increased funding levels. It could be part of a larger oceanography program such as ORION. However, the short and mid-term objectives lack concrete ideas that will lead to new data. The short-term recommendation (1-2 year time horizon) of *mapping the location of marine laboratories and who is doing what* is valuable but this activity will not demonstrate the collaborative strengths of the microbial oceanography community. If the ICoMM community does not believe that it is possible to collect new data over the short-term horizon, then at a minimum its members can work together to organize existing data sets and derive a new synthesis that will point the way towards new projects. If this is to become the short-term goal, we must identify data sets that are most worthy of integration.

We had aspirations that it would be possible to identify a community-based project that might be accomplished for something less than 1,500,000 dollars. Such a project would require identification of an important site or perhaps up to three sites and the analysis of samples with some agreed upon technology. The alternative would be to select a larger number of sites that are under investigation by different laboratories and to contribute samples or data for analyses using a community-based resource such as sequencing power or DNA micro-arrays in concert with the collection of contextual data. It might be possible to distribute a common set of tools to individual laboratories for making measurements at many sites. The question is should we be more restrictive in our scale of sampling by increasing the frequency and/or density of sampling? Further discussions are necessary to make decisions that will allow us to go forward with a credible short-term plan.

Group 2. Measurements. (Jonathan Zehr- Group Leader, John Paul-Rapporteur, Heidi Sosik, Claire Mahaffey, Michael Zubkov, Grieg Steward, Michael Rappé, Robert Bidigare, Valerie Franck, Karen Selph, Craig Carlson)

To make informed recommendations about optimal measurements for ICOMM, the Measurements breakout group considered the questions that ICOMM should address. An

important goal would be to gain new insights about how the ocean works, information about evolution, and discoveries of novel microbial forms. In addressing the seven specific questions outline below, they considered the counterpoint position that knowledge about diversity might not lead to improved understanding of biogeochemical cycles and underlying processes. Perhaps the concept of a functional or physiologic census is just as important as a taxonomic census. The working group did not elaborate on this thread, but the inference of metabolic function from genomic surveys provides a way to infer general physiologic and functional properties of a microbial population. At the same time, the genetic basis of these phenotypic traits often reveals phylogenetic affinities. A census based upon sequence data offers a means to organize and interpret information from functional and physiologic investigations. Taxonomic or genomic surveys versus a physiologic census may require disparate kinds of measurements. The density and efficiencies of measurement techniques coupled with the sampling plan will ultimately determine the cost of ICoMM related activities. The Measurements breakout group offered answers to each of seven specific questions.

***1)** When is it important to quantify numbers of a kind of organism and what is the best way to do the quantification?*

If the goal is to detect and identify pathogens such as harmful algal blooms (HABS) or other pathogens such as *Vibrio*, knowing the identity of the organisms in a sample is essential. If the goal is to associate discrete biogeochemical changes with a list of processes, knowledge about the organisms responsible for catalyzing transformations and the presence or absence in a microbial population becomes important. Finally, if new genes or phenotypic characteristics are discovered e.g. discovery of proteorhodopsins, knowledge about the organisms that harbor these genes offers a means to infer their presence in other populations and to understand the evolution of these physiological capabilities.

***2)** When is it more important to survey diversity and what are the optimal measurement criteria?*

Early in a census it is important to obtain an initial measure of diversity because this can guide the ultimate sampling design and measurement technology. As part of this effort, simultaneous analysis of processes can be very informative when evaluating diversity of a population and/or how that diversity relates to functional capabilities of the population. One significant challenge is, "How do we detect/quantify the rare organisms in a survey?" One answer is to use subtractive hybridization techniques that remove the dominant organisms from a population. The group discussed the importance of minor members in microbial populations. These organisms might serve key functions but represent very limited biomass. Alternatively their low numbers might reflect patchiness in space and time.

***3)** What is the role of genomics and molecular evolution in assessing diversity and how should this be accomplished?*

Molecular evolution provides an important tool for identifying microbial diversity in a studied environment. It also provides a qualitative genetic tool for directly comparing microbial populations in different samples. Microbial ecologists commonly use comparisons of Ribosomal RNAs for such analyses but other genes can also be valuable – e.g. comparisons of genes involved in nitrogen fixation, sulfur metabolism, etc. Yet there are

limitations to single gene comparisons and much more information about a particular organisms emerges from studies of large DNA inserts that contain homologous genes. By surveying the sequence neighborhood of a functional or structural gene used to select a large DNA insert for analysis, it becomes possible to discover new functional processes and/or understand evolutionary history of a functional or structural gene. In other words the study of large DNA inserts provides genomic contextual information required to understand how individual genes or gene families have evolved. Genome sequences assembled from pure cultures or from properly executed metagenomic investigations opens the door for understanding function and discovery of novel biochemical diversity.

4) *Let's assume we have 1 million reads available. What fraction of these reads might we use to look broadly at genetic diversity? What fraction might be just 16S? What fraction might be used for genes responsible for specific metabolic activities and which genes are most important? What fraction should be devoted to metagenomic studies and what should be the format – short insert clones? Long insert clones? Fosmids? Etc.*

One million reads is equivalent to shot-gun sequencing 8-10 X coverage of 100 microbial genomes. Single-cell sequencing of uncultivated microbes may soon become tractable but the selection of which single cells to sequence would be challenging. Should we select the 100 most abundant organisms? Should we focus on the 100 organisms that serve the most important biogeochemical functions? From an experimental or technical perspective, how would we make decisions about which single cells should be selected for genome analysis? The most likely method would be to construct single-cell genome libraries and to make decisions according to annotations of a small but preliminary set of sequencing reactions.

5) *What physiological measurements can be made to assess diversity?*

Physiological measurements can be made on single cells or in batches. However there is a need to develop and carry out experimental manipulation studies.

6) *Genome-enabled sensing role via arrays?*

Arrays would be a powerful tool for the community. Whole genome expression arrays could be developed for individual organisms but arrays for functional genes would also be of considerable value for assessing diversity of biogeochemical function.

7) *How important is culturing? And how can we improve our ability to culture marine microbes?*

Culturing capabilities are very important – It makes possible genomics for single organisms and the ability to link genomics to physiological and biogeochemical functions.

The Measurement breakout group considered additional topics that ranged from sampling strategies to methods. There was enthusiasm for sampling along natural gradients including trophic gradients, salinity gradients, depth, and oxia/anoxia). These functional gradients are forcings that will influence the compositions of microbial communities. It is also important to consider fine scaling questions such as the distribution of microbes in organic matter continuums and gels that provide an organic structure to microbial communities. A more general issue for sampling is size-fractionation to separate eukaryotes from prokaryotes.

The group discussed the need for standardization methods and common tools for use by microbial oceanographers. For example Denaturing Gradient Gel Electrophoresis (DGGE) is capable of concurrently detecting single base differences over a few hundred base pairs for many homologous genes but the technology is highly variable in a single laboratory. If the community is to continue use of this technology, there would be merit in establishing common standards.

The breakout group discussed the potential use of DNA macro arrays as an alternative to more expensive microarrays. It is theoretically possible to produce functional gene macroarrays with standard target genes such as nitrogen geochemical genes. Such arrays could be used to interrogate communities across large spatial scales (polar to tropics). The use of arrays in general could lead to finer temporal/spatial scale resolution). It is possible to produce large numbers of DNA microarrays by using lithographic, printing or spotting technologies. Densities can be greater than 50,000-probes on a chip. When the cost of a well-designed chip is amortized over the research activities of many laboratories, genomic and expression profiling becomes an affordable tool for microbial oceanography laboratories. Through further discussion, the working group could determine which targets (phylogenetic probes versus functional probes versus diversity probes) would be of the greatest value to the microbial oceanography community. Scanners for interpreting the results could be operated in a small number of internationally distributed laboratories. Automated scanners could also be operated and managed as virtual instruments.

The ICoMM community must also consider the archiving of samples from time series measurements and/or the storage of extracted nucleic acids. Resources would be required for maintaining and distributing archived samples and nucleic acids. Culture collections such as the ATCC carryout similar activities for the genome community but given the number of samples that might be generated by the census, the costs will rapidly escalate.

If ICoMM were to focus upon "activity" as a definition of microbial diversity, it will be necessary to develop high throughput assays for productivity. Perhaps specialized laboratories that focus upon standardization of activity measurements could train researchers in particular techniques. In this regard, there was enthusiasm for the development of single-cell activity assays. It would also be desirable to link activity to genotype. This might be accomplished by radio-labeling cells and seeing what genes are tagged through measurements of radioactive RNA. Non radioactive labeling techniques would also be of considerable interest.

What measurements should be made-time series vs. survey?

Time Series	**Survey**
JGOFS-type measurements	JGOFS-type measurements
1°Production	1°Production
Bacterial prod	Bacterial prod
Abundance(viruses-to eukaryotic)	Abundance(viruses-to eukaryotic)
N-fix	N-fix
HPLC pigs	

Metabolic indicators (ATP)	
16SrDNA /TRFLP/etc	16SrDNA /TRFLP/etc
FISH samples	
DNA samples (size fractionation)/large insert lib	DNA samples (size fractionation)/ large insert lib
RNA/Transcriptome Analysis	
Organic Matter characterizations.	
Functional gene/array analyses.	
Extinction Culture (occasionally)	

Group 3. Data analysis and training. Molecular sequence data deposited in public archives will serve a key role for molecular studies in all areas of molecular microbial ecology and for ICoMM related activities. The datasets for microbial oceanography are both interdisciplinary and very large. For example the Sorcerer Cruise is generating enormous trace archives, metagenomic assemblies in the form of sequence bins and super scaffolds, and annotations that may change in response to improvements in both data quantity and our ability to more accurately assemble environmental shotgun sequences. These data are collected along with measurements of cell counts, nutrients, temperature, oxygen, GPS coordinates and time. The Data group considered the following questions to guide their discussions.

1) *What data elements are most critical to developing a census of marine microbes in the open ocean and coastal systems environments?*

In order to compare samples, we need to coordinate measurements in terms of latitude, longitude, time and depth. The molecular databases do not currently support these attributes. The ICoMM community must convince the database community about the importance of these fields. The alternative will be to construct parallel databases that capture both Gene identifiers (GID numbers) and the "big four" parameters – Latitude, Longitude, Depth and Time. As outlined by the Benthic and Technology working groups, scientific return from ICoMM investments will depend upon the capturing of additional data elements including biogeochemical information and contextual data that describes the studied ecosystem. For example, if the census includes studies of commensal populations associated with metazoans, it will be essential to include taxonomic descriptors for the host.

2) *What modes for data analysis are optimal for a census of marine microbes?*

The collection of molecular data by the ICoMM community will require many kinds of analysis including annotation of microbial genomes, phylogenetic analysis of genes and genomes, and integration with contextual and biogeochemical databases. Because of the large amounts of information and interdisciplinary nature of the scientific questions, the databases are likely to be specialized yet interdependent. They will contain both archival and dynamic databases. MICROBIS is an example of how federated databases can be integrated across different laboratories. The underlying structure of MICROBIS provides a "turn key" solution for multiple laboratories that gather taxonomic data. It provides the ability to link both in and out to other databases containing molecular, positional, biogeochemical etc. information.

MICROBIS forms linkages to other taxonomic databases that contain all of the information in MICROBIS but only serves selected data elements. In return, MICROBIS harvests information from the partner database in order to expand its total information content. This organization of federated databases not only provides for the sharing of software capabilities but also provides for redundancy if the databases are kept in a synchronized state. Data elements within MICROBIS and between its partner databases are linked by a taxonomic naming system served by the uBIO project at the MBL. This allows for alternative taxonomic assignments that reduce to a common numerical identification. The structure can also accommodate other conventions such as the use of GPS, depth and time of sampling coordinates to link together sample information. For such databases to function its data elements and identifiers must be portable to other systems. There was a general consensus that the databases should be distributed and that a "monster" database is not a viable concept.

The GMOD (Genome Model Organism Database) consortium offers another example of federated databases that have the potential to share functional capabilities and information. The combination of databases such as MICROBIS (image rich microbial diversity descriptions), GMOD (graphical displays of genome annotations) Micro-Mar (graphical displays of global distributions of rRNA sequences in marine environments) and ARB (multiple sequence alignments and phylogenetic representations) could provide a basis for a powerful network of federated databases for molecular genomic data linked to archival information from GenBank, cruise data and other sources of contextual information. Currently MICROBIS is designed to function as a "switching" or "integrating" system for linking together databases that share common data descriptors e.g. taxon ids, gene ID's or sampling coordinates. Analytical platforms such as GMOD can be readily integrated into MICROBIS but current phylogenetic content of ARB will require significant changes to its database design. To maintain and operate this federation of databases it will be necessary to establish common procedures of synchronization and data sharing, to train researchers in the use of these databases and associated tools, and to establish a biology-driven community of software developers for implementation of new analytical tools.

3) *How should census data be integrated with contextual databases?*

The ICoMM community must encourage NCBI to either host appropriate molecular databases and/or provide fields for integration into other data repositories that collect relevant non-molecular data. The simplest mode for achieving this goal would be to require the submission of GPS coordinates, depth and time elements when environmental molecular data is submitted. For all of the databases, information must be released within time frames agreed to by the community of microbial oceanographers.

4) *How important is quantitative data for building ecological models?* Not discussed in detail but modeling is an important outcome if ICoMM and microbial oceanography are to have significant impacts on interpretations of Ocean Observing systems and management activities.

5) *Can we assume there will be a succession of bacterial groups under specific environmental conditions, for example a phytoplankton bloom?* This issue was not discussed.

6) *How can we integrate new information with what is already known and is worth assembling into a database?* The capturing of existing information into on-line databases was

discussed but represents an open-ended challenge that the entire biological community currently faces. There are efforts to digitize journals that appeared before the age of digital publishing. We anticipate that web-crawlers, and other agents will play important roles in extracting valuable information from these future resources.

7) *Training issues: How can we move forward with integrating the next generation of marine microbiologists with the challenge of bioinformatics? – What are the optimal training environments? Should they be oriented around genome centers or around marine microbiology laboratories?*

For ICoMM related activities, the key to education will be "cross-training". There are a variety of options that range from post-doctoral rotations across fields to special graduate and post-doctoral summer courses that include an equal number of faculty and students. The emphasis should be bioinformatics but should also include laboratory or ship time in order to maintain connections for the enormous scope of field and biochemical work required for ICoMM related efforts in microbial oceanography. There was a strong suggestion to develop a workshop on molecular ecology – or a mixed community workshop. The group discussed other training models including long-term rotations (3-6 months) for post doctoral students at labs with significant bioinformatics activities. The idea is to train them to go beyond the mere use of menu-driven data bases. They need to learn how to use powerful software tools in concert with scripting skills. To facilitiate training, a valuable community resource that could be established today would be a database of Powerpoint and PDF slides for lectures and talks in microbial oceanography.

With respect to the databases, there will be a meeting in Woods Hole on September 25th and 26th, 2005 sponsored by ICoMM to discuss databases in much greater detail.

International Census of Marine Microbes: Open Ocean and Coastal Systems Workshop

May 10th and 11th , 2005

Imin International Conference Center, Jefferson Hall,

University of Hawai'i at Manoa, East West Road, Honolulu, HI 96822

This agenda is only a guide for discussions by the Open Ocean and Coastal Systems workshop participants on May 10th and 11th. The times for discussion topics will shift in response to specific interests, questions, etc. At the end of this meeting we would like to outline a consensus "Roadmap" for ICoMM's future directions.

Tuesday, May 10th, 2005

07:45-08:00: Hotel pick up (We will furnish van transportation to the East West Center.)

0815: Coffee and pastries at the Imin International Conference Center

0830: Welcome to the University of Hawaii

0900: **Dave Karl** - Brief Introductions by workshop participants

0920: **Mitchell L. Sogin** - The Ocean's hidden majority and ICoMM

0950: **Dave Karl** - Microbial Oceanography: A time for action

1020: **Coffee Break**

1045: **John Heidelberg** - Shotgun genome sequencing of marine microbial communities

1100: **Forest Rohwer** - Molecular diversity of uncultured viral communities

1115: **Heidi Sosik** - Continuous flow cytometry of picophytoplankton.

1130: **William Li** - Geographic variations in microbial cytometric diversity

1145: **Carlos Pedrós-Alió** - Marine microbiology from space

1200: **Lunch Buffet** (East West Center Dining Room)

1330: **Osvaldo Ulloa -** Microbial biogeochemistry of oxygen minimum zones off Chile

1345: **Gordon Taylor** - Microbial processes in the Cariaco Basin

1400: **Dave Caron** - The SPOTS Microbial Observatory

1415: **Mike Zubkov:** Atlantic Meridional Transect (AMT)

1430: **Matt Church** - Hawaii Ocean Time-series: 17 years and counting

1445: **Craig Carlson** - BATS/ Bermuda Microbial Observatory

1500: **Open Mic Session** - Part 1

1530: **Break:** Cookie/fruit, juice/water

1550: **Open Mic Session** - Part 2, Open discussion

1630: **Linda Amaral-Zettler** - Formation of Break-out Groups

1730: End of Day 1--- Dinner at the Willows, an "old Hawai'ian style" restaurant in the residential section of Moiliili, 817 Hausten Street (roughly half way between the University and Waikiki)

May 11th, 2005:

0800: Hotel pick up

0815: Coffee and pastries at the Imin International Conference Center

0840: **Dave Karl:** Progress summaries, General discussion

0900: Break-out group meetings

1100: **Dave Karl:** Plenary summary of initial group discussions:

1200: **Lunch Buffet** (East West Center Dining Room)

1330: Break-out group meetings

1530: Cookie/fruit break, juice/water

1600: **Mitch Sogin:** Plenary summaries and next steps

1700: END

Questions relevant for all ICoMM working groups

Diversity:

1) What metric can be used to describe microbial diversity?
2) For molecular measures, what are the strengths and weaknesses of single-gene, genomic, and populations-level perspectives? Which of these will have the greatest long-term benefits?
3) How should the dynamics of diversity be handled, and should the approach be based on species (phylotypes) or populations?
4) How will approaches to microbial diversity differ from those used in the Census of Marine Life, which focuses mainly on metazoans?

Integration:

11) What level of biodiversity is necessary to interpret ecological, physiological and process-related observations?
12) How can process and ecological data inform us about diversity and what levels of information are required?
13) What are the key scientific questions that a census can address?
14) Where are the gaps in the investigative framework?
15) How is diversity related to process stability? Can predictive frameworks be defined?

Sampling, prioritization and coordination with other programs:

19) Schedules, locations and priorities.
20) Relationship to sampling program in the Census of Marine Life and related subprojects.
21) Are there mileposts that will logically define phases of the project?
22) What observations are needed at each sampling site?
23) How should we address temporal variations?
24) How should we address spatial heterogeneity, particularly with regard to commensal populations and chemosynthetic environments (e.g. seeps, whale falls, wood falls)?

Databases:

8) What is the structure of the information that will be produced?
9) What are the preferred techniques for carrying out a census?
10) How can databases be structured to facilitate communication?

Relationships with other programs:

1) RIDGE, ODP, Genomes to Life, Microbial Observatories, existing CoML field projects
2) Biodiversity Organization issues: Centralized, Coordinated, Distributed, Combinational Models

How to Proceed:

1) Funding Strategies
    i. What are the highest priorities?
    ii. Where could ICoMM seed monies be most effective towards initiating programs and collaborations in benthic, open ocean and coastal systems?
    iii. What are the key programs for benthic, open ocean and coastal studies? How can benthic, open ocean and coastal diversity research be better advocated within these or other programs for multi-institutional and international programs?
2) White Papers
    i. What audiences should be targeted?
    ii. What are the most important messages?

Sampling questions:	**Measurement questions:**	**Data questions:**
Group 1	**Group 2**	**Group 3**
William Li - Group Leader	Jonathan Zehr- Group Leader	Anthony Michaels- Group Leader
David Kirchman- Rapporteur	John Paul-Rapporteur	Forest Rohwer-Rapporteur
Matthew Church	Linda Amaral-Zettler	Mitchell L. Sogin
David Caron	Heidi Sosik	Lita Proctor
Gordon Taylor	Claire Mahaffey	Oscar Schofield
John Waterbury	Michael Zubkov	Carlos Pedrós-Alió
Osvaldo Ulloa	Grieg Steward	Chris Winn
Daniel Vaulot	Michael Rappé	Alexandra Worden
Karin Björkman	Robert Bidigare	Zachary Johnson
James Cowen	Valerie Franck	Ricardo Letelier
David Karl	Karen Selph	John Heidelberg
Lucas Stal	Craig Carlson	

Meeting Participants

Robert Bidigare, University of Hawai'i at Manoa, Honolulu, HI

Karin Björkman, University of Hawai'i at Manoa, Honolulu, HI

Craig Carlson, University of California Santa Barbara, Santa Barbara, CA

David Caron, University of Southern California, Los Angeles, CA

Matthew Church, University of Hawai'i at Manoa, Honolulu, HI

James Cowen, University of Hawai'i at Manoa, Honolulu, HI

Valerie Franck, Hawai'i Pacific University, Kaneohe, HI

John Heidelberg, The Institute for Genomic Research, Rockville, MD

Zachary Johnson, University of Hawai'i at Manoa, Honolulu, HI

David Karl*, University of Hawai'i at Manoa, Honolulu, HI

David Kirchman, University of Delaware, Lewes, DE

Ricardo Letelier, Oregon State University, Corvallis, OR

William Li*, Bedford Institute of Oceanography, Dartmouth, CANADA

Claire Mahaffey, University of Hawai'i at Manoa, Honolulu, HI

Anthony Michaels, USC Wrigley Institute for Environmental Studies, Los Angeles, CA

John Paul, University of South Florida, St. Petersburg, FL

Carlos Pedrós-Alió, Institut de Ciencies del Mar (ICM), Barcelona, Spain

Lita Proctor, The Moore Foundation, San Francisco, CA

Michael Rappé, University of Hawai'i at Manoa, Honolulu, HI

Forest Rohwer*, San Diego State University, San Diego, CA

Oscar Schofield, Rutgers University, New Brunswick, NJ

Karen Selph, University of Hawai'i at Manoa, Honolulu, HI

Mitchell L. Sogin, ICoMM, Marine Biological Laboratory, Woods Hole, MA

Heidi Sosik, Woods Hole Oceanographic Institution, Woods Hole, MA

Lucas Stal, ICoMM, Netherlands Institute of Ecology (NIOO-KNAW), The Netherlands

Grieg Steward, University of Hawai'i at Manoa, Honolulu, HI

Gordon Taylor, Stony Brook University, Stony Brook, NY

Osvaldo Ulloa, Universidad de Concepción, Concepción, CHILE

Daniel Vaulot*, Station Biologique, Roscoff, FRANCE

John Waterbury, Woods Hole Oceanographic Institution, Woods Hole, MA

Chris Winn, Hawai'i Pacific University, Kaneohe, HI

Alexandra Worden, University of Miami, Miami, FL

Jonathan Zehr, University of California, Santa Cruz, CA

Michael Zubkov, Southampton Oceanography Centre, Southampton, THE UNITED KINGDOM

*ICoMM Open Ocean and Coastal Systems Working Group Members (Bess Ward and Peter Burkhill were unable to attend)

Working Group Report
Informatics and Data Management Working Group

International Census of Marine Microbes: Informatics and Data Management Working Group Agenda
Marine Biological Laboratory, Woods Hole, MA 02543
September 25th- 26th, 2005

Meeting Participants:

Dr. Linda Amaral-Zettler, ICoMM Secretariat, Marine Biological Laboratory (MBL)
Dr. Adorian Ardelean, Taxonomic Bioinformatician, MBL
Ms. Cyndy Chandler, Information Systems Associate II, Woods Hole Oceanographic Institution (WHOI)
Dr. Vishwas Chavan, Scientist, Information Division, National Chemical Laboratory in India
Professor Jan W. de Leeuw, ICoMM Principal Investigator, The Royal Netherlands Institute for Sea Research (NIOZ)
Dr. Terry Gaasterland, Professor, Computational Biology Director, Scripps Genome Center, Scripps Institution of Oceanography
Professor Frank Oliver Glöckner, Max Planck Institut für Marine Mikrobiologie (MIP)
Mr. Robert Groman, Information Systems Specialist, (WHOI)
Dr. Victor M. Markowitz, Department Head, Biological Data Management & Technology Center, Lawrence Berkeley National Lab
Mr. Phillip R. Neal, ICoMM IT Specialist, Senior Research Assistant, (MBL)
Dr. William Nelson, Bioinformatics Manager, The Institute for Genomic Research, (TIGR)
Dr. David J. Patterson, ICoMM Scientific Organizing Committee, (MBL)
Dr. William Piel, Assistant Professor, University at Buffalo, Department of Biological Sciences
Dr. Francisco Rodriguez-Valera, Division de Microbiología and Evolutionary Genomics Group, Universidad Miguel Hernandez
Mr. Wade M. Sheldon, Jr., GCE-LTER Information Manager/SIMO Database Administrator, University of Georgia, School of Marine Science
Dr. Mitchell L. Sogin, ICoMM Principal Investigator, (MBL)
Dr. Edward Vanden Berghe, Manager, Flanders Marine Data and Information Center, Flanders Marine Institute
Dr. Alexandra Z. Worden, Assistant Professor, Marine Biology & Fisheries
Rosenstiel School of Marine & Atmospheric Science, University of Miami

Workshop Summary:
The International Census of Marine Microbes (ICoMM) Informatics and Data Management Working Group met for two days in Woods Hole to provide feedback and recommendations on ICoMM's draft schema (v. 2.1) for MICROBIS and on its immediate future development.

The following recommendations emerged:

1. MICROBIS should seek to differentiate itself from other microbial databases – and can do so through (a) its taxonomic scope (inclusive of viruses, bacteria, archaea and

protists), (b) the combination of molecular and other identifiers with geospatial and environmental data, and (c) its purpose in assisting scientists to answer questions of a global nature.

2. MICROBIS will be a sample-centric system that mainly stores molecular, genomic, proteomic, transcriptomic, lipidomic, taxonomic and legacy data (collectively referred to as identifiers).

3. MICROBIS will be a data repository, will provide services and share data with other data services and users through an array of input and export services.

4. MICROBIS is a node of OBIS and must share content with OBIS. The dominance of molecular data in MICROBIS, the lack of named taxa, and the number of datum points will present challenges for harvesting data. MICROBIS will require services that overcome impediments to interoperability, especially services that provide taxonomic identifiers for un-named sequences. Various models, in house and external, exist or can be developed.

5. Members of the IDM working group regarded the scope of the draft MICROBIS schema as a secure foundation for on-going development and interoperability.

6. MICROBIS should add capacity to include additional information, such as information about complementary programs, or information about 'cruises' ('platforms') and methodology.

7. ICoMM should design the MICROBIS structure so that it can continue to evolve over time and so that the changes should be in response to the needs of suppliers and consumers of data.

8. The form of the tables in the MICROBIS database should be amenable to "weak typing" when possible (i.e. generic structures that don't require fixed columns for every attribute).

9. MICROBIS should set its initial goals within the context of cataloging marine microbial diversity AND on representing the geographic distribution of entities. These are interim goals. MICROBIS should mature into an environment that will enable the community to address larger scientific questions to support a larger community.

10. Accommodating small bodies of sample data from a variety of sources (Bay Paul Center, Wade Sheldon, Bill Nelson, and the application of 454 technology to TRANSAT samples) should drive the next phases of MICROBIS development. These interactions should explore the input format for data (in the case of sequence data, the GenBank SEQUIN format, various Fasta formats, etc.).

11. ICoMM should compile legacy data for MICROBIS partly because of the priority given to this in responses to the questionnaire, but also as a device to engage a large international community of non-molecular marine microbiologists.

12. ICoMM should defer development of XML schema and XSD until the MICROBIS structure is sufficiently stable to allow projects with their own RDBMS or format standards to map to MICROBIS for ongoing data submissions with no or minimal reformatting.

13. Initially, MICROBIS should include an array of templates to ease the input of data.

14. ICoMM should see that MICROBIS identifies and continues to work with other initiatives to avoid reinventing wheels and to optimize interoperability. Possible targets for strategic links should include, the Joint Genome Institute/eGenome survey regarding metadata extensions to NCBI databases, Egenomics, the Knowledge Network for Biocomplexity (KNB), Science Environment for Ecological Knowledge (SEEK) for general ecological/geospatial metadata, and the MMI (Marine Metadata Initiative).

15. MICROBIS should start simple, add additional structures to accommodate new data as required (i.e. don't try to build a universal structure for every possible contingency – start with the high priority data targets and then scale from there).

16. ICoMM must plan for a team that includes MICROBIS data managers. These managers will be required to bring incoming data to a common format, perform quality evaluation, and analyze and convert raw data. Data managers can also develop input aids and provide data on CDs and DVDs.

17. MICROBIS should include an array of "marshalling" tables to lodge incoming data. In-house data managers will parse the data and subject it to Quality Assurance checks before passing the data into the main data tables.

18. MICROBIS will service a diversity of users, mostly scientists, OBIS, genomicists, oceanographers, biodiversity experts, ecologists, policy makers, educators, environmentalists and conservationists. This mandates the emergence of flexible and diverse access tools – such as a Navigational Tree of Life, as well as different means of representing data – especially graphical interfaces.

19. Harvesting options and interoperability with particular collaborating environments, data visualization tools (specifically geographic mapping services or output tables combining identifiers with distributional data) should be priorities for data output. ICoMM should defer from providing other services in MICROBIS, such as export functions to allow general users to upload data into their own analysis tools or to allow other websites/data systems to build MICROBIS-like functionality into their

systems, until it achieves stabilization of the data model and acquires a reasonable body of data from different sources.

20. Sequences in MICROBIS should be BLASTable.

21. ICoMM should offer DNA barcoding services for marine microbes.

22. MICROBIS should plan to deliver both primary data (raw and amassed data) – such as sequences (with links to GenBank / EMBL / gi numbers), lipids, proteins, metabolites, trees, environmental data, and 'secondary' data – being information that has been 'worked up using various on the fly and resident services – such as phylogenetic trees from sample sequences, diversity measures, abundance measures, images and metadata of the images.

23. MICROBIS needs to allow for alternatives to Internet delivery because of the potential for incompatibility of dataset size and slow internet connections

24. ICoMM should develop an explicit data policy for MICROBIS.

25. ICoMM will invite Cyndy Chandler, Vishwas Chavan, Frank Oliver Glöckner, William Nelson, Wade M. Sheldon, and Edward Vanden Berghe to serve on the permanent ICoMM Informatics and Data Management Working Group.

26. *ICoMM recommends that the next meeting of the working group, not scheduled, take place at the Flanders Marine Data and Information Center, Flanders Marine Institute.*

Recommendations for the ICoMM Scientific Advisory Council (SAC):

27. The ICoMM Scientific Advisory Council should consider that if MICROBIS is to fulfill a destiny that will allow questions of a global scale to be addressed, then MICROBIS will need to be supported by Data Management personnel. Such individuals need to be familiar with the data and possess discipline expertise.

28. The ICoMM Scientific Advisory Council should advise on mechanisms by which MICROBIS can identify, disseminate, and respond to new needs.

29. The OBIS names-requirement precludes sharing of most of MICROBIS data with OBIS. Closed dialogue with OBIS and ICoMM have lead to an agreement to accept GenBank numbers in lieu of taxonomic names but second-hand accession numbers will not be accepted. Funds should be sought for the development of in-house services that will assign identities to un-named sequences by running BLAST-based and phylogenetic analyses on the samples. The ICoMM SAC should advise on these matters.

30. The SAC should provide ICoMM with a short list of initial questions to be addressed. ICoMM recommends that MICROBIS' initial focus be on the development of a simple catalog and the delivery of data in the context of geographical location. Emphasis should be given to sequence data that is combined with environmental data, and the capture of legacy data.

31. The ICoMM working groups and SAC should assess this report and provide direction as to subsequent priorities for content and functionality.