An integrative academic-industry collaboration for biotechnology driven discovery of new functions in marine environmental metagenomes

Renzo Kottmann
Microbial Genomics and Bioinformatics Research Group
Max Planck Institute for Marine Microbiology

Marine Biotechnology in the European Research Area: Challenges and Opportunities for Europe
March 2013
Partners: 25 academic & 7 SMEs

January 1st 2012
4 years
Transferring Data into Ecosystems Knowledge and Biotechnological Applications

Data is key

Software Service Infrastructure
Samples

- Long Term Ecological Research Sites
  - L4 Site Western Channel Observatory
  - BBMQ Blanes Bay Microbial Observatory
  - LTER-MC Bay of Naples
  - M3A HCMR Marine Monitoring Station Crete

- Extreme habitats
  - Iceland
  - Mediterranean deep-sea anoxic hypersaline lakes (DHALs)

- Ocean Sampling Day
Ocean Sampling Day

- Global
- Standardized
- Orchestrated
- Sampling event
  fixed in time
  - June 21st 2014

www.oceansamplingday.org
Current OSD participants
Ocean Sampling Day App

- Smartphone Application (Android, iPhone)
  - Early consistent digital metadata acquisition

InterWorks
Bioinformatics and Data Integration

Data Processing  Data Integration  Community Services

Processing results

Integrated data

Visualization & Access

- Sequence Data Archives
- Micro B3 Catalogue
- MegDb
- Genomic Data
- Environmental Data
- Marine Data Archives

- Social Web Technologies
- Visualization
- Community Annotation
- Ecological Analysis Tools

- Web Browsing
- Web Services
Data Access: Software Services

megx
marine ecological genomics

informatics + microbiology = quality & efficiency

Megx.net allows access to integrated environmental & (meta)genomic data intended for use in marine microbial ecology

Map Server
View georeferenced sampling sites in their environmental context...

GeoBlast
BLAST your sequences against georeferenced (meta)genomes...

Envo-lite
Browse genome projects classified according to sampling environment...
Discovery: knowns, known *unknowns* and *unknown unknowns*

*Pelagibacter ubique* proteome centered subnetwork
Antonio Fernandez, submitted
Data Access: Visualization of *unknown* networks
Potential new targets

- Laboratory experiments
  - Confirmation of functions
  - Pre-enrichments
  - Functional screening of large-insert libraries
  - Micro-colonies and micro-cultures
  - Heterologous expression
  - Bioassays

Cluster1800572 Unknown unknown
SAR11_0487 Tryptophan synthase
SAR11_1266 hypothetical protein
SAR11_0686 hypothetical protein
SAR11_1277 aspartate racemase
Summary & Conclusion

- Micro B3 is on a good way to establish data flow from sampling to the discovery of potential new targets

- Key for collaboration
  - open access to data
  - open source software products
    - Permissive industry friendly license
  - clear communication of intellectual property rights
  - training
Thanks for your attention

PROCEEDINGS

Hypotheses come and go but data remain
Santiago Ramón y Cajal,
Nobel Prize 1906

1st Marine Board Forum: Marine data Challenges: from Observation to Information
JBL Test-Set Mg+Ca / Magnesium Calzium
statt 22,95 €
jetzt ab 5,99 €*

» alle Varianten
zur Produktseite
18 Bewertungen

JBL Test-Set NO3 / Nitrat
statt 16,95 €
jetzt ab 8,39 €*

» alle Varianten
zur Produktseite
24 Bewertungen

JBL Test-Set SiO2 / Silikat
statt 15,95 €
jetzt ab 7,98 €*

» alle Varianten
zur Produktseite
14 Bewertungen

JBL Test-Set Fe / Eisen
statt 15,45 €
jetzt ab 8,19 €*

» alle Varianten
zur Produktseite
27 Bewertungen

JBL EasyTest 6 in 1
statt 15,49 €
jetzt nur 13,19 €*

» alle Varianten
zur Produktseite
15 Bewertungen
Current “Sample Workflow” for Pilot-OSD

Pilot-OSD Samples

Sample/DNA

Argonne

Sequences

ENA
European Nucleotide Archive

Smithsonian Institution

Sample/Environmental data

SeaDataNet

Sample/Material

App
Micro B3

- Micro B3 will develop innovative bioinformatic approaches and a legal framework to make large-scale data on marine viral, bacteria, archaeal and protists genomes and metagenomes accessible for marine ecosystems biology and to define new targets for biotechnological applications.

- Micro B3 will build upon a highly interdisciplinary consortium of 32 academic and industrial partners comprising world-leading experts in bioinformatics, computer science, biology, ecology, oceanography, bioprospecting and biotechnology, as well as legal aspects.

- Budget: 9 Mio Euro for 4 years
In silico predictions

Total sequence data set

Oceanographic & environmental data

Standards & data integration

Hypotheses

Candidates

Function & application

Small and large scale study sites

Bioinformatics Laboratory

Process & filter

Test & learn

Intellectual property rights

Ecosystems biology & biotechnology
small and large scale study sites

bioinformatics

process & filter

lab

test & learn

ecosystems biology & biotechnology

intellectual property rights

function & application

hypotheses candidates

total sequence data set

in silico predictions

oceanographic & environmental data

standards & data integration
Exploring Ecosystems Biology

Key parameters

Statistics

Organisms

Ecosystems Biology

Environment

Modelling

Predictions

Function
Helgoland Roads

Gunnar Gerdts (AWI)
Predicted diversity hotspots
Tara-Oceans

Slide modified from Chris Bowler
fig. S1: Workflow of the initial sample processing. Details are provided in the Supporting Online Material description of the materials and methods.

Teeling et al., Science under revision
10 μm prefiltration

3 μm prefiltration

0.2 μm „harvesting“
Diversity

Pedros-Alio, Trends in Microbiology, 2006, vol. 12, issue 6, page 257
Oceanographic Data
Involvement of stakeholders, esp. companies
PharmaMar, MATIS, BIO-Illiberis, Bio-Prodict, Interworks, Ribocon, EuropaBio, CIESM
Function & Biotechnology

- Explore integrated datasets
  - Genes of unknown functions
  - Biosynthetic gene clusters in marine organisms
  - New genes from metagenomes

- Laboratory experiments
  - Confirmation of functions
  - Pre-enrichments
  - Functional screening of large-insert libraries
  - Micro-colonies and micro-cultures
  - Heterologous expression
  - Bioassays
Intellectual Property Management for Marine Bioprospecting

- Model agreements for pre-competitive access to microbial materials and exchange of materials and data within research networks.
- IP model agreements for pre-competitive access to large-scale microbial genomic research databases
- Outreach workshop on topics like IPR, scientific potential, biotech options, environmental impact to inform and discuss application and protection options for marine microbial diversity with stakeholders from regional, national and international policy-making bodies
Dissemination and Outreach

- Timely dissemination of relevant information
  - to industrial sector including SMEs
  - to policy makers & advisors in field of biodiversity protection

- Capacity building through training of next generation of scientists
  - By establishing a dedicated bioinformatics training pipeline consisting of three courses

- Outreach to general/interested public on our three topics
Highlights 2012
Tasks

- Sampling site registry
  - Survey
  - Commitments

- Consistency
  - Best practices
  - Standards

- Legal framework
  - Access and benefit sharing
  - Model agreements
Pilot OSD June 20 2012, next 21 December 2012

1. L4
2. Roscoff
3. Helgoland
4. Naples
5. Greece
6. Blanes
7. Moorea
8. BATS
9. SPOTS
10. HOTS
11. Rothera
12. Churchill
13. Thames
14. Banyuls
15. Villefranche
16. Kristineberg
17. VLIZ
18. North Cyprus
19. Red Sea
20. Iceland
Ocean Sampling Day - Summer solstice 2012-06-20 – Roscoff

Sampling was performed at the SOMLIT-Aslan long-term observatory site (48° 48' 40 N, 3° 56' 15 W) at 10:30 local time during the 2012 summer solstice. Water was collected from a depth of 1 meter using a Niskin bottle. Water samples were transported to the lab in an acid cleaned plastic bottle within 15 minutes and filtered. Within 30 minutes, 4 replicates of 1 liter were filtered for metagenomic analyses using the Sterivex 0.22μm filters. Filters were sealed and immediately frozen at -80°C until extraction.

Ocean Sampling Day – Rothera

Rothera is a British Antarctic Survey research station, located on Adelaide Island on the Antarctic Peninsula at 67° South http://www.antarctica.ac.uk/living_and_working/research_stations/rothera. The plan was to get out and sample on our Winter Solstice (20th June), which for most of OSD partners was the Summer Solstice. Unfortunately sea ice and wind were against us and we didn’t manage to get out to sample until the 31st of July. The sea ice formed the week before the Solstice and was too thick to get one of our RIBs (Rigid inflatable boat) through, but too thin to walk on. The ice then thickened up and we were able to go for short walks and even dive through it near to the shore; however it was too dangerous to go too far offshore and get to our sampling sites. The winds picked up towards the middle of July and blew the sea ice out – but the winds remained too high to launch the boats, sometimes even blowing the ice back in. This continued for a few weeks until finally we had a window where there was an area of open water leading toward the sample site and no wind! The Rothera Oceanographic and Biological Time Series (RaTS) has been running since 1997 and continues year round – collecting data through weekly water sampling and CTD events. http://www.antarctica.ac.uk/staff-profiles/webpage/mm/myRaTS.html
Call for Participation

WEDNESDAY, 21 NOVEMBER 2012

Call for Participation in Ocean Sampling Day

A Call to Join Ocean Sampling Day

You can use the simple form below to register an expression of interest in joining OSD 2014 (or prior pilot activities). Please fill in form and we will contact you.

Register your Expression of Interest in Joining OSD:  http://www.surveymonkey.com/s/Z6SX289

Expression of Interest in joining Ocean Sampling Day

Please briefly description your interest in Ocean Sampling Day. All the fields are optional. After you make your expression of interest in participating you will be contacted by the OSD Consortium. Eventually, if you chose to join OSD, you will need to formally register to participate.

1. Contact Name
2. Contact Email
3. Contact Address (Institution)
4. Your ideas - how would you like to participate in OSD?  
5. How would your contributed samples be contextualized/unique?
6. When would you want to participate
7. Do you have funding to participate?

OSD will become what the community makes it....
Contextual Data Flow – Mobile App

Use Case: Mobile Field Information Application

Mobile App
- Change Settings
- Create Sample
- Push Data

Lab App
- Show Sample Data
- Synchronize
- Pull Data

Sampler
- 1..* 1

Server
- Receive Data

Lab User
Mobile App
Film about OSD & Financing of Sequencing

- Crowd Funding Project

- Trailer by mediomix

- Target 50,000 Euro
  - Covers the film
  - Additional money -> Micro B3 sequencing

- Start October 1st 2013 to November 10th 2013 (UNESCO World Science Day)
Micro B3 Connected

GLOBAL BIODIVERSITY INFORMATION FACILITY

Marine Genomics 4 Users

EUROMARINE - INTEGRATION OF EUROPEAN MARINE RESEARCH NETWORKS OF EXCELLENCE

ICOMM

MIRADA

LIFEWATCH