Microbial diversity and activity in the Southern Ocean - a coupled ‚omics' approach

Pavla Debeljak

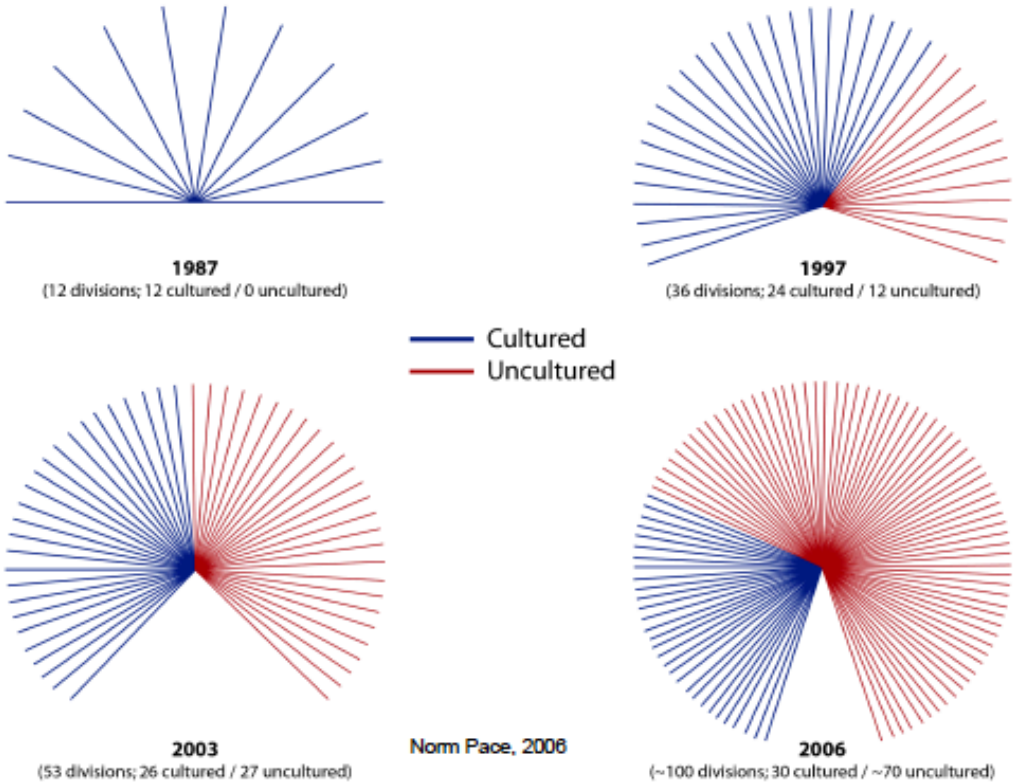
Post – cruise meeting

18./19. September 2017

Villefranche sur Mer

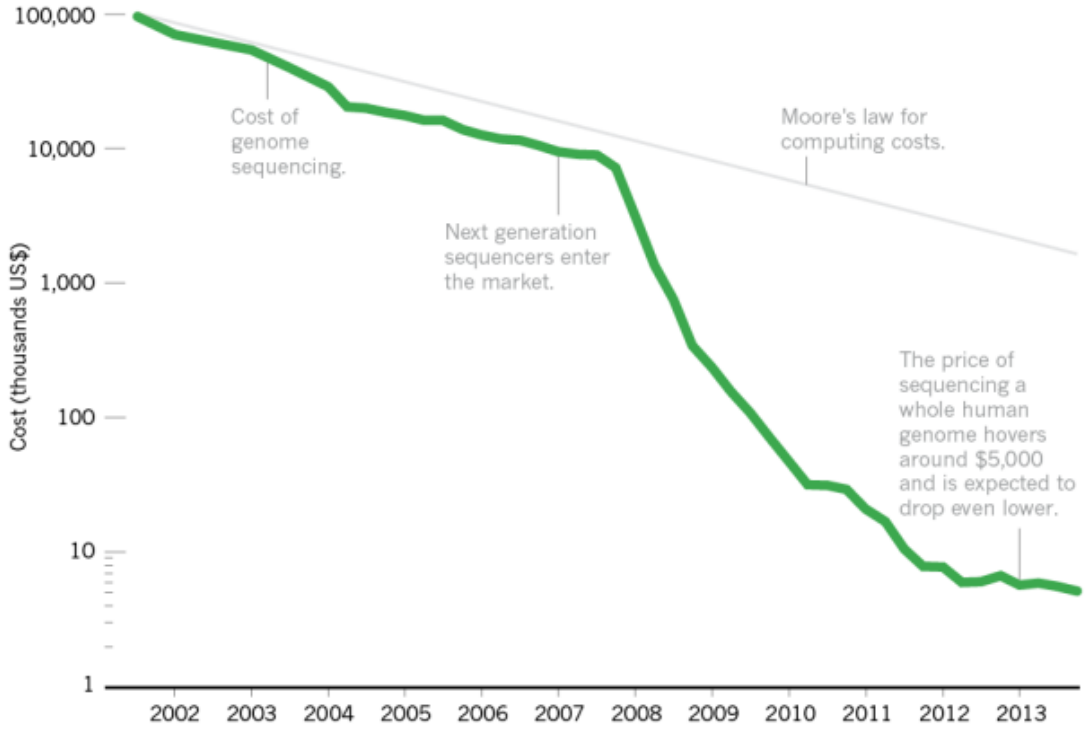
Why ‘omics’ ?

Known Bacterial Phylogenetic Divisions

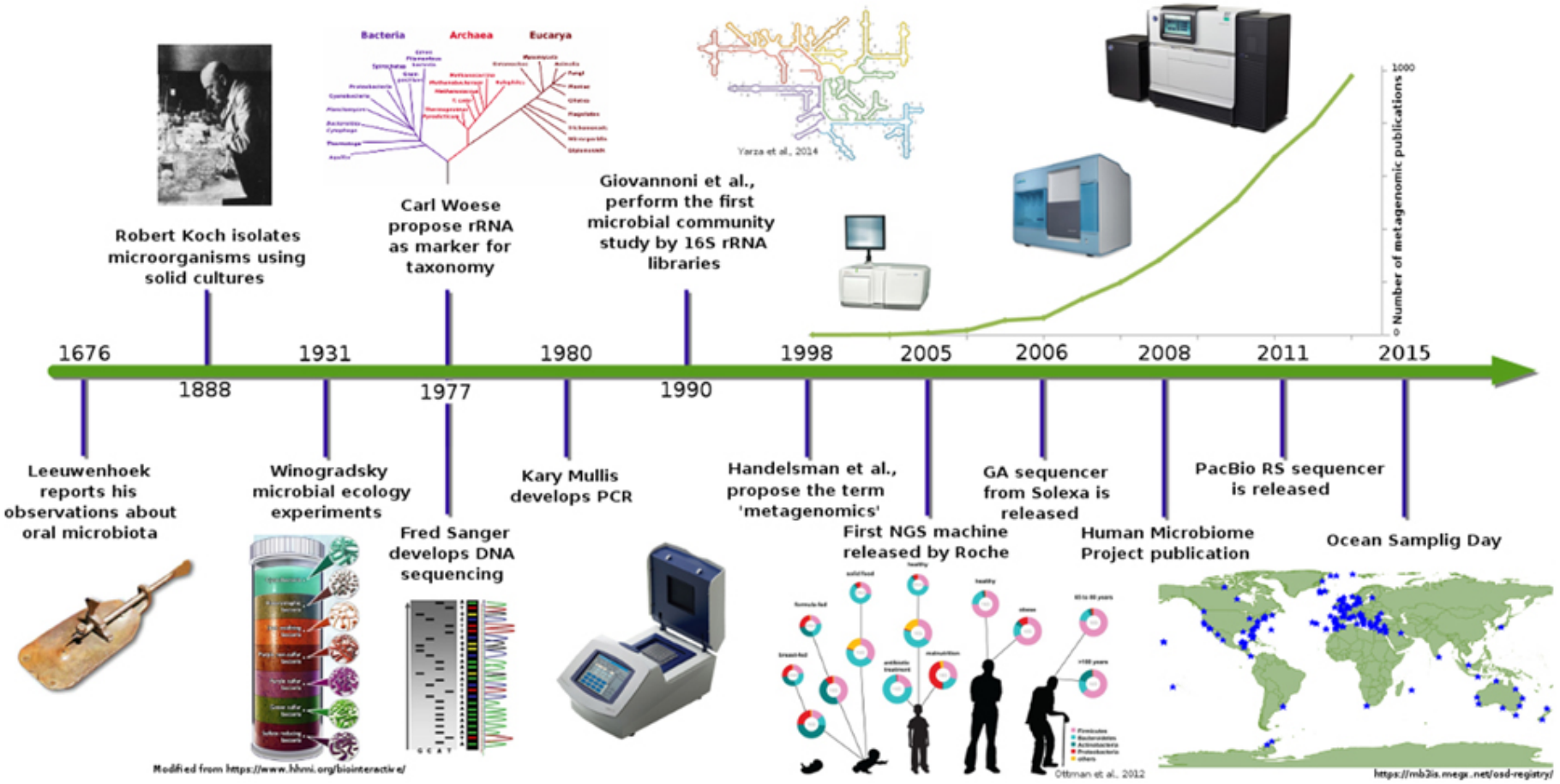


Falling fast

In the first few years after the end of the Human Genome Project, the cost of genome sequencing roughly followed Moore's law, which predicts exponential declines in computing costs. After 2007, sequencing costs dropped precipitously.



Possibilities for Microbial Ecology



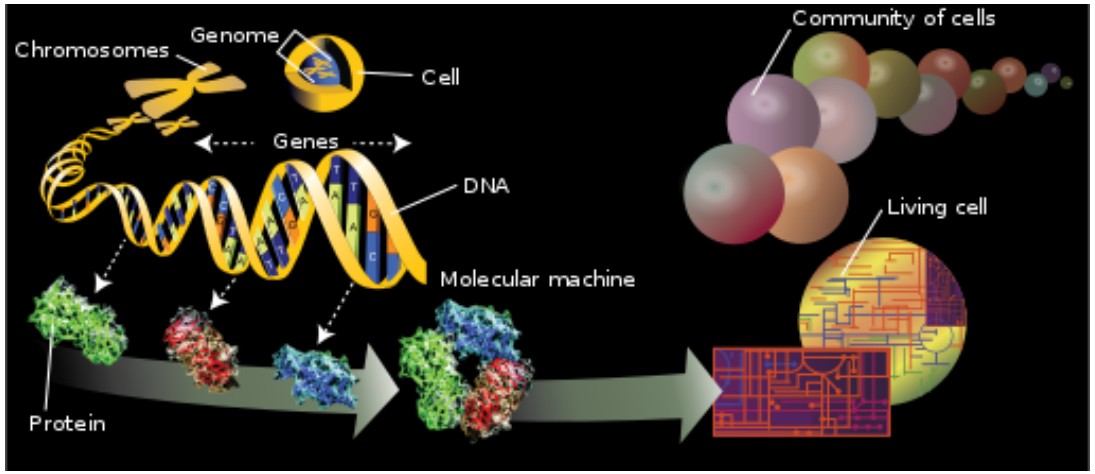
RESEARCH ARTICLE

Environmental Genome Shotgun Sequencing of the Sargasso Sea

J. Craig Venter[1,*], Karin Remington[1], John F. Heidelberg[3], Aaron L. Halpern[2], Doug Rusch[2], Jonathan A. Eisen[3], Do...

+ See all authors and affiliations

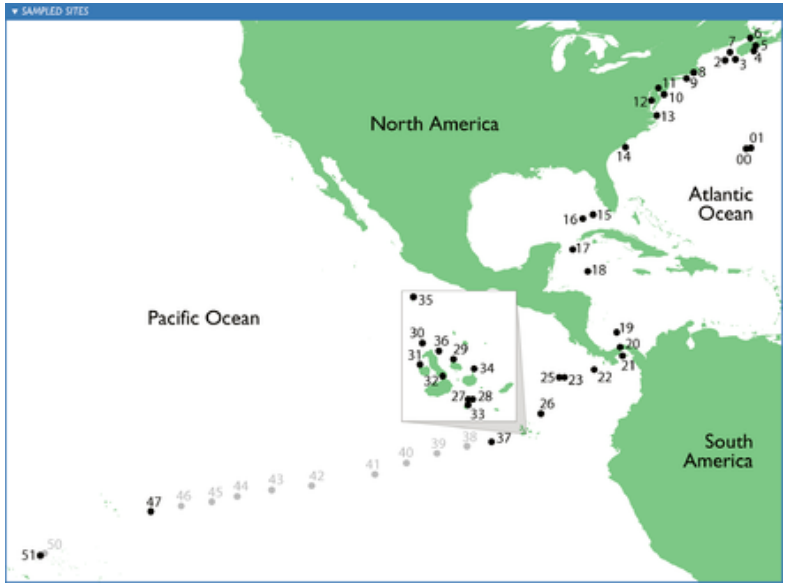
Science 02 Apr 2004:
Vol. 304, Issue 5667, pp. 66-74
DOI: 10.1126/science.1093857



- ‘Whole-genome shotgun sequencing’

Total of 1.045 billion base pairs of nonredundant sequences was generated, annotated and analyzed to elucidate the gene content, diversity, and relative abundance of organisms within these environmental samples.

Identifying over 1.2 million previously unknown genes

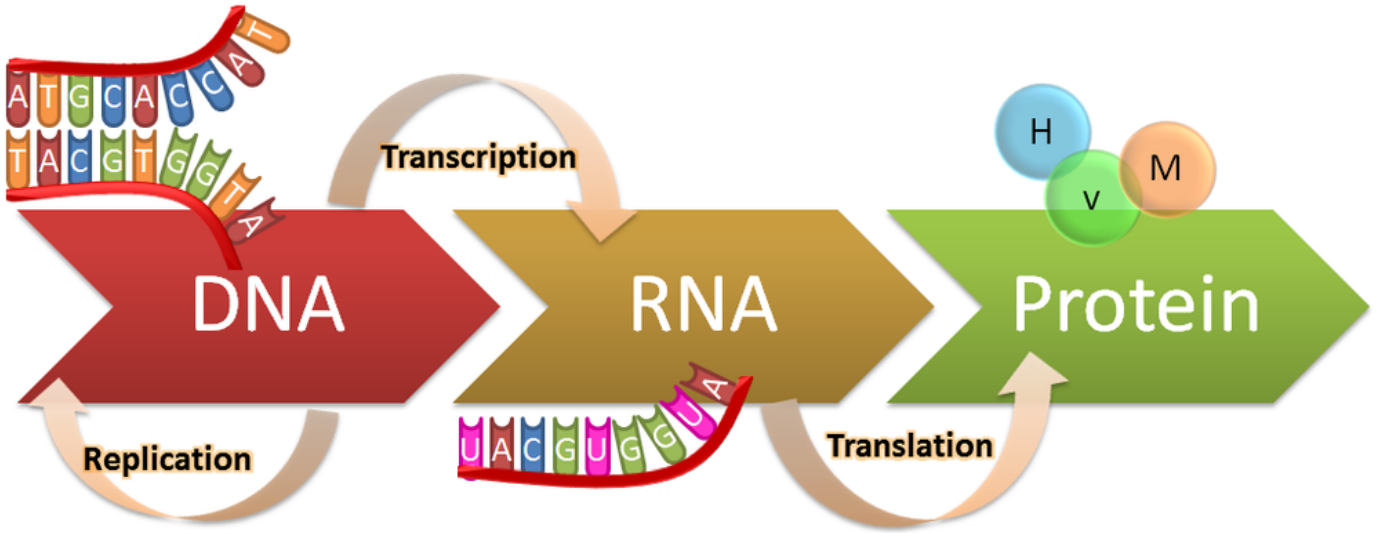


‘Omics’ - Levels

Metagenome

Metatranscriptome

Metaproteome

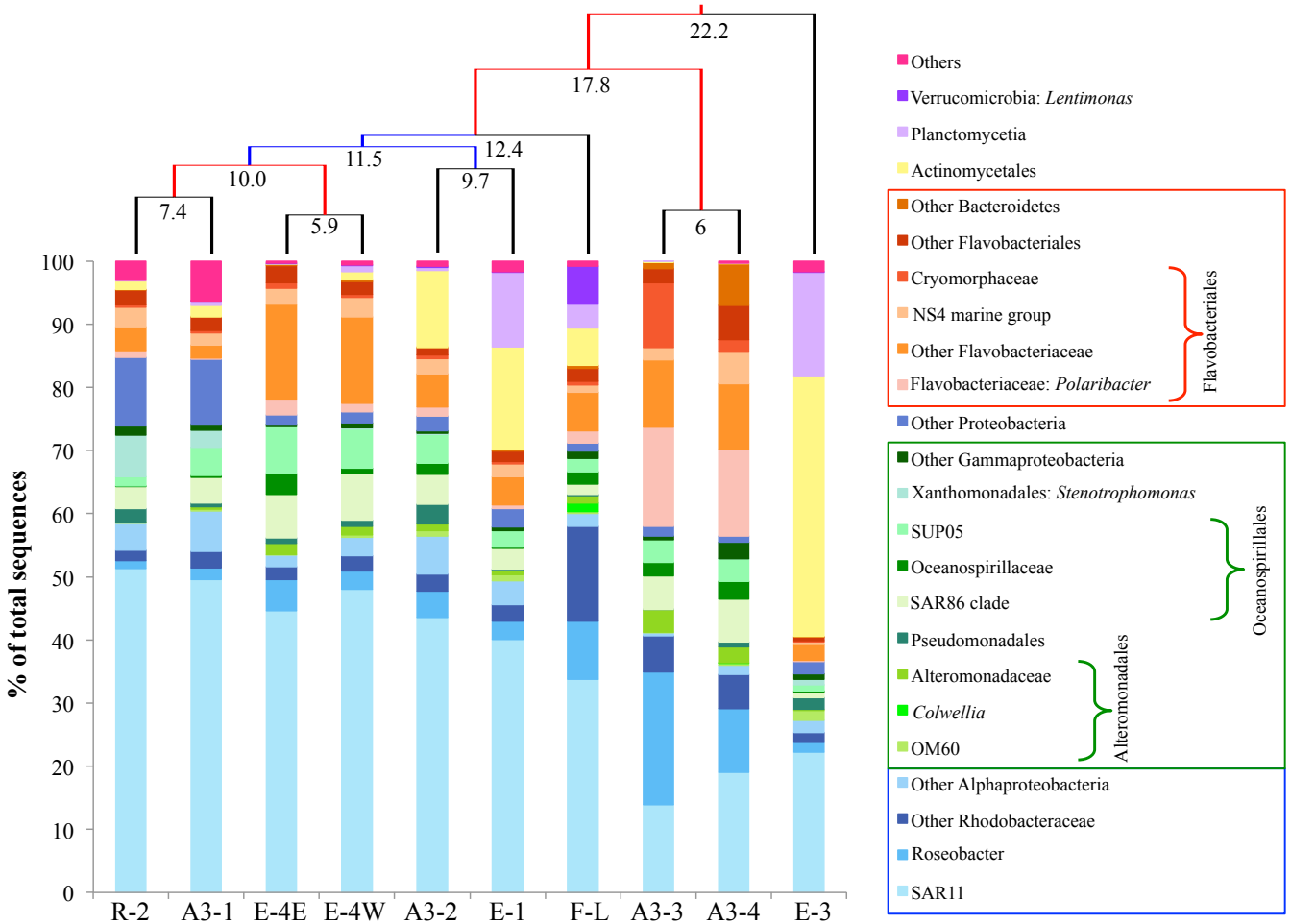


Who?

What?

How?

Kerguelen and ‘omics’

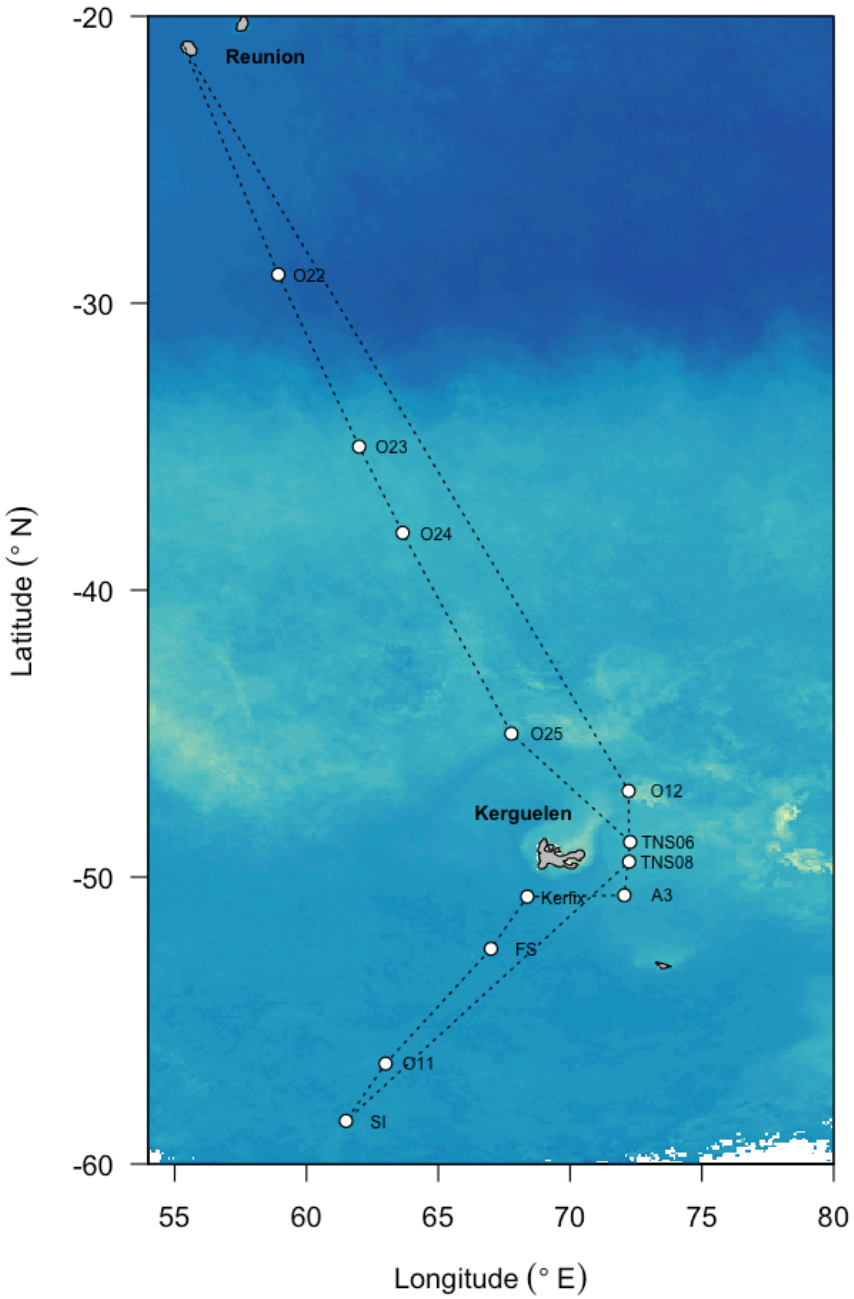


KEOPS2 Project Data published in Landa et al. Shifts in bacterial community composition associated with increased carbon cycling in a mosaic of phytoplankton blooms. The ISME Journal (2016) 10, 39-50

Thesis Outline

- Southern Ocean HNLC area – major nutrients present in high concentrations but phytoplankton biomass remains low
- Biological pump of carbon not effective due to limitation of Fe
- **How does Fe affect heterotrophic microbial metabolism?**
- Study the physiological strategies of heterotrophic microorganisms as an adaptation to low carbon and iron availability

→ Gene expression patterns in different nutrient regimes in the Southern Ocean



Sampled stations

- Metagenom - A3
- Metatranscriptomes - A3, Kerfix
- Metaproteomes - A3, Kerfix, O11, SI, TNS06, O12, O25, O24, O22

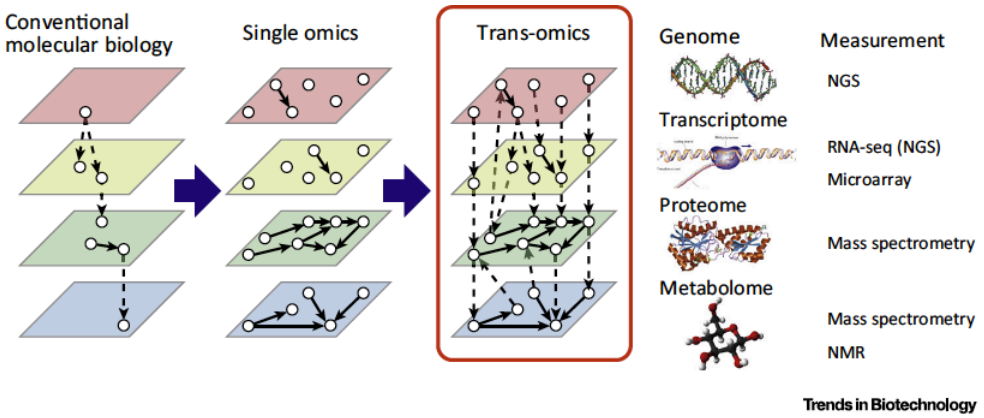
Trans-Omics Approach SOCLIM & MOBYDICK

From early spring (SOCLIM) and summer (MOBYDICK)

2 Stations : A3 (naturally fertilized) and KERFIX (HNLC)

- **Metagenome of A3**
- **Metatranscriptomes A3 & KERFIX**
- **Metaproteomes A3 & KERFIX**

→ Surface waters of the Southern Ocean around Kerguelen



DNA → RNA → Protein

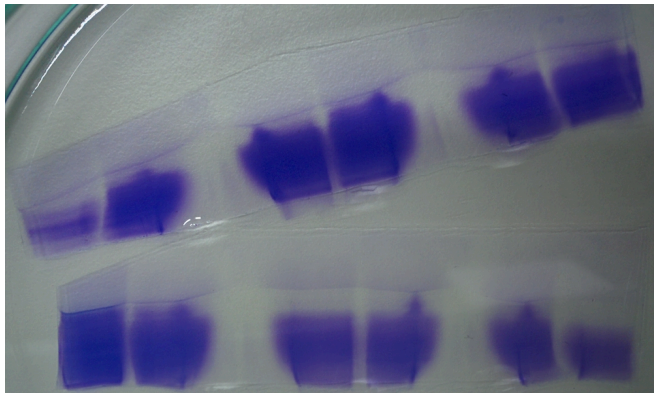
Yugi et al. 2016, Trends in Biotechnology

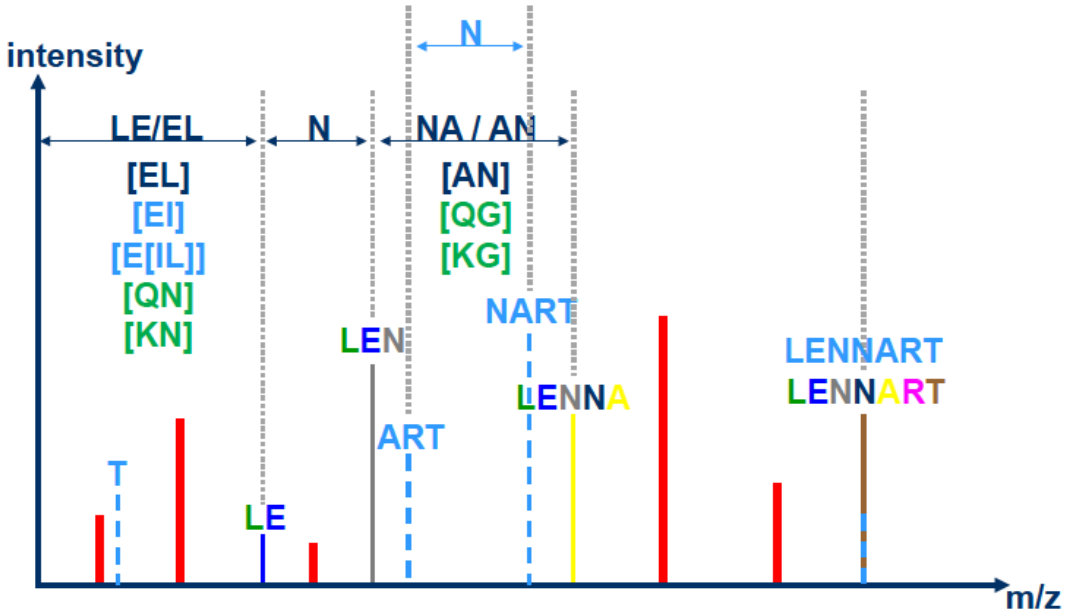
Metaproteomics

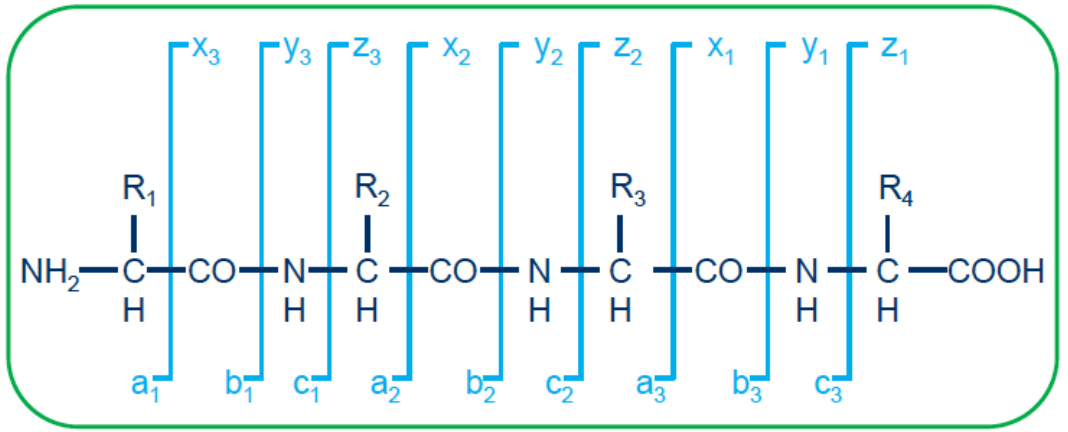
Proteomics Workflow

1. Protein extraction from filter with LN
2. Protein clean-up and fractionation different sizes
3. Protein digestion with Trypsin
4. Peptide extraction and desalting
5. Measurement of peptides
6. MS Spectra
7. Protein identification

Peptides distinguished by proteins on the basis of size,
And as an arbitrary contain approx. 50 or fewer AA







There are several other ion types that can be annotated, as well as 'internal fragments'. The latter are fragments that no longer contain an intact terminus. These are harder to use for 'ladder sequencing', but can still be interpreted.

Protein expression in microbial communities and their in situ physiological states

The ISME Journal (2009) 3, 93–105

www.nature.com/ismej

ORIGINAL ARTICLE

Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea

Sarah M Sowell[1], Larry J Wilhelm[2], Angela D Norbeck[3], Mary S Lipton[3], Carrie D Nicora[3], Douglas F Barofsky[4], Craig A Carlson[5], Richard D Smith[3] and Stephen J Giovanonni[2]
[1]*Molecular and Cellular Biology Program, Oregon State University, Corvallis, OR, USA;* [2]*Department of Microbiology, Oregon State University, Corvallis, OR, USA;* [3]*Pacific Northwest National Laboratory, Richland, WA, USA;* [4]*Department of Chemistry, Oregon State University, Corvallis, OR, USA and* [5]*Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, CA, USA*

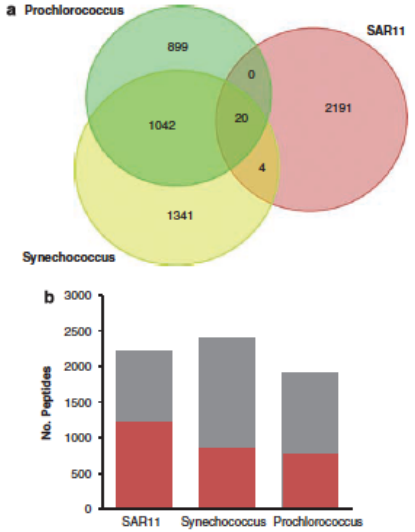
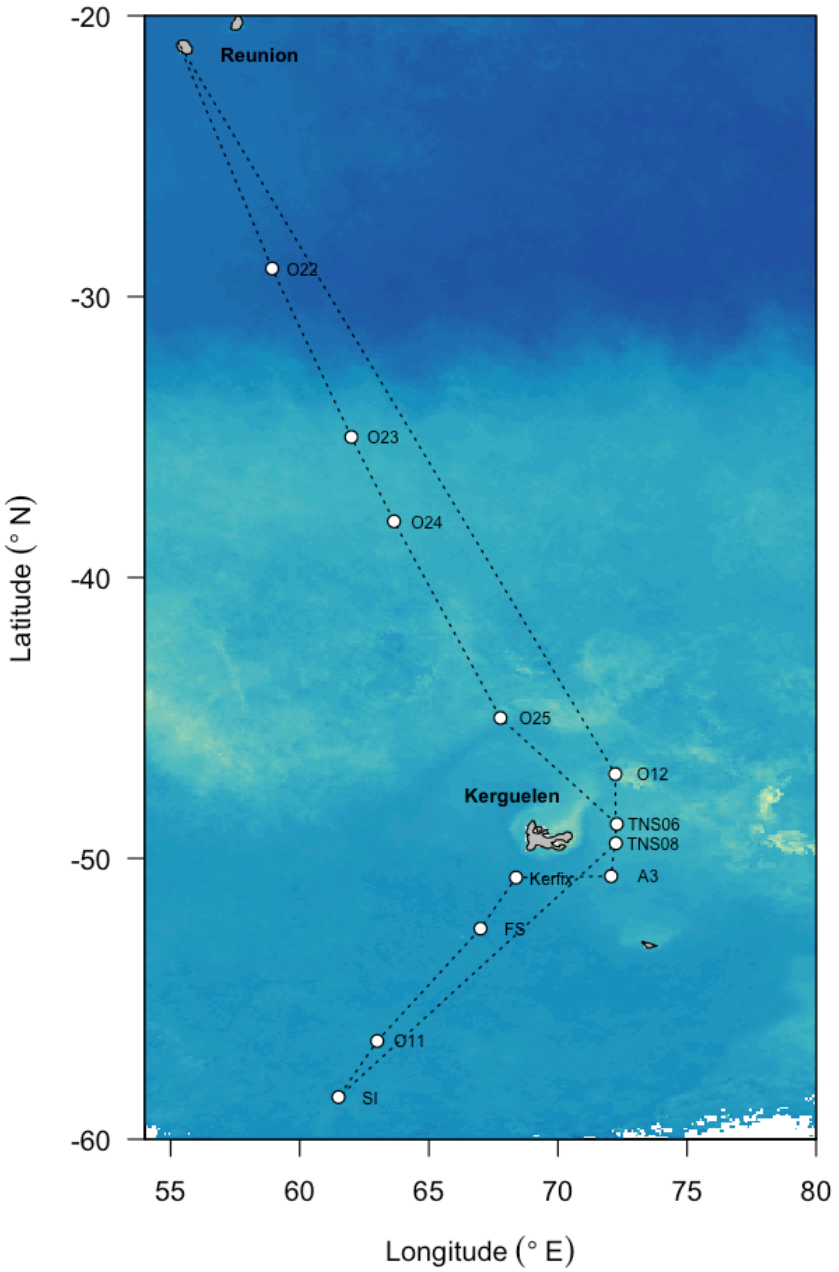
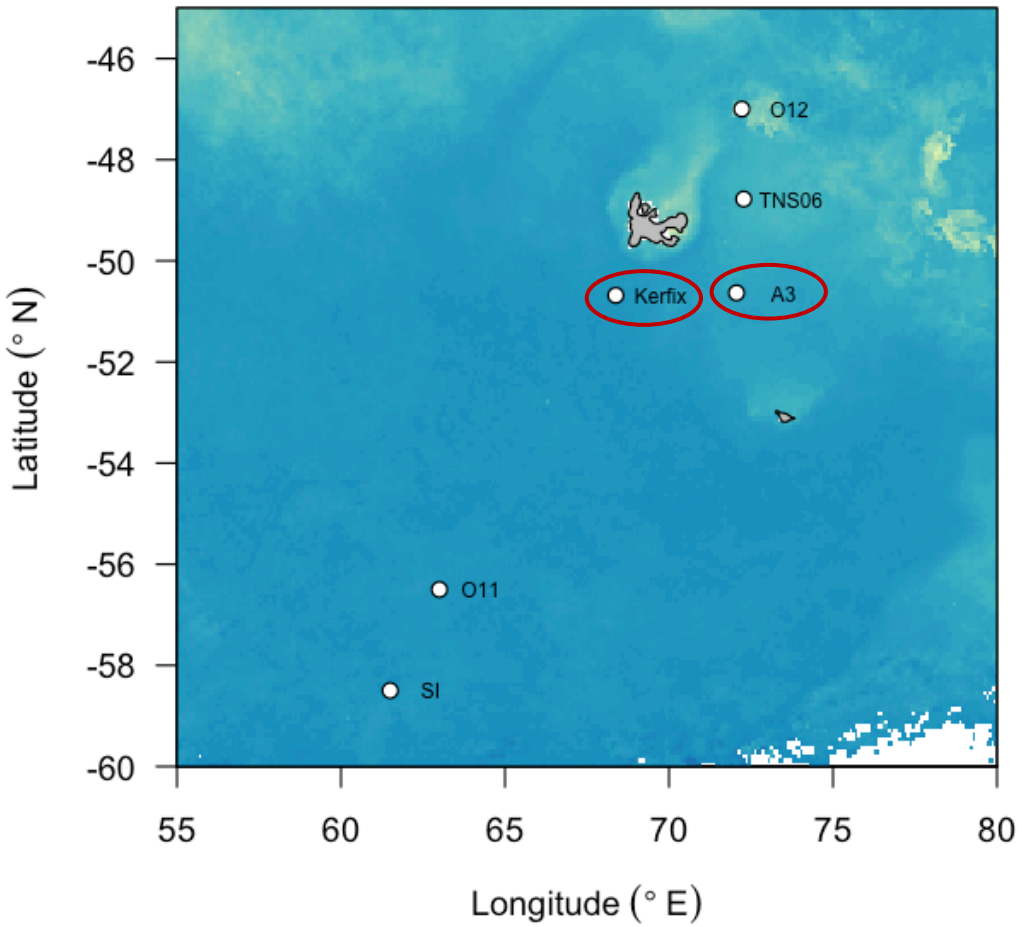


Figure 1 A large proportion of the detected peptides were unique to either SAR11, *Prochlorococcus* or *Synechococcus* (**a**). SAR11, *Prochlorococcus* and *Synechococcus* peptides detected in a Sargasso Sea protein sample and their distribution between the SAR11 eCDS, the *Prochlorococcus* eCDS and the *Synechococcus* eCDS databases. (**b**) Total number of peptides detected from each database and the fraction (red) that was also seen in the remainder of the Sargasso Sea eCDS database (Sargasso Sea eCDS database−(SAR11 eCDS database+*Prochlorococcus* eCDS database+*Synechococcus* eCDS database)) determined using BLASTP with a requirement of 100% identity over the length of the query peptide. eCDS,environmental protein-coding sequence.

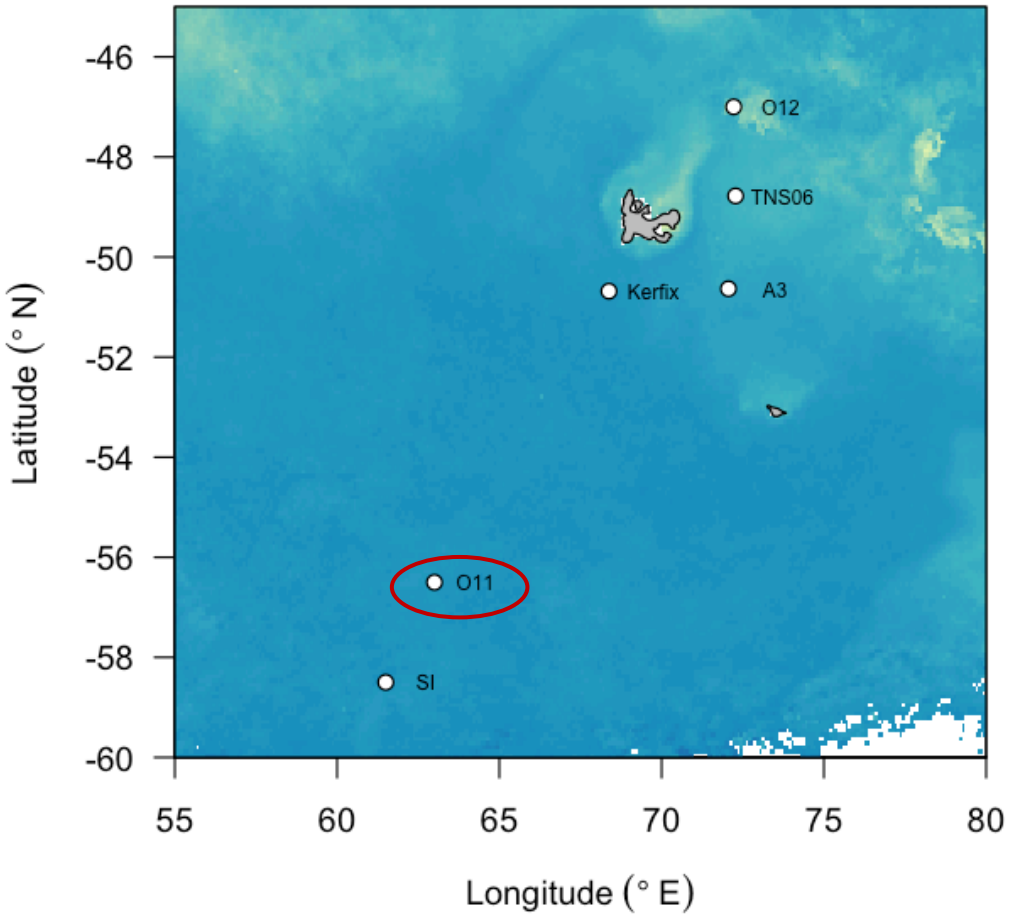


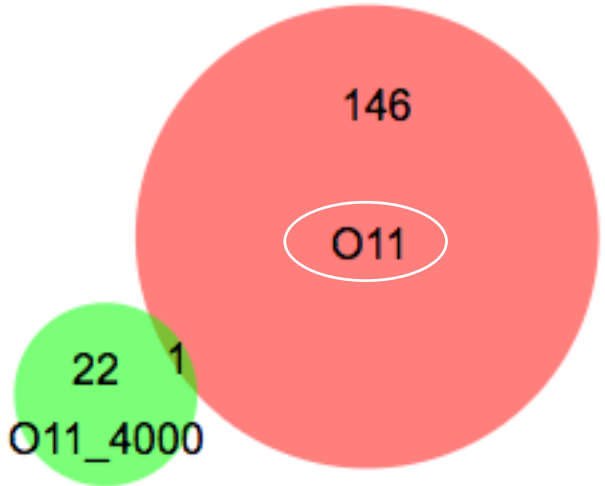
Sampled stations

- Metagenom - A3
- Metatranscriptomes - A3, Kerfix
- Metaproteomes - A3, Kerfix, O11, SI, TNS06, O12, O25, O24, O22

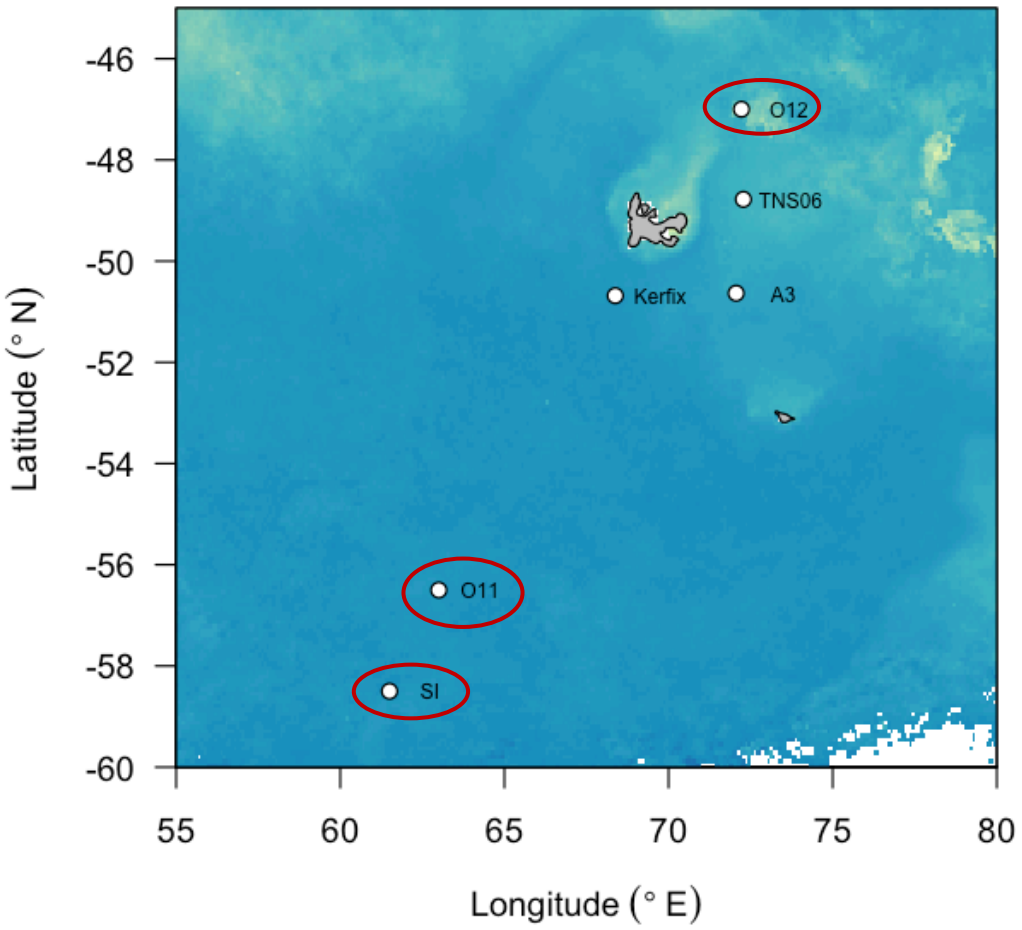


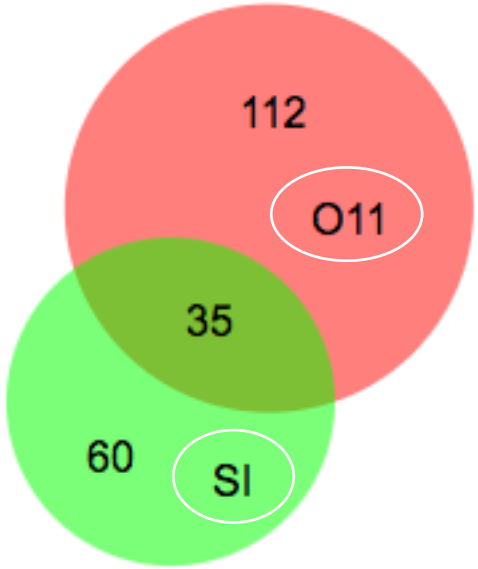
→ Shared Proteins mostly associated with cellular transport functions



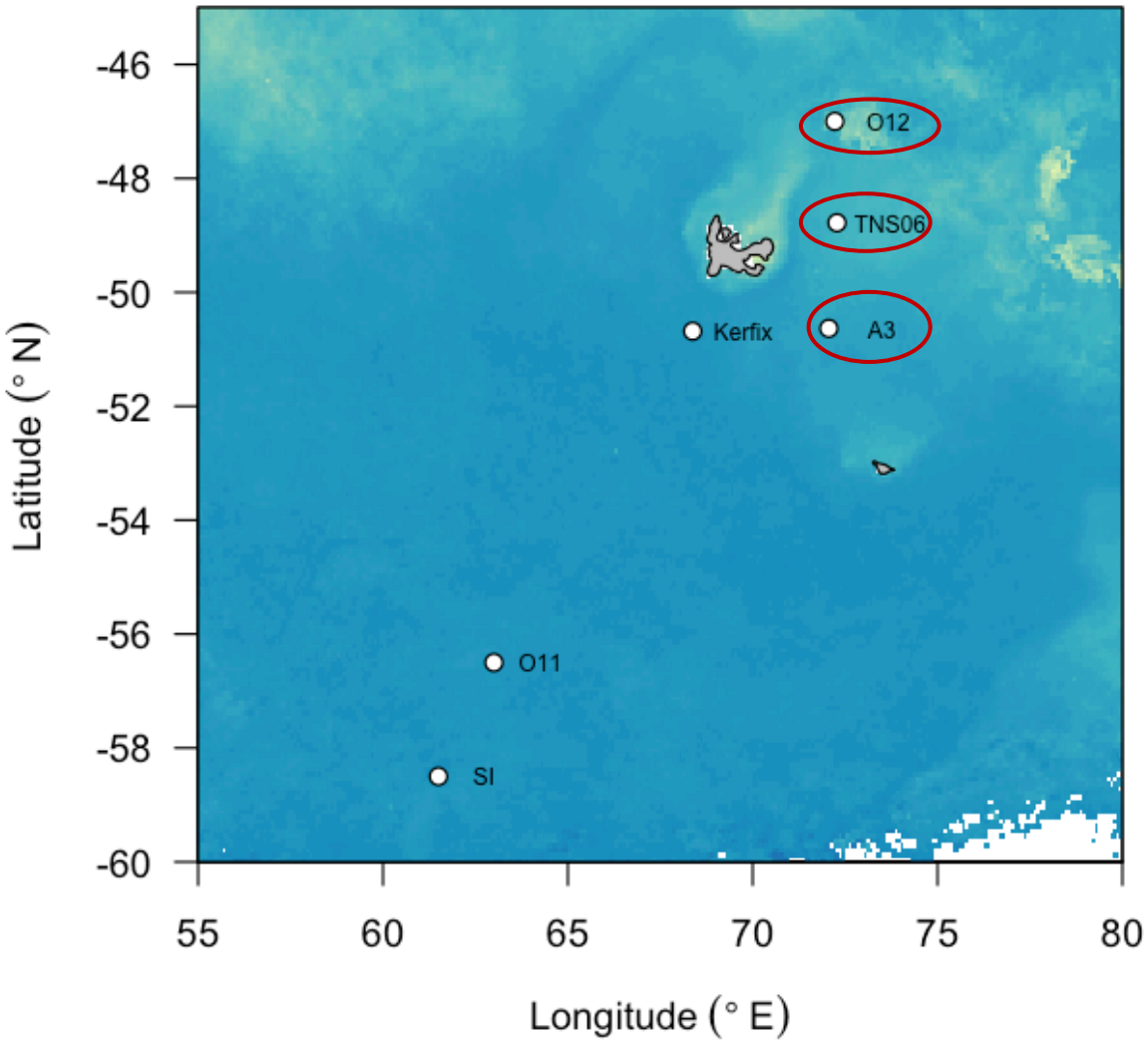


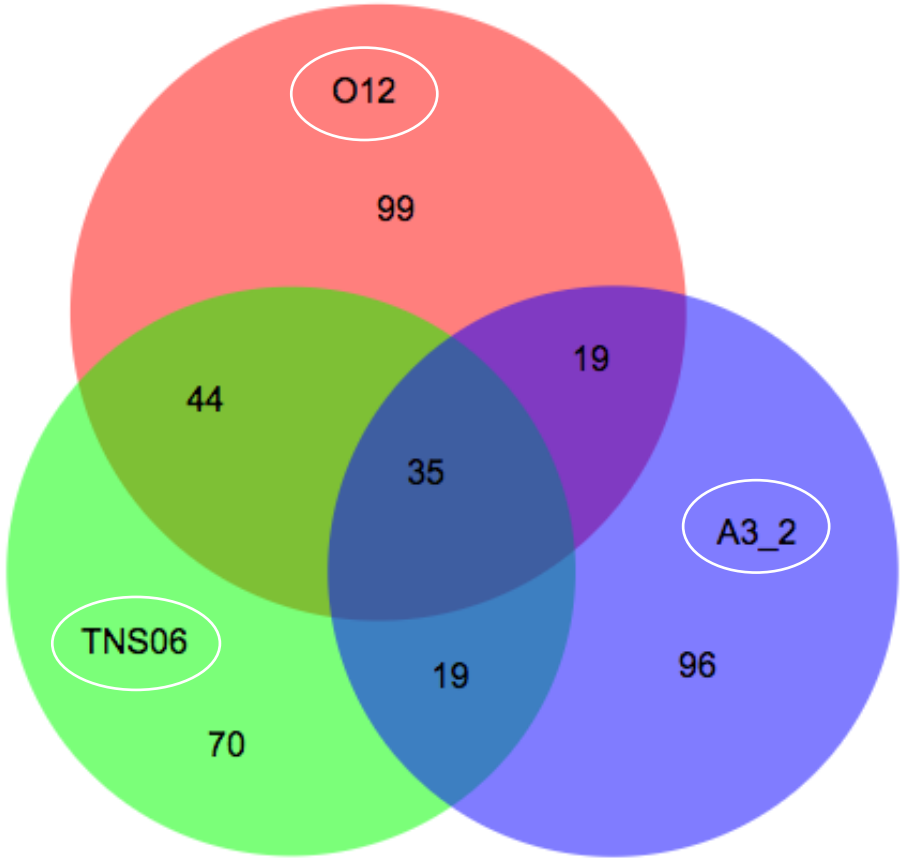
→ Only one shared sequence: Amino Acid transport system substrate binding protein of SAR324 cluster

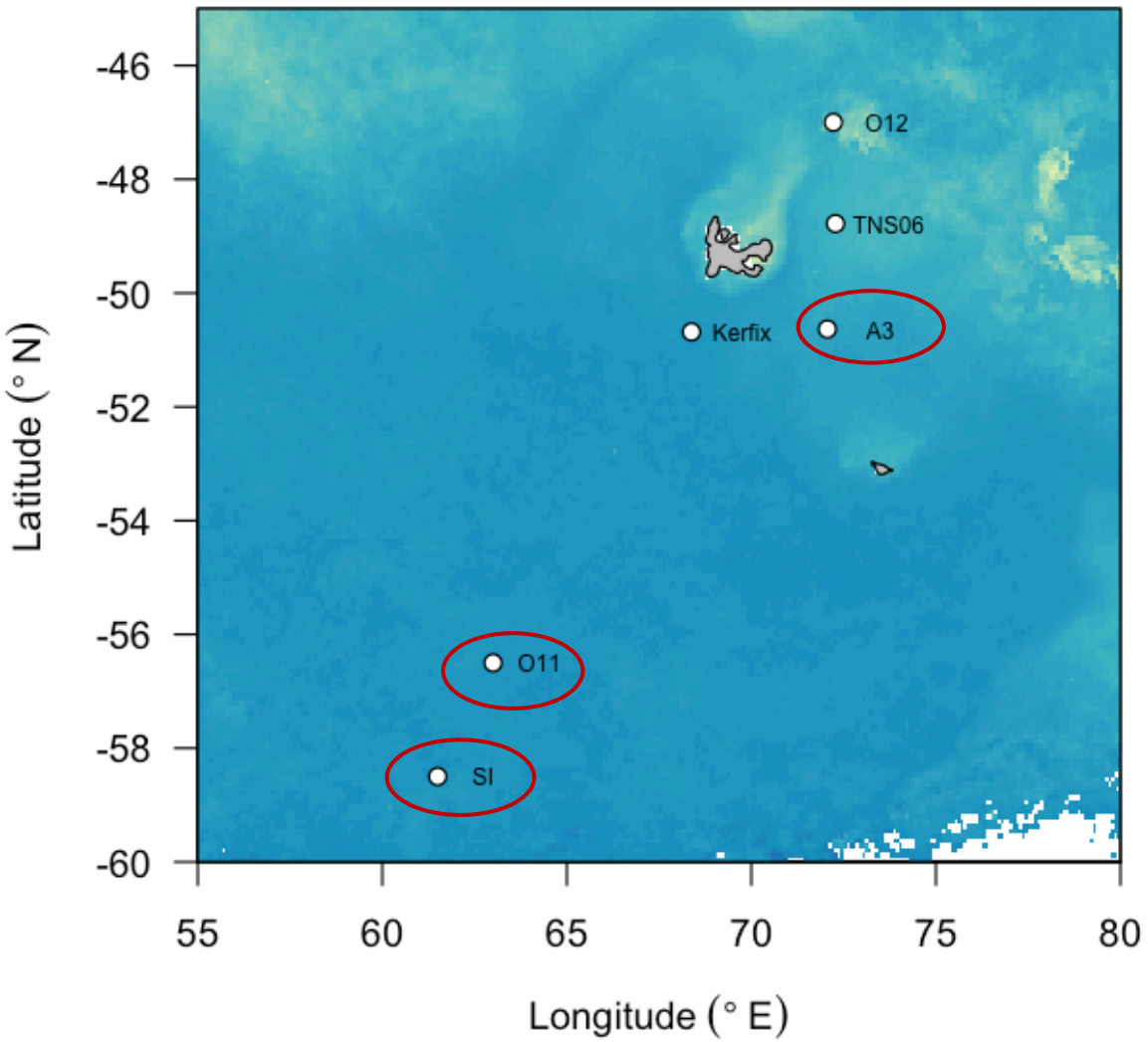

-46
-48
-50
-52
-54
-56
-58
-60
Latitude (° N)
55
60
65
70
75
80
Longitude (° E)
O12
TNS06
Kerfix
A3
O11
SI

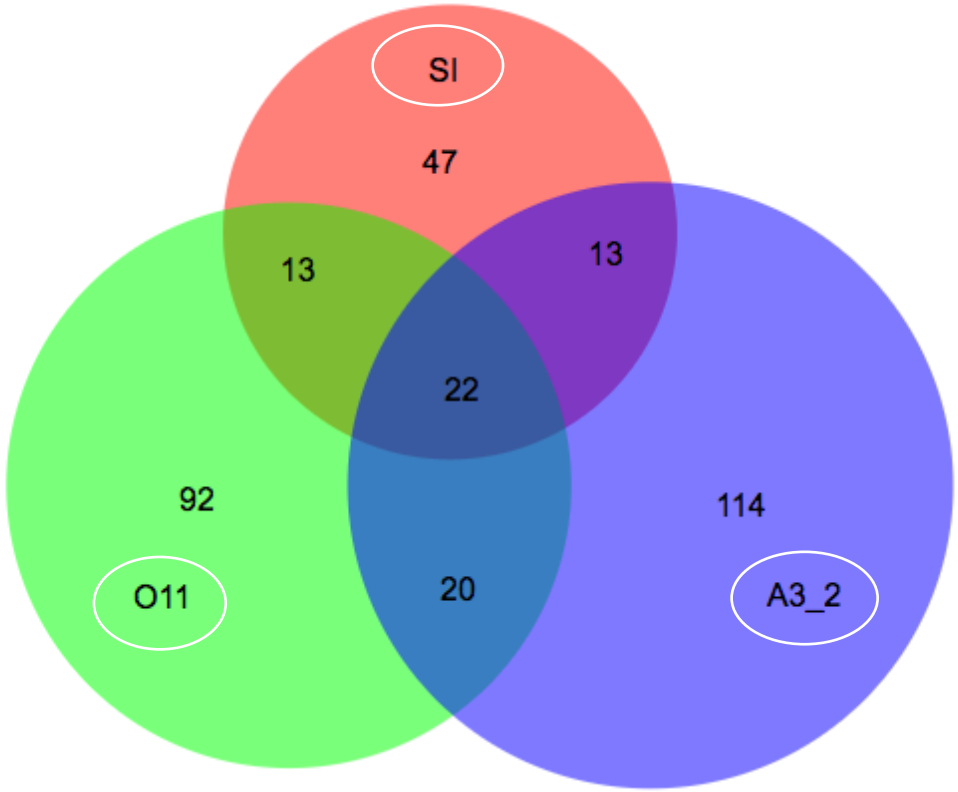

112
O11
35
60
SI


160
O12
37
58
SI


-46
-48
-50
-52
-54
-56
-58
-60
Latitude (° N)
55
60
65
70
75
80
Longitude (° E)
O12
TNS06
Kerfix
A3
O11
SI


O12
99
44
19
35
A3_2
TNS06
19
96
70


-46
-48
-50
-52
-54
-56
-58
-60
Latitude (° N)
55
60
65
70
75
80
Longitude (° E)
O12
TNS06
Kerfix
A3
O11
SI


SI
47
13
13
22
92
114
O11
20
A3_2

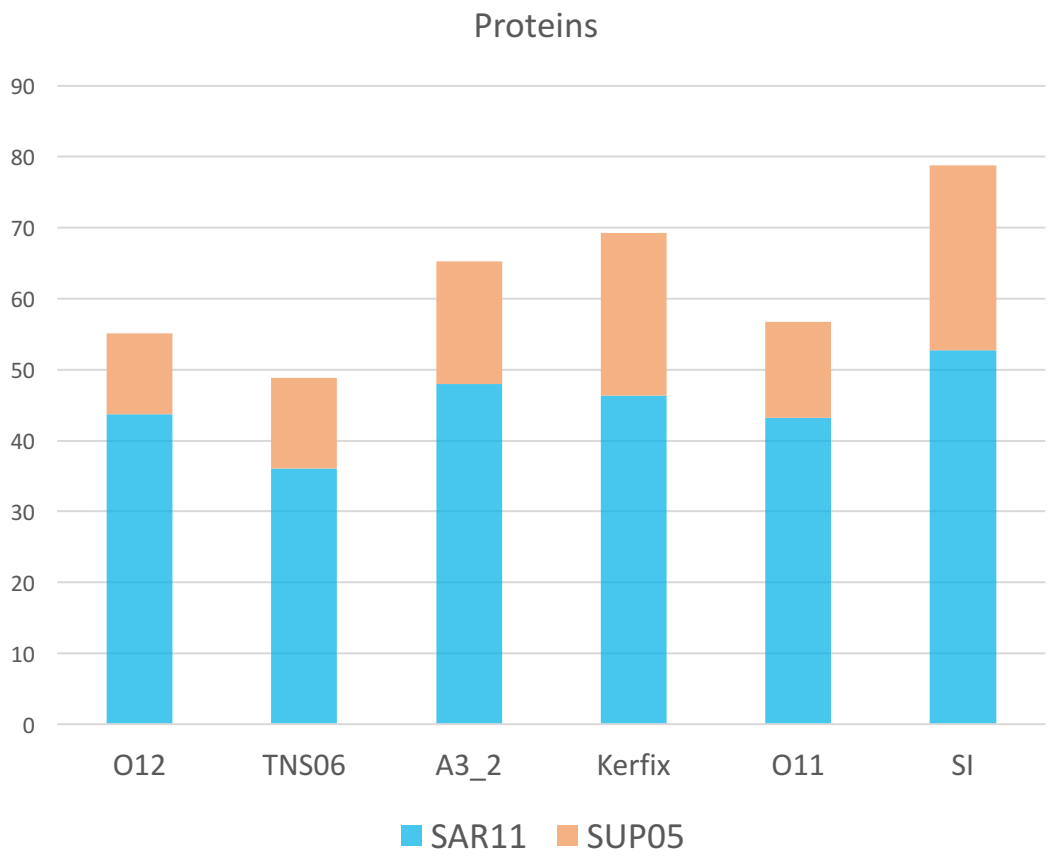
→Transport proteins dominate all Metaproteomes

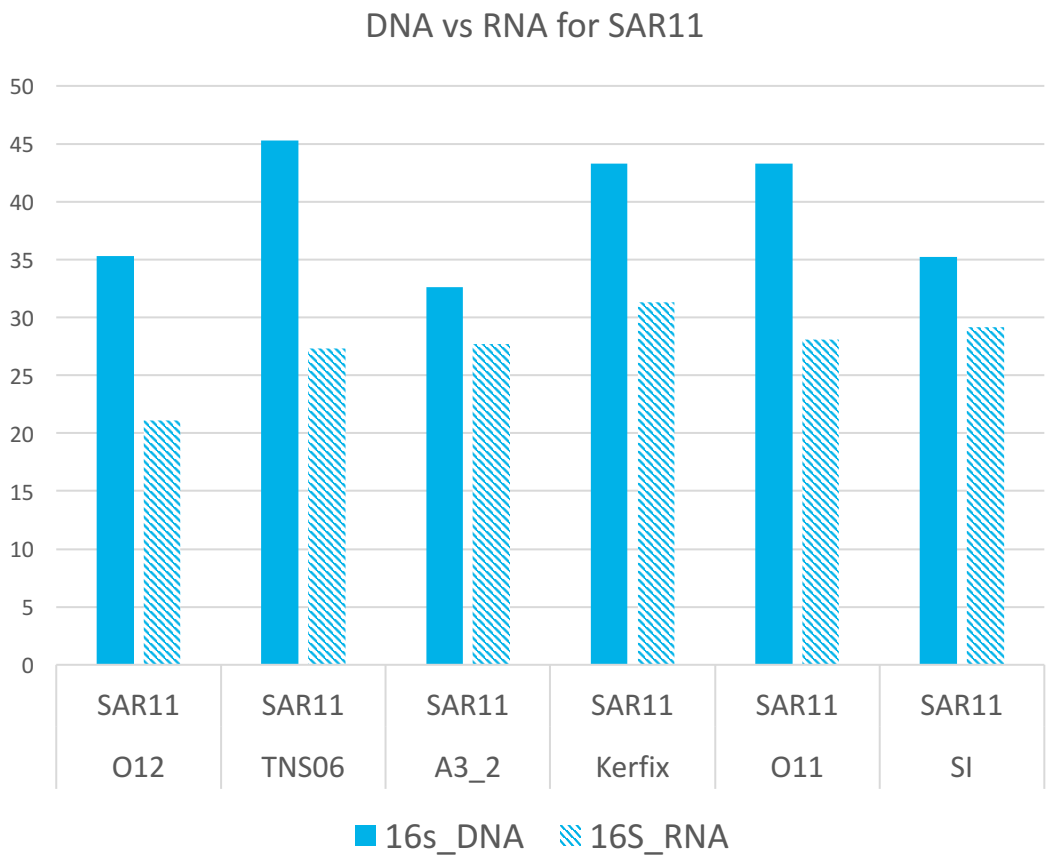
→Many SAR11 Transporters such as periplasmic substrate-binding proteins for phosphate, amino acids, sugars and spermidine highly abundant

→Supporting the view of extreme competition for multiple nutrients

→Several outer membrane receptor proteins, mostly for Fe transport

Combining datasets





→SAR11 & SUP05 at least 50% presence in Metaproteomes (TNS06) even up to 80% relative abundance (SI)

→Metagenomic and Metatranscriptomic data needed to interpret results for taxonomy

→Absence of SUP05 and *Gammaproteobacteria* abundance highlights importance on proteomics analysis for functional description of an ecosysteme

Thank you for your attention !