FRIAS

FRIAS
Lunch
Lecture
Series
2017/18

Quantitative vs. qualitative methods across sciences: mutual reinforcement, (un)happy co-existence, or source of schisms?

UNI FREIBURG

FRIAS

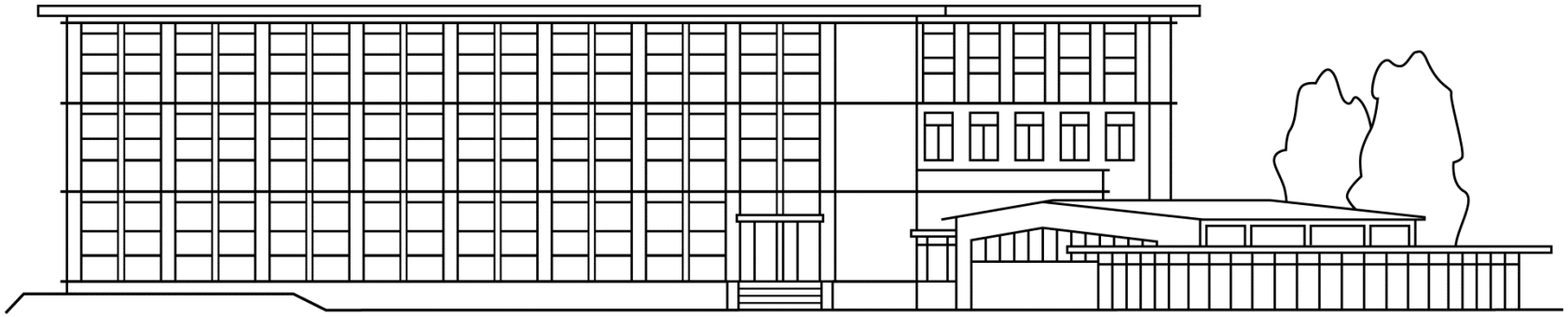
Freiburg Institute for Advanced Studies

Life in the marine realm – counting microbes…and what else?

Prof. Dr. Wolfgang Hess

University of Freiburg
Genetics and Experimental Bioinformatics

pdf-download: https://www.frias.uni-freiburg.de/de/veranstaltungen/lunch-lectures/lunch-lecture-hess



UNI FREIBURG

Quantitative vs. qualitative methods across sciences: mutual reinforcement, (un)happy co-existence, or source of schisms?

At a time when in most sciences (including the humanities and social sciences) quantitative methods have come to play a central role, it should be explored which role qualitative methods (still) play in different disciplines, in terms of research questions, trends and schools of research, the publication of research results and, not least, in the training of Master and PhD students.

Questions to be addressed include the following:

What is it that qualitative methods can do that quantitative methods cannot, and vice versa?

What does *qualitative* and *quantitative* mean concerning the methodologies of different academic disciplines in the first place?

……

Quantitative approaches have always played a role in the life sciences

Variazioni e fluttuazioni del numero d'individui…. *Vito Volterra*

PARTE PRIMA

Associazione biologica di du

§ 1. – DUE SPECIE CHE SI DISPU
STESSO NUTRIMENT

1. Supponiamo di avere due speci
stesso ambiente: i numeri degli individ
N_1 e N_2 e siano ε_1 e ε_2 i valori che avre
ficienti di accrescimento se il nutrime
in quantità sempre tale da soddisfare p
voracità. Avremo

$$\frac{dN_1}{dt} = \varepsilon_1 N_1 \quad , \quad \frac{dN_2}{dt} = \varepsilon_2 N_2$$

Volterra, V. (1926). "Variazioni e fluttuazioni del

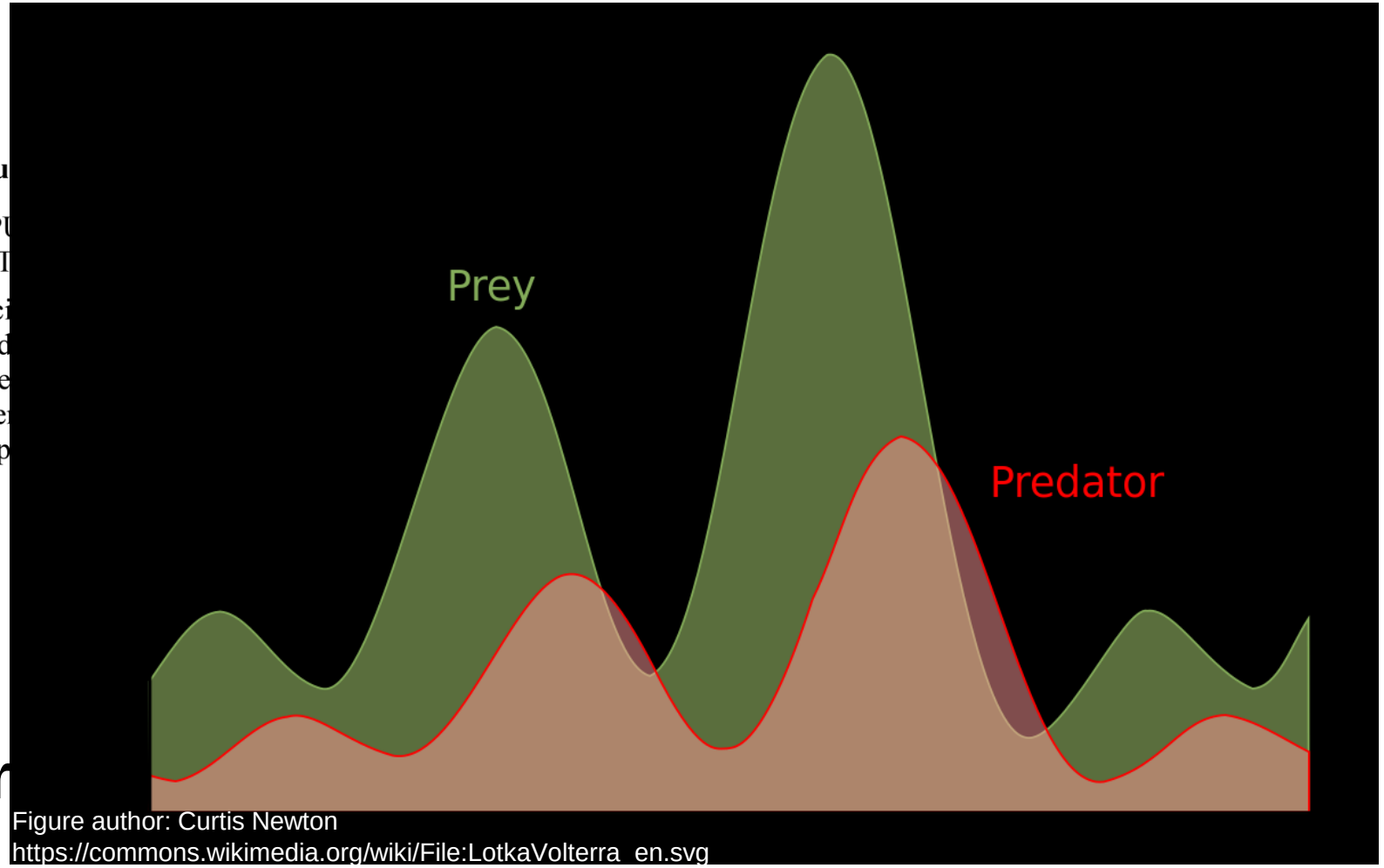
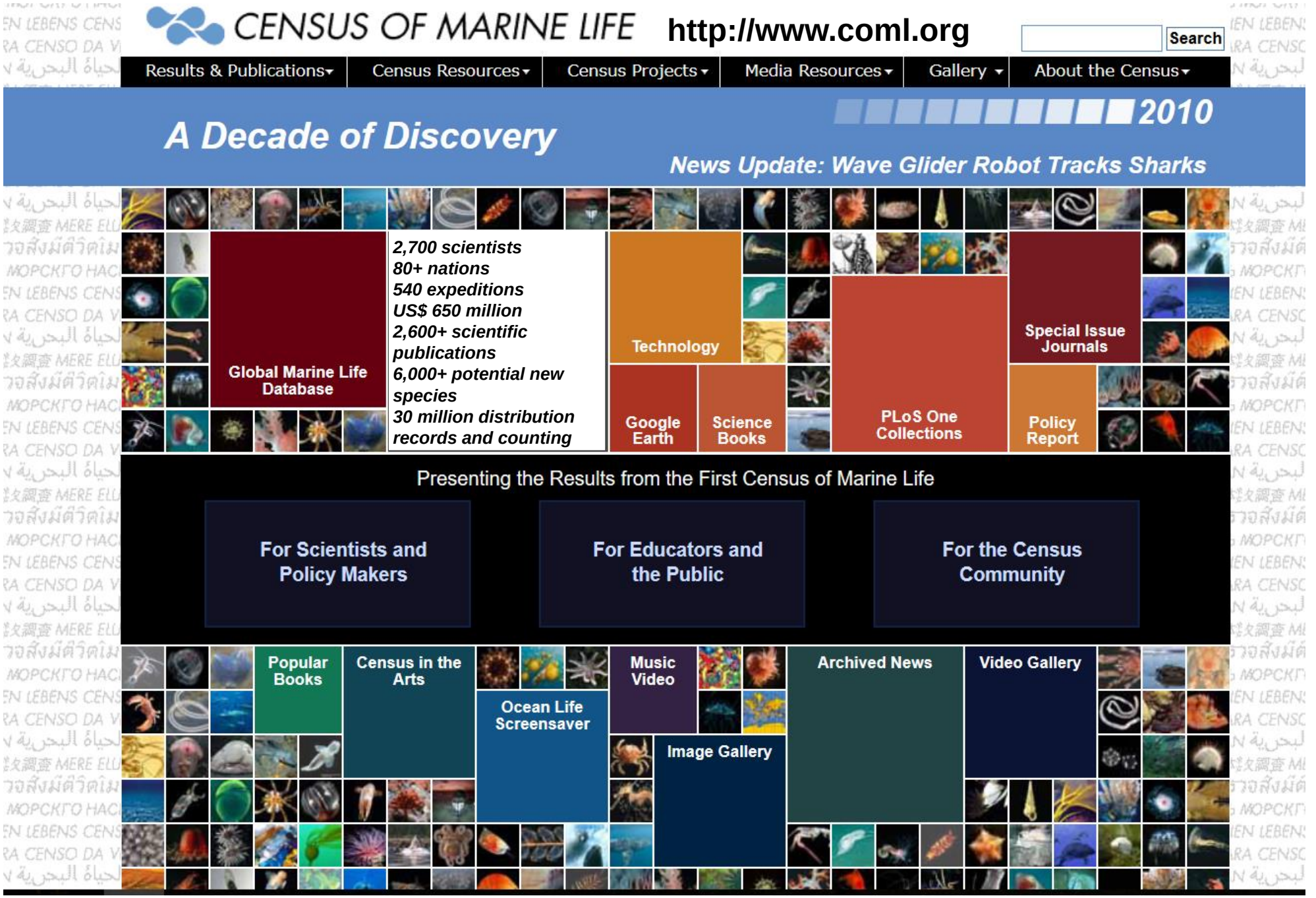


Figure author: Curtis Newton
https://commons.wikimedia.org/wiki/File:LotkaVolterra_en.svg

Lotka–Volterra model predator

Today, we are „counting“ all forms of life... and not only that



http://www.coml.org/investigating/identifying/traditional_identification_methods.html

CENSUS OF MARINE LIFE **Investigating Marine Life**

Search

Investigating Marine Life

- Exploring
- Collecting Organisms
- Observing and Counting
- Measuring Physical Properties
- Studying Movement
- Identifying
  - Molecular Techniques
  - Traditional Identification Methods
- Photo Gallery
- About
- Glossary
- SCOR Tech Panel
- Site Map

Traditional Identification Methods

The traditional procedure for identifying organisms involves comparing the physical characteristics of a collected specimen with the characteristics for a known species. There are numerous taxonomic books that describe the physical appearance, both externally and internally, of millions of species, as well as what is known about their habitats and general biology. Census researchers study collected specimens, often through microscopes, to distinguish features such as the number of tentacles on a jellyfish or the length of spines on a deep-sea anglerfish, and match what they find with existing species descriptions. The Census of Marine Life has collected a large number of species previously unknown to science and there is a backlog of species to be described by taxonomists. Another relevant effort by the Census of Marine Life is the World Registry of Marine Species (WoRMS), which is attempting to provide an authoritative and comprehensive list of names of marine organisms, including information on valid species names, synonyms and vernacular names. While highest priority goes to valid names, other names in use are included so that this register can serve as a guide to interpret taxonomic literature.

The image above contains species of worms. The diversity of marine species demonstrates the difficulty of taxonomic identification. (Natural Geography in Shore Areas - NaGISA. Tetsuya Kato)

In order to „count“ we must know „who is who“

http://www.coml.org/investigating/identifying/molecular_techniques.html

CENSUS OF MARINE LIFE

Investigating Marine Life

Search

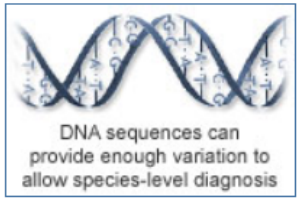
Investigating Marine Life

- Exploring
- Collecting Organisms
- Observing and Counting
- Measuring Physical Properties
- Studying Movement
- Identifying
  - Molecular Techniques
  - Traditional Identification Methods
- Photo Gallery
- About
- Glossary
- SCOR Tech Panel
- Site Map

Molecular Techniques

A new field of research has allowed for more accurate naming of species. Molecular techniques utilize the unique genetic code found in an individual organism's DNA as the best way to determine the species of a specimen. This is a more accurate way to identify organisms than traditional methods as it is independent of an individual taxonomist's opinion. It doesn't rely on researchers trying to categorize physical features that can sometimes be unclear. Additionally, scientists use molecular techniques to determine how different species relate to each other, allowing taxonomists to build a more thorough and accurate "tree of life" than had previously been possible.

One important recent advance in the field of molecular techniques is the development of DNA bar-coding. This approach characterizes a small segment of an organism's DNA to connect to its species name. This technology affords Census scientists an advantage when trying to identify large numbers of collected organisms and offers the possibility in the future of quickly identifying organisms when taxonomists and others with expertise in identifying organisms by sight are not available.



DNA sequences can provide enough variation to allow species-level diagnosis

A drawing of a small sequence of DNA showing the different paired bases that allow for identification of species. (Biodiversity Institute of Ontario)

The identification of organisms by DNA analysis is a rationale and accurate way to find out „who is who“

My thesis is, quantitative methods have been there all the time but they now percolate life sciences like never before.

A major driving force for that has been the broad introduction of DNA sequencing techniques.

In the following, I will use **Prochlorococcus** as an example to illustrate this development over the last years.

FRIAS

„*Prochlorococcus marinus*“

A novel free-living prochlorophyte abundant in the oceanic euphotic zone

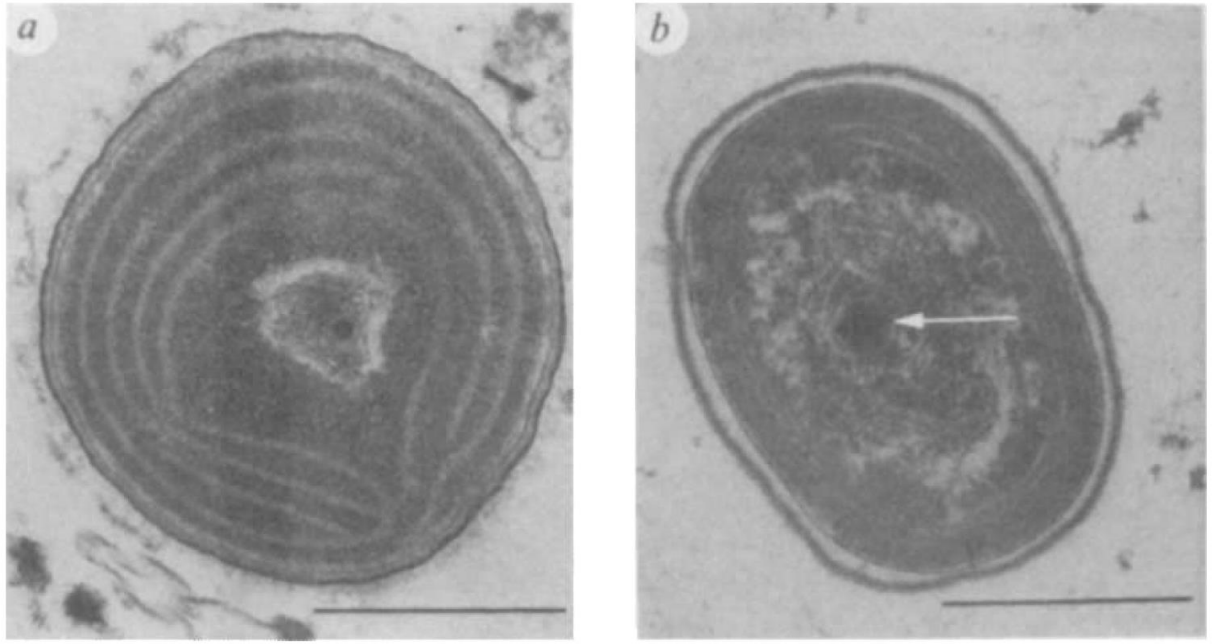
Sallie W. Chisholm, Robert J. Olson*, Erik R. Zettler*, Ralf Goericke†, John B. Waterbury* & Nicholas A. Welschmeyer†

48-425 Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
* Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA
† Harvard University, Cambridge, Massachusetts 02138, USA

Qualitative analyses are important, and remain so.

Example: Identification and description of a new organism – here of *Prochlorococcus* when most people thought of the oceans as not having much microscopic life at all.

LETTERS TO NATURE NATURE VOL. 334 28 JULY 1988



UNI FREIBURG

FRIAS

Prochlorococcus in the real world: The Red Sea

UNI FREIBURG

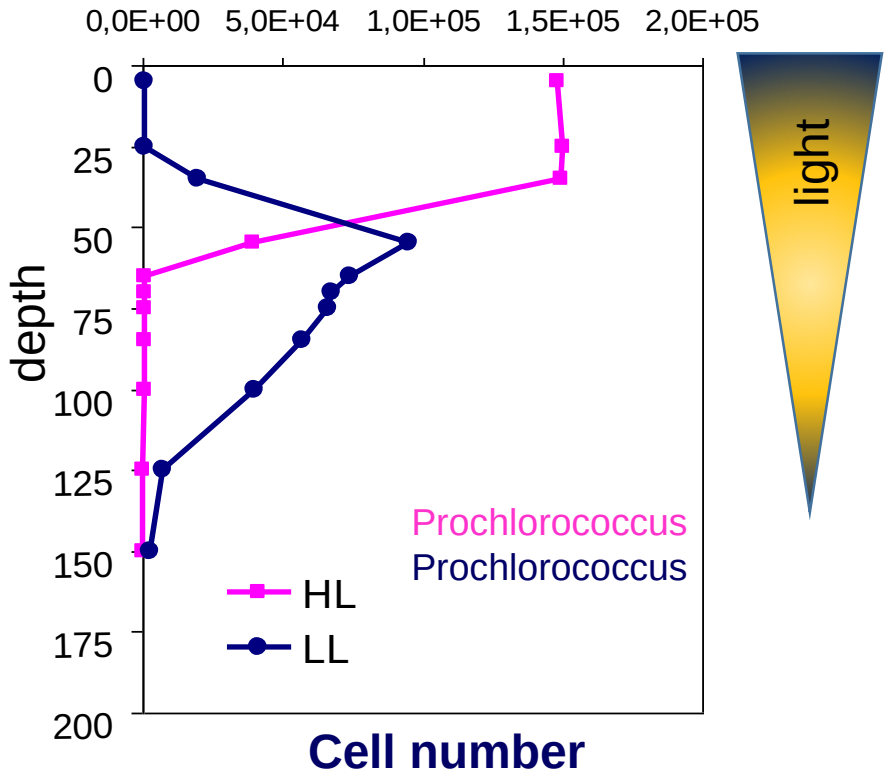
FRIAS

UNI FREIBURG

FRIAS

Cell counts of *Prochlorococcus* in the Northern Red Sea

smallest (0.5 - 0.7 µm) and most abundant photosynthetic organism in the ocean; 0-200 m (>2000-0.5 µE)



Gulf of Aqaba 29° N 34° E, Red Sea

Estimated Total: 10^{25} cells on Earth

Steglich C., Post A.F., Hess W.R. (2003) Analysis of natural populations of *Prochlorococcus* spp. in the northern Red Sea using phycoerythrin gene sequences. Environ. Microbiol. *5,* 681-690.
Holtzendorff J., Marie D., Post A.F., Partensky F., Rivlin A., Hess W.R. (2002): Synchronized expression of *ftsZ* in natural *Prochlorococcus* populations of the Red Sea. Environ. Microbiol**.** *4,* 644-653.
Pfreundt U., Miller D., Adusumilli L., Stambler N., Berman-Frank I., Hess W. R. (2014) Depth dependent metatranscriptomes of the marine pico-/nanoplanktonic communities in the Gulf of Aqaba/Eilat during seasonal deep mixing. Marine Genomics*18,* 93–95.

UNI FREIBURG

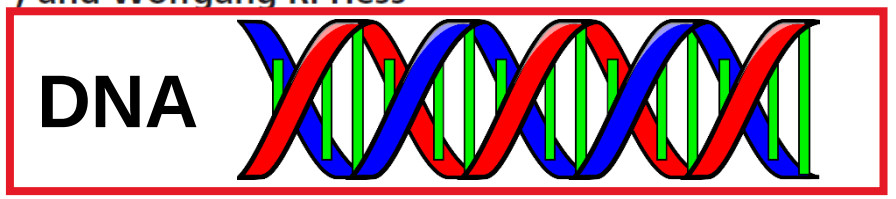
So, we can count the cells and estimate population sizes.

But why are they so abundant? What is it that makes *Prochlorococcus* so ‚successful'?

So, we thought we sequence its genome and will learn that.

Genome sequence of the cyanobacterium *Prochlorococcus marinus* SS120, a nearly minimal oxyphototrophic genome

Alexis Dufresne*, Marcel Salanoubat†, Frédéric Partensky*‡, François Artiguenave†, Ilka M. Axmann§, Valérie Barbe†, Simone Duprat†, Michael Y. Galperin¶, Eugene V. Koonin¶, Florence Le Gall*, Kira S. Makarova¶, Martin Ostrowski‖, Sophie Oztas†, Catherine Robert†, Igor B. Rogozin¶, David J. Scanlan‖, Nicole Tandeau de Marsac**, Jean Weissenbach†, Patrick Wincker†, Yuri I. Wolf¶, and Wolfgang R. Hess§††



Counting genes:

Only 1,884 genes, that is a small number, right?

This makes ~10% of the human protein-coding gene count.

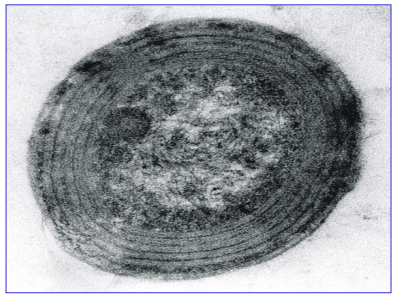
But *Prochlorococcus* is also unicellular and much smaller, tiny actually.

RNA

protein

Functions

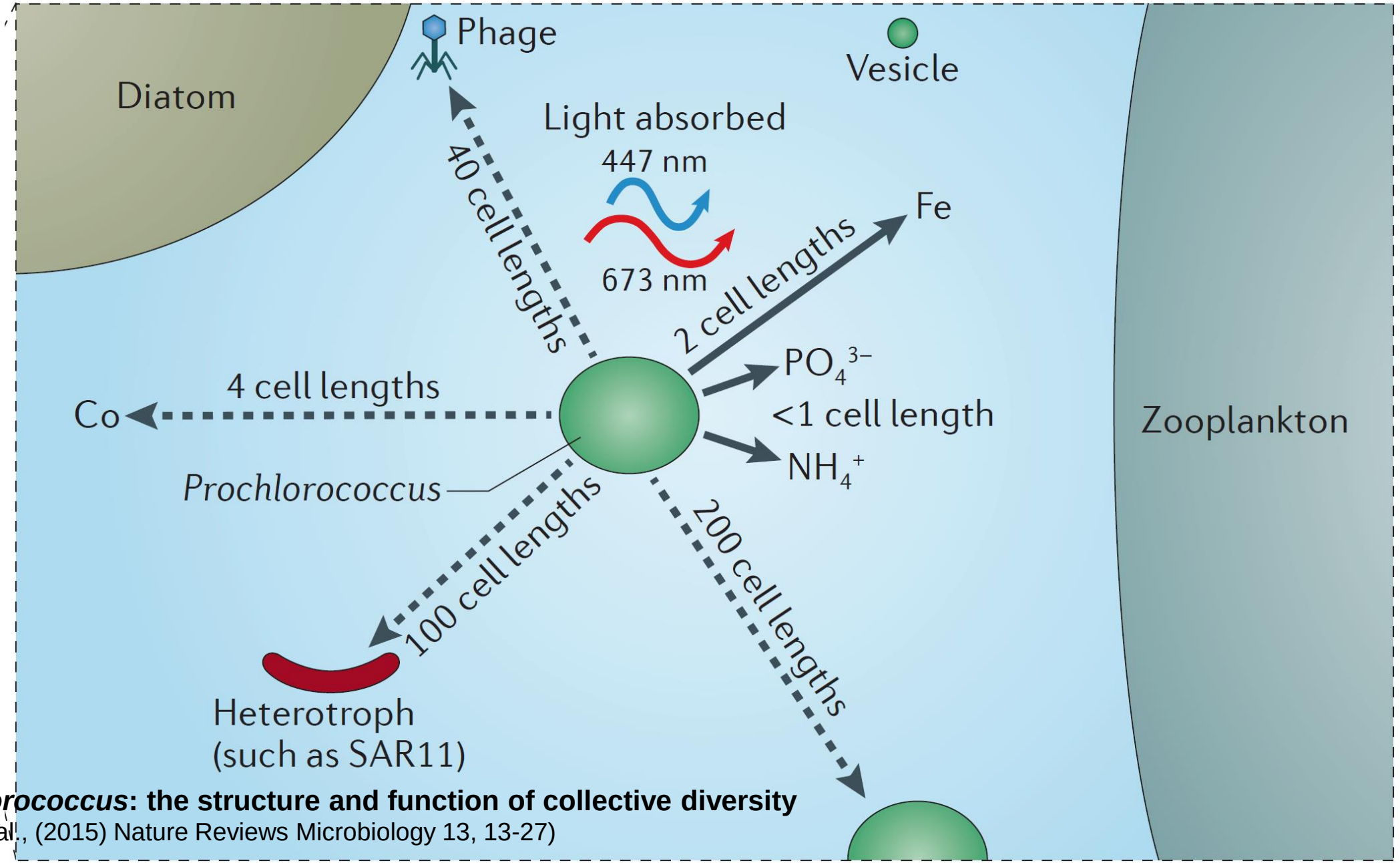
Prochlorococcus



0.4 µm

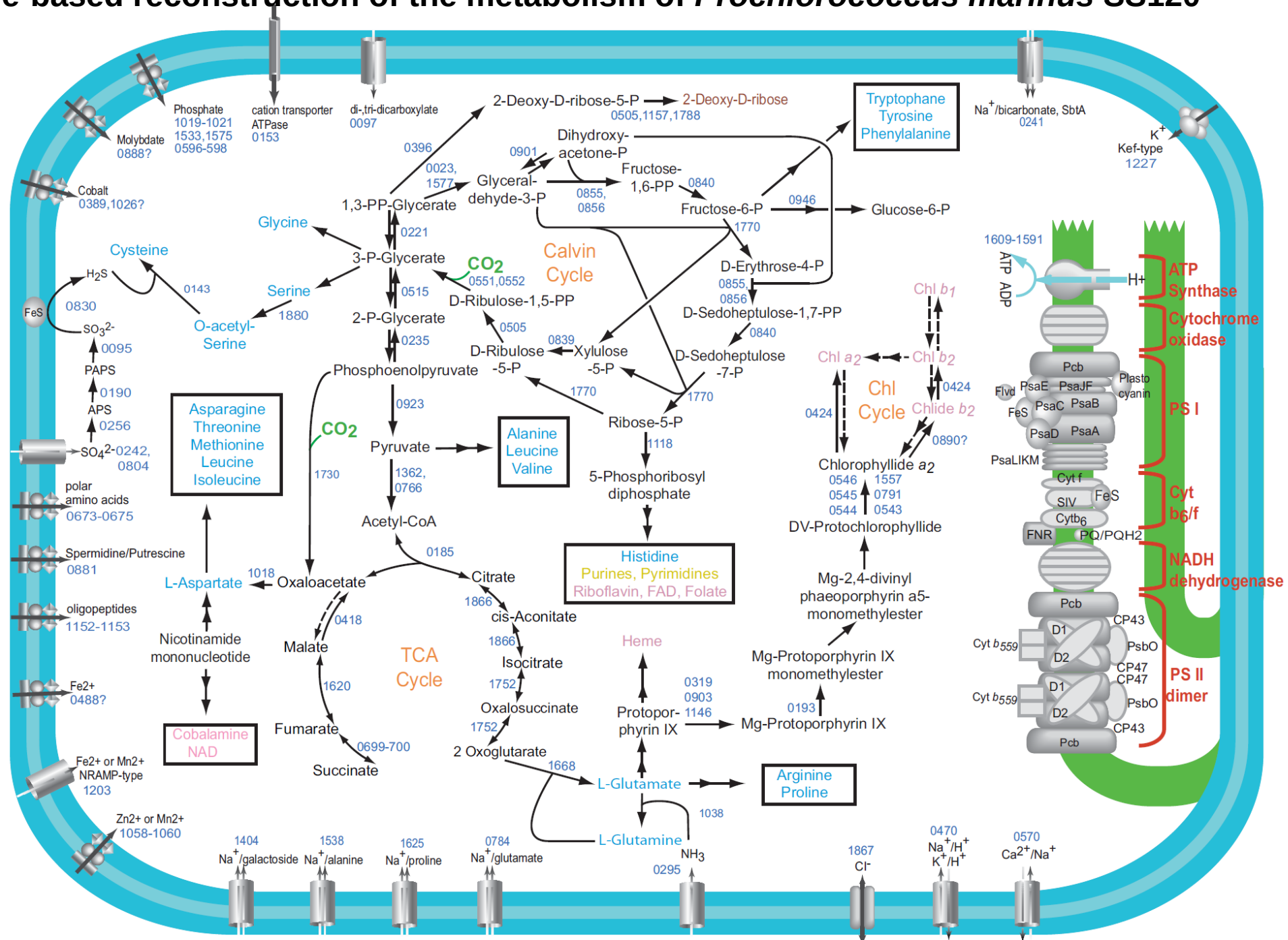
0.6 µm

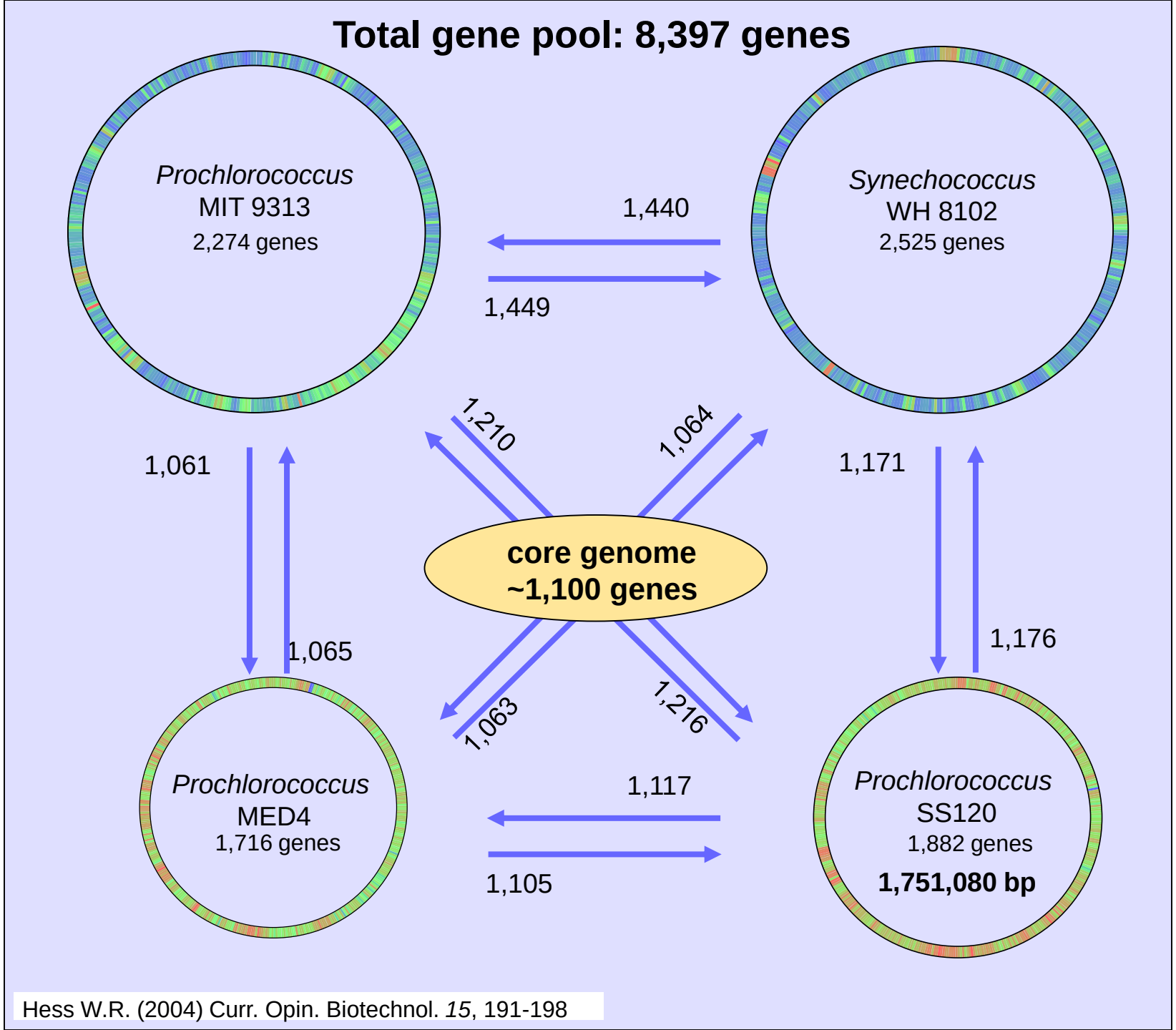
Dimensions: What does „small“ mean for a unicellular microorganism?



***Prochlorococcus*: the structure and function of collective diversity**
(Biller et al., (2015) Nature Reviews Microbiology 13, 13-27)

Genome-based reconstruction of the metabolism of *Prochlorococcus marinus* SS120

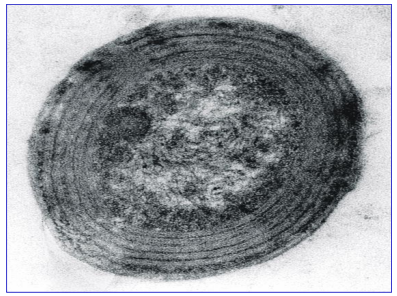




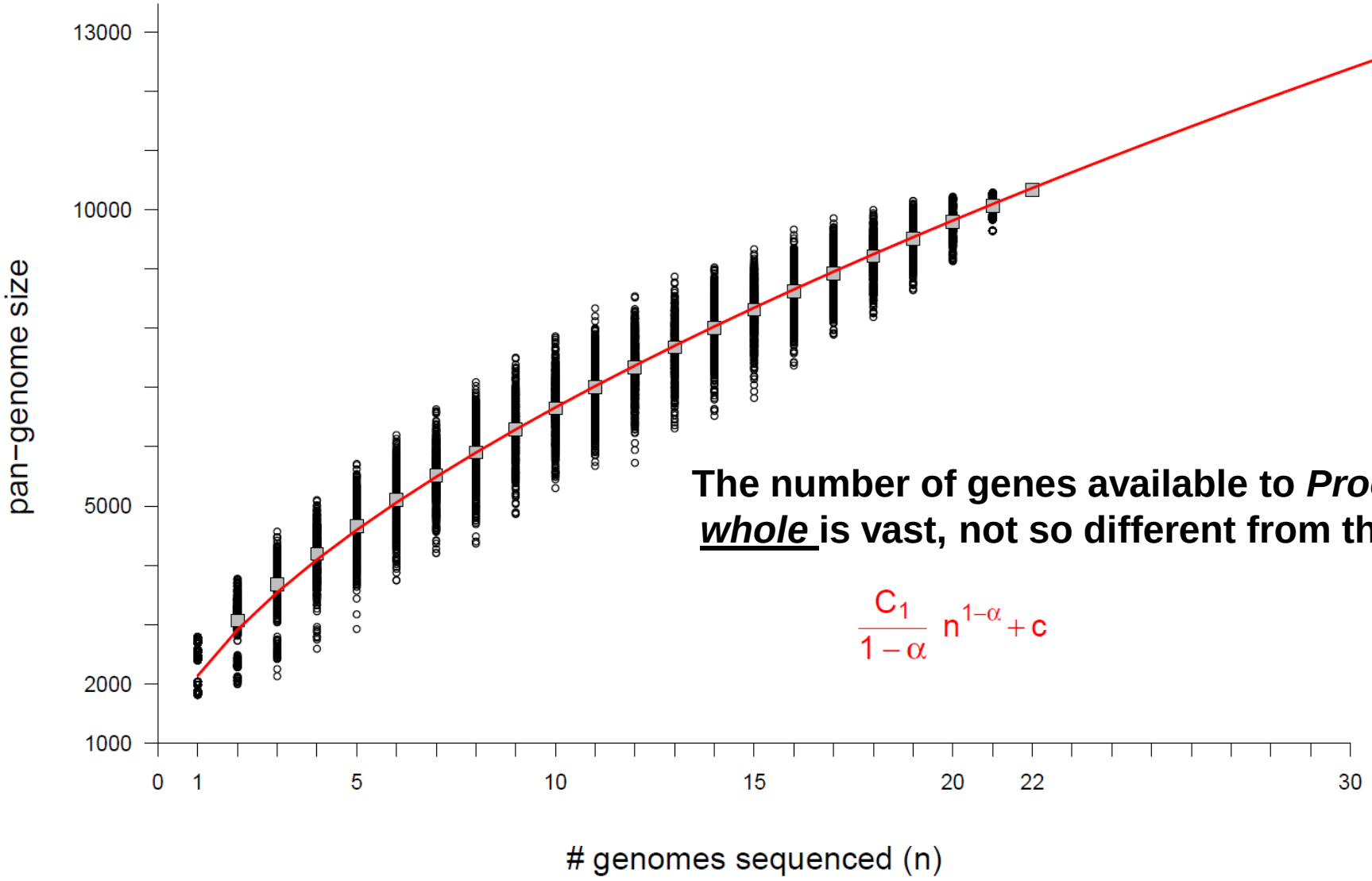
Hess W.R. (2004) Curr. Opin. Biotechnol. *15*, 191-198

Surprise:
Additional genome sequences showed that different isolates had a quite divergent gene complement

Prochlorococcus



How many new genes can we find?



The number of genes available to *Prochlorococcus* <u>*as a whole*</u> is vast, not so different from the human genome.

$$\frac{C_1}{1-\alpha}\, n^{1-\alpha} + c$$

Baumdicker F., Hess W.R., Pfaffelhuber P. (2010): The diversity of a distributed genome in bacterial populations. Ann. Appl. Probability (2010) *20,* 1567–1606

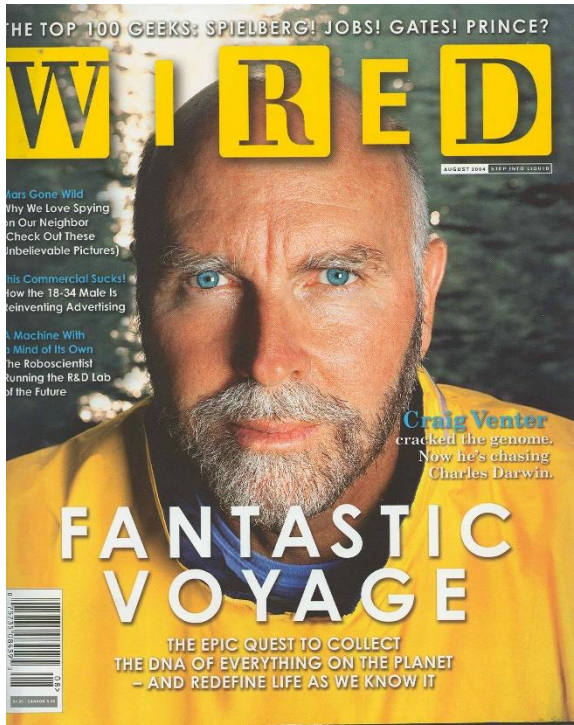
FRIAS

Are these microbes genetically really so different from each other?

What would we learn if we sequenced many more than 3 or 4 isolates?

Or, why not sequence all cells that are out there?

Craig Venter was one of the first people asking this question



UNI FREIBURG

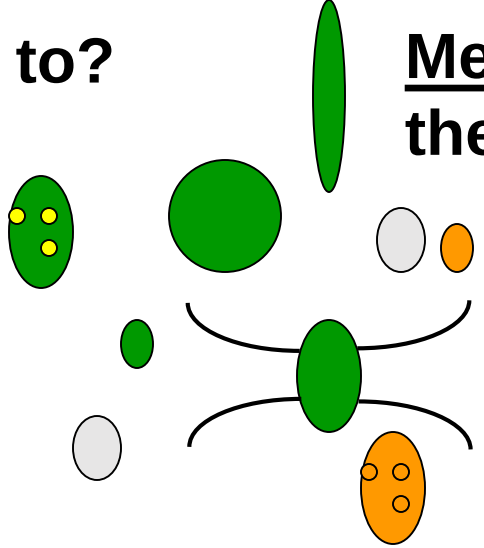
FRIAS

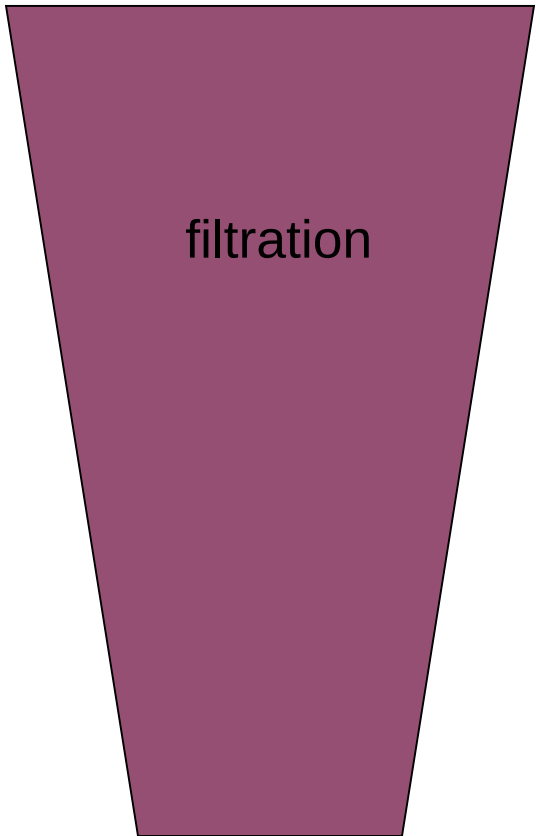
Where to go? Bermuda

UNI FREIBURG

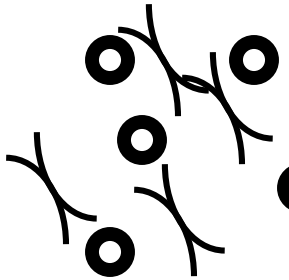
How to?

Metagenomics: Direct extraction of DNA from the environment & sequencing





cell concentration



cell lysis with detergents,
inactivation of DNases,
removal of proteins, RNA and debris,
precipitation of DNA with ethanol

Next Generation DNA sequencing

Sciencexpress Research Article

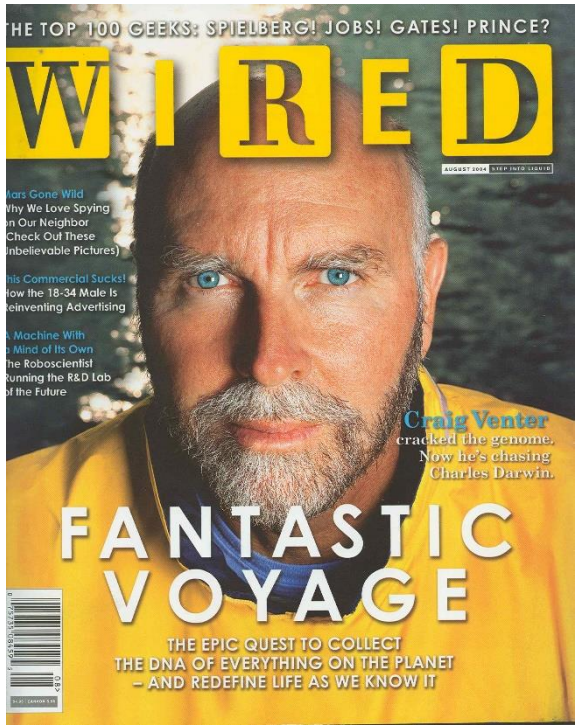
Environmental Genome Shotgun Sequencing of the Sargasso Sea

J. Craig Venter,[1]* Karin Remington,[1] John F. Heidelberg,[3] Aaron L. Halpern,[2] Doug Rusch,[2] Jonathan A. Eisen,[3] Dongying Wu,[3] Ian Paulsen,[3] Karen E. Nelson,[3] William Nelson,[3] Derrick E. Fouts,[3] Samuel Levy,[2] Anthony H. Knap,[6] Michael W. Lomas,[6] Ken Nealson,[5] Owen White,[3] Jeremy Peterson,[3] Jeff Hoffman,[1] Rachel Parsons,[6] Holly Baden-Tillson,[1] Cynthia Pfannkoch,[1] Yu-Hui Rogers,[4] Hamilton O. Smith[1]

[1]The Institute for Biological Energy Alternatives, [2]The Center for the Advancement of Genomics, 1901 Research Boulevard, Rockville, MD 20850, USA. [3]The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA. [4]The J. Craig Venter Science Foundation Joint Technology Center, 5 Research Place, Rockville, MD 20850, USA. [5]University of Southern California, 223 Science Hall, Los Angeles, CA 90089–0740, USA. [6]Bermuda Biological Station for Research, Inc., 17 Biological Lane, St George GE 01, Bermuda.

*To whom correspondence should be addressed. E-mail: jcventer@tcag.org

Sciencexpress / www.sciencexpress.org / 4 March 2004 / Page 1/ 10.1126/science.1093857



Science. 2004 Apr 2;304(5667):66-74

What did they find in 2004?

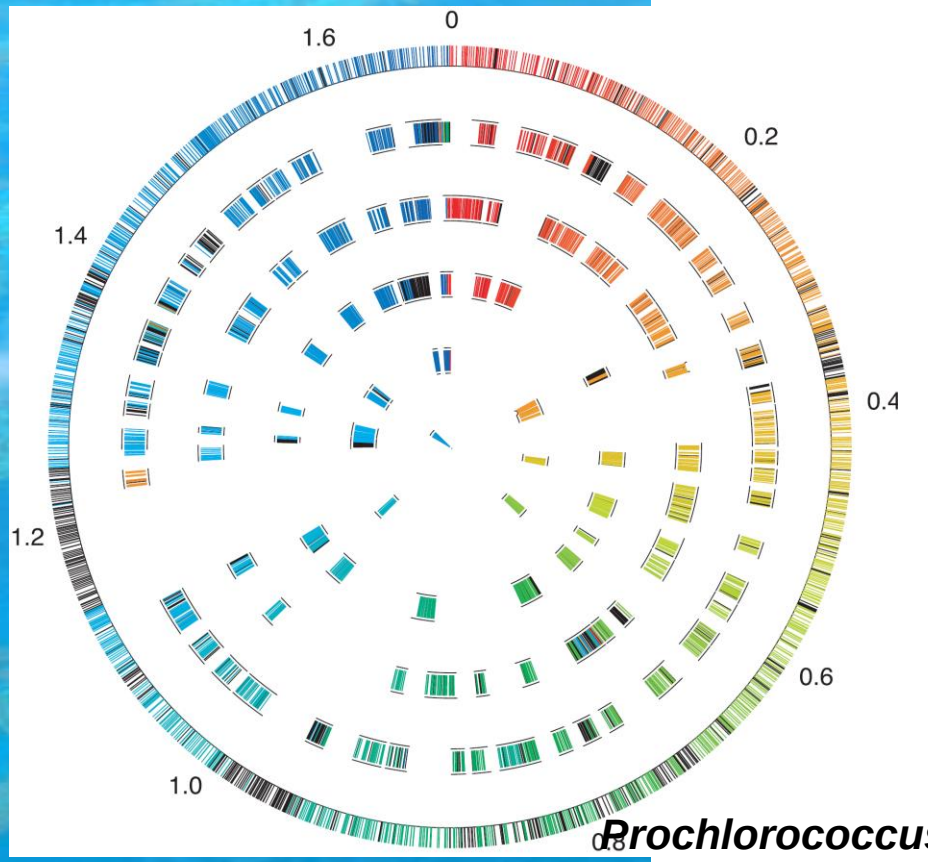
Environmental Metagenomics
- from 900 L Sea water <3 µm:

1.045 billion nucleotides

1.2 million new genes

1,800 different species

148 unknown species



Prochlorococcus

Estimation: ~48,000 species were in a single sample of seawater

→ Biodiversity much higher than expected

Science. 2004 Apr 2;304(5667):66-74

Underwater reefs close to Bermuda

FRIAS

This has led to a new concept: Relationships between the core-, Pan-, and metagenome

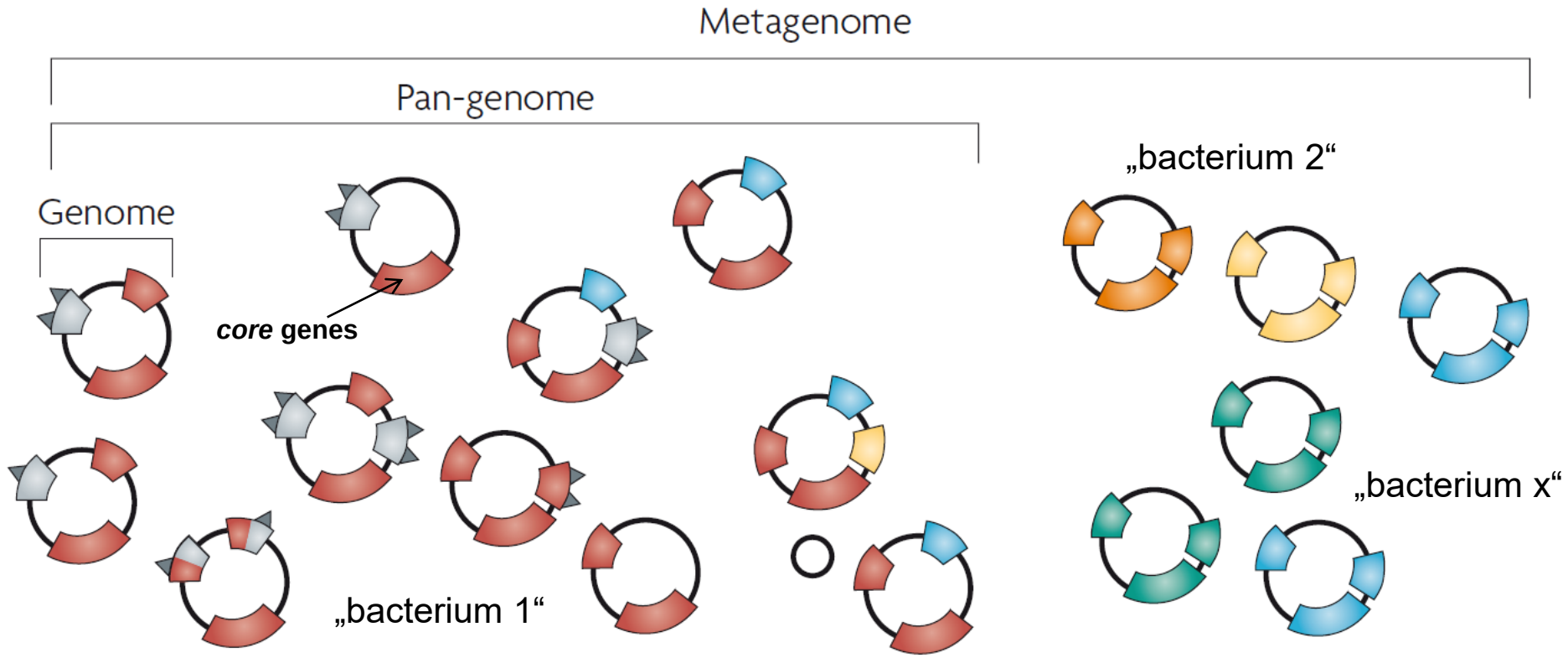


Illustration: Medini et al., Nature Reviews Microbiology 6, 419–430 (2008)

UNI FREIBURG

FRIAS

These discoveries had a tremendous impact.

What is a ‚species' in view of these findings?

ARTICLE

doi:10.1038/nature11711

Genomic variation landscape of the human gut microbiome

Siegfried Schloissnig[1]*, Manimozhiyan Arumugam[1]*, Shinichi Sunagawa[1]*, Makedonka Mitreva[2], Julien Tap[1], Ana Zhu[1], Alison Waller[1], Daniel R. Mende[1], Jens Roat Kultima[1], John Martin[2], Karthik Kota[2], Shamil R. Sunyaev[3], George M. Weinstock[2] & Peer Bork[1,4]

Whereas large-scale efforts have rapidly advanced the understanding and practical impact of human
the practical impact of variation is largely unexplored in the human microbiome. We therefore develo
metagenomic variation analysis and applied it to 252 faecal metagenomes of 207 individuals from
America. Using 7.4 billion reads aligned to 101 reference species, we detected 10.3 million single n
phisms (SNPs), 107,991 short insertions/deletions, and 1,051 structural variants. The average ratio of
synonymous polymorphism rates of 0.11 was more variable between gut microbial species than a
Subjects sampled at varying time intervals exhibited individuality and temporal stability of SNP variation patterns, despite considerable composition changes of their gut microbiota. This indicates that individual-specific strains are not easily replaced and that an individual might have a unique metagenomic genotype, which may be exploitable for personalized diet or drug intake.

Create a global microbiome effort

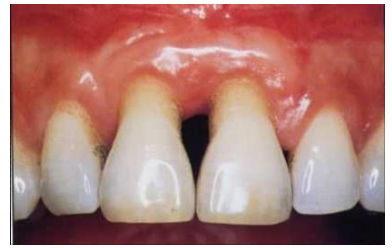
Understanding how microbes affect health and the biosphere requires an international initiative, argue **Nicole Dubilier**, **Margaret McFall-Ngai** and **Liping Zhao**.

Opinion

CellPress

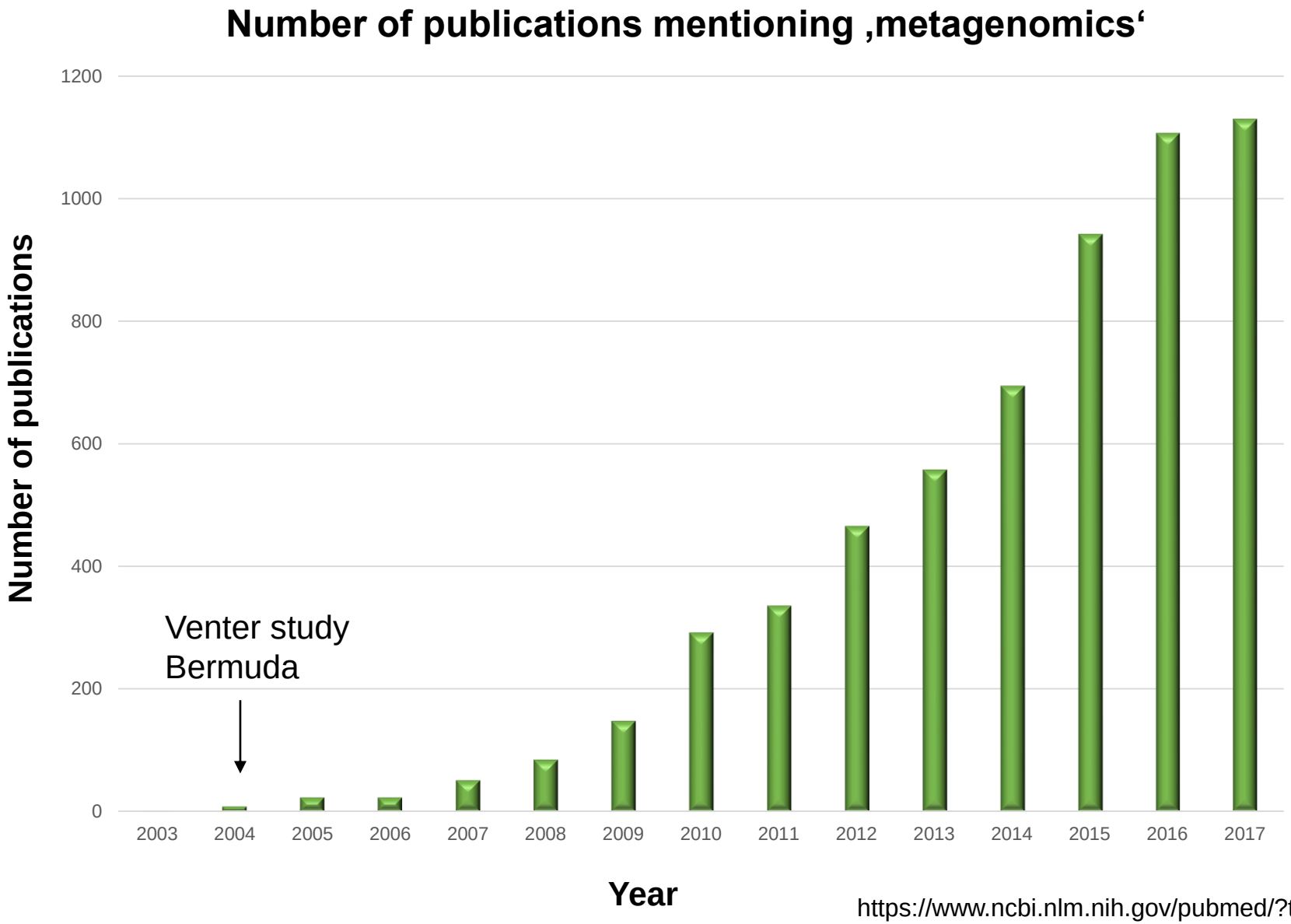
Solving the etiology of dental caries

Aurea Simón-Soro and Alex Mira



UNI FREIBURG

These discoveries had a tremendous impact.



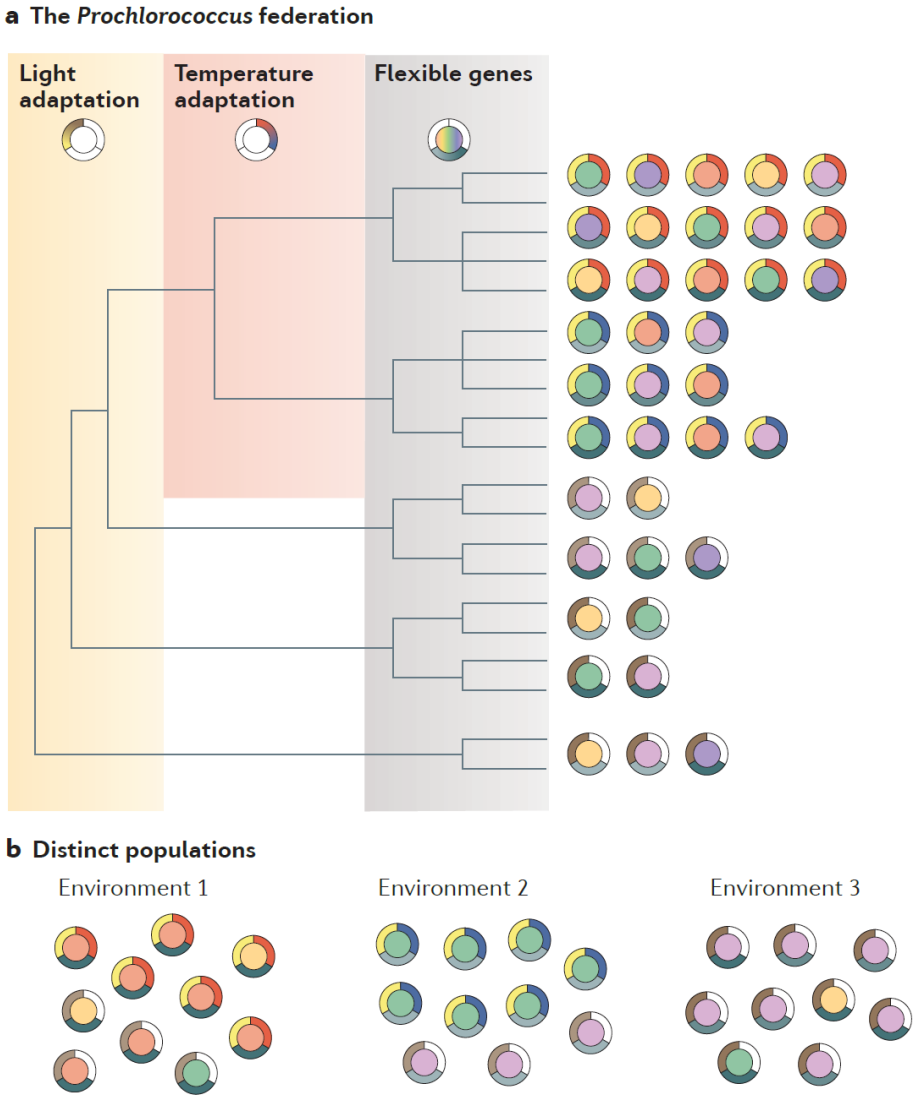
https://www.ncbi.nlm.nih.gov/pubmed/?term=metagenomics

10 years later:

Single-Cell Genomics Reveals Hundreds of Coexisting Subpopulations in Wild *Prochlorococcus*

Nadav Kashtan,[1*] Sara E. Roggensack,[1] Sébastien Rodrigue,[1,2] Jessie W. Thompson,[1] Steven J. Biller,[1] Allison Coe,[1] Huiming Ding,[1,3] Pekka Marttinen,[4] Rex R. Malmstrom,[5] Roman Stocker,[1] Michael J. Follows,[6] Ramunas Stepanauskas,[7] Sallie W. Chisholm[1,3*]

25 APRIL 2014 VOL 344 SCIENCE



The *Prochlorococcus* federation
(Biller et al., (2015) Nature Reviews Microbiology 13, 13-27)

FRIAS

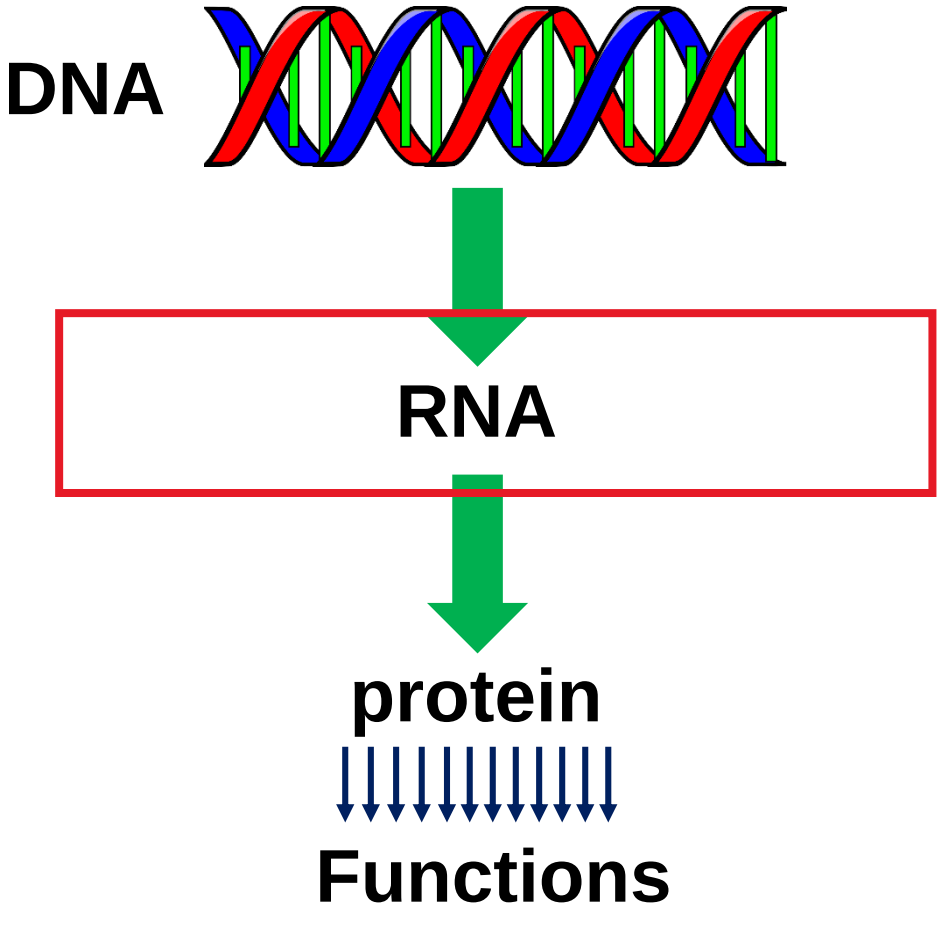
Metagenomic analyses tell us for a certain biotope, environment, potentially a whole ecosystem:

- „who is there“ (qualitative)
- how many are there (quantitative)
- what could they possibly be doing, based on genomic reconstruction (qualitative)
- what could be the consequences for the respective environment (what ever it is..... ocean, biofilms, human gut, dental plaques.... (qualitative and quantitative)

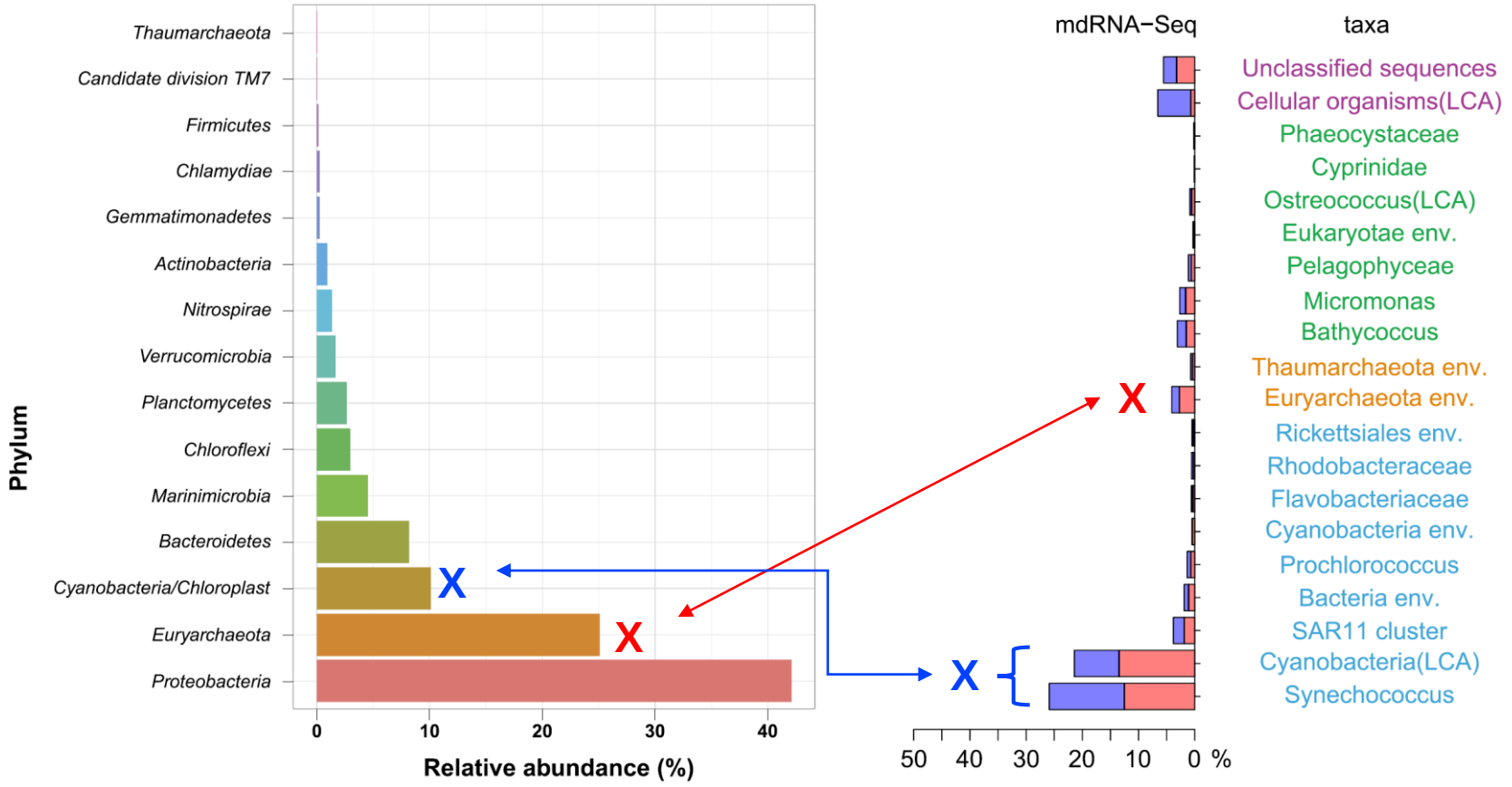
.....but such analyses don‘t tell us whether this is actually happening, answering „what are they doing“

UNI FREIBURG

Therefore, metatranscriptomics was developed, in which, after conversion to cDNA, the RNA is sequenced – telling us about the specific activity of specific microbes and quantifying this activity



Metatranscriptomics tells us about the activity of certain types of organisms – which can differ substantially from their sheer number



In this example (from the Red Sea), Cyanobacteria (**X**) are much more active than expected due to their abundance. Euryarchaeota (**X**) much less.

SCIENTIFIC REPORTS

OPEN

mdRNA-Seq analysis of marine microbial communities from the northern Red Sea

Shengwei Hou[1], Ulrike Pfreundt[1], Dan Miller[2], Ilana Berman-Frank[2] & Wolfgang R. Hess[1]

Received: 25 May 2016
Accepted: 28 September 2016

Metatranscriptomics has become a popular tool, too.

OPEN ACCESS Freely available online

PLoS one

Detection of Large Numbers of Novel Sequences in the Metatranscriptomes of Complex Marine Microbial Communities

Jack A. Gilbert[1]*, Dawn Field[2], Ying Huang[3], Rob

jove Journal of Visualized Experiments

Video Article

Analyzing Gene Expres… Environmental Transcri…

Rachel S. Poretsky, Scott Gifford, Joha…
Department of Marine Sciences, Univers…

Correspondence to: Mary Ann Moran at

URL: http://www.jove.com/index/Details.
DOI: 10.3791/1086
Citation: Poretsky R.S., Gifford S., Rinta-Kanto J., Transcriptomics. JoVE. 24. http://www.jove.com/in

The ISME Journal (2013) 7, 2315–2329
© 2013 International Society for Microbial Ecology All rights reserved 1751-7362/13
www.nature.com/ismej

ORIGINAL ARTICLE

Transcriptional response of bathypelagic marine bacterioplankton to the Deepwater Horizon oil spill

Adam R Rivers[1], Shalabh Sharma[1], Susannah G Tringe[2], Jeffrey Martin[2], Samantha B Joye[1] and Mary Ann Moran[1]
[1]*Department of Marine Sciences, University of Georgia, Athens, GA, USA and* [2]*DOE Joint Genome Institute, Walnut Creek, CA, USA*

nature Vol 459|14 May 2009|doi:10.1038/nature08055

LETTERS

Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column

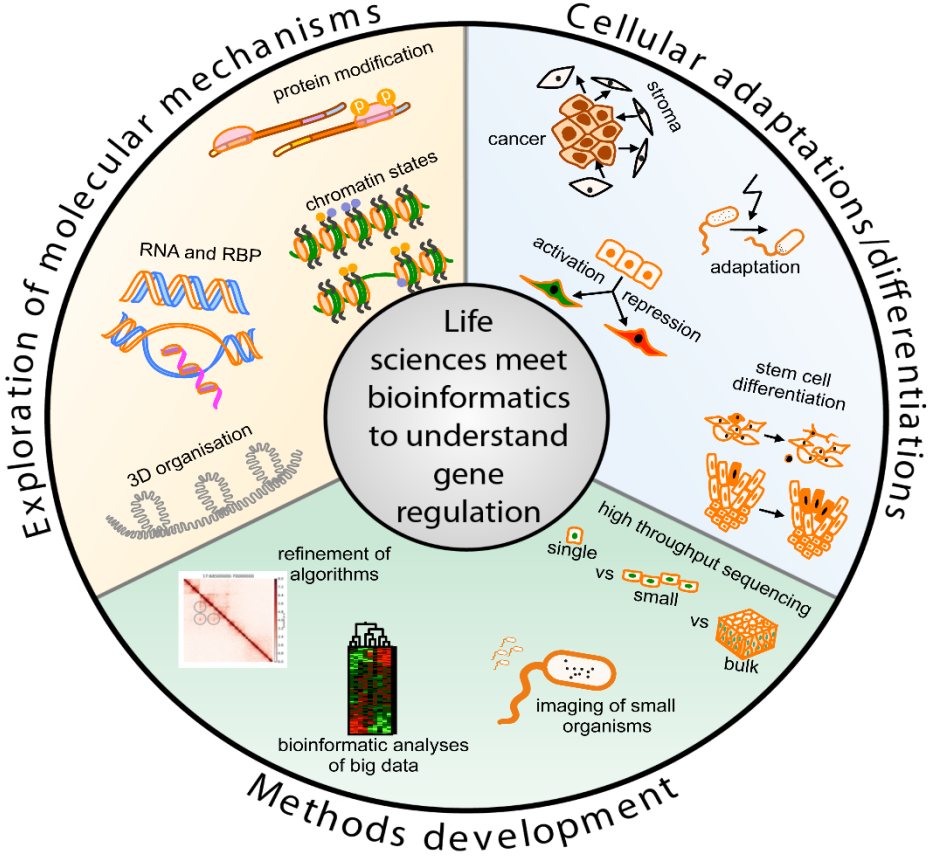
Yanmei Shi[1], Gene W. Tyson[1] & Edward F. DeLong[1,2]

FRIAS

Life scientists are used to apply qualitative and quantitative approaches alike, but the “big data” based techniques have been becoming crucial analytical tools

UNI FREIBURG

New DFG-funded program in Freiburg to meet the challenge



GRK 2344 started in Fall 2017

SUMMARY

- Life scientists use qualitative and quantitative approaches alike.
- Sequencing-based techniques are ‘big data’ and provide quantitative and qualitative information.
- Microbial populations are huge (exp. *Prochlorococcus*)
- Metagenomics and metatranscriptomics approaches are productive for characterizing large microbiomes (e.g., the surface or deep ocean), as well as smaller ‘ecosystems’, e.g., the human microbiome.
- Sequencing-based metagenomics and metatranscriptomics approaches lead to ‘extra big data’
- These tendencies need to be considered in our curricula, bioinformatics skills are crucial for life scientists in the 21st century
- Understanding the human microbiome is helpful for our health (exp. caries).
- More than 10,000 microbial species have been identified as living in human bodies thus far. Many of these cannot be cultivated!
- Microbial cells outnumber human cells in a human body 10:1! Yet, microbes account for only 1-3% of human body weight.
- The human microbiome contributes ~8 million different protein-coding genes.
- The microbiome contributes 99.7% of the genetic material that encodes protein in our bodies.

THANKS: HESS lab, all collaborators and funders.

FRIAS

UNI FREIBURG

Deutsche Forschungsgemeinschaft DFG

SEVENTH FRAMEWORK PROGRAMME

Bundesministerium für Bildung und Forschung

FRIAS

Next lecture on **January 25th, 2018**

Data collection and data analysis in the Social Sciences

Diana Panke (University of Freiburg)

All lectures are available as a video podcast at
www.frias.uni-freiburg.de/en/media-library

UNI FREIBURG